



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

(21) **PI0621906-3 A2**

(22) Data de Depósito: 21/07/2006
(43) Data da Publicação: 16/08/2011
(RPI 2119)



(51) *Int.Cl.:*
C12N 15/29 2006.01
C12N 15/82 2006.01
A01H 5/00 2006.01

(54) Título: **MOLÉCULA DE DNA QUE CODIFICA UM POLIPEPTÍDEO DE HISTIDINA QUINASE MUTANTE E SEU USO RELACIONADO, POLIPEPTÍDEO DE HISTIDINA QUINASE MUTANTE, BEM COMO MÉTODO DE PRODUÇÃO DE UMA PLANTA GENETICAMENTE MODIFICADA**

(73) Titular(es): Plant Bioscience Limited

(72) Inventor(es): Elena Simona Radutoiu, Jens Stougaard, Leila Tirichine, Lene H. Madsen, Niels Norgaard Sandal

(74) Procurador(es): Dannemann, Siemsen, Bigler & Ipanema Moreira

(86) Pedido Internacional: PCT DK2006050031 de 21/07/2006

(87) Publicação Internacional: WO 2008/009287 de 24/01/2008

(57) Resumo: MOLÉCULA DE DNA QUE CODIFICA UM POLIPEPTÍDEO DE HISTIDINA QUINASE MUTANTE E SEU USO RELACIONADO, POLIPEPTÍDEO DE HISTIDINA QUINASE MUTANTE, BEM COMO MÉTODO DE PRODUÇÃO DE UMA PLANTA GENETICAMENTE MODIFICADA. A presente invenção refere-se à formação de nódulos radiculares fixadores de nitrogênio em leguminosas que é induzida pela captação de moléculas sinalizadoras de oligossacarídeo de lipoquitina secretadas por bactérias *Rhizobium* compatíveis, que desencadeiam uma via simbiótica comum. A presente invenção provê um mutante de formação espontânea de nódulo (*snf2*), no qual a formação de nódulos simbióticos é espontânea, levando ao desenvolvimento nodular na ausência bem como na presença de bactérias *Rhizobium* e/ou sinais rizobiais exógenos. A invenção ainda provê uma sequência de DNA isolada que codifica uma histidina quinase mutante independente de citocinina, cuja atividade resulta neste fenótipo dominante de 'ganho de função' da formação espontânea de nódulo. Além disso, mostra-se que o gene *snf2* confere um fenótipo caracterizado pela organogênese regulada de nódulos espontâneos em plantas que têm um contexto genético de nodulação deficiente. Um gene da invenção, que confere este fenótipo de nodulação espontânea, tem utilidade para transferência e estabelecimento da capacidade de fixação de nitrogênio em plantas não-noduladas, e por meio disso reduz a dependência de fertilizantes nitrogenados das culturas de plantas não-noduladas.

72-0621906-3

PEI-020090072727

DATA - 29/07/09

Relatório Descritivo da Patente de Invenção para "**MOLÉCULA DE DNA QUE CODIFICA UM POLIPEPTÍDEO DE HISTIDINA QUINASE MUTANTE E SEU USO RELACIONADO, POLIPEPTÍDEO DE HISTIDINA QUINASE MUTANTE, BEM COMO MÉTODO DE PRODUÇÃO DE UMA PLANTA GENETICAMENTE MODIFICADA**".

Antecedentes da Invenção

O crescimento das culturas agrícolas é geralmente limitado pela disponibilidade de nitrogênio, e pelo menos 50% da demanda global é suprida pela aplicação de fertilizantes sintéticos na forma de amônia, nitrato ou ureia. Contudo, há uma necessidade crescente de explorar uma das fontes naturais de nitrogênio mais importantes para agricultura, ou seja, a fixação biológica de nitrogênio.

A fonte primária de fixação biológica de nitrogênio são *Rhizobium* ou *Rhizobia* spp e a actinobactéria *Frankia* spp, que são um pequeno grupo de procariotos que produz nitrogenases e forma associações endossimbióticas com plantas conferindo a capacidade de fixar nitrogênio. Embora muitas plantas possam associar-se com bactérias fixadoras de nitrogênio, somente algumas plantas, todas membros do Clado Rosídea I, formam uma associação endossimbiótica com a *Rhizobia* spp e *Frankia* spp, que são as únicas em que a maior parte do nitrogênio é transferida e assimilada pela planta hospedeira. As plantas da família *Leguminosae*, que inclui soja, feijão, ervilha, amendoim, grão-de-bico, feijão-caupi, lentilha, guandu, alfafa e trevo, são os membros agronomicamente mais importantes deste pequeno grupo de plantas fixadoras de nitrogênio. A fixação biológica de nitrogênio através da associação endossimbiótica reduz a necessidade de fertilizantes nitrogenados caros na cultura de leguminosas e é uma importante característica da agricultura sustentável. As leguminosas também podem utilizar o nitrogênio disponível no solo, tal que quando os níveis de nitrato do solo são altos, a formação de nódulo é suprimida e a planta substitui o metabolismo de nitrogênio por crescimento em nitrato (Wopereis *et al.*, 2000).

A simbiose *Rhizobium*-leguminosa inclui a interação de uma série de genes de planta e bacterianos em um processo complexo que leva à

iniciação e desenvolvimento de nódulos radiculares. A organogênese dos nódulos é desencadeada pelo microssimbionte rizobial, mas a planta leguminosa hospedeira codifica o programa de desenvolvimento responsável por construir os tecidos do nódulo e por regular o processo. O baixo peso molecular dos oligossacarídeos de lipoquitina (fatores Nod), sintetizados e secre-

5 tados pelo *Rizóbio*, são as principais moléculas sinalizadoras que desencadeiam este processo. O principal fator Nod secretado pelo *Mesorhizobium loti*, microssimbionte do *Lotus*, é uma N-acetilglicosamina pentamérica que porta um ácido cis-vacênico e um grupo carbamoíla no resíduo terminal não

10 reduzido junto a uma 4-O-acetilfucose no resíduo terminal reduzido. A captação do fator Nod no *Lotus* é mediada pelas quinases receptoras NFR1 e NFR5 (Radutoiu *et al.*, 2003 *Nature* 425: 585-592; Madsen *et al.*, 2003 *Nature* 425: 637-640), que em conjunto com um gene da quinase receptora LRR, *SymRK*, comunicam-se com uma via comum de transdução de sinal, com-

15 partilhada com a simbiose micorrizal (Oldroyd e Downie, 2004 *Mol. Cell Biology* 5: 566-576). Essa via comum é codificada por sete genes, *SymRK*, *Castor*, *Pollux*, *Nup133*, *CCaMK [Sym15]*, *Sym6* e *Sym24*. A análise de mutantes tem mostrado que receptor (es) NFR1/NFR5, *SymRK* que codifica a proteína quinase LRR, canal(is) de cátion CASTOR/POLLUX e nucleoporina133 são

20 necessários para a variação de cálcio, uma das mais precoces respostas fisiológicas detectáveis nos pelos radiculares expostos ao fator Nod purificado.

Para estabelecer a simbiose, o *Rizóbio* tem acesso a células individuais da planta por endocitose onde são instalados em simbiossomas rodeados por uma membrana peribacterioide. Em *Lotus*, a infecção ocorre

25 através de um filamento de infecção que leva as bactérias através dos pelos radiculares no córtex radicular e as distribui às células, que se tornam infectadas com o simbiossoma que contém células fixadoras de nitrogênio. Em resposta às bactérias introduzidas, os pelos radiculares se deformam e enrolam, montando uma pequena bolsa que provê um sítio para recepção do

30 filamento de infecção (Geurts *et al.*, 2005 *Curr. Opinion Plant Biol.*, 8: 346-352). Os filamentos de infecção são estruturas derivadas das plantas, originados da invaginação da membrana plasmática, acompanhada pela deposi-

ção externa de material da parede celular. Frente ao progresso interno do filamento intracelular, células corticais radiculares desdiferenciam e reentram no ciclo celular para iniciar o primórdio de nódulo. Mais tarde, no processo, a formação padrão e a diferenciação celular especificam tipos de tecido e célula inclusive as células infectadas que endocitaram os *Rizóbios*. No nódulo maduro funcional, os feixes vasculares periféricos estão conectados à vasculatura radicular e os principais tipos de tecidos/célula podem ser distinguidos (Pawlowski e Bisseling, 1996, *Plant Cell* 8: 1899-1913).

A análise de um grupo de mutantes de nodulação, inclusive alguns que falham em mostrar oscilações de cálcio em resposta aos sinais do fator Nod, revelou que além da falta de nodulação, esses mutantes são incapazes de formar endossimbioses com fungos micorrizais arbusculares. Isso implica que uma via comum de transdução de sinal simbiótica é compartilhada por dois tipos de relações endossimbióticas, ou seja, simbiose de nódulo radicular, que é basicamente restrita à família das leguminosas, e fungos micorrizais arbusculares, que é comum à maioria das espécies de plantas terrestres. Isto sugere que possa haver alguns genes-chave dos quais as leguminosas dispõem para empregar na nodulação, e que estão faltando em cultura de plantas tais como cereais. A identificação desses genes-chave, que codificam funções que são indispensáveis para o estabelecimento de um sistema de fixação de nitrogênio em leguminosas, e sua transferência e expressão em plantas não-noduladas, tem sido por muito tempo um objetivo de melhoradores moleculares de plantas. Isto poderia ter um impacto agrônomo significativo no cultivo de cereais como arroz, onde a produção de duas colheitas por ano pode necessitar de fertilização com até 400 kg de nitrogênio por hectare.

A simbiose de nódulo radicular depende de uma interação bem-sucedida entre a planta hospedeira e o seu simbionte cognato que inclui a etapa do reconhecimento do fator nod pela planta hospedeira. A identificação de genes que regulam a nodulação em Leguminosas forneceria ferramentas para otimizar e modificar este processo em benefício da agricultura.

Em resumo, há uma necessidade de melhorar a capacidade de

formação de nódulos e as propriedades de fixação de nitrogênio nas culturas de leguminosas, bem como transferir essa via para a cultura não-nodulada para suprir as necessidades nutricionais de uma população global crescente, minimizando o uso futuro de fertilizantes nitrogenados e o seu impacto ambiental negativo associado.

Sumário da Invenção

Uma primeira modalidade da invenção é uma molécula de DNA que codifica um polipeptídeo de histidina quinase mutante compreendendo uma sequência de aminoácidos selecionada entre: SEQ ID NO: 6, 9, 12, 14-20, em que o resíduo de aminoácido correspondente a Xaa é selecionado entre isoleucina, serina, treonina, valina, metionina, alanina, fenilalanina, tirosina, triptofano, arginina, lisina, glicina, histidina, aspartato, asparagina, glutamato, glutamina, prolina e cisteína,

um ortólogo de (a) e um truncamento de (a) ou (b), capaz de induzir formação espontânea de nódulo em uma planta.

Uma segunda modalidade da invenção é um polipeptídeo de histidina quinase mutante composto de uma sequência de aminoácidos selecionada entre: (a) SEQ ID NO: 6, 9, 12, 14-20, em que o resíduo de aminoácido correspondente a Xaa é selecionado entre isoleucina, serina, treonina, valina, metionina, alanina, fenilalanina, tirosina, triptofano, arginina, lisina, glicina, histidina, aspartato, asparagina, glutamato, glutamina, prolina e cisteína; (b) uma variante alélica de (a) independente de citocinina; (c) um ortólogo de (a) independente de citocinina; e (d) um truncamento de (a), (b) ou (c).

Uma modalidade adicional da invenção é uma planta geneticamente modificada caracterizada por ter uma sequência de nucleotídeos que codifica um polipeptídeo compreendendo um mutante de histidina quinase consistindo em uma sequência de aminoácidos selecionada entre: (a) SEQ ID NO: 6, 9, 12, 14-20, em que o resíduo de aminoácido correspondente a Xaa é selecionado entre isoleucina, serina, treonina, valina, metionina, alanina, fenilalanina, tirosina, triptofano, arginina, lisina, glicina, histidina, aspartato, asparagina, glutamato, glutamina, prolina e cisteína; (b) uma variante alé-

lica de (a) independente de citocinina; (c) um ortólogo de (a) independente de citocinina; e (d) um truncamento de (a), (b) ou (c), em que a dita planta é capaz da formação espontânea de nódulo.

5 A invenção é ainda direcionada ao uso de uma molécula de ácido nucleico que codifica uma histidina quinase mutante consistindo em uma sequência de aminoácidos selecionada entre: (a) SEQ ID NO: 7, 8, 9, 10, 11, 15, 26 e 27, em que o resíduo de aminoácido correspondente a Xaa é selecionado entre isoleucina, serina, treonina, valina, metionina, alanina, fenilalanina, tirosina, triptofano, arginina, lisina, glicina, histidina, aspartato, asparagina, glutamato, glutamina, prolina e cisteína; (b) um ortólogo de (a); e (c) 10 um truncamento de (a) ou (b), como um transgene para produzir a planta geneticamente modificada da invenção, capaz de nodulação espontânea de acordo com suas várias modalidades.

15 A invenção é ainda direcionada a um método de produção de uma planta geneticamente modificada de acordo com a invenção nas suas várias modalidades, caracterizada pela introdução de um cassete de gene compreendendo a dita sequência de nucleotídeos que codifica o dito polipeptídeo e selecionando uma planta transgênica e sua progênie que expressa o dito polipeptídeo.

20 A invenção ainda inclui uma planta geneticamente modificada produzida de acordo com um processo de mutagênese de DNA e seleção de uma planta capaz de formação espontânea de nódulo, ou por um método de transformação com um transgene que codifica uma histidina quinase mutante da invenção.

25 A invenção ainda inclui uma semente ou uma cultura obtida da planta geneticamente modificada da invenção. Além disso, a invenção é direcionada ao uso de uma planta geneticamente modificada de acordo com a invenção em um programa de melhoramento, e uma planta selecionada no programa de melhoramento compreendendo uma sequência de nucleotídeos 30 que codifica um polipeptídeo compreendendo uma histidina quinase mutante consistindo em uma sequência de aminoácidos selecionada entre: (a) SEQ ID NO: 6, 9, 12, 14-20 em que o resíduo de aminoácido correspondente a

4. Xaa é selecionado entre isoleucina, serina, treonina, valina, metionina, alanina, fenilalanina, tirosina, triptofano, arginina, lisina, glicina, histidina, aspartato, asparagina, glutamato, glutamina, prolina e cisteína; (b) uma variante alélica de (a) independente de citocinina; (c) um ortólogo de (a) independente de citocinina, e (d) um truncamento de (a), (b) ou (c), em que a dita planta é capaz de formação espontânea de nódulo.

Breve Descrição dos Desenhos

Figura 1. Caracterização fenotípica do mutante *snf2*.

- (A) rizóbios do tipo selvagem induziram o nódulo radicular (B) nódulo radicular espontâneo na raiz *snf2* 5 semanas após a germinação. As cabeças de seta indicam nódulos. Secção transversal do (C) tipo selvagem e (D) da raiz *snf2* no tempo zero e (E) tipo selvagem e (F) raiz *snf2* após 6 dias de crescimento no meio B5 sem hormônio. A cabeça de seta indica células se dividindo no periciclo. As setas indicam células do xilema. (G) e (H) crescimento do calo a partir de hipocótilos do tipo selvagem e *snf2* em diferentes concentrações de auxina e citocinina. Os calos foram fotografados depois de 21 dias de crescimento a 26°C. (I) e (J) segmentos radiculares do tipo selvagem e *snf2* incubados por três semanas em meio sem hormônio. Barras de escala: (C); (D); (E); (F) 50 µm.

- 20 Figura 2. Clonagem baseada no mapa do gene *Lhk1*

- A: Figura do cromossomo IV de *Lotus japonicus* com mapa genético estendido da região abaixo do *snf2* (*locus Lhk1*), compreendendo 6 clones de BAC/TAC (sublinhado em negrito). As linhas verticais indicam microssatélites ou marcadores de polimorfismo de nucleotídeo único. O número das plantas recombinantes obtidas por um marcador de microssatélite (TM) é indicado em negrito abaixo do marcador.

- B: A estrutura éxon-intron do gene *Lhk1*, onde os 11 éxons de *Lhk1* são indicados por caixas abertas. A transição C→T no éxon 4 do alelo mutante *snf2* (indicado com uma cabeça de seta preta) codifica um polipeptídeo mutante LHK1 que compreende uma substituição de F266 para L266.

Figura 3. Transformação do alelo *snf2* no tipo selvagem nas raízes cabeludas leva à organogênese controlada e fenótipo de nodulação es-

pontânea.

Plantas *Lotus japonicus* do tipo selvagem foram transformadas com *Agrobacterium rhizogenes* portando o gene *snf2*. Nódulos espontâneos nas raízes cabeludas estão identificados por cabeças de seta brancas. A raiz principal está identificada por uma seta branca.

Figura 4. Estrutura da proteína LHK1 de *Lotus*.

(A) Representação esquemática dos domínios da proteína LHK1. (B) A sequência de aminoácidos de LHK1 arranjada em domínios proteicos. O domínio extracelular do receptor predito é dado em itálico e o domínio CHASE dentro do domínio extracelular do receptor está sublinhado; o domínio histidina quinase é dado em negrito e sublinhado; o domínio ATPase His Quinase é dado em negrito; e o domínio do receptor está em negrito e itálico. O asterisco no domínio CHASE marca a posição da substituição do aminoácido no alelo *snf2*.

Figura 5. Ensaios *in vivo* de sinalização de citocinina mediada pelo receptor.

(A) Ensaio em placa da atividade de β -galactosidase expressa por um gene repórter *cps::lacZ* em *E. coli*. Uma cepa de *E. coli* SRC122 que porta o repórter de *cps::lacZ*, transformada com um construto de plasmídeo compreendendo o DNAC de *snf2* ou o DNAC do tipo selvagem, foi cultivada em placas na ausência de citocinina ou em presença de quatro citocininas diferentes. A cor azul mostra a conversão do substrato X-Gal pela β -galactosidase.

(B) Indução da atividade de β -galactosidase pela citocinina em culturas líquidas de SRC122 *cps::lacZ* transformada com o DNAC de *snf2* ou com DNAC do tipo selvagem. T-z: Trans-zeatina.

(C) Modelo de trabalho do papel funcional de *Lhk1* na nodulação. O reconhecimento de um fator Nod rizobial corretamente adornado pelo NFR1 e NFR5 induz transdução de sinal através da via comum, incluindo variação de cálcio e a interpretação da variação de cálcio pela proteína CCaMK. Um aumento localizado na biossíntese de citocinina captado pelo receptor LHK1 leva então à desdiferenciação celular e ativação do ciclo celu-

lar. *snf2* está constitutivamente ativo mas ainda necessita dos genes *Nin* e *Sym35* para a organogênese do nódulo.

Figura 6. Efeito do tratamento com fito-hormônio no crescimento de calo em explantes de raiz.

5 O crescimento de calo no ecotipo Gifu do *L. japonicus* do tipo selvagem **(A)** e **(B)** fragmentos de raiz *snf2* em diferentes concentrações de auxina e citocinina. Os calos foram fotografados após 21 dias de crescimento a 26°C.

10 Figure 7. Expressão do gene *Lhk1* em órgãos e *Lhk1*, *Lrr5* e *nin* em resposta à citocinina.

(A) Expressão de *Lhk1* em órgãos diferentes. **(B)** Expressão do gene *Lrr5* em explantes de raiz do tipo selvagem e *snf2* incubados em placa de ágar-ágar B5 com ou sem 0,5 µg/ml de BAP durante 10 dias. **(C)**, **(D)**, **(E)** Expressão dos genes *Lrr5*, *Lhk1* e *Nin*, em raízes intactas do tipo selvagem e *snf2* em resposta à citocinina. **(F)** Expressão do gene *Nin* em explantes de raiz do tipo selvagem e *snf2* incubados em placas de ágar-ágar B5 com ou sem 0,5 µg/ml de BAP durante 10 dias.

Figura 8: Efeito da citocinina no comprimento do broto e da raiz de *L. japonicus* do tipo selvagem e no mutante *snf2-2*.

20 Plantas foram cultivadas em ¼ de B&D contendo diferentes concentrações de BAP. O comprimento do broto e da raiz foi medido após 3 semanas de crescimento. **(A)** comprimento do broto **(B)** comprimento da raiz.

Figure 9. Fenótipo de nodulação de *snf2* e do duplo mutante *snf2 har1* na ausência dos rizóbios.

25 **(A)** *snf2* mutante e **(B)** duplo mutante *snf2 har1* 5 semanas após a germinação.

Descrição Detalhada da Invenção

A presente invenção provê um gene isolado que codifica um polipeptídeo mutante, cuja expressão em plantas confere um fenótipo de nodulação espontânea. Mais especificamente o gene foi isolado de mutantes de nodulação espontânea *snf2* do *Lotus japonicus* que desenvolvem nódulos sem rizóbios brancos na ausência do microssimbionte *M. loti*, enquanto as

plantas *Lotus* do tipo selvagem somente são noduladas depois da indução com os seus cognatos fatores nod ou simbiote rizobial (figura 1a, b). Análises histológicas detalhadas de secções de nódulos demonstram que os nódulos espontâneos *snf2* são nódulos genuínos com uma ontogenia e fisiologia semelhantes aos nódulos rizobiais induzidos. Assim, os nódulos radiculares espontâneos formados em mutantes *snf2* originam-se de células corticais e têm uma vasculatura periférica característica de nódulos do tipo selvagem, embora eles sejam desprovidos de filamentos de infecção e rizóbios. O alelo *snf2* é monogênico dominante e a inoculação de mutantes *snf2* com o *M. loti* resulta no desenvolvimento de nódulos radiculares normais fixadores de nitrogênio compatível com a presença de uma mutação de ganho de função neste alelo.

O alelo *snf2* localiza-se no cromossomo 4 em uma posição aproximadamente a 1 cM do fim do braço longo (figura 2). O gene, correspondente ao alelo *snf2*, codifica uma forma mutante de uma *Lotus* histidina quinase (LHK1) que é um homólogo dos genes *Arabidopsis histidina quinase* (AHK) que codificam as proteínas receptoras de citocinina. Uma comparação do gene *Lotus* histidina quinase do tipo selvagem (*Lhk1*) [SEQ ID NO:1] com o gene correspondente do alelo *snf2* [SEQ ID NO:4] revela uma transição de nucleotídeo único (C para T) na sequência de codificação do alelo *snf2* que resulta na substituição de uma leucina conservada [L266 em *Lhk1*; SEQ ID NO:3] por uma fenilalanina (F266 no alelo *snf2* do *Lhk1*; SEQ ID NO:6) no polipeptídeo codificado. A invenção ainda provê DNAs isolados correspondentes ao transcrito do gene *Lhk1* do tipo selvagem [SEQ ID NO:2] e ao alelo *snf2* [SEQ ID NO:5]. O alinhamento das sequências genômicas do *Lhk1* e sequências de DNAc ou o respectivo gene do alelo *snf2* e sequências de DNAc, define uma estrutura primária de *Lhk1* e o seu alelo *snf2* como composto de 11 éxons (figura 2).

A invenção ainda provê a proteína receptora de citocinina (LHK1) isolada, do tipo selvagem e mutante, codificada, respectivamente, pelo gene *Lhk1* e seu alelo *snf*, compreendendo 993 aminoácidos com uma massa predita de 110 kD (figura 4). Ambas as proteínas compreendem dois

segmentos transmembrana no N-terminal, localizados entre os aminoácidos 37 e 57 e entre os aminoácidos 328 e 357. Localizado entre esses segmentos estão os motivos característicos de ciclases/histidina quinases associadas ao domínio sensorial extracelular (CHASE). Este domínio extracelular predito é seguido por uma histidina quinase intracelular putativa entre os aminoácidos 379 e 693 e um domínio receptor entre os aminoácidos 852 e 985. Esses domínios são característicos de sistemas reguladores de dois componentes que funcionam por uma fosforrealocação. A fenilalanina 266 no alelo *snf2*, substituída pela leucina 266, está localizada em um domínio conservado compartilhado entre os domínios extracelulares CHASE de receptores histidina quinase (figura 4). As propriedades funcionais da histidina quinase (proteína receptora de citocinina) codificada pelo alelo *snf2* incluem a atividade independente de citocinina, tal que o alelo *snf2*, em contraste com alelo do tipo selvagem *Lhk1* pode induzir a nodulação na ausência de sinalização de citocinina.

Assim sendo, a presente invenção provê uma proteína histidina quinase mutante isolada (proteína mutante receptora de citocinina) capaz de induzir a nodulação espontânea quando expressa em uma planta, em que o resíduo de aminoácido correspondente a L266 na histidina quinase LHK1 do tipo selvagem codificado pelo gene *Lhk1* do *Lotus japonicus* é substituído por um aminoácido selecionado entre isoleucina, serina, treonina, valina, metionina, alanina, fenilalanina, tirosina, triptofano, arginina, lisina, glicina, histidina, aspartato, asparagina, glutamato, glutamina, prolina e cisteína. Alternativamente, o resíduo de aminoácido correspondente à L266 na histidina quinase LHK1 do tipo selvagem (Xaa) é conservativamente substituído com um aminoácido selecionado entre fenilalanina, tirosina, triptofano, histidina ou prolina. O dito último grupo pode ser aumentado para incluir qualquer dos 3 grupos seguintes de aminoácidos (arginina, lisina, aspartato, glutamato) ou (asparagina, glutamina, serina, treonina) ou (cisteína, metionina, isoleucina, valina, glicina, alanina). Essa histidina quinase mutante da invenção inclui homólogos de várias plantas nos quais o resíduo de aminoácido correspondente a L266 na histidina quinase LHK1 do tipo selvagem (com base no ali-

nhamento como mostrado na Tabela 2) é deletado ou substituído com qualquer outro aminoácido diferente de leucina. Mais especificamente, a invenção inclui os homólogos de histidina quinase mutante do *Lotus japonicus* (forma mutante de Lhk1; SEQ ID NO:6,15), *M. truncatula* (forma mutante de ABE94286; SEQ ID NO:16), *Arabidopsis* (forma mutante de BAB33311; SEQ ID NO:9,17), arroz (forma mutante de XP_469566; SEQ ID NO:18), milho (forma mutante de BAE80688; SEQ ID NO:12,19) e *Cucurbita maxima* (forma mutante de CAF31355.1; O SEQ ID NO:14,20) ou ortólogos ou variantes alélicas das mesmas, em que Xaa é qualquer outro aminoácido diferente de leucina. Em uma modalidade, a histidina quinase mutante é aquela na qual Xaa é selecionada a partir do grupo da fenilalanina, tirosina, triptofano, histidina ou prolina, ou alternativamente Xaa é fenilalanina. O ortólogo ou variante alélica da histidina quinase mutante da invenção compartilha uma porcentagem de identidade com a sequência da dita histidina quinase mutante, selecionada do grupo composto de pelo menos 60, 65, 70, 75, 80, 85, 90, 95, e 98 por cento, ou truncamento desses, em que a dita quinase tem todas as propriedades funcionais da histidina quinase mutante da invenção, tal que quando expressa em uma planta da invenção ela confere a capacidade de formar nódulos espontâneos. Em uma modalidade adicional o alelo ortólogo/variante é uma histidina quinase mutante que é uma proteína receptora de citocinina que é independente citocinina.

Esta substituição única de aminoácido ou deleção codificada pelo alelo mutante *snf2* do gene *Lhk1* ou o seu homólogo é suficiente para conferir um fenótipo de nodulação espontânea nas raízes das plantas transformadas do tipo selvagem que expressam um transgene compreendendo o gene correspondente ao alelo *snf* mutante (Tabela 1 e figura 3). O alelo *snf2* atua como um gene dominante de ganho de função, visto que a nodulação espontânea é vista em um contexto genético do tipo selvagem, e está ausente em plantas transformadas com gene *Lhk1* do tipo selvagem. Portanto, a expressão de uma única cópia da histidina quinase mutante da invenção em uma planta, como no caso de uma planta heterozigota é suficiente para conferir o fenótipo *snf2* de nodulação espontânea.

A presente invenção ainda provê uma molécula de DNA que codifica a dita histidina quinase mutante homóloga (em que Xaa é Phe) do *Lotus japonicus* (SEQ ID NO:4,5), *M. truncatula* (forma mutante da Acesso AC141922.19; SEQ ID NO:7), *Arabidopsis* (forma mutante de AB049935.1; SEQ ID NO:8), arroz (forma mutante de NT_079916.2; SEQ ID NO:10), milho (forma mutante de AB206392.1; SEQ ID NO:11) e *Cucurbita maxima* (forma mutante de AJ628045.1; SEQ ID NO:13), ou fragmentos desses que codificam uma histidina quinase mutante funcional inteira ou truncada capaz de causar nodulação espontânea quando expressa em uma planta.

De acordo com a presente invenção, a sequência de DNA que codifica a proteína histidina quinase mutante da invenção está operavelmente ligada a uma sequência de DNA promotora capaz de dirigir a expressão da dita histidina quinase em uma planta, e a uma sequência terminadora 3'. O promotor pode ser um promotor que dirige a expressão da dita histidina quinase em tecidos radiculares de uma planta e/ou em células destinadas a tornar-se o primórdio do nódulo e desenvolver-se em nódulos. Os exemplos convenientes de um promotor e um terminador incluem o promotor e a sequência terminadora correspondente ao gene *Lhk1* do tipo selvagem. Em uma modalidade da invenção, o promotor usado para dirigir a expressão da histidina quinase mutante é um regulado (por exemplo, promotor de tecido ou específico do tipo celular) que inclui o promotor nativo do gene *Lhk1* ou o seu homólogo como definido na presente invenção. Um exemplo de um promotor constitutivo heterólogo inclui o promotor do 35SCaMV (Acc. No: V00141, J02048). Um transgene (cassete de gene) compreendendo uma sequência de DNA que codifica uma histidina quinase mutante da invenção operavelmente fundida a uma sequência promotora e opcionalmente a uma sequência terminadora pode ser construído por técnicas de DNA recombinante.

De acordo com a presente invenção, um transgene compreendendo uma sequência de DNA que codifica a histidina quinase mutante pode ser usado para gerar uma planta que expressa a histidina quinase mutante da invenção. O transgene pode estar integrado de maneira estável no genoma de uma planta hospedeira por técnicas de transformação bem-

conhecidas pelo versado na técnica. Além disso, os vetores binários e métodos baseados em *Agrobacterium tumefaciens* para a integração estável dos transgenes em todas as principais plantas de cereal são conhecidos, como descrito por exemplo para arroz (Hiei *et al.*, 1994, *The Plant J.* 6; 271-282), e milho (Yuji *et al.*, 1996, *Nature Biotech.* 14: 745-750). Uma sequência de DNA que codifica uma histidina quinase mutante também pode ser introduzida em outra planta pelo cruzamento com uma planta geneticamente modificada que expressa a histidina quinase mutante da invenção.

A planta geneticamente modificada da invenção, se gerada por mutagênese, transformação com um transgene da invenção, ou introdução do dito transgene, pode ser usada em um programa de melhoramento, para selecionar as plantas com a capacidade de fixar nitrogênio, ou capacidade de fixação de nitrogênio aumentada, que herdaram o gene que codifica a histidina quinase mutante. A invenção assim inclui uma planta geneticamente modificada, produzida pela transformação de uma planta natural, que é capaz de nodulação espontânea. A expressão de um gene que codifica uma histidina quinase mutante em uma planta fixadora de nitrogênio, como um membro da *Leguminosae* (tal como soja, feijão, ervilha, amendoim, grão-de-bico, feijão-caupi, lentilha, guandu, alfafa e trevo), tem utilidade particularmente com respeito ao aumento da capacidade fixadora de nitrogênio da dita planta sob uma ou mais condições ambientais de crescimento. A expressão de um gene que codifica uma histidina quinase mutante em uma cultura que não fixa naturalmente o nitrogênio, como uma planta dicotiledônea ou uma planta monocotiledônea inclusive um membro dos cereais (como trigo, centeio, aveia, cevada, sorgo, painço, milho, grama *Poaceae* e arroz), tem particularmente a utilidade com respeito à concessão de habilidade de fixar nitrogênio. As plantas, bem como a progênie da planta, selecionadas em tal programa de melhoramento podem ser cultivadas com o objetivo da colheita de uma cultura, onde o cultivo pode ser as partes vegetativas da planta, por exemplo folha, tronco ou tubérculo, ou partes reprodutivas, inclusive flores, semente, grãos, espiga ou fruto.

Os exemplos dados abaixo servem para ilustrar várias modali-

dades da invenção e as suas respectivas características. Eles demonstram que uma planta, por exemplo o *Lotus*, que é homozigoto ou heterozigoto para um gene que codifica o polipeptídeo de histidina quinase mutante da invenção, forma nódulos espontâneos que podem ser infectados por bactérias simbióticas fixadoras de nitrogênio, por exemplo, o *M. loti*, e que são capazes de fixar o nitrogênio e dar suporte ao crescimento da planta sob condições com nitrogênio limitado. A formação de nódulos fixadores de nitrogênio depois da inoculação com bactérias fixadoras de nitrogênio na dita planta nodulada espontaneamente não é dependente da produção do fator nod pelas bactérias infectantes ou captação do fator nod pela planta infectada. Isso indica que somente um subconjunto de genes relacionados à nodulação é necessário para a fixação de nitrogênio ocorrer em uma planta nodulada espontaneamente que expressa uma histidina quinase mutante depois da inoculação com uma bactéria rizóbio. As propriedades únicas da histidina quinase mutante da invenção podem ser exploradas para aumentar a fixação de nitrogênio nas plantas fixadoras de nitrogênio existentes, bem como no estabelecimento de fixação de nitrogênio em plantas não-noduladas.

Exemplo 1. Clonagem posicional e identificação do gene mutante *snf-2* em *Lotus japonicus*

Mutantes de *Lotus* (*snf2-1* e *snf2-2*) têm um fenótipo de nodulação espontânea, originado de uma triagem das sementes de *Lotus japonicus* do ecotipo Gifu com EMS. O gene mutante *snf-2* em *Lotus japonicus*, que dá origem à nodulação espontânea, foi localizado no braço longo do cromossomo IV, a aproximadamente 1 cM do final, em um locus denominado *Lhk1*.

A posição do gene *snf2* foi determinada pelo mapeamento fino usando marcadores de microssatélite (marcadores TM) e marcadores polimórficos de nucleotídeo único desenvolvidos em clones BAC ou TAC ancorados ao mapa genético da região *Lhk1*. O mapeamento foi realizado em uma população F2 estabelecida de um híbrido dos ecotipos de *Lotus japonicus*, Miyakojima e MG 20. O mapa fino foi usado para construir um "contig" físico de TAC/BAC compreendendo seis clones de BAC/TAC de MG20 que foram reunidos para cobrir a região *Lhk1* flanqueando entre os dois marca-

dores TM1146 e TM0069 (Figure 2). Uma vez que *snf2* é uma mutação dominante, plantas F2 do tipo selvagem, e não as plantas mutantes, foram marcadas para mapeamento dos alvos. A genotipagem de 853 plantas do tipo selvagem que não desenvolveram nódulos radiculares espontaneamente foi usada para identificar marcadores que delimitam o locus correspondente à mutação *snf2* a uma região de 120 kb. Quinze genes foram preditos dentro desta região, um dos quais foi predito para codificar um receptor de citocinina. A sequência dos alelos tanto do tipo selvagem [SEQ ID NO:1] quanto do mutante *snf2* [SEQ ID NO:4] deste gene candidato foi determinada, a qual revelou uma substituição do nucleotídeo C para T em ambos os alelos *snf2*. O gene de tipo selvagem foi denominado *gene* histidina quinase de *Lotus* (*Lhk1*) que codifica um polipeptídeo LHK1 [SEQ ID NO:3]. A transição de nucleotídeo único (C para T) em uma sequência do éxon do alelo mutante *snf2* codifica um polipeptídeo que tem uma fenilalanina (F266 em SEQ ID NO:6) em substituição da leucina conservada (L266 em SEQ ID NO:3) no polipeptídeo codificado pelo gene de tipo selvagem, e identifica *snf2* como um alelo de um gene histidina quinase de *Lotus* (*Lhk1*).

Exemplo 2. Clonagem e identificação do DNAc de *Lhk1* correspondente ao transcrito de tipo selvagem do gene *Lhk1* em *Lotus japonicus*

Um DNAc inteiro de *Lhk1* (3568 pb) foi isolado de uma biblioteca λ ZAPII de DNAc preparada a partir de RNAm isolado de *M. loti* inoculada nas raízes de *Lotus japonicus*. O DNAc de *Lhk1* foi sequenciado [SEQ ID NO:2], no qual o sítio de início de transcrição foi determinado estando situado a pelo menos 137 nucleotídeos a montante do códon de iniciação e a sequência de codificação foi seguida por uma região 3' não traduzida de aproximadamente 445 nucleotídeos. O alinhamento das sequências genômicas e de DNAc definiu uma estrutura primária de *Lhk1* composta de 11 éxons (Figura 2B). A sequência de nucleotídeos do alelo mutante *snf2* do DNAc do *Lhk1* é indicada SEQ ID NO:5.

Exemplo 3: Expressão do alelo *snf2* em raízes de *Lotus* de tipo selvagem confere um fenótipo de nodulação espontânea.

A expressão transgênica do alelo *snf2* de tipo selvagem ou mu-

tante em raízes de *Lotus* foi realizada para confirmar *in planta* a base genética do fenótipo de nodulação espontânea. Um fragmento genômico de 12,7 kb do clone de BAC 1K18 foi usado para clonar o gene *Lhk1* de tipo selvagem incluindo uma região promotora de 2,2 kb, e o seu alelo *Snf2* compreendendo a transição C→T, que então foram subclonados no plasmídeo pIV10. Os construtos foram integrados em *A. rhizogenes* da cepa AR12 pelo cruzamento triparental. A transformação foi realizada como descrito por Stougaard (1995) *Method Mol Biology* 49: Plant Gene Transfer and Expression Protocols, p. 49-63, e incluída a transformação com um vetor de controle vazio. Nodulação radicular, registrada na ausência da indução rizobial, só foi detectada em raízes transformadas com o alelo *snf2*, enquanto o gene *Lhk1* de tipo selvagem falhou em conferir a nodulação espontânea em raízes de tipo selvagem (Tabela 1, figura 3).

Tabela 1

15 Indução de nódulos espontâneos em raízes cabeludas de tipo selvagem transformadas com o gene *Lhk1* de tipo selvagem ou seu alelo de ganho de função *snf2*

Construto de transformação	% de Nodulação espontânea*	Número de nódulos/planta*	Número de nódulos/planta nodulada*
vetor vazio	0 (0/15)	0	0
alelo <i>Lhk1</i> de tipo selvagem	0 (0/41)	0	0
alelo <i>snf2</i>	37 (30/81)	0,87	2,36

* Raízes transgênicas foram registradas na ausência de *M. loti*.

20 A capacidade do alelo *snf2* conferir nodulação espontânea em raízes transformadas que têm o contexto genético de tipo selvagem serve para confirmar a natureza dominante do alelo *snf2*. Além disso, é surpreendentemente mostrado que a histidina quinase independente de citocinina (*snf2*), quando sob controle de um promotor regulado, por exemplo, promotor do gene *Lhk1*, confere a organogênese controlada de nódulos radiculares, em contraste com a proliferação celular descontrolada levando ao desenvolvimento de nódulos volumosos. A formação controlada de nódulos
 25 espontâneos, capazes de infecção e fixação de nitrogênio em plantas trans-

formadas com o alelo *snf2*, aumenta sua capacidade fixadora de nitrogênio durante o cultivo. Uma vez que a única diferença estrutural entre o gene *Lhk1* de tipo selvagem e o alelo *snf2* está situada em uma mutação pontual que codifica uma substituição única de aminoácido, esses dados demonstram que a nodulação espontânea é o resultado de uma substituição L/F266 em um polipeptídeo LHK1.

Exemplo 3: Lotus LHK1 é um membro da família de receptor de citocinina.

A observação do clone de DNAC do *Lhk1* revela uma fase aberta de leitura de 2979 nucleotídeos que é predita para codificar uma proteína receptora de citocinina (LHK1) consistindo em 993 aminoácidos com uma massa predita de 110 kD (figura 4). No término N, dois segmentos transmembrana estão localizados entre os aminoácidos 37 e 57 e entre os aminoácidos 328 e 357. Localizado entre esses segmentos estão os motivos característicos das ciclases/histidina quinases associadas ao domínio sensorial extracelular (CHASE). Este domínio extracelular predito é seguido por uma histidina quinase intracelular putativa entre os aminoácidos 379 e 693 e um domínio receptor entre os aminoácidos 852 e 985. Esses domínios são característicos de sistemas reguladores de dois componentes que funcionam por uma fosforrealocação.

A análise comparativa define a proteína Lotus LHK1 como um membro da família de receptor de citocinina que inclui proteínas de *Medicago truncatula*, *Arabidopsis*, arroz e milho (tabela 2). A proteína Lotus LHK1 compartilha uma identidade de sequência de aminoácidos de 83%, 68%, 58% e 49%, respectivamente, com as proteínas homólogas de *M. truncatula* (ABE94286), *Arabidopsis* (BAB33311), arroz (XP_469566) e milho (BAD01584). Dentre os três receptores de citocinina *Arabidopsis*, LHK1 é o mais semelhante a AHK4/(Cre1) que é importante para o desenvolvimento radicular normal e apresenta uma função na captação da citocinina fornecida externamente (Mahonen *et al.* (2000), *Genes Dev.* 4: 2938). A leucina 266 substituída por uma fenilalanina 266 no alelo *snf2* está localizada em um motivo conservado compartilhado entre os domínios CHASE extracelulares dos receptores de histidina quinase (figura 4).

Tabela 2

Alinhamento das Histidina quinases de *Lotus*, *Arabidopsis*, milho, arroz e *Medicago*.

[illegible]

[illegible]

Tabela 2 - continuação

Os-HK	725: HLRGLTSEVGLPTFGV... 820	810	820	830	840	850	860	870	880
Zm-HK1	697: HLRGLTSEVGLPTFGV... 820	810	820	830	840	850	860	870	880
At-AHK4	780: HLRGLTSEVGLPTFGV... 820	810	820	830	840	850	860	870	880
Mt-cre	722: HLRGLTSEVGLPTFGV... 820	810	820	830	840	850	860	870	880
LHK1	721: HLRGLTSEVGLPTFGV... 820	810	820	830	840	850	860	870	880
Os-HK	802: A... 890	890	900	910	920	930	940	950	960
Zm-HK1	770: LV... 890	890	900	910	920	930	940	950	960
At-AHK4	858: FA... 890	890	900	910	920	930	940	950	960
Mt-cre	796: LA... 890	890	900	910	920	930	940	950	960
LHK1	785: LA... 890	890	900	910	920	930	940	950	960
Os-HK	878: KVNLRVAA... 970	970	980	990	1000	1010	1020	1030	1040
Zm-HK1	838: KVNLRVAA... 970	970	980	990	1000	1010	1020	1030	1040
At-AHK4	931: VNRVAA... 970	970	980	990	1000	1010	1020	1030	1040
Mt-cre	870: VNRVAA... 970	970	980	990	1000	1010	1020	1030	1040
LHK1	858: VNRVAA... 970	970	980	990	1000	1010	1020	1030	1040
Os-HK	958: TAK... 1050	1050	1060	1070	1080	1090	1100	1110	1120
Zm-HK1	917: TTA... 1050	1050	1060	1070	1080	1090	1100	1110	1120
At-AHK4	1001: ...E... 1050	1050	1060	1070	1080	1090	1100	1110	1120
Mt-cre	945: KSE... 1050	1050	1060	1070	1080	1090	1100	1110	1120
LHK1	936: KSE... 1050	1050	1060	1070	1080	1090	1100	1110	1120

* Indica a substituição de aminoácido na posição 266 de uma Leucina para uma Fenilalanina no mutante.

Exemplo 4. Atividade independente de Citocinina

O fenótipo de nodulação espontânea conferido pelo alelo *snf2* é mostrado ser devido a uma função independente de citocinina do receptor de citocinina LHK1 mutante codificado por este alelo, que é caracterizado por uma substituição L/F266 no domínio CHASE (figura 4). A atividade *in vivo* dos receptores de citocinina *Lotus* LHK1 de tipo selvagem e mutante e a sua resposta à citocinina foi demonstrada por meio de um ensaio de fosforrealocação de dois componentes desenvolvido em *E. coli* (Suzuki *et al.*, (2001) *Plant Cell Physiol.* 42: 107). Neste ensaio a indução de citocinina pode ser determinada sem interferência de níveis de citocinina endógena, penetração de citocinina e vias metabólicas interferentes presentes em plantas. A fosforrealocação explorada neste ensaio está normalmente envolvida na regulação da síntese de polissacarídeo extracelular bacteriano através da ativação do *operon cps* em resposta a um sensor (RcsC), o qual tem uma estrutura de domínio semelhante aos receptores de citocinina de planta. A expressão funcional de um receptor de citocinina em uma cepa de *E. coli* de sensor RcsC deletado permite a captação de citocinina ser lida como atividade de β -galactosidase pela expressão de uma fusão *cps::lacZ*.

Consequentemente, os DNAs *Lhk1* correspondentes ao alelo de tipo selvagem e *snf2*, foram clonados no vetor de expressão pIN-III e transformados na cepa de sensor negativo SRC122 (Δ RcsC) abrigando o gene de *cps::lacZ*. Os transformantes foram cultivados durante a noite em Luria-ágar-ágar líquido e então colocados em placas LB, suplementadas com e sem t-zeatina, quinetina, BAP (200 μ M) cada um e tidiazuron (50 μ M). Os meios LB foram ainda suplementados com X-Gal (5-bromo-4-cloro-3-indolil- β -D-galactosidase), um substrato para β -galactosidase. Após 40 horas de incubação à temperatura ambiente (25°C), as placas foram fotografadas. A cepa de *E. coli* (SRC122) foi tomada como um controle negativo. A expressão do alelo *snf2* (receptor de citocinina LHK1 mutante) é vista por induzir a atividade de β -galactosidase na ausência de citocinina (figura 5A). Ao contrário, a LHK1 de tipo selvagem induziu a atividade de β -galactosidase de uma maneira dependente de citocinina (figura 5A). O nível de atividade de β -

galactosidase induzida pela expressão dos alelos de tipo selvagem e *snf2* foi quantitativamente determinado nos transformantes depois do crescimento durante a noite a 37°C em meio LB de acordo com o protocolo seguinte. As culturas durante a noite foram diluídas 1:1000 em 50 ml de LB e, além disso, cultivadas durante 24 horas a 37°C. Depois do período de incubação, 1 ml da cultura foi centrifugado por 10 minutos a 13000 rpm. A DO₆₀₀ da cultura restante foi medida. O pélete foi ressuspensão em 1 ml de NaCl 0,85 %. 200 µl da suspensão foram ajustados para 1 ml com o tampão Z (Guarentee (1983) *Methods Enzymol.*, 101: 183). Duas gotas de SDS 0,1 % e 1 gota de clorofórmio foram adicionadas. As amostras foram agitadas em vórtex e colocadas em repouso a 30°C por 5 minutos. 200 µl de o-Nitrofenil-β-D-galactopiranosídeo ONPG (4 mg/ml) foram então adicionados à amostra, que foi ainda incubada a 30°C. Quando as amostras obtiverem uma cor amarela apropriada (correspondente a DO₄₂₀ = 0,300-0,600), a reação é parada pela adição de 500 µl de Na₂CO₃ 1M. O tempo de parada foi registrado. A DO₄₂₀ foi medida dentro de 1 hora por espectrofotometria. A atividade de β-galactosidase foi calculada como se segue: Unidades Relativas = $(DO_{420} \cdot 1000) / (DO_{600} \cdot \text{min} \cdot \text{ml})$ (Guarentee, (1983) *acima*).

Como mostrado na figura 5B, a expressão do alelo *snf2* de LHK1 em células de *E. coli* induziu um nível três vezes maior da atividade de β-galactosidase do que o controle, células de *E. coli* não transformadas e células que expressam a LHK1 de tipo selvagem. A adição de citocinina resulta em uma indução de duas vezes da atividade de β-galactosidase em células que expressam LHK1 enquanto células que expressam o alelo *snf2* respondem apenas com um aumento marginal na atividade. Esses resultados demonstram que LHK1 é um receptor de citocinina, enquanto o receptor mutante codificado pelo alelo *snf2* é constitutivamente ativo, cuja atividade é comparável com a atividade induzida pela citocinina do receptor selvagem. Embora não pretenda ser vinculado pela teoria, é proposto que o domínio extracelular CHASE, que normalmente liga a citocinina para ativar a quinase (Kakimoto, (2001) *Plant Cell Physiol.* 42: 677; Anantharaman e Aravind, (2001) *Trends Biochem Sci.* 26: 579; Pas et al., (2004), *FEBS Lett.* 576: 287)

é trancado em uma conformação ativa no receptor mutante (*snf2*). Essas propriedades da LHK1 mutante explicariam tanto natureza genética dominante do alelo *snf2* como os resultados do ensaio de fosforrealocação.

Exemplo 5. Impacto do alelo *snf2* na regulação do crescimento e da diferenciação celular

- A conversão espontânea das células corticais da raiz em nódulos radiculares iniciais vista em mutantes *snf2* é provavelmente devido à atividade constitutiva em uma ou mais etapas da via que controla a diferenciação celular. A realização de cultura de tecido *in vitro* de células *snf2* provê um meio de
- 10 avaliar o impacto global do alelo *snf2* em processos de diferenciação celular e/ou interferência com respostas normais a fito-hormônios. Assim sendo, hipocótilo e explantes de raiz (Figura 1 e 6) de 10 dias de idade de plantas de tipo selvagem e *snf2* (cultivados em ½ de meio salino B5) foram cultivados em
- 15 36 proporções de ácido naftaleno acético (NAA) e 6-benzilaminopurina (BAP) durante três semanas. A faixa de concentrações de citocinina e auxina empregadas inclui níveis normalmente usados para a mais eficiente cultura *in vitro* de explantes de tipo selvagem. Os explantes foram movidos para um meio novo toda semana. Explantes representativos de cada combinação hormonal foram arranjados em uma placa de Petri e fotografados.
- 20 Como mostrado na figura 1G, H e na figura 6, há pouca diferença observável no crescimento do calo em hipocótilo ou em explantes de raiz de plantas com alelos de tipo selvagem e *snf2*. Geralmente, a resposta à dose hormonal é semelhante. Somente em altas concentrações de auxina (2 µg/ml NAA), explantes *snf2* sobreviveram ligeiramente melhor do que explantes
- 25 de tipo selvagem. De modo inverso, em acréscimo de citocinina (BAP) na ausência de auxina, explantes *snf2* desenvolveram menos calo do que os de tipo selvagem. Quando cultivado em placas de cultura de tecido sem qualquer fito-hormônio, o calo verde cresceu a partir do final dos segmentos basais da raiz *snf2*, de uma maneira similar a uma resposta de tecido induzida pela
- 30 citocinina, e claramente diferente de segmentos de raiz de tipo selvagem (figura 1I e J). A cossegregação deste fenótipo do calo verde e do alelo *snf2* mostram que essa formação não induzida de calo radicular foi causada pelo

ganho de função do receptor de citocinina codificado pelo alelo *snf2*.

A microscopia em secções finas de raízes de tipo selvagem e *snf2* (figura1 C,D e E,F) mostra pelo menos uma camada extra de célula de vasos do xilema de raízes *snf2*. Depois de 6 dias da incubação em meio sem hormônio, camadas adicionais de célula do periciclo originadas de divisões periclinais foram observadas em conjunto com um aumento no número de células da vasculatura da raiz *snf2* enquanto nenhuma modificação foi observada em raízes de tipo selvagem (figura 1E, F).

Exemplo 6. Expressão do gene *Lhk1* em órgãos e dos genes *Lhk1*, *Lrr5* e *Nin* em plantas *Lotus* de tipo selvagem e *snf2* em resposta à citocinina

Os níveis no estado de equilíbrio dos transcritos *Lhk1* e *ARR* foram determinados em raízes e outros órgãos da planta pela análise por RT-PCR quantitativa de acordo com o protocolo seguinte.

Sementes de tipo selvagem e *snf2* foram esterilizadas superficialmente e germinadas como descrito anteriormente (Handberg, e Stougaard, (1992) *Plant J.* 2: 487). As plantas foram cultivadas por um período de 16h no claro/8h no escuro a 21°C. O RNA total foi isolado de raízes, nódulos, folhas, flores e vagens de *L. japonicus* de tipo selvagem inoculado. O RNAm foi extraído de raízes de plantas de tipo selvagem e *snf2* cultivadas durante 2 semanas em ¼ de B&D (Broughton e Dilworth, (1971) *Biochem. J.* 125: 1075) suplementado com 0,5 µM de nitrato de potássio. As raízes foram tratadas com 10 µM de BAP em três pontos de tempo, 30 minutos, 3 horas e 8 horas. As raízes não-tratadas foram tomadas como controle. RNAm também foi extraído de segmentos de raiz de tipo selvagem e plantas *snf2* cultivadas durante 10 dias em meio salino B5 sem hormônio e meio salino B5 suplementado com 0,5 µg/ml de BAP. O kit Dynabeads mRNA direct (Invitrogen) foi usado para a extração de RNAm.

DNAc fita simples foi preparado usando transcriptase reversa (Fermentas). PCR quantitativa foi realizada em um Light Cycler (Roche Molecular Biochemicals) usando um Fast Start DNA master SYBR green kit (Roche) para amplificar os transcritos alvo a partir de 5 µl de DNAc diluído. Quatro genes constitutivos (Czechowski *et al.*, (2005) *Genoma analysis* 139:

5) foram usados para determinar a expressão relativa dos genes alvo (veja abaixo). Para cada tratamento, as razões relativas normalizadas dos genes alvo e os quatro genes constitutivos independentes foram calculadas usando o Software de Quantificação Relativa da Roche (Quant). A média geométrica (Vandesompele *et al.*, (2002) *Genome Biology* 3:1) das razões da expressão relativa para as três replicatas técnicas e biológicas e os intervalos de confiança correspondentes a 95% foram calculados como descrito anteriormente (Radutoiu *et al.*, (2003) *Nature* 425: 585).

Os seguintes pares de iniciadores foram usados para a PCR quantitativa:

Sequências dos iniciadores do gene constitutivo independente:

1. PP2A-TC9878-homólogo de AT1G13320-proteína fosfatase 2A

5'-GTAAATGCGTCTAAAGATAGGGTCC-3'

5'-ACTAGACTGTAGTGCTTGAGAGGC-3'

2. UBC-CB828248 homólogo de AT5G25760- enzima conjugadora de ubiquitina

5'-ATGTGCATTTTAAGACAGGG-3'

5'-GAACGTAGAAGATTGCCTGAA-3'

3. TB2C-BI418560- cadeia beta da tubulina

5'-GCTCACCACCCCAAGCTTTGG-3'

5'-TGTC AATGGAGCAAACCCAACC-3'

4. ATP-AW719841-ATP sintase-

5'-AACACCACTCTCGATCATTCTCTG-3'

5'-CAATGTCGCCAAGGCCCATGGTG-3'

Sequências de iniciadores do gene alvo:

Lhk1

5'-AATTTGGTGAACCGAAGGGTCGCCG-3'

5'-TCGACGAGTGGCCTCAAACCCATCC-3'

alelo snf2 de Lhk1

Snf2LC3 Fwd: 5'-AGAGGTCTTAAAGCCATTGTG-3'

Snf2LC6Rev: 5'-TATCAGGCTGAAATAATGCCG-3'

Nin-AJ239041.

5'-AGGAGCCCAAGTGAGTGCTA-3'

5'-GCCATCAAGGTATATGACGAG-3'

ARR5-CB827384-homólogo de AT3G48100-ARR5

5'-TCTTGACTCGAATTGATAGGTGC-3'

5'-GATAGAGATGGCCTGCAACTACTG-3'

A análise do transcrito mostra *Lhk1* a ser expresso em níveis mais altos em raízes, nódulos e folhas, mas os transcritos estiveram presentes em todos os órgãos testados (figura 7A).

- 10 Modificações induzidas pela citocinina em processos celulares em plantas são acompanhadas pela expressão aumentada de genes reguladores de resposta (ARR) que pertencem à classe Tipo-A (Hutchison e Kieber, (2002) *The plant Cell* 14: 47). Dez genes que codificam ARR em *Arabidopsis* e os seus transcritos foram encontrados em todos os tecidos adultos.
- 15 Os genes *ARR* Tipo-A são transcricionalmente induzidos pela citocinina e *ARR4* e *ARR5* são genes de resposta primária rapidamente induzidos. Uma vez que a proteína receptora mutante codificada pelo alelo *snf2* mostrou atividade constitutiva no sistema de teste em *E. coli*, a ativação de um *Lotus* *ARR5* homólogo (denominado *Lrr5*) foi determinado em raízes *snf2* e de tipo selvagem (figura 7). Um nível duas vezes maior do transcrito *Lrr5* foi encontrado em explantes de raiz de mutantes *snf2* incubados no meio B5 sem hormônio em comparação com explantes de tipo selvagem, enquanto a adição de citocinina [BAP] ao meio B5 aumentou duas a três vezes o nível de transcrito *Lrr5* em ambos os explantes, *snf2* e de tipo selvagem (figura 7B).
- 20 O tratamento direto das raízes com citocinina também aumentou o nível de transcrito *Lrr5* em raízes *snf2* e de tipo selvagem, mas nesse experimento, uma diferença na expressão entre raízes *snf2* não-tratadas e de tipo selvagem não foi encontrada (figura 7C).

- 25 Os transcritos do gene *Lhk1* também são vistos aumentando rapidamente em resposta ao tratamento com citocinina no tipo selvagem e mutantes *snf2* (figura 7D), que é compatível com a inducibilidade da citocinina de seu homólogo em *Arabidopsis* (gene *AHK4*). O gene *Nin* em *Lotus*,
- 30

que é necessário para a iniciação do nódulo primordial (Figura 5C), é também transcricionalmente super-regulado pela citocinina (figura 7E), mas os níveis do transcrito em raízes *snf2* não-tratadas e tratadas não foram significativamente diferentes daqueles de raízes de tipo selvagem. A expressão ectópica de *Nin* em explantes mutantes de raiz *snf2* incubados em meio B5 sem hormônio não foi detectada (Fig. 2F).

Em Arabidopsis, a via de resposta *ARR* é dessensibilizada depois da exposição prolongada à citocinina (Rashotte et al., (2003), *Plant Physiol.* 132: 1998), que pode explicar a relativamente pequena ou a falta de super-regulação de *Lrr5* e *Nin* em raízes de mutante *snf2*.

Exemplo 7. O hábito de crescimento da planta *Lotus Snf2* é sensível à citocinina

Embora as modificações transcricionais na planta *Lotus snf2* sejam limitadas, o crescimento da planta foi visto ser fortemente afetado pela aplicação de citocinina fornecida externamente. Assim sendo, as plantas de tipo selvagem e *snf2* foram cultivadas em ½ de salina B5 com e sem aumento das concentrações de BAP durante 3 semanas. O comprimento do broto e da raiz de pelo menos 60 plantas de cada tratamento foram medidos. De acordo com a atividade constitutiva do receptor de citocinina mutante codificado pelo alelo *snf2*, o crescimento do broto e da raiz da planta *snf2* foram hipersensíveis à citocinina em comparação ao de tipo selvagem (figura 8).

Exemplo 8. O fenótipo de raiz *snf2* cossegregado com a mutação *snf2*

As raízes da plantas *snf2*, cultivada em ½ de meio salino B5 sem hormônio, são caracterizados pelo aumento da divisão celular no periciclo e tecido vascular (figura 1 D,F). Este padrão de crescimento, que leva a um fenótipo de raiz inchada, é atribuído à expressão do alelo *snf2*. Para confirmar que o fenótipo de raiz inchada de fato co-segrega com o fenótipo de nodulação espontânea, o alelo mutante *snf2-2* foi retrocruzada com o ecotipo Gifu de tipo selvagem e as plantas F2 [50] foram cultivadas em ¼ de meio B&D durante tempo suficiente para permitir o registro do fenótipo de nodulação espontânea (~5 semanas). As raízes das plantas noduladas e não-noduladas foram cultivadas durante 3 semanas em meio sem hormônio,

e então registradas para a presença e a ausência de inchaço radicular. As plantas de tipo selvagem e *snf2* foram tomadas como controles. Somente as raízes das plantas F2 que desenvolveram nódulos espontâneos mostraram o inchaço dos segmentos radiculares que foi comparável ao inchaço dos segmentos da raiz *snf2*. Isto indica que a expressão da proteína Lhk1 independente de citocinina mutante nas plantas *Lotus* confere à raiz e ao broto *snf2* o fenótipo de crescimento e o fenótipo de nodulação espontânea.

Exemplo 9. *snf2* e a sinalização de citocinina atuam a jusante na via de transdução de sinal induzida pelo do fator Nod.

O fenótipo dos mutantes *snf2* sugere que a sinalização de citocinina seja parte de, ou atue a jusante, da via de transdução de sinal induzida pelo fator Nod (figura 5C). Para testar esta hipótese, o construto do gene *snf2* foi transformado nas plantas mutantes que transportam mutações em genes que constituem a via comum de transdução de sinal compartilhada com fungos micorrizais ou nas plantas mutantes defeituosas em genes a montante ou a jusante.

Assim sendo, sete *L. japonicus* único e duplo mutantes de nodulação foram transformados usando *Agrobacterium rhizogenes* que transportam o construto mutante do gene *snf2* ou um vetor vazio como descrito anteriormente (Stougaard, 1995 *acima*). Todos os mutantes exceto *sym35* foram genotipados por sequenciamento dos produtos de PCR que cobrem a mutação usando os seguintes iniciadores:

nfr1

Nfr1-1 MseI CAPSfw(56) 5'-

CGCTGGTTTACCCATAAACGTGTTC-3'

Nfr1-1 MseI CAPSrv(55) 5'-GGGCAAATGCATTTGTGCTGAG-3'

nfr5

K2fwB 5'-CCAGCTAGGTGATAGCTACG-3'

K2revC 5'-CCAGAAGATGAATGCTGCTTT-3'

S5R rev5 (64) 5'-GGTATTAGAACGCCCCCTGG-3'

symRK (tipo selvagem)

symRK fw1 5'-CTGAGTTTGGACCCCTTTTG-3'

symRK rev1 5'-ACGCCCTTATGAAAATGTGG-3'

symRK (gene mutante)

Lore2 5' LTR fora P 5'-GGAGCTCTGATACCAATGTTAGG-3'

cac41.5F 5'-CGGCAATAGAGCGCTGGAGAGTTG-3'

5

ccamk

LTsym15 (60) fw 5'-TATGACACAGATAGATCAGGG-3'

LTsym15 (60) rev 5'-GAGAGCGGCTCAATGAATGT-3'

nin

Resc 5'-CTCAGAGCACGCTTCTTGGA-3' futuros'

10

Resc rev 5'-ATCATGTGTGCAATCCATGATG-3'

har1

#3fw2 - 5'-CCTGAAATGCCTATTCTTGAG-3'

#3rev2 5'-CACAGCTTCTTCTGCATGCG-3'

snf1

15

Ca2fwd 5'-TGGCTTGCATCCAAACGGC-3'

Ca3-3rev 5'-ACTATTGTTGTCTCACTTTAGTG-3'

snf2

C3fwd 5'-TGGGATAATTGGTTGCTTGACA-3'

C3Brev 5'-TGACAATGTGAGTTCCAGCAG-3'

20

A raízes cabeludas transgênicas foram então monitoradas para a nodulação espontânea na ausência do rizóbio.

Tabela 3

Nodulação espontânea em plantas Lotus não-nodulada mutantes transformadas com o gene *snf2*.

Mutante*	% Nodulação espontânea
<i>nfr1</i>	43 (3/7)
<i>nfr5</i>	41 (24/58)
<i>nfr1nfr5</i>	28 (33/118)
<i>symrk</i>	36 (31/87)
<i>ccamk</i>	19 (18/94)
<i>nin</i>	0 (0/112)
<i>sym35</i>	0 (0/136)

25 * As raízes cabeludas de mutantes transformados foram registradas na au-

sência da *M. loti*. As raízes das plantas mutantes controle transformadas com um vetor vazio não desenvolveram nódulos espontâneos.

A nodulação espontânea observada em *nfr1-1*, *nfr5-2* e *nfr1-1 nfr5-2* receptor de fator Nod, único e duplo mutantes perdendo a resposta eletrofisiológica precoce ao fator Nod (Radutoiu *et al.*, (2003) *acima*; Madsen *et al.*, (2003) *acima*), demonstra que *Lhk1* funciona a jusante da captação de sinalização do fator Nod. As vias comuns mutantes *symRK* desprovidas de rápida oscilação citoplasmática de Ca^{2+} (Stracke *et al.*, (2002) *Nature*, 417: 959; Niwa *et al.*, (2001) *Mol Plant Microbe Interact.* 14: 848) e os *ccamk* mutantes sugerem ser incapazes de interpretar a rápida oscilação citoplasmática de Ca^{2+} (Tirichine *et al.*, (2006) *Nature* 441: 1153), também desenvolvem nódulos radiculares espontâneos em raízes transgênicas transformadas com o construto do gene *snf2*. Nos mutantes *nin* e *sym35* que são suprimidos antes da iniciação da divisão celular induzida através da via comum, nenhum nódulo espontâneo foi observado em raízes transgênicas, mostrando que a captação da sinalização de citocinina atua a montante do início da divisão celular, ou funciona em uma via paralela (figura 5C). Evidências adicionais de um papel central da citocinina e da captação de citocinina na via de transdução de sinal induzida pelo fator Nod amplificada através de uma via comum vêm do efeito aditivo das mutações *snf1-1* e *snf2* em duplo mutantes. Os mutantes *snf1-1* sintetizam uma proteína CCaMK defeituosa em autofosforilação (Tirichine *et al.*, (2006) *acima*) e desenvolvem uma média de sete nódulos espontâneos. Mutantes *snf2* desenvolvem uma média de três nódulos espontâneos, enquanto *snf1-1 snf2* duplo mutantes excedem ambos e desenvolvem aproximadamente dezessete nódulos espontâneos. A conversão de células corticais em células tronco de nódulo ou o subsequente desenvolvimento de órgão parecem, por isso, rigidamente controlados. Essa teoria foi testada pelo cruzamento do alelo *snf2* em um contexto genético de hipernodulação *har1-1* (Krusell *et al.*, (2002) *Nature* 420: 422). Na ausência do rizóbio, *snf2 har1-1* duplo mutantes homozigotos desenvolveram uma média de quatorze nódulos espontâneos, enquanto os mutantes *snf2* desenvolveram uma média de três, e *har1-1* nenhum (figura 9). Esses resultados

indicam que somente algumas poucas células se desdiferenciaram ou que somente poucas células desdiferenciadas sustentam as divisões celulares durante o processo de iniciação do nódulo *snf2*. A autorregulação controlada pelo broto do número de nódulos radiculares (Krusell *et al.*, (2002) *Nature* 420: 422) está assim atuando a jusante da sinalização de citocinina e a citocinina induziu a ativação das células iniciadoras do nódulo radicular (figura 5C). Os dados claramente demonstram que a sinalização de citocinina é necessária e suficiente para a desdiferenciação e a proliferação celular levando à formação de nódulo radicular e que este processo de desenvolvimento pode ser provocado independentemente da citocinina pela expressão do alelo *snf2* dominante. Surpreendentemente, o fenótipo conferido pela expressão do alelo *snf2*, expresso sob controle de um promotor regulado (por exemplo, promotor do gene *Lhk1*) é caracterizado pela organogênese controlada de nódulos radiculares, em vez de proliferação celular descontrolada, independente da citocinina.

Exemplo 10. Kit de detecção para distinguir o alelo mutante *snf2* do alelo de tipo selvagem que codifica a histidina quinase LKH1

Um kit para a detecção do alelo mutante *snf2* útil para o rastreamento de genótipos alvo compreende pelo menos:

Dois grupos alternativos de dCAPS (sequência polimórfica amplificada clivada derivada) iniciadores são providos para a detecção do tipo selvagem versus a mutação:

Grupo um:

Snf2 XmnI/Asp700 dCAPS fw(74)

5'-CACCAAAATTGCTTGGTTACCAGCAAGTTGACCGA

Snf2 XmnI/Asp700 dCAPS rv (69) 5'

CCCTTCATGTGGCCCTTACCCAAC

Os iniciadores são convenientes para o uso em um teste PCR realizado com DNA genômico como molde, onde a amplificação é realizada a 60°C com 30 segundos de alongamento e 35-40 ciclos. O produto de 225 pb é clivado pela *XmnI/Asp700* no mutante, e não é clivado no tipo selvagem. Os produtos de PCR clivados podem ser identificados depois de sua

separação em gel de agarose a 1%.

Grupo dois:

Snf2 RsaI dCAPS fw (72)

5'-GATCCTTTGATGTTGAGTCCCTTGTGGAGAATGTA

5 Snf2 RsaI dCAPS rv (62)5'-

CTGAAACTCAGAAAATGTACTACAAC

Os iniciadores são convenientes para o uso em um teste PCR realizada com DNA genômico como molde, onde a amplificação é realizada a 48°C com 30 segundos de alongamento e 35-40 ciclos. O produto de pb é clivado por bu RsaI no tipo selvagem, mas não é clivado no mutante.

Os produtos de PCR clivados podem ser identificados depois de sua separação em gel de agarose a 1%.

Exemplo 11. Iniciadores de amplificação para clonagem da sequência de codificação do DNAC do *snf2*.

Os iniciadores para amplificar a sequência de codificação da partida para parada (DNAC) são:

Snf2 DNAC fw ATGGGTCTTGGGTTCAAGATGCA

Snf2 DNAC rev TCATGAGTCTGAAGCAGGCTTGG

A PCR foi realizada com o anelamento a 58°C e tempo de alongamento de 3 minutos e 20 segundos por 25 ciclos com o clone de DNAC como molde.

LISTAGEM DE SEQUÊNCIA

<110> Aarhus Universitet, Ledelsessekretariatet
 Tirichine, Leila
 <120> MOLÉCULA DE DNA QUE CODIFICA UM POLIPEPTÍDEO DE HISTIDINA QUI-
 5 NASE MUTANTE E SEU USO RELACIONADO, POLIPEPTÍDEO DE HISTIDINA
 QUINASE MUTANTE, BEM COMO MÉTODO DE PRODUÇÃO DE UMA PLANTA GE-
 NETICAMENTE MODIFICADA
 <130> P80600597DK00
 <160> 20
 10 <170> PatentIn version 3.2
 <210> 1
 <211> 12738
 <212> DNA
 <213> Lotus japonicus
 15 <220>
 <221> promoter
 <222> (1)..(2359)
 <223> promoter + 5' UTR of transcript
 <220>
 20 <221> exon
 <222> (2360)..(2704)
 <223> start of CDS in 1st exon
 <220>
 <221> Intron
 25 <222> (2705)..(3024)
 <220>
 <221> exon
 <222> (3025)..(3258)
 <220>
 30 <221> Intron
 <222> (3259)..(3390)
 <220>

<221> exon
 <222> (3391)..(3560)
 <220>
 <221> Intron
 5 <222> (3561)..(3834)
 <220>
 <221> exon
 <222> (3835)..(4053)
 <220>
 10 <221> Intron
 <222> (4054)..(4145)
 <220>
 <221> exon
 <222> (4146)..(4332)
 15 <220>
 <221> Intron
 <222> (4333)..(4420)
 <220>
 <221> exon
 20 <222> (4421)..(4472)
 <220>
 <221> Intron
 <222> (4473)..(4586)
 <220>
 25 <221> exon
 <222> (4587)..(4816)
 <220>
 <221> Intron
 <222> (4817)..(5037)
 30 <220>
 <221> exon
 <222> (5038)..(5136)

```

<220>
<221> Intron
<222> (5137)..(5308)
<220>
5  <221> exon
    <222> (5309)..(6024)
    <220>
    <221> Intron
    <222> (6025)..(7042)
10 <220>
    <221> exon
    <222> (7043)..(7513)
    <220>
    <221> Intron
15 <222> (7514)..(8306)
    <220>
    <221> exon
    <222> (8307)..(8564)
    <220>
20 <221> Intron
    <222> (8565)..(12738)
    <223> 3'UTR + 3 prime of transcript
    <400> 1

cccgggattg ttcatatccc catgtgtgtg tttggtttgg acacaagcaa gcttgccaca    60
25 gtgggtgttta caactcattc tgaattatat atatatatat atatatataa ctaatctcta    120
   tcaataatTT agttagtttt atgggataat atgattcccg tcaaattagc ttgtaggtta    180
   aaaatccata tttatataat aaattagact tatctaaatc acttaagtaa atatgatgta    240
   aaatcacaat agcttaggta gttaacaat gttttatcat gaatttaaaa aaaattagtt    300
   gtttcaaaat attattggac caaataagca aagaggagct aatccaatta taaacagcac    360
30 atctttataa ttgttttcta ctcgatctaa tattggttcc ttctttcctt gagaaaaaaaa    420
   ggggttaaaaa tgcgatttgt tttgtaaaat taatttgtga cattaacaag agaaaactct    480
   acttttttaa tttttaagaa gttacttaaa tatagaaaga aagaagttac ttaaatatga    540

```

	aacttgtcta accaaagtca accgatcaaa gcgttttggtt agcaaagaga acttaaaaaa	600
	ataactaaat taactgtttg aacagtgatg taacagagaa atagaatggc ttacatatgg	660
	attatggagt taatggaagg cacactagtt ttagttcacg gcaacaagac agaattatcc	720
	tcctatattt taggaggttt tcttatttat ttttaatgta agacgtgtat ctttcttttg	780
5	aggtaataat gttcgagctt gaaacattca cattatagtg tgactaattg cttaatatga	840
	atcaagcatg taccacttga aacactcact catcgtttat aatataatta ctttatagta	900
	ctataaagta tgtaattatt ttctcctttt taagtatctc aatcattttt ttgacccaaa	960
	tattatttta gagaagaaac atgtcttaag accaatacat caacatatat ggagagtagt	1020
	ttgacaacct ctcatgagct ttctaaagcc aaacaaatgt actcacataa aaaaattaaa	1080
10	aatttacagc aagtgagtta taaagagtga ttaataaatg ttgaataacg atggagggta	1140
	gttatgaaat tccataaata aagcaaagga tgtttatggc aattgacatg ggaataagat	1200
	ccgcgcgttg tcgccgtgtc ataatcgctc agatttgtga tagcgagaga tttccattct	1260
	tttcctcttt tttcgaacga acgaacgaac gaacacagta gcagctgtat taggattcag	1320
	attgcatgat acattgatat tgatattgat attgatattg atattgttgt ggggtcaagtc	1380
15	tctactctac tatccaaagt aagcatatat atagagagag agagagagct tgggactagg	1440
	gaggatatta gcttatgtga ctgtgaagtt gaaggagagac aagagcgtat ctggcaaat	1500
	cctcaaaata aaatactagt actagtacag agaaaagaga ctaagagaga gagtgtgtct	1560
	gcacatcaag acccattgtg atttgtgatt tgtgatttgt gatttgtgat ttgtgagtga	1620
	gttcattgta caggatttat tgtttgttgt ttctctcttc aaccaccctc taaagtctaa	1680
20	tctaactcat tgggctctgt gcttagctgg ttgtgttttg tgtatggtga attagggggg	1740
	caatctctgg ttttcatcat tattatatta tatatgggaa taccgtgctc tcttctttgt	1800
	ccttcaccaa aactagtttc acgcctacac aacatgatta gagcctcttc atttttttaa	1860
	tctcatcctt taagtgtatt tctatttcta ttggctattg ggacaagggg aagggtggtgc	1920
	ttcttaggaa cttgagctgt tttcatctt ttgagaccca tgctttgtct ctctcatttt	1980
25	taattctggg tctctttctt ctcttgcctt gattttttta atgtgcttct tttttgcttc	2040
	ttacaaccac cctctaaacc attcatcatg cttggtttgc ttttgcttct cctttcacag	2100
	gtttcaatca cgcaaaacaa tgctgcaatg atgctgtact tggagtttct tctgtgaccc	2160
	cttttttcct tccttcaaca atcaaccac cagagaaaag tgtctcagat ttgagacta	2220
	ctttcaactt tcaaaacaaa gtggatggga tcttcatctt atataaccac acatcaatca	2280
30	tttgtgctac ttctccaatt ttctttagag atgaaatgaa gagctaagca gacaagacaa	2340
	gtttatttgt ttgttgctg atg ggt ctt ggg ttc aag atg cag cag agc cac	2392

Met Gly Leu Gly Phe Lys Met Gln Gln Ser His

	1	5	10	
	cac cct gtg gct ttg aag tta cat gag caa gct ggg agc cag aga aag	2440		
	His Pro Val Ala Leu Lys Leu His Glu Gln Ala Gly Ser Gln Arg Lys			
	15	20	25	
5	ttc act ttc att cag aac ttc aga aac tgg ttt cta ccc ctt ctg ttt	2488		
	Phe Thr Phe Ile Gln Asn Phe Arg Asn Trp Phe Leu Pro Leu Leu Phe			
	30	35	40	
	gta tgg ttc att gtt atg gct gca ttt ggt gcc tgc atc tac cat aaa	2536		
	Val Trp Phe Ile Val Met Ala Ala Phe Gly Ala Cys Ile Tyr His Lys			
10	45	50	55	
	atg gat gct gaa act aaa gtc aga agg aaa gag gtg ctg ggt agc ctc	2584		
	Met Asp Ala Glu Thr Lys Val Arg Arg Lys Glu Val Leu Gly Ser Leu			
	60	65	70	75
	tgt gat caa agg gct aga atg cta caa gac caa ttc agt gtc agt gtc	2632		
15	Cys Asp Gln Arg Ala Arg Met Leu Gln Asp Gln Phe Ser Val Ser Val			
	80	85	90	
	aac cat gtc cat gcc ctt gcc atc ctt gtt tca acc ttc cat tac tac	2680		
	Asn His Val His Ala Leu Ala Ile Leu Val Ser Thr Phe His Tyr Tyr			
	95	100	105	
20	aga aat act tca gcc att gac cag gtttgtgctt gattttcctt tccttgaagc	2734		
	Arg Asn Thr Ser Ala Ile Asp Gln			
	110	115		
	atttttttagt tggaggctca atttcttttt ctgatttgat tctggcctta aaaattagaa	2794		
	tcaattgtag aaggatttcc aaacatgccc atttttgaaa ttggtgcatc tgatagtatc	2854		
25	atgttttagat cagtttcttt ttctcagaa ttgattttgg gcttaaaatc aattgtggaa	2914		
	ggatattcat tagtaatttg gatattgttg catcatatgg ttctatctag ttacatcatt	2974		
	tttttccact ctgattgcat gtatctttct cctgttcttt tccctatcag gaa acc	3030		
			Glu Thr	
30	ttt gca gaa tac acg gcc agg aca gca ttt gaa cgg cca tta atg agt	3078		
	Phe Ala Glu Tyr Thr Ala Arg Thr Ala Phe Glu Arg Pro Leu Met Ser			
	120	125	130	

	ggg gtg gcc tat gca cag aga gtg gtt cac tca gag aga gaa aga ttt	3126
	Gly Val Ala Tyr Ala Gln Arg Val Val His Ser Glu Arg Glu Arg Phe	
	135 140 145	
	gag aag caa cat ggg tgg gtt ata aag aca atg gaa aga gtg cct tca	3174
5	Glu Lys Gln His Gly Trp Val Ile Lys Thr Met Glu Arg Val Pro Ser	
	150 155 160 165	
	ggg gtt agg gat gag tat gca gca gtg ata ttt gca cag gaa act gtc	3222
	Gly Val Arg Asp Glu Tyr Ala Ala Val Ile Phe Ala Gln Glu Thr Val	
	170 175 180	
10	tct tac ctt gaa tct att gat atg atg tct ggg gag gtaaatgtca	3268
	Ser Tyr Leu Glu Ser Ile Asp Met Met Ser Gly Glu	
	185 190	
	acacttgta attaatgta aaactcagaa gctactcaga gaagctcttc cccagaattg	3328
	gttctgcctt tagaataaat tgtacatgga tttgaccaca ttttctcatt tgcattgatgc	3388
15	ag gag gac cga gag aac att ttg agg gct aga gcc act ggg aaa gct	3435
	Glu Asp Arg Glu Asn Ile Leu Arg Ala Arg Ala Thr Gly Lys Ala	
	195 200 205	
	gtt ctg act agc cct ttc aga ctg ctg gat tct cat cat ctt ggc gtg	3483
	Val Leu Thr Ser Pro Phe Arg Leu Leu Asp Ser His His Leu Gly Val	
20	210 215 220	
	gtt cta aca ttt cct gtt tat aaa tct aag ctc cct cca gag cca acg	3531
	Val Leu Thr Phe Pro Val Tyr Lys Ser Lys Leu Pro Pro Glu Pro Thr	
	225 230 235 240	
	acg gaa gag gtc att aaa gcc ata gca gg gtatgtcctc atttcacttt	3580
25	Thr Glu Glu Val Ile Lys Ala Ile Ala Gly	
	245	
	tcttgccaaa accagacttc tatttggttg tgtttccgta ggctatgact gatattgtagt	3640
	ttcaactcag ttagactata atataaaccc ttcatgtggc ccttaccxaa cagcttaagc	3700
	ttttgggata attggttgct tgacaaactc cttccgtaga aaacttggtt agctttggtt	3760
30	ctatgtgggc tttatgtttt ccctgagctt atgtaatagc atgatgtgtt taatgtactt	3820
	tttaatggaa acag a tat att gga gga tcc ttt gat gtt gag tcc ctt gtg	3871
	Tyr Ile Gly Gly Ser Phe Asp Val Glu Ser Leu Val	

		255	260	
		gag aat tta ctt ggt caa ctt gct ggt aac caa gca att ttg gtg aag		3919
		Glu Asn Leu Leu Gly Gln Leu Ala Gly Asn Gln Ala Ile Leu Val Lys		
		265	270	275
5		gta tat gat ata aca aac tct agc gac ccc cta atc atg tat ggc agc		3967
		Val Tyr Asp Ile Thr Asn Ser Ser Asp Pro Leu Ile Met Tyr Gly Ser		
		280	285	290
		caa tat gaa gag ggt gat atg tct ctt gtc cat gaa agt aag ctt gat		4015
		Gln Tyr Glu Glu Gly Asp Met Ser Leu Val His Glu Ser Lys Leu Asp		
10		295	300	305 310
		ttt gga gat cca tac agg aaa cat cac atg atc tgt ag gtgggtgctt		4063
		Phe Gly Asp Pro Tyr Arg Lys His His Met Ile Cys Arg		
		315	320	
		ctagttattg ttgtagtaca ttttctgagt ttcagtggtt tatcaattat cagcagattc		4123
15		ttatgatcaa tttttttaac ag a tat cac caa cag gca cca aca aat tgg		4173
		Tyr His Gln Gln Ala Pro Thr Asn Trp		
		325	330	
		ata gca tat acc acg gca ttc cta ttc ttt gtg att ctt tgt tta gtg		4221
		Ile Ala Tyr Thr Thr Ala Phe Leu Phe Phe Val Ile Leu Cys Leu Val		
20		335	340	345
		ggt tac att tta tat gct gct gga act cac att gtc aag gta gaa gat		4269
		Gly Tyr Ile Leu Tyr Ala Ala Gly Thr His Ile Val Lys Val Glu Asp		
		350	355	360
		gat tac aat gca atg cag gat tta aaa gtc aaa gca gaa gca gct gat		4317
25		Asp Tyr Asn Ala Met Gln Asp Leu Lys Val Lys Ala Glu Ala Ala Asp		
		365	370	375 380
		att gcc aag tca cag gtacttttca tgacatgtta gcaactgttcg ttatttcctt		4372
		Ile Ala Lys Ser Gln		
		385		
30		gaattgcata ctgatcacta gaaactgaaa atttggttatt aatgtcag ttt cta gct		4429
		Phe Leu Ala		

	acc gtc tct cat gaa att aga act ccc atg aat gga att tta g	4472
	Thr Val Ser His Glu Ile Arg Thr Pro Met Asn Gly Ile Leu	
	390 395 400	
	gtaactttaa gattctctct cgttctcttt ccactgaaaa gcaacacatg ctttcattcc	4532
5	atacctgata ctttcccatt agtgatgcta tcgttaaact ccttgtcact gtag ga	4588
	Gly	
	atg ctt ggt ctg ctt tta cgc aca gaa ttg agt tca aca caa aga gac	4636
	Met Leu Gly Leu Leu Leu Arg Thr Glu Leu Ser Ser Thr Gln Arg Asp	
10	405 410 415	
	tat gct cag act gct caa gca tgt ggg aag gca cta ata gca tta ata	4684
	Tyr Ala Gln Thr Ala Gln Ala Cys Gly Lys Ala Leu Ile Ala Leu Ile	
	420 425 430 435	
	aat gag gtg ctt gac cga gct aaa att gaa gca ggc aaa tta gag cta	4732
15	Asn Glu Val Leu Asp Arg Ala Lys Ile Glu Ala Gly Lys Leu Glu Leu	
	440 445 450	
	gaa gca gtt cca ttt gac ctt cgt tcc ata ctt gac gat gtc ctt tct	4780
	Glu Ala Val Pro Phe Asp Leu Arg Ser Ile Leu Asp Asp Val Leu Ser	
	455 460 465	
20	ctt ttt tct gag aag tca aga cac aaa ggc tta gag gtacgtttag	4826
	Leu Phe Ser Glu Lys Ser Arg His Lys Gly Leu Glu	
	470 475	
	tcattgctaa atctgttggtg aagtctgtga caagtggcgt aatttcatag tcctatttcc	4886
	ttttcttcaa ttgatgatct tatttcatat tcctctcgtg tttctctctt tctatgttgc	4946
25	catgtcgttg ggcggtgatg ttctactct gatcccaaatt tcctgatgtg caattattcc	5006
	gacttggact tcaatggttt gggaagtata g ctg gca gtg ttt gtt tct gac	5058
	Leu Ala Val Phe Val Ser Asp	
	480 485	
	aaa gtt ccg gat ata gtt atg ggc gat cct ggg cga ttc aga caa ata	5106
30	Lys Val Pro Asp Ile Val Met Gly Asp Pro Gly Arg Phe Arg Gln Ile	
	490 495 500	
	gtg aca aat ctt gtt gga aac tct gtt aag gttagtggaa ttttcaaact	5156

	Val Thr Asn Leu Val Gly Asn Ser Val Lys			
	505	510		
	ttatttgcct aatgttgtgt gcaagttgtg tgttggaaat gcgtcctttt aacgttataa	5216		
	aatcgtacaa gttcgtattc tccattgtat acaataactt attagcaaag tacttgttga	5276		
5	tatcattact gattaacttt aatatcttgc ag ttc act gag cga ggt cat ata	5329		
	Phe Thr Glu Arg Gly His Ile			
	515			
	ttt gtt aaa gtc cat tta gct gaa aaa aga cag tgc aca atg aat gga	5377		
	Phe Val Lys Val His Leu Ala Glu Lys Arg Gln Cys Thr Met Asn Gly			
10	520	525	530	535
	aaa tgt gag act ttt cta aat gga ggc tgt gat gat gtt ttg cat gta	5425		
	Lys Cys Glu Thr Phe Leu Asn Gly Gly Cys Asp Asp Val Leu His Val			
	540	545	550	
	tct ggc agt tat aat ttg aaa acc ctt agt gga tat gaa gcc gct gat	5473		
15	Ser Gly Ser Tyr Asn Leu Lys Thr Leu Ser Gly Tyr Glu Ala Ala Asp			
	555	560	565	
	gaa cgg aac agc tgg gat aat ttt aag cat cat att gct gac gaa gaa	5521		
	Glu Arg Asn Ser Trp Asp Asn Phe Lys His His Ile Ala Asp Glu Glu			
	570	575	580	
20	ttt ttc ttt gat gct tcg gtt aaa aag ttg gcc tct agt gaa tct tat	5569		
	Phe Phe Phe Asp Ala Ser Val Lys Lys Leu Ala Ser Ser Glu Ser Tyr			
	585	590	595	
	gag caa gtc acc ttg atg gtc agc gtg gag gac act gga att ggg att	5617		
	Glu Gln Val Thr Leu Met Val Ser Val Glu Asp Thr Gly Ile Gly Ile			
25	600	605	610	615
	tct ttc tct gcc caa gat agt att ttc atg cct ttt gtg cag gct gac	5665		
	Ser Phe Ser Ala Gln Asp Ser Ile Phe Met Pro Phe Val Gln Ala Asp			
	620	625	630	
	agc tca acc tct cga aac tat ggg ggt acc ggg atc ggc ttg agt atc	5713		
30	Ser Ser Thr Ser Arg Asn Tyr Gly Gly Thr Gly Ile Gly Leu Ser Ile			
	635	640	645	
	agt aag tgc ttg gtt gaa ctg atg ggc ggt cag ata aac ttc ata agc	5761		

	Ser Lys Cys Leu Val Glu Leu Met Gly Gly Gln Ile Asn Phe Ile Ser	
	650 655 660	
	cga ccc cag gtc ggg agc acg ttt tca ttc act gca gat ttc gga aca	5809
	Arg Pro Gln Val Gly Ser Thr Phe Ser Phe Thr Ala Asp Phe Gly Thr	
5	665 670 675	
	ttt aag aaa aac tca aca act gac atg aag aaa ctt aac ttt gaa gat	5857
	Phe Lys Lys Asn Ser Thr Thr Asp Met Lys Lys Leu Asn Phe Glu Asp	
	680 685 690 695	
	cta cct tct agt ttt aga ggt ctt aaa gcc att gtg gtt gat gga aaa	5905
10	Leu Pro Ser Ser Phe Arg Gly Leu Lys Ala Ile Val Val Asp Gly Lys	
	700 705 710	
	cct gtt aga gct gca gtg act aga tac cat ttg aag aga cta ggg ata	5953
	Pro Val Arg Ala Ala Val Thr Arg Tyr His Leu Lys Arg Leu Gly Ile	
	715 720 725	
15	caa gct aaa gtt gca att agc atc aat aag gct gtt tct tta tgt ggg	6001
	Gln Ala Lys Val Ala Ile Ser Ile Asn Lys Ala Val Ser Leu Cys Gly	
	730 735 740	
	aaa aat ggt tct ttg acc tcg gc gtaagtcttt aattaacctt tttggtttca	6054
	Lys Asn Gly Ser Leu Thr Ser Ala	
20	745 750	
	attatgtaga aatgtattga atgttatgat aatcagtagc attatcaact tttagtaatt	6114
	gttcttaaca tatgctaata gtcatatctt tctataatac tacaatactg tagccatata	6174
	atatctttcc tgtattggag tgagttttca aatgtttttc tgtgatattt tgggaagttat	6234
	cttcagtttg agaactcatt tgtcattttt gcattttgtt attggatatt tggatggatc	6294
25	tttaciaaagg atgtgtggat ttgacttgt tatacacatt tcttctccat tttatattgt	6354
	ttgtgttatt ctttttactc ataaagaaat ttagaaactg cattgactgg ttctttttaa	6414
	ttacttacag atattgacat tgatattttt tgtaaagtct gtcttgacat ggtttaatta	6474
	cttacagact aggtttttct ttccttttct aacatgcata tccatttact tttttgacca	6534
	accaacatcc tcatgagtca tgacatgttg atgatttata tggttgactt gagactatct	6594
30	agacattaaa taaccgcaaa ttccatgttg tttgtgtgtt tggttccctg ttgggtaatc	6654
	tcagaatcaa ttatgataga gtaaaatcaa ttttgatga gatgtgtggg tgtcattttg	6714
	taaacctaaa cccaaaatca attctgctat aagctagaga gagtagttga acataatcaa	6774

	ttgtgagatt ttgcaagtgg attgcacaac attgccctat gaaaatcact ttttgttcac	6834
	aaaatttatc taaacataca taacttcatt ttcaaccttt actataatca attttacaat	6894
	aattaatttt acccaaaatc aattgtgaca atgagtttcc aaacacacac ttaaagacta	6954
	ccatttgcag aaaatatgtg atagaagact tatgtttatg tagtgtgttt cagttcattc	7014
5	actgatttaa actactcgga ttttgcag a tta ttt cag cct gat att att ttt	7067
	Leu Phe Gln Pro Asp Ile Ile Phe	
	755	
	gtt gag aag gac tct tgg gtt tct gga gag gat ggt ggt atc ttc aat	7115
	Val Glu Lys Asp Ser Trp Val Ser Gly Glu Asp Gly Gly Ile Phe Asn	
10	760 765 770 775	
	gcg ttt aag atg cct caa atg atc ctt ctt gca acc aat atc tgt aac	7163
	Ala Phe Lys Met Pro Gln Met Ile Leu Leu Ala Thr Asn Ile Cys Asn	
	780 785 790	
	gct gaa ttt gat aaa gcc aaa gct gca ggt ttc agt gat aca gtg atc	7211
15	Ala Glu Phe Asp Lys Ala Lys Ala Ala Gly Phe Ser Asp Thr Val Ile	
	795 800 805	
	atg aag cca ctg aga gct agt atg ctg gct gct tgt ctt cag caa gtt	7259
	Met Lys Pro Leu Arg Ala Ser Met Leu Ala Ala Cys Leu Gln Gln Val	
	810 815 820	
20	ttc ggg act ggc aag acg agg cag ttt ggg aaa gac atg tcg aat ggt	7307
	Phe Gly Thr Gly Lys Thr Arg Gln Phe Gly Lys Asp Met Ser Asn Gly	
	825 830 835	
	tct tca gta cga agc ctt ctt tgc gga aag aaa atc tta gtg gtt gat	7355
	Ser Ser Val Arg Ser Leu Leu Cys Gly Lys Lys Ile Leu Val Val Asp	
25	840 845 850 855	
	gat aat ttg gtg aac cga agg gtc gcc gcc ggc gcg ttg aaa aac ttt	7403
	Asp Asn Leu Val Asn Arg Arg Val Ala Ala Gly Ala Leu Lys Asn Phe	
	860 865 870	
	gga gct gat gtc aaa tgt gca gca agt ggc aaa gct gct ctt gaa atg	7451
30	Gly Ala Asp Val Lys Cys Ala Ala Ser Gly Lys Ala Ala Leu Glu Met	
	875 880 885	
	ctt caa tat cct cac gat ttc gat gct tgc ttc atg gat att caa atg	7499

	Leu Gln Tyr Pro His Asp Phe Asp Ala Cys Phe Met Asp Ile Gln Met	
	890 895 900	
	cca gaa atg gat gg gtatgcttac tggcactgac taatacatgt tttttgccaa	7553
	Pro Glu Met Asp Gly	
5	905	
	cttaatatat tactctttca atattcgttg tggtattaga agatcatata gattaattta	7613
	taaattttct tttagcaaaa ctttatcaat taagtgtgta gaaaagtcag tctcacatta	7673
	tggtcaaata agtgtttaggg caagcttcac ctcaaagcta gctatttggg tagatttagg	7733
	cctaaccgga attctaagat ggtatcagag tctatcctag atctttttat tggaaaccac	7793
10	ccgtatatga gcaactcgta gatattcatt cttgaaagtt gcacgctcca tatgtccatt	7853
	cctaggtgcg agagagaagt ctactttga ctagagatat gattaataaaa atatttataa	7913
	agggttgagc aatcctcacc tcagagctaa gcttttgggg taaagttagg cctaactcga	7973
	actctaataa agtgtttagc tggtgtgtca actgtcaata tgaaatcttt tgcaatttac	8033
	tatgcattca cttacctact ttattgaagc ttattgacaa tttgtgcaga agcatcatta	8093
15	attaggaaca tgtagctat acaagttatg atgtttttgt atagcatatc atgttccaac	8153
	cttccaataa caaaatatgt ggttcaagtg tgagaatata taggttaaac aataaagtat	8213
	tgagttaaca gaaatctaaa cacacgctgt cactagctct tcatattgag acatgcatgg	8273
	gatttgacaa aacatctgaa taaatatttg cag g ttt gag gcc act cgt cga	8325
	Phe Glu Ala Thr Arg Arg	
20	910	
	att cgg atg atg gaa aga gag gca agt gag cag ctg aaa agt gaa tct	8373
	Ile Arg Met Met Glu Arg Glu Ala Ser Glu Gln Leu Lys Ser Glu Ser	
	915 920 925 930	
	ggg gaa gaa aat ggt aag aaa agt gag ttc cac atg cct ata ttg gcc	8421
25	Gly Glu Glu Asn Gly Lys Lys Ser Glu Phe His Met Pro Ile Leu Ala	
	935 940 945	
	atg aca gct gat gta atc cat gct aca tat gat aag tgc tta aat tgt	8469
	Met Thr Ala Asp Val Ile His Ala Thr Tyr Asp Lys Cys Leu Asn Cys	
	950 955 960	
30	ggg atg gat gga tac gtc tca aag cct ttt gaa gaa gag aat ctc tat	8517
	Gly Met Asp Gly Tyr Val Ser Lys Pro Phe Glu Glu Glu Asn Leu Tyr	
	965 970 975	

	caa gca gtt gca aag ttc ttc aag tcc aag cct gct tca gac tca tg	8564
	Gln Ala Val Ala Lys Phe Phe Lys Ser Lys Pro Ala Ser Asp Ser	
	980 985 990	
	acactgctta ttctgcagaa cagggtcaacc aacttttgat tgagaaacat ttagtgtag	8624
5	catgtttgga tcaacttctc ccagcatcaa ttctgaaact cagaagctac tcatagaagc	8684
	ttctagccag aattgatttt ggcttcagaa gctactaatg gttttatgta gagagcaaaa	8744
	gttggtttcc aagatatgcg aggatccatg atgatacaca cgtctgaagt gatcatttct	8804
	aaccagttga agtttcactc gacgtgattt gaatccaagt aaatgcatac cacataatta	8864
	tccatcccca tcttgtgtac agattctccc aagggataag aaaatttatg taaattcaat	8924
10	ttttttcttt tgcattctca tacttccctg ttagaacttt tccctatga ttatttccac	8984
	ttttcatttt caattatatt ttttgtaaaa ttggtctcca tcctataagg ttgtcttagt	9044
	tttttttaat tcagatgaga aggttggtgc ttataatgtg tatacctttt tagcagtact	9104
	tgtgatataa atatttgtct tatttttaga acttaggaga taataagatc gtaggagtaa	9164
	atagcatttt tggagaatga tatcataatt ataatcatct acattctgca gttataaaaa	9224
15	aaatcaatat taatgcatag caagtagctt taagaagtat gatctattta agtattgaat	9284
	tttttttcta aactaattac aatgtattat tgctgggata tgcgcttttt actcgataga	9344
	tatcatgttt ggattagttt ttcttagaat ccattttgga accgaaaaaa ttatcacata	9404
	atttattctt caaagttgat ttgggcaatc aaggtaaaca ctcataaaag cttaaaggag	9464
	ttgtttctta ttattgtgat tttagctcat tgtggttttt caccgtgtat ataaacatgc	9524
20	atttgatgtg acgttgactt gtgaagcacg tggctatcca ccagcaatct caagtaacac	9584
	ttgtcagcac tatgttgcta ttataagtga caaggcagtg accacacccc tctatatatg	9644
	cataaacttt tttattaaaa gtaacagaat caaaagggaa gttttctaca aattaattaa	9704
	tggaaaatta tcagaaccag tatggagctg ttctgtctag cagggctccg gttcagagcc	9764
	agagtgaat attggaacc accaatgttg gcggtgctac acaagtcata agaaatgatt	9824
25	atgatggaag aaccactgct gcagatgcta ctgttggttg tgagagcgat ctgcagcatc	9884
	accacaagaa aggaatatta gagaagatca aggagaagct tcctggaaca caccaccatc	9944
	aagatcacia atagctagct agctttgttt agtttatttt gcttctatca ttataaatgt	10004
	aataataata gtatgtgtgt tgctacatgc atgtgtatgt atcgtttagt cagctagttt	10064
	gttaataact tgctttcaat tttgcttcct tacaagccta tttatgcaat gtatttggtc	10124
30	atttcactat tggtttctaa tcaatctaaa ttagtacatg gtcctcttcc ttttagtgcc	10184
	ttttaattaa agagtaaagtg ttgaatattg tttttagtca ttctatttac atttttattt	10244
	tttttgcttc atcgaaaaat tatttacatt tttattaagt gtgtggttag attttcgttt	10304

4	ttcgttaaaa atttcataat taatTTTTTA taacagtgat gtatgcagct aaggcatcct	10364
	aacagacctt tctttgtatg tatggcattc tgggcgggtc ggattttattt ttagcccat	10424
	aattccccac tactatacat aaaatgatac ctgctagtaa aataatttaa agtgtttcga	10484
	agtcataattg gtaaattgat ttcagttaaa aatcgaagtt tgtagctac taaagacggt	10544
5	agcttctatg tcaaaatagt taattcctat attttataat tggatcattt tttctaaaat	10604
	ttattcaaac atctcaactc actttgacta taatcacttt ataattaatt ttgttgaatt	10664
	aaataaatta tattcaaatt aaatgtgata ataccactg aatgtgattg gaatgccatc	10724
	agggaaatat ggaacaaaga gagctgcatg tggaagggct tgttatgtct gcaaaacaga	10784
	caaagaagtt gtaatcacia ttaggacaat tgccaaagat gtggggtcac attgacttgt	10844
10	cagggaaagt ctatagatcc actgtgctct taaagctaga acaattaaca taaaatttct	10904
	gccaattaag gctttcactt cctatcccta tctgctgtat attttaagtt tgagtaaagg	10964
	gacagttgtg gtgatgttat aatcaactgc aggctgaggg aaaagttggg ggtcaagtgg	11024
	tgtctatata aactTTTTTg tcattattaa tctttaatct ttttggtttt gtaaaatgac	11084
	agcttttgtc ttgtgctgaa agttgactgt gctagctagt aaagtcatgc tggtttctta	11144
15	attgattgtg gtgtttttat gcgacaactg tagacattgc aatgcaggca cacaagtcaa	11204
	gaggcaatta caaacaaaat acagctatta ataattgacc tactaactac taaggaggga	11264
	atttaaattgt gatTTTgaaa tcttatattg gatgtaaaga ggcaatactc atcagactgt	11324
	aaagtgcag aacaagtaga tggtaaacia gtagaagtag aagcagcaac aaattgagta	11384
	aagaccatca atgtcaacac ttccttgag aagaatttca taagtgtcat gtgaaaataa	11444
20	ctagacaaaa tgtagccaaa atgaagtatg aaactcagtt attgaaacga gtgacactga	11504
	ctagattcca agttattgaa ggaaccagca aaaggTTaag tgacagtga gtgtgtgggt	11564
	tgaagagga ttgaataaaa gctgaaccaa tggactgaat aggcaactt gacaatttcc	11624
	cattagaatc tgttttaact tccttgata aaagtttctt cggtgaaatt tccagcattg	11684
	ttggtaacat gatgagtagc tcatgaatct ggaaaccaa cagttttgtt tgttcaagct	11744
25	gtagagaaga atgagaacta catgtgagca ttgttatgta ccaaaaaaac ctgtgagcat	11804
	tgttgctgca atcctggatc agtacgata gcctgagcac aaagatgttg aattgaagta	11864
	gttgctctaa tatgttatga atattcaaga ttaaaacatt ttttaagttat catgtgacaa	11924
	tcacataact attaaattat tttaatctta actattaatc attagatcta acgatcatat	11984
	tctcatataa aataattatt ctgatacaat aatatatata tatatatata tgtatgtagg	12044
30	ggatctgagg atgtatcata tgagaaatgt gatattatgt gagaacatga gaataaatca	12104
	caaccgtag atttaatcaa cggaccagat taaaacaatt taatgctcac atgaccactt	12164
	aaaaatgtca tgtgactttc ttaaaaagtc aaatgacttc atgttttaatt tttgaccggt	12224

```

catgattaga tttaacggtc atgatttatt ctcatgttct cacataaaaa catttttcat 12284
ttttctgata tgtatatttg taacataaaa cacaacataa aatactatgt aaatcagcca 12344
atttcaccaa atacaatatg atgtgaaaaa tataatatta ctattgatat tcaactcaaat 12404
gcacaatcgg ccctaaaaaa tgggtcaattt ttctgggttaa tatatttctt ttttaaaaca 12464
5  tgagtattca tgcaatatat gatttcaaaa tatataattt ctttcaacat ttaaccaatt 12524
   tgcaccatgt gtaatatgta agagatcaaa aattatatgt agttttattt atcttttaac 12584
   tgatatggga aacattttta tagataattt ttattcaagt tctaagttac ttcattttta 12644
   tttcggcgaa gttggtgctt ttttttcttt tcaaaccggt tgaagatggg ttcctcatgg 12704
   attggagttt aatccatttt tctatcatga gctc 12738
10 <210> 2
   <211> 3568
   <212> DNA
   <213> Lotus japonicus
   <220>
15 <221> 5'UTR
   <222> (1)..(137)
   <220>
   <221> CDS
   <222> (138)..(3119)
20 <220>
   <221> 3'UTR
   <222> (3120)..(3568)
   <400> 2

ttcaactttc aaaacaaagt ggatgggatc ttcattttat ataaccacac atcaatcatt 60
25 tgtgctactt ctccaatttt ctttagagat gaaatgaaga gctaagcaga caagacaagt 120
   ttatttgttt gttgctg atg ggt ctt ggg ttc aag atg cag cag agc cac 170
           Met Gly Leu Gly Phe Lys Met Gln Gln Ser His
                1             5             10
cac cct gtg gct ttg aag tta cat gag caa gct ggg agc cag aga aag 218
30 His Pro Val Ala Leu Lys Leu His Glu Gln Ala Gly Ser Gln Arg Lys
           15             20             25
ttc act ttc att cag aac ttc aga aac tgg ttt cta ccc ctt ctg ttt 266

```

	Phe Thr Phe Ile Gln Asn Phe Arg Asn Trp Phe Leu Pro Leu Leu Phe	
	30 35 40	
	gta tgg ttc att gtt atg gct gca ttt ggt gcc tgc atc tac cat aaa	314
	Val Trp Phe Ile Val Met Ala Ala Phe Gly Ala Cys Ile Tyr His Lys	
5	45 50 55	
	atg gat gct gaa act aaa gtc aga agg aaa gag gtg ctg ggt agc ctc	362
	Met Asp Ala Glu Thr Lys Val Arg Arg Lys Glu Val Leu Gly Ser Leu	
	60 65 70 75	
	tgt gat caa agg gct aga atg cta caa gac caa ttc agt gtc agt gtc	410
10	Cys Asp Gln Arg Ala Arg Met Leu Gln Asp Gln Phe Ser Val Ser Val	
	80 85 90	
	aac cat gtc cat gcc ctt gcc atc ctt gtt tca acc ttc cat tac tac	458
	Asn His Val His Ala Leu Ala Ile Leu Val Ser Thr Phe His Tyr Tyr	
	95 100 105	
15	aga aat act tca gcc att gac cag gaa acc ttt gca gaa tac acg gcc	506
	Arg Asn Thr Ser Ala Ile Asp Gln Glu Thr Phe Ala Glu Tyr Thr Ala	
	110 115 120	
	agg aca gca ttt gaa cgg cca tta atg agt ggg gtg gcc tat gca cag	554
	Arg Thr Ala Phe Glu Arg Pro Leu Met Ser Gly Val Ala Tyr Ala Gln	
20	125 130 135	
	aga gtg gtt cac tca gag aga gaa aga ttt gag aag caa cat ggg tgg	602
	Arg Val Val His Ser Glu Arg Glu Arg Phe Glu Lys Gln His Gly Trp	
	140 145 150 155	
	gtt ata aag aca atg gaa aga gtg cct tca ggg gtt agg gat gag tat	650
25	Val Ile Lys Thr Met Glu Arg Val Pro Ser Gly Val Arg Asp Glu Tyr	
	160 165 170	
	gca gca gtg ata ttt gca cag gaa act gtc tct tac ctt gaa tct att	698
	Ala Ala Val Ile Phe Ala Gln Glu Thr Val Ser Tyr Leu Glu Ser Ile	
	175 180 185	
30	gat atg atg tct ggg gag gag gac cga gag aac att ttg agg gct aga	746
	Asp Met Met Ser Gly Glu Glu Asp Arg Glu Asn Ile Leu Arg Ala Arg	
	190 195 200	

	gcc act ggg aaa gct gtt ctg act agc cct ttc aga ctg ctg gat tct	794
	Ala Thr Gly Lys Ala Val Leu Thr Ser Pro Phe Arg Leu Leu Asp Ser	
	205 210 215	
	cat cat ctt ggc gtg gtt cta aca ttt cct gtt tat aaa tct aag ctc	842
5	His His Leu Gly Val Val Leu Thr Phe Pro Val Tyr Lys Ser Lys Leu	
	220 225 230 235	
	cct cca gag cca acg acg gaa gag gtc att aaa gcc ata gca gga tat	890
	Pro Pro Glu Pro Thr Thr Glu Glu Val Ile Lys Ala Ile Ala Gly Tyr	
	240 245 250	
10	att gga gga tcc ttt gat gtt gag tcc ctt gtg gag aat tta ctt ggt	938
	Ile Gly Gly Ser Phe Asp Val Glu Ser Leu Val Glu Asn Leu Leu Gly	
	255 260 265	
	caa ctt gct ggt aac caa gca att ttg gtg aag gta tat gat ata aca	986
	Gln Leu Ala Gly Asn Gln Ala Ile Leu Val Lys Val Tyr Asp Ile Thr	
15	270 275 280	
	aac tct agc gac ccc cta atc atg tat ggc agc caa tat gaa gag ggt	1034
	Asn Ser Ser Asp Pro Leu Ile Met Tyr Gly Ser Gln Tyr Glu Glu Gly	
	285 290 295	
	gat atg tct ctt gtc cat gaa agt aag ctt gat ttt gga gat cca tac	1082
20	Asp Met Ser Leu Val His Glu Ser Lys Leu Asp Phe Gly Asp Pro Tyr	
	300 305 310 315	
	agg aaa cat cac atg atc tgt aga tat cac caa cag gca cca aca aat	1130
	Arg Lys His His Met Ile Cys Arg Tyr His Gln Gln Ala Pro Thr Asn	
	320 325 330	
25	tgg ata gca tat acc acg gca ttc cta ttc ttt gtg att ctt tgt tta	1178
	Trp Ile Ala Tyr Thr Thr Ala Phe Leu Phe Phe Val Ile Leu Cys Leu	
	335 340 345	
	gtg ggt tac att tta tat gct gct gga act cac att gtc aag gta gaa	1226
	Val Gly Tyr Ile Leu Tyr Ala Ala Gly Thr His Ile Val Lys Val Glu	
30	350 355 360	
	gat gat tac aat gca atg cag gat tta aaa gtc aaa gca gaa gca gct	1274
	Asp Asp Tyr Asn Ala Met Gln Asp Leu Lys Val Lys Ala Glu Ala Ala	

	365	370	375	
	gat att gcc aag tca cag ttt cta gct acc gtc tct cat gaa att aga	1322		
	Asp Ile Ala Lys Ser Gln Phe Leu Ala Thr Val Ser His Glu Ile Arg			
	380	385	390	395
5	act ccc atg aat gga att tta gga atg ctt ggt ctg ctt tta cgc aca	1370		
	Thr Pro Met Asn Gly Ile Leu Gly Met Leu Gly Leu Leu Leu Arg Thr			
	400	405	410	
	gaa ttg agt tca aca caa aga gac tat gct cag act gct caa gca tgt	1418		
	Glu Leu Ser Ser Thr Gln Arg Asp Tyr Ala Gln Thr Ala Gln Ala Cys			
10	415	420	425	
	ggg aag gca cta ata gca tta ata aat gag gtg ctt gac cga gct aaa	1466		
	Gly Lys Ala Leu Ile Ala Leu Ile Asn Glu Val Leu Asp Arg Ala Lys			
	430	435	440	
	att gaa gca ggc aaa tta gag cta gaa gca gtt cca ttt gac ctt cgt	1514		
15	Ile Glu Ala Gly Lys Leu Glu Leu Glu Ala Val Pro Phe Asp Leu Arg			
	445	450	455	
	tcc ata ctt gac gat gtc ctt tct ctt ttt tct gag aag tca aga cac	1562		
	Ser Ile Leu Asp Asp Val Leu Ser Leu Phe Ser Glu Lys Ser Arg His			
	460	465	470	475
20	aaa ggc tta gag ctg gca gtg ttt gtt tct gac aaa gtt ccg gat ata	1610		
	Lys Gly Leu Glu Leu Ala Val Phe Val Ser Asp Lys Val Pro Asp Ile			
	480	485	490	
	gtt atg ggc gat cct ggg cga ttc aga caa ata gtg aca aat ctt gtt	1658		
	Val Met Gly Asp Pro Gly Arg Phe Arg Gln Ile Val Thr Asn Leu Val			
25	495	500	505	
	gga aac tct gtt aag ttc act gag cga ggt cat ata ttt gtt aaa gtc	1706		
	Gly Asn Ser Val Lys Phe Thr Glu Arg Gly His Ile Phe Val Lys Val			
	510	515	520	
	cat tta gct gaa aaa aga cag tgc aca atg aat gga aaa tgt gag act	1754		
30	His Leu Ala Glu Lys Arg Gln Cys Thr Met Asn Gly Lys Cys Glu Thr			
	525	530	535	
	ttt cta aat gga ggc tgt gat gat gtt ttg cat gta tct ggc agt tat	1802		

	Phe	Leu	Asn	Gly	Gly	Cys	Asp	Asp	Val	Leu	His	Val	Ser	Gly	Ser	Tyr	
	540					545						550				555	
	aat	ttg	aaa	acc	ctt	agt	gga	tat	gaa	gcc	gct	gat	gaa	cgg	aac	agc	1850
	Asn	Leu	Lys	Thr	Leu	Ser	Gly	Tyr	Glu	Ala	Ala	Asp	Glu	Arg	Asn	Ser	
5						560						565				570	
	tgg	gat	aat	ttt	aag	cat	cat	att	gct	gac	gaa	gaa	ttt	ttc	ttt	gat	1898
	Trp	Asp	Asn	Phe	Lys	His	His	Ile	Ala	Asp	Glu	Glu	Phe	Phe	Phe	Asp	
						575						580				585	
	gct	tcg	gtt	aaa	aag	ttg	gcc	tct	agt	gaa	tct	tat	gag	caa	gtc	acc	1946
10	Ala	Ser	Val	Lys	Lys	Leu	Ala	Ser	Ser	Glu	Ser	Tyr	Glu	Gln	Val	Thr	
						590						595				600	
	ttg	atg	gtc	agc	gtg	gag	gac	act	gga	att	ggg	att	tct	ttc	tct	gcc	1994
	Leu	Met	Val	Ser	Val	Glu	Asp	Thr	Gly	Ile	Gly	Ile	Ser	Phe	Ser	Ala	
						605						610				615	
15	caa	gat	agt	att	ttc	atg	cct	ttt	gtg	cag	gct	gac	agc	tca	acc	tct	2042
	Gln	Asp	Ser	Ile	Phe	Met	Pro	Phe	Val	Gln	Ala	Asp	Ser	Ser	Thr	Ser	
						620						625				630	
	cga	aac	tat	ggg	ggg	acc	ggg	atc	ggc	ttg	agt	atc	agt	aag	tgc	ttg	2090
	Arg	Asn	Tyr	Gly	Gly	Thr	Gly	Ile	Gly	Leu	Ser	Ile	Ser	Lys	Cys	Leu	
20						640						645				650	
	gtt	gaa	ctg	atg	ggc	ggg	cag	ata	aac	ttc	ata	agc	cga	ccc	cag	gtc	2138
	Val	Glu	Leu	Met	Gly	Gly	Gln	Ile	Asn	Phe	Ile	Ser	Arg	Pro	Gln	Val	
						655						660				665	
	ggg	agc	acg	ttt	tca	ttc	act	gca	gat	ttc	gga	aca	ttt	aag	aaa	aac	2186
25	Gly	Ser	Thr	Phe	Ser	Phe	Thr	Ala	Asp	Phe	Gly	Thr	Phe	Lys	Lys	Asn	
						670						675				680	
	tca	aca	act	gac	atg	aag	aaa	ctt	aac	ttt	gaa	gat	cta	cct	tct	agt	2234
	Ser	Thr	Thr	Asp	Met	Lys	Lys	Leu	Asn	Phe	Glu	Asp	Leu	Pro	Ser	Ser	
						685						690				695	
30	ttt	aga	ggg	ctt	aaa	gcc	att	gtg	gtt	gat	gga	aaa	cct	gtt	aga	gct	2282
	Phe	Arg	Gly	Leu	Lys	Ala	Ile	Val	Val	Asp	Gly	Lys	Pro	Val	Arg	Ala	
	700					705						710				715	

	gca gtg act aga tac cat ttg aag aga cta ggg ata caa gct aaa gtt	2330
	Ala Val Thr Arg Tyr His Leu Lys Arg Leu Gly Ile Gln Ala Lys Val	
	720 725 730	
	gca att agc atc aat aag gct gtt tct tta tgt ggg aaa aat ggt tct	2378
5	Ala Ile Ser Ile Asn Lys Ala Val Ser Leu Cys Gly Lys Asn Gly Ser	
	735 740 745	
	ttg acc tcg gca tta ttt cag cct gat att att ttt gtt gag aag gac	2426
	Leu Thr Ser Ala Leu Phe Gln Pro Asp Ile Ile Phe Val Glu Lys Asp	
	750 755 760	
10	tct tgg gtt tct gga gag gat ggt ggt atc ttc aat gcg ttt aag atg	2474
	Ser Trp Val Ser Gly Glu Asp Gly Gly Ile Phe Asn Ala Phe Lys Met	
	765 770 775	
	cct caa atg atc ctt ctt gca acc aat atc tgt aac gct gaa ttt gat	2522
	Pro Gln Met Ile Leu Leu Ala Thr Asn Ile Cys Asn Ala Glu Phe Asp	
15	780 785 790 795	
	aaa gcc aaa gct gca ggt ttc agt gat aca gtg atc atg aag cca ctg	2570
	Lys Ala Lys Ala Ala Gly Phe Ser Asp Thr Val Ile Met Lys Pro Leu	
	800 805 810	
	aga gct agt atg ctg gct gct tgt ctt cag caa gtt ttc ggg act ggc	2618
20	Arg Ala Ser Met Leu Ala Ala Cys Leu Gln Gln Val Phe Gly Thr Gly	
	815 820 825	
	aag acg agg cag ttt ggg aaa gac atg tcg aat ggt tct tca gta cga	2666
	Lys Thr Arg Gln Phe Gly Lys Asp Met Ser Asn Gly Ser Ser Val Arg	
	830 835 840	
25	agc ctt ctt tgc gga aag aaa atc tta gtg gtt gat gat aat ttg gtg	2714
	Ser Leu Leu Cys Gly Lys Lys Ile Leu Val Val Asp Asp Asn Leu Val	
	845 850 855	
	aac cga agg gtc gcc gcc ggc gcg ttg aaa aac ttt gga gct gat gtc	2762
	Asn Arg Arg Val Ala Ala Gly Ala Leu Lys Asn Phe Gly Ala Asp Val	
30	860 865 870 875	
	aaa tgt gca gca agt ggc aaa gct gct ctt gaa atg ctt caa tat cct	2810
	Lys Cys Ala Ala Ser Gly Lys Ala Ala Leu Glu Met Leu Gln Tyr Pro	

	880	885	890	
	cac gat ttc gat gct tgc ttc atg gat att caa atg cca gaa atg gat	2858		
	His Asp Phe Asp Ala Cys Phe Met Asp Ile Gln Met Pro Glu Met Asp			
	895	900	905	
5	ggg ttt gag gcc act cgt cga att cgg atg atg gaa aga gag gca agt	2906		
	Gly Phe Glu Ala Thr Arg Arg Ile Arg Met Met Glu Arg Glu Ala Ser			
	910	915	920	
	gag cag ctg aaa agt gaa tct ggt gaa gaa aat ggt aag aaa agt gag	2954		
	Glu Gln Leu Lys Ser Glu Ser Gly Glu Glu Asn Gly Lys Lys Ser Glu			
10	925	930	935	
	ttc cac atg cct ata ttg gcc atg aca gct gat gta atc cat gct aca	3002		
	Phe His Met Pro Ile Leu Ala Met Thr Ala Asp Val Ile His Ala Thr			
	940	945	950	955
	tat gat aag tgc tta aat tgt ggg atg gat gga tac gtc tca aag cct	3050		
15	Tyr Asp Lys Cys Leu Asn Cys Gly Met Asp Gly Tyr Val Ser Lys Pro			
	960	965	970	
	ttt gaa gaa gag aat ctc tat caa gca gtt gca aag ttc ttc aag tcc	3098		
	Phe Glu Glu Glu Asn Leu Tyr Gln Ala Val Ala Lys Phe Phe Lys Ser			
	975	980	985	
20	aag cct gct tca gac tca tga cactgcttat tctgcagaac aggtcaacca	3149		
	Lys Pro Ala Ser Asp Ser			
	990			
	acttttgatt gagaaacatt tagtgtagc atgtttggat caacttctcc cagcatcaat	3209		
	tctgaaactc agaagctact catagaagct tctagccaga attgattttg gcttcagaag	3269		
25	ctactaatgg ttttatgtag agagcaaaag ttggtttcca agatatgcga ggatccatga	3329		
	tgatacacac gtctgaagtg atcatttcta accagttgaa gtttcactcg acgtgatttg	3389		
	aatccaagta aatgcatacc acataattat ccatcccat cttgtgtaca gattctccca	3449		
	agggataaga aaatttatgt aaattcaatt tttttctttt gcatctcaat acttccctgt	3509		
	tagaactttt tccctatgat tatttccact tttcattttc aattatatatt tttgtaaaa	3568		
30	<210> 3			
	<211> 993			
	<212> PRT			

<213> Lotus japonicus

<400> 3

```

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

```

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

Gly Arg Phe Arg Gln Ile Val Thr Asn Leu Val Gly Asn Ser Val Lys
 500 505 510
 Phe Thr Glu Arg Gly His Ile Phe Val Lys Val His Leu Ala Glu Lys
 515 520 525
 5 Arg Gln Cys Thr Met Asn Gly Lys Cys Glu Thr Phe Leu Asn Gly Gly
 530 535 540
 Cys Asp Asp Val Leu His Val Ser Gly Ser Tyr Asn Leu Lys Thr Leu
 545 550 555 560
 Ser Gly Tyr Glu Ala Ala Asp Glu Arg Asn Ser Trp Asp Asn Phe Lys
 10 565 570 575
 His His Ile Ala Asp Glu Glu Phe Phe Phe Asp Ala Ser Val Lys Lys
 580 585 590
 Leu Ala Ser Ser Glu Ser Tyr Glu Gln Val Thr Leu Met Val Ser Val
 595 600 605
 15 Glu Asp Thr Gly Ile Gly Ile Ser Phe Ser Ala Gln Asp Ser Ile Phe
 610 615 620
 Met Pro Phe Val Gln Ala Asp Ser Ser Thr Ser Arg Asn Tyr Gly Gly
 625 630 635 640
 Thr Gly Ile Gly Leu Ser Ile Ser Lys Cys Leu Val Glu Leu Met Gly
 20 645 650 655
 Gly Gln Ile Asn Phe Ile Ser Arg Pro Gln Val Gly Ser Thr Phe Ser
 660 665 670
 Phe Thr Ala Asp Phe Gly Thr Phe Lys Lys Asn Ser Thr Thr Asp Met
 675 680 685
 25 Lys Lys Leu Asn Phe Glu Asp Leu Pro Ser Ser Phe Arg Gly Leu Lys
 690 695 700
 Ala Ile Val Val Asp Gly Lys Pro Val Arg Ala Ala Val Thr Arg Tyr
 705 710 715 720
 His Leu Lys Arg Leu Gly Ile Gln Ala Lys Val Ala Ile Ser Ile Asn
 30 725 730 735
 Lys Ala Val Ser Leu Cys Gly Lys Asn Gly Ser Leu Thr Ser Ala Leu
 740 745 750

Phe Gln Pro Asp Ile Ile Phe Val Glu Lys Asp Ser Trp Val Ser Gly
 755 760 765
 Glu Asp Gly Gly Ile Phe Asn Ala Phe Lys Met Pro Gln Met Ile Leu
 770 775 780
 5 Leu Ala Thr Asn Ile Cys Asn Ala Glu Phe Asp Lys Ala Lys Ala Ala
 785 790 795 800
 Gly Phe Ser Asp Thr Val Ile Met Lys Pro Leu Arg Ala Ser Met Leu
 805 810 815
 Ala Ala Cys Leu Gln Gln Val Phe Gly Thr Gly Lys Thr Arg Gln Phe
 10 820 825 830
 Gly Lys Asp Met Ser Asn Gly Ser Ser Val Arg Ser Leu Leu Cys Gly
 835 840 845
 Lys Lys Ile Leu Val Val Asp Asp Asn Leu Val Asn Arg Arg Val Ala
 850 855 860
 15 Ala Gly Ala Leu Lys Asn Phe Gly Ala Asp Val Lys Cys Ala Ala Ser
 865 870 875 880
 Gly Lys Ala Ala Leu Glu Met Leu Gln Tyr Pro His Asp Phe Asp Ala
 885 890 895
 Cys Phe Met Asp Ile Gln Met Pro Glu Met Asp Gly Phe Glu Ala Thr
 20 900 905 910
 Arg Arg Ile Arg Met Met Glu Arg Glu Ala Ser Glu Gln Leu Lys Ser
 915 920 925
 Glu Ser Gly Glu Glu Asn Gly Lys Lys Ser Glu Phe His Met Pro Ile
 930 935 940
 25 Leu Ala Met Thr Ala Asp Val Ile His Ala Thr Tyr Asp Lys Cys Leu
 945 950 955 960
 Asn Cys Gly Met Asp Gly Tyr Val Ser Lys Pro Phe Glu Glu Glu Asn
 965 970 975
 Leu Tyr Gln Ala Val Ala Lys Phe Phe Lys Ser Lys Pro Ala Ser Asp
 30 980 985 990
 Ser

<210> 4
 <211> 12738
 <212> DNA
 <213> Lotus japonicus
 5 <220>
 <221> promoter
 <222> (1)..(2359)
 <223> promoter + 5' UTR
 <220>
 10 <221> exon
 <222> (2360)..(2704)
 <220>
 <221> Intron
 <222> (2705)..(3024)
 15 <220>
 <221> exon
 <222> (3025)..(3258)
 <220>
 <221> Intron
 20 <222> (3259)..(3390)
 <220>
 <221> exon
 <222> (3391)..(3560)
 <220>
 25 <221> Intron
 <222> (3561)..(3834)
 <220>
 <221> exon
 <222> (3835)..(4053)
 30 <223> C/T mutation in exon 4
 <220>
 <221> Intron

<222> (4054)..(4145)
 <220>
 <221> exon
 <222> (4146)..(4332)
 5 <220>
 <221> Intron
 <222> (4333)..(4420)
 <220>
 <221> exon
 10 <222> (4421)..(4472)
 <220>
 <221> Intron
 <222> (4473)..(4586)
 <220>
 15 <221> exon
 <222> (4587)..(4816)
 <220>
 <221> Intron
 <222> (4817)..(5037)
 20 <220>
 <221> exon
 <222> (5038)..(5136)
 <220>
 <221> Intron
 25 <222> (5137)..(5308)
 <220>
 <221> exon
 <222> (5309)..(6024)
 <220>
 30 <221> Intron
 <222> (6025)..(7042)
 <220>

<221> exon
 <222> (7043)..(7513)
 <220>
 <221> Intron
 5 <222> (7514)..(8306)
 <220>
 <221> exon
 <222> (8307)..(8564)
 <220>
 10 <221> Intron
 <222> (8565)..(12738)
 <223> 3' UTR + 3' intron
 <400> 4

 cccgggattg ttcatatccc catgtgtgtg tttggtttgg acacaagcaa gcttgccaca 60
 15 gtgggtgttta caactcattc tgaattatat atatatatat atatatataa ctaatctcta 120
 tcaataattt agttagtttt atgggataat atgattcccg tcaaattagc ttgtaggtta 180
 aaaatccata tttatataat aaattagact tatctaaatc acttaagtaa atatgatgta 240
 aaatcacaat agcttaggta gtttaacaat gttttatcat gaatttaaaa aaaattagtt 300
 gtttcaaaat attattggac caaataagca aagaggagct aatccaatta taaacagcac 360
 20 atctttataa ttgttttcta ctcgatctaa tattggttcc ttctttcctt gagaaaaaaa 420
 ggggttaaaaa tgcgatttgt tttgtaaaat taatttgtga cattaacaag agaaaactct 480
 acttttttaa tttttaagaa gttacttaaa tatagaaaga aagaagttac ttaaatatga 540
 aacttgtcta accaaagtca accgatcaaa gcgtttgggt agcaaagaga acttaaaaaa 600
 ataactaaat taactgtttg aacagtgatg taacagagaa atagaatggc ttacatatgg 660
 25 attatggagt taatggaagg cacactagtt ttagttcacg gcaacaagac agaattatcc 720
 tcctatatatt taggaggttt tcttatttat ttttaatgta agacgtgtat ctttcttttg 780
 aggtaataat gttcgagctt gaaacattca cattatagtg tgactaattg cttaatatga 840
 atcaagcatg taccacttga aacactcact catcgtttat aatataatta ctttatagta 900
 ctataaagta tgtaattatt ttctcctttt taagtatctc aatcattttt ttgacccaaa 960
 30 tattatttta gagaagaaac atgtcttaag accaatacat caacatatat ggagagtagt 1020
 ttgacaacct ctcatgagct ttctaaagcc aaacaaatgt actcacataa aaaaattaaa 1080
 aatttacagc aagtgagtta taaagagtga ttaataaatg ttgaataacg atggagggtg 1140

	gttatgaaat tccataaata aagcaaagga tgtttatggc aattgacatg ggaataagat	1200
	ccgcgcgttg tcgccgtgtc ataatcgctc agatttgtga tagcgagaga tttccattct	1260
	tttcctcttt tttcgaacga acgaacgaac gaacacagta gcagctgtat taggattcag	1320
	attgcatgat acattgatat tgatattgat attgatattg atattgttgt ggggtcaagtc	1380
5	tctactctac tatccaaagt aagcatatat atagagagag agagagagct tgggactagg	1440
	gaggatatta gcttatgtga ctgtgaagtt gaaggagagac aagagcgtat ctggcaaat	1500
	cctcaaaata aaatactagt actagtacag agaaaagaga ctaagagaga gagtgtgtct	1560
	gcacatcaag acccattgtg atttgtgatt tgtgatttgt gatttgtgat ttgtgagtga	1620
	gttcattgta caggtattat tgttgttgtt ttctctctc aaccaccctc taaagtctaa	1680
10	tctaactcat tgggctctgt gcttagctgg ttgtgttttg tgtatggtga attagggggg	1740
	caatctctgg ttttcatcat tattatatta tatatgggaa taccgtgctc tcttctttgt	1800
	ccttcaccaa aactagtttc acgcctacac aacatgatta gagcctcttc atttttttaa	1860
	tctcatcctt taagtgtatt tctatttcta ttggctattg ggacaagggg aagggtgtgc	1920
	ttcttaggaa cttgagctgt tttccatctt ttgagacca tgctttgtct ctctcatttt	1980
15	taattctggg tctctttctt ctctgtcct gattttttaa atgtgcttct ttttgcctc	2040
	ttacaaccac cctctaaacc attcatcatg cttggtttgc ttttgcttct ctttccacag	2100
	gtttcaatca cgcaaaacaa tgctgcaatg atgctgtact tggagtttct tctgtgacct	2160
	cttttttctt tccttcaaca atcaaccac cagagaaaag tgtctcagat tttgagacta	2220
	ctttcaactt tcaaaacaaa gtggatggga tttcatctt atataaccac acatcaatca	2280
20	tttgtgtac ttctcaatt ttctttagag atgaaatgaa gagctaagca gacaagacaa	2340
	gtttatttgt ttgttgctg atg ggt ctt ggg ttc aag atg cag cag agc cac	2392
	Met Gly Leu Gly Phe Lys Met Gln Gln Ser His	
	1 5 10	
	cac cct gtg gct ttg aag tta cat gag caa gct ggg agc cag aga aag	2440
25	His Pro Val Ala Leu Lys Leu His Glu Gln Ala Gly Ser Gln Arg Lys	
	15 20 25	
	ttc act ttc att cag aac ttc aga aac tgg ttt cta ccc ctt ctg ttt	2488
	Phe Thr Phe Ile Gln Asn Phe Arg Asn Trp Phe Leu Pro Leu Leu Phe	
	30 35 40	
30	gta tgg ttc att gtt atg gct gca ttt ggt gcc tgc atc tac cat aaa	2536
	Val Trp Phe Ile Val Met Ala Ala Phe Gly Ala Cys Ile Tyr His Lys	
	45 50 55	

	atg gat gct gaa act aaa gtc aga agg aaa gag gtg ctg ggt agc ctc	2584
	Met Asp Ala Glu Thr Lys Val Arg Arg Lys Glu Val Leu Gly Ser Leu	
	60 65 70 75	
	tgt gat caa agg gct aga atg cta caa gac caa ttc agt gtc agt gtc	2632
5	Cys Asp Gln Arg Ala Arg Met Leu Gln Asp Gln Phe Ser Val Ser Val	
	80 85 90	
	aac cat gtc cat gcc ctt gcc atc ctt gtt tca acc ttc cat tac tac	2680
	Asn His Val His Ala Leu Ala Ile Leu Val Ser Thr Phe His Tyr Tyr	
	95 100 105	
10	aga aat act tca gcc att gac cag gttgtgctt gattttcctt tccttgaagc	2734
	Arg Asn Thr Ser Ala Ile Asp Gln	
	110 115	
	atttttttagt tggaggctca atttcttttt ctgatttgat tctggcctta aaaattagaa	2794
	tcaattgtag aaggatttcc aaacatgccc attttggaaa ttggtgcatc tgatagtatc	2854
15	atgttttagat cagtttcttt ttcctcagaa ttgattttgg gcttaaaatc aattgtggaa	2914
	ggatattcat tagtaatttg gatattgttg catcatatgg ttctatctag ttacatcatt	2974
	tttttccact ctgattgcat gtatctttct cctgttcttt tccctatcag gaa acc	3030
	Glu Thr	
20	ttt gca gaa tac acg gcc agg aca gca ttt gaa cgg cca tta atg agt	3078
	Phe Ala Glu Tyr Thr Ala Arg Thr Ala Phe Glu Arg Pro Leu Met Ser	
	120 125 130	
	ggg gtg gcc tat gca cag aga gtg gtt cac tca gag aga gaa aga ttt	3126
	Gly Val Ala Tyr Ala Gln Arg Val Val His Ser Glu Arg Glu Arg Phe	
25	135 140 145	
	gag aag caa cat ggg tgg gtt ata aag aca atg gaa aga gtg cct tca	3174
	Glu Lys Gln His Gly Trp Val Ile Lys Thr Met Glu Arg Val Pro Ser	
	150 155 160 165	
	ggg gtt agg gat gag tat gca gca gtg ata ttt gca cag gaa act gtc	3222
30	Gly Val Arg Asp Glu Tyr Ala Ala Val Ile Phe Ala Gln Glu Thr Val	
	170 175 180	
	tct tac ctt gaa tct att gat atg atg tct ggg gag gtaaagtca	3268

	Ser Tyr Leu Glu Ser Ile Asp Met Met Ser Gly Glu			
	185		190	
	acacttgatga attaatgtga aaactcagaa gctactcaga gaagctcttc cccagaattg			3328
	gttctgcctt tagaataaat tgtacatgga ttgaccaca ttttctcatt tgcattgatgc			3388
5	ag gag gac cga gag aac att ttg agg gct aga gcc act ggg aaa gct			3435
	Glu Asp Arg Glu Asn Ile Leu Arg Ala Arg Ala Thr Gly Lys Ala			
	195	200	205	
	gtt ctg act agc cct ttc aga ctg ctg gat tct cat cat ctt ggc gtg			3483
	Val Leu Thr Ser Pro Phe Arg Leu Leu Asp Ser His His Leu Gly Val			
10	210	215	220	
	gtt cta aca ttt cct gtt tat aaa tct aag ctg cct cca gag cca acg			3531
	Val Leu Thr Phe Pro Val Tyr Lys Ser Lys Leu Pro Pro Glu Pro Thr			
	225	230	235	240
	acg gaa gag gtc att aaa gcc ata gca gg gtatgtcctc atttcacttt			3580
15	Thr Glu Glu Val Ile Lys Ala Ile Ala Gly			
	245			
	tcttgccaaa accagacttc tatttggttg tgtttccgta ggctatgact gatatgtagt			3640
	ttcaactcag ttagactata atataaacc ttcatgtggc ccttacccaa cagcttaagc			3700
	ttttgggata attggttgct tgacaaactc cttccgtaga aaacttggtt agctttggtt			3760
20	ctatgtgggc tttatgtttt ccctgagctt atgtaatagc atgatgtggt taatgtactt			3820
	tttaatggaa acag a tat att gga gga tcc ttt gat gtt gag tcc ctt gtg			3871
	Tyr Ile Gly Gly Ser Phe Asp Val Glu Ser Leu Val			
	255		260	
	gag aat tta ttt ggt caa ctt gct ggt aac caa gca att ttg gtg aag			3919
25	Glu Asn Leu Phe Gly Gln Leu Ala Gly Asn Gln Ala Ile Leu Val Lys			
	265	270	275	
	gta tat gat ata aca aac tct agc gac ccc cta atc atg tat ggc agc			3967
	Val Tyr Asp Ile Thr Asn Ser Ser Asp Pro Leu Ile Met Tyr Gly Ser			
	280	285	290	
30	caa tat gaa gag ggt gat atg tct ctt gtc cat gaa agt aag ctt gat			4015
	Gln Tyr Glu Glu Gly Asp Met Ser Leu Val His Glu Ser Lys Leu Asp			
	295	300	305	310

	ttt gga gat cca tac agg aaa cat cac atg atc tgt ag gtgggtgctt	4063
	Phe Gly Asp Pro Tyr Arg Lys His His Met Ile Cys Arg	
	315 320	
	ctagttattg ttgtagtaca ttttctgagt ttcagtgggtt tatcaattat cagcagattc	4123
5	ttatgatcaa tttttttaac ag a tat cac caa cag gca cca aca aat tgg	4173
	Tyr His Gln Gln Ala Pro Thr Asn Trp	
	325 330	
	ata gca tat acc acg gca ttc cta ttc ttt gtg att ctt tgt tta gtg	4221
	Ile Ala Tyr Thr Thr Ala Phe Leu Phe Phe Val Ile Leu Cys Leu Val	
10	335 340 345	
	ggg tac att tta tat gct gct gga act cac att gtc aag gta gaa gat	4269
	Gly Tyr Ile Leu Tyr Ala Ala Gly Thr His Ile Val Lys Val Glu Asp	
	350 355 360	
	gat tac aat gca atg cag gat tta aaa gtc aaa gca gaa gca gct gat	4317
15	Asp Tyr Asn Ala Met Gln Asp Leu Lys Val Lys Ala Glu Ala Ala Asp	
	365 370 375 380	
	att gcc aag tca cag gtacttttca tgacatgtta gcactgttcg ttatttcctt	4372
	Ile Ala Lys Ser Gln	
	385	
20	gaattgcata ctgatcacta gaaactgaaa atttggttatt aatgtcag ttt cta gct	4429
	Phe Leu Ala	
	acc gtc tct cat gaa att aga act ccc atg aat gga att tta g	4472
	Thr Val Ser His Glu Ile Arg Thr Pro Met Asn Gly Ile Leu	
25	390 395 400	
	gtaactttaa gattctctct cgttctcttt ccaactgaaaa gcaacacatg ctttcattcc	4532
	atacctgata ctttccatt agtgatgcta tcgttaaact ccttgtcact gtag ga	4588
	Gly	
30	atg ctt ggt ctg ctt tta cgc aca gaa ttg agt tca aca caa aga gac	4636
	Met Leu Gly Leu Leu Leu Arg Thr Glu Leu Ser Ser Thr Gln Arg Asp	
	405 410 415	

	tat gct cag act gct caa gca tgt ggg aag gca cta ata gca tta ata	4684
	Tyr Ala Gln Thr Ala Gln Ala Cys Gly Lys Ala Leu Ile Ala Leu Ile	
	420 425 430 435	
	aat gag gtg ctt gac cga gct aaa att gaa gca ggc aaa tta gag cta	4732
5	Asn Glu Val Leu Asp Arg Ala Lys Ile Glu Ala Gly Lys Leu Glu Leu	
	440 445 450	
	gaa gca gtt cca ttt gac ctt cgt tcc ata ctt gac gat gtc ctt tct	4780
	Glu Ala Val Pro Phe Asp Leu Arg Ser Ile Leu Asp Asp Val Leu Ser	
	455 460 465	
10	ctt ttt tct gag aag tca aga cac aaa ggc tta gag gtacgtttag	4826
	Leu Phe Ser Glu Lys Ser Arg His Lys Gly Leu Glu	
	470 475	
	tcattgctaa atctgttgtg aagtcctgta caagtggcgt aatttcatag tcctatttcc	4886
	ttttcttcaa ttgatgatct tatttcatat tcctctcgtg tttctctctt tctatgttgc	4946
15	catgtcgttg ggcgggtgatg ttctactct gatcccaaatt tcctgatgtg caattattcc	5006
	gacttggact tcaatggttt gggaagtata g ctg gca gtg ttt gtt tct gac	5058
	Leu Ala Val Phe Val Ser Asp	
	480 485	
	aaa gtt ccg gat ata gtt atg ggc gat cct ggg cga ttc aga caa ata	5106
20	Lys Val Pro Asp Ile Val Met Gly Asp Pro Gly Arg Phe Arg Gln Ile	
	490 495 500	
	gtg aca aat ctt gtt gga aac tct gtt aag gttagtggaa ttttcaaact	5156
	Val Thr Asn Leu Val Gly Asn Ser Val Lys	
	505 510	
25	ttatttgcct aatgttgtgt gcaagttgtg tgttggaaat gcgtcctttt aacgttataa	5216
	aatcgtacaa gttcgtattc tcattgtat acaataactt attagcaaag tacttgttga	5276
	tatcattact gattaacttt aatatcttgc ag ttc act gag cga ggt cat ata	5329
	Phe Thr Glu Arg Gly His Ile	
	515	
30	ttt gtt aaa gtc cat tta gct gaa aaa aga cag tgc aca atg aat gga	5377
	Phe Val Lys Val His Leu Ala Glu Lys Arg Gln Cys Thr Met Asn Gly	
	520 525 530 535	

	aaa tgt gag act ttt cta aat gga ggc tgt gat gat gtt ttg cat gta	5425
	Lys Cys Glu Thr Phe Leu Asn Gly Gly Cys Asp Asp Val Leu His Val	
	540 545 550	
	tct ggc agt tat aat ttg aaa acc ctt agt gga tat gaa gcc gct gat	5473
5	Ser Gly Ser Tyr Asn Leu Lys Thr Leu Ser Gly Tyr Glu Ala Ala Asp	
	555 560 565	
	gaa cgg aac agc tgg gat aat ttt aag cat cat att gct gac gaa gaa	5521
	Glu Arg Asn Ser Trp Asp Asn Phe Lys His His Ile Ala Asp Glu Glu	
	570 575 580	
10	ttt ttc ttt gat gct tcg gtt aaa aag ttg gcc tct agt gaa tct tat	5569
	Phe Phe Phe Asp Ala Ser Val Lys Lys Leu Ala Ser Ser Glu Ser Tyr	
	585 590 595	
	gag caa gtc acc ttg atg gtc agc gtg gag gac act gga att ggg att	5617
	Glu Gln Val Thr Leu Met Val Ser Val Glu Asp Thr Gly Ile Gly Ile	
15	600 605 610 615	
	tct ttc tct gcc caa gat agt att ttc atg cct ttt gtg cag gct gac	5665
	Ser Phe Ser Ala Gln Asp Ser Ile Phe Met Pro Phe Val Gln Ala Asp	
	620 625 630	
	agc tca acc tct cga aac tat ggg ggt acc ggg atc ggc ttg agt atc	5713
20	Ser Ser Thr Ser Arg Asn Tyr Gly Gly Thr Gly Ile Gly Leu Ser Ile	
	635 640 645	
	agt aag tgc ttg gtt gaa ctg atg ggc ggt cag ata aac ttc ata agc	5761
	Ser Lys Cys Leu Val Glu Leu Met Gly Gly Gln Ile Asn Phe Ile Ser	
	650 655 660	
25	cga ccc cag gtc ggg agc acg ttt tca ttc act gca gat ttc gga aca	5809
	Arg Pro Gln Val Gly Ser Thr Phe Ser Phe Thr Ala Asp Phe Gly Thr	
	665 670 675	
	ttt aag aaa aac tca aca act gac atg aag aaa ctt aac ttt gaa gat	5857
	Phe Lys Lys Asn Ser Thr Thr Asp Met Lys Lys Leu Asn Phe Glu Asp	
30	680 685 690 695	
	cta cct tct agt ttt aga ggt ctt aaa gcc att gtg gtt gat gga aaa	5905
	Leu Pro Ser Ser Phe Arg Gly Leu Lys Ala Ile Val Val Asp Gly Lys	

	700	705	710	
	cct gtt aga gct gca gtg act aga tac cat ttg aag aga cta ggg ata	5953		
	Pro Val Arg Ala Ala Val Thr Arg Tyr His Leu Lys Arg Leu Gly Ile			
	715	720	725	
5	caa gct aaa gtt gca att agc atc aat aag gct gtt tct tta tgt ggg	6001		
	Gln Ala Lys Val Ala Ile Ser Ile Asn Lys Ala Val Ser Leu Cys Gly			
	730	735	740	
	aaa aat ggt tct ttg acc tcg gc gtaagtcttt aattaacctt tttggtttca	6054		
	Lys Asn Gly Ser Leu Thr Ser Ala			
10	745	750		
	attatgtaga aatgtattga atgttatgat aatcagtagc attatcaact tttagtaatt	6114		
	gttcttaaca tatgctaata gtcatatctt tctataatac tacaatactg tagccatata	6174		
	atatctttcc tgtattggag tgagttttca aatgtttttc tgtgatattt tggaagttat	6234		
	cttcagtttg agaactcatt tgtcattttt gcatttttgtt attggatatt tggaatggatc	6294		
15	tttacaaagg atgtgtggat tttgacttgt tatacacatt tcttctccat tttatattgt	6354		
	ttgtgttatt ctttttactc ataaagaaat ttagaaactg cattgactgg ttctttttaa	6414		
	ttacttacag atattgacat tgatattttt tgtaaagtct gtcttgacat ggtttaatta	6474		
	cttacagact aggtttttct ttccttttct aacatgcata tccatttact tttttgacca	6534		
	accaacatcc tcatgagtca tgacatgttg atgatttata tggttgactt gagactattt	6594		
20	agacattaaa taaccgcaaa ttccatgttg tttgtgtgtt tggttccctg ttgggtaatc	6654		
	tcagaatcaa ttatgataga gtaaaatcaa ttttgatga gatgtgtggg tgtcattttg	6714		
	taaacctaaa cccaaaatca attctgctat aagctagaga gagtagttga acataatcaa	6774		
	ttgtgagatt ttgcaagtgg attgcacaac attgccctat gaaaatcact ttttgttcac	6834		
	aaaattttatc taaacataca taacttcatt ttcaaccttt actataatca attttacaat	6894		
25	aattaatttt acccaaatc aattgtgaca atgagtttcc aaacacacac ttaaagacta	6954		
	ccatttgcag aaaatatgtg atagaagact tatgtttatg tagtgtgttt cagttcattc	7014		
	actgatttaa actactcgga ttttgcag a tta ttt cag cct gat att att ttt	7067		
	Leu Phe Gln Pro Asp Ile Ile Phe			
	755			
30	gtt gag aag gac tct tgg gtt tct gga gag gat ggt ggt atc ttc aat	7115		
	Val Glu Lys Asp Ser Trp Val Ser Gly Glu Asp Gly Gly Ile Phe Asn			
	760	765	770	775

	gcg ttt aag atg cct caa atg atc ctt ctt gca acc aat atc tgt aac	7163
	Ala Phe Lys Met Pro Gln Met Ile Leu Leu Ala Thr Asn Ile Cys Asn	
	780 785 790	
	gct gaa ttt gat aaa gcc aaa gct gca ggt ttc agt gat aca gtg atc	7211
5	Ala Glu Phe Asp Lys Ala Lys Ala Ala Gly Phe Ser Asp Thr Val Ile	
	795 800 805	
	atg aag cca ctg aga gct agt atg ctg gct gct tgt ctt cag caa gtt	7259
	Met Lys Pro Leu Arg Ala Ser Met Leu Ala Ala Cys Leu Gln Gln Val	
	810 815 820	
10	ttc ggg act ggc aag acg agg cag ttt ggg aaa gac atg tcg aat ggt	7307
	Phe Gly Thr Gly Lys Thr Arg Gln Phe Gly Lys Asp Met Ser Asn Gly	
	825 830 835	
	tct tca gta cga agc ctt ctt tgc gga aag aaa atc tta gtg gtt gat	7355
	Ser Ser Val Arg Ser Leu Leu Cys Gly Lys Lys Ile Leu Val Val Asp	
15	840 845 850 855	
	gat aat ttg gtg aac cga agg gtc gcc gcc ggc gcg ttg aaa aac ttt	7403
	Asp Asn Leu Val Asn Arg Arg Val Ala Ala Gly Ala Leu Lys Asn Phe	
	860 865 870	
	gga gct gat gtc aaa tgt gca gca agt ggc aaa gct gct ctt gaa atg	7451
20	Gly Ala Asp Val Lys Cys Ala Ala Ser Gly Lys Ala Ala Leu Glu Met	
	875 880 885	
	ctt caa tat cct cac gat ttc gat gct tgc ttc atg gat att caa atg	7499
	Leu Gln Tyr Pro His Asp Phe Asp Ala Cys Phe Met Asp Ile Gln Met	
	890 895 900	
25	cca gaa atg gat gg gtatgcttac tggcactgac taatacatgt tttttgccaa	7553
	Pro Glu Met Asp Gly	
	905	
	cttaatatat tactctttca atattcggttg tgttattaga agatcatata gattaattta	7613
	taaattttct tttagcaaaa cttatcaat taagtgtgta gaaaagtcag tctcacatta	7673
30	tggtcaaata agtgtaggg caagcttcac ctcaaagcta gctatttggg tagatttagg	7733
	cctaaccgga attctaagat ggtatcagag tctatcctag atctttttat tggaaaccac	7793
	ccgtatatga gcaactcgta gatattcatt cttgaaagtt gcacgctcca tatgtccatt	7853

	cctaggtgcg agagagaagt ctcactttga ctagagatat gattaataaaa atatttataa	7913
	agggttgagc aatcctcacc tcagagctaa gcttttgggg taaagttagg cctaactcga	7973
	actctaataa agtgtttagc tgggtgtgtca actgtcaata tgaaatcttt tgcaatttac	8033
	tatgcattca cttacctact ttattgaagc ttattgacaa tttgtgcaga agcatcatta	8093
5	attaggaaca tgtagctat acaagttatg atgtttttgt atagcatatc atgttccaac	8153
	cttccaataa caaaatatgt gggtcaagtg tgagaatata taggttaaac aataaagtat	8213
	tgagttaaca gaaatctaaa cacacgtgt cactagctct tcatattgag acatgcatgg	8273
	gatttgacaa aacatctgaa taaatatttg cag g ttt gag gcc act cgt cga	8325
	Phe Glu Ala Thr Arg Arg	
10	910	
	att cgg atg atg gaa aga gag gca agt gag cag ctg aaa agt gaa tct	8373
	Ile Arg Met Met Glu Arg Glu Ala Ser Glu Gln Leu Lys Ser Glu Ser	
	915 920 925 930	
	ggg gaa gaa aat ggt aag aaa agt gag ttc cac atg cct ata ttg gcc	8421
15	Gly Glu Glu Asn Gly Lys Lys Ser Glu Phe His Met Pro Ile Leu Ala	
	935 940 945	
	atg aca gct gat gta atc cat gct aca tat gat aag tgc tta aat tgt	8469
	Met Thr Ala Asp Val Ile His Ala Thr Tyr Asp Lys Cys Leu Asn Cys	
	950 955 960	
20	ggg atg gat gga tac gtc tca aag cct ttt gaa gaa gag aat ctc tat	8517
	Gly Met Asp Gly Tyr Val Ser Lys Pro Phe Glu Glu Glu Asn Leu Tyr	
	965 970 975	
	caa gca gtt gca aag ttc ttc aag tcc aag cct gct tca gac tca tg	8564
	Gln Ala Val Ala Lys Phe Phe Lys Ser Lys Pro Ala Ser Asp Ser	
25	980 985 990	
	acactgctta ttctgcagaa caggtcaacc aacttttgat tgagaaacat ttagtgtag	8624
	catgtttgga tcaacttctc ccagcatcaa ttctgaaact cagaagctac tcatagaagc	8684
	ttctagccag aattgatttt ggcttcagaa gctactaatg gttttatgta gagagcaaaa	8744
	gttggtttcc aagatatgcg aggatccatg atgatacaca cgtctgaagt gatcatttct	8804
30	aaccagttga agtttcactc gacgtgattt gaatccaagt aaatgcatac cacataatta	8864
	tccatcccca tcttgtgtac agattctccc aagggataag aaaatttatg taaattcaat	8924
	ttttttcttt tgcattctcaa tacttccctg ttagaacttt ttccctatga ttatttccac	8984

	ttttcatttt	caattatatt	ttttgtaaaa	ttgggtctcca	tcctataagg	tttgcttagt	9044
	ttttttta	tcagatgaga	aggttgggtgc	ttataatgtg	tatacctttt	tagcagtact	9104
	tgtgatataa	atatttgtct	tattttttaga	acttaggaga	taataagatc	gtaggagtaa	9164
	atagcatttt	tggagaatga	tatcataatt	ataatcatct	acattctgca	gttataaaaa	9224
5	aaatcaatat	taatgcatag	caagtagctt	taagaagtat	gatctattta	agtattgaat	9284
	tttttttcta	aactaattac	aatgtattat	tgctgggata	tgcgcttttt	actcgataga	9344
	tatcatgttt	ggattagttt	ttcttagaat	ccattttgga	accgaaaaaa	ttatcacata	9404
	atttattctt	caaagttgat	tttggcaatc	aaggtaaaca	ctcataaaaag	ctaaagggag	9464
	ttgtttctta	ttattgtgat	tttagctcat	tgtggttttt	caccgtgtat	ataaacatgc	9524
10	atttgatgtg	acgttgactt	gtgaagcacg	tggctatcca	ccagcaatct	caagtaacac	9584
	ttgtcagcac	tatgttgcta	ttataagtga	caaggcagtg	accacacccc	tctatatatg	9644
	cataaacttt	tttattaaaa	gtaacagaat	caaaagggaa	gttttctaca	aattaattaa	9704
	tggaaaatta	tcagaaccag	tatggagctg	ttcctgctag	cagggctccg	gttcagagcc	9764
	agagtgaat	atttgaaacc	accaatgttg	gcggtgctac	acaagtcata	agaaatgatt	9824
15	atgatggaag	aaccactgct	gcagatgcta	ctgttggttg	tgagagcgat	ctgcagcatc	9884
	accacaagaa	aggaatatta	gagaagatca	aggagaagct	tcctggaaca	caccaccatc	9944
	aagatcacia	atagctagct	agctttgttt	agtttatttt	gcttctatca	ttataaatgt	10004
	aataataata	gtatgtgtgt	tgctacatgc	atgtgtatgt	atcgtttagt	cagctagttt	10064
	gttaataact	tgctttcaat	tttgcttcct	tacaagccta	tttatgcaat	gtatttggtc	10124
20	atttcactat	tggtttctaa	tcaatctaaa	ttagtacatg	gtcctcttcc	ttttagtgcc	10184
	ttttaattaa	agagtaaatg	ttgaatattg	tttttagtca	ttctatttac	atttttattt	10244
	tttttgcttc	atcgaaaaat	tatttacatt	tttattaagt	gtgtgttttag	attttcgttt	10304
	ttcgtaaaaa	atttcataat	taatttttta	taacagtgat	gtatgcagct	aaggcatcct	10364
	aacagaccta	tctttgtatg	tatggcattc	tgggcgggtc	ggatttattt	ttagcccat	10424
25	aattccccac	tactatacat	aaaatgatac	ctgctagtaa	aataatttaa	agtgtttcga	10484
	agtcataattg	gtaaattgat	ttcagttaaa	aatcgaagtt	tgtagctac	taaagacggt	10544
	agcttctatg	tcaaaatagt	taattcctat	attttataat	tggatcattt	tttctaaaat	10604
	ttattcaaac	atctcaactc	actttgacta	taatcacttt	ataattaatt	ttgttgaatt	10664
	aaataaatta	tattcaaatt	aatgtgata	ataccactg	aatgtgattg	gaatgccatc	10724
30	agggaaatat	ggaacaaaga	gagctgcatg	tgggaagggt	tggtatgtct	gcaaaacaga	10784
	caaagaagtt	gtaatcacia	ttaggacaat	tgccaaagat	gtgggggtcac	attgacttgt	10844
	cagggaaagt	ctatagatcc	actgtgctct	taaagctaga	acaattaaca	taaaatttct	10904

gccaattaag gctttcactt cctatcccta tctgctgtat attttaagtt tgagtaaagg 10964
 gacagttgtg gtgatgttat aatcaactgc aggctgaggg aaaagttggg ggtcaagtgg 11024
 tgtctatata aacttttttg tcattattaa tctttaatct ttttggtttt gtaaaatgac 11084
 agcttttgtc ttgtgctgaa agttgactgt gctagctagt aaagtcatgc tggtttctta 11144
 5 attgattgtg gtgtttttat gcgacaactg tagacattgc aatgcaggca cacaagtcaa 11204
 gaggcaatta caaacaaaat acagctatta ataattgacc tactaactac taaggaggga 11264
 atttaaatgt gattttgaaa tcttatattg gatgtaaaga ggcaatactc atcagactgt 11324
 aaagtgacag aacaagtaga tggtaaacaa gtagaagtag aagcagcaac aaattgagta 11384
 aagaccatca atgtcaacac ttccttggag aagaatttca taagtgtcat gtgaaaataa 11444
 10 ctagacaaaa ttagccaaa atgaagtatg aaactcagtt attgaaacga gtgacactga 11504
 ctagattcca agttattgaa ggaaccagca aaaggttaag tgacagtga gtgtgtgggt 11564
 tgaaagagga ttgaataaaa gctgaaccaa tggactgaat aggcaaactt gacaatttcc 11624
 cattagaatc tgttttaact tccttgatca aaagtctctt cggtgaaatt tccagcattg 11684
 ttggtaacat gatgagtagc tcatgaatct ggaaaccaa cagttttgtt tgttcaagct 11744
 15 gtagagaaga atgagaacta catgtgagca ttgttatgta ccaaaaaaac ctgtgagcat 11804
 tgttgctgca atcctggatc agtacgcata gcctgagcac aaagatgttg aattgaagta 11864
 gttgctctaa tatgttatga atattcaaga ttaaaacatt ttttaagttat catgtgacaa 11924
 tcacataact attaaattat tttaatctta actattaatc attagatcta acgatcatat 11984
 tctcatataa aataattatt ctgatacaat aatatatata tatatatata tgtatgtagg 12044
 20 ggatctgagg atgtatcata tgagaaatgt gatattatgt gagaacatga gaataaatca 12104
 caaccgtag atttaatcaa cggaccagat taaaacaatt taatgctcac atgaccactt 12164
 aaaaatgtca tgtgactttc ttaaaaagtc aaatgacttc atgttttaatt tttgaccggt 12224
 catgattaga tttacggtc atgatttatt ctcatgttct cacataaaaa catttttcat 12284
 ttttctgata tgtatatttg taacataaaa cacaacataa aatactatgt aaatcagcca 12344
 25 atttcaccaa atacaatatg atgtgaaaa tataatatta ctattgatat tcaactcaaat 12404
 gcacaatcgg ccctaaaaaa tggtaattt ttctgggttaa tatatttctt ttttaaaaca 12464
 tgagtattca tgcaatatat gatttcaaaa tatataattt ctttcaacat ttaaccaatt 12524
 tgcaccatgt gtaatatgta agagatcaaa aatttatattg agttttattt atcttttaac 12584
 tgatatggga aacattttta tagataattt ttattcaagt tctaagttac ttcattttta 12644
 30 tttcggcgaa gttggtgctt ttttttctt tcaaaccggt tgaagatggt ttcctcatgg 12704
 attggagttt aatccatttt tctatcatga gctc 12738

```

<211> 3568
<212> DNA
<213> Lotus japonicus
<220>
5 <221> 5'UTR
   <222> (1)..(137)
   <220>
   <221> CDS
   <222> (138)..(3119)
10 <220>
    <221> 3'UTR
    <222> (3120)..(3568)
    <400> 5
ttcaactttc aaaacaaagt ggatgggatc ttcattcttat ataaccacac atcaatcatt      60
15 tgtgctactt ctccaatttt ctttagagat gaaatgaaga gctaagcaga caagacaagt      120
   ttatttgttt gttgctg atg ggt ctt ggg ttc aag atg cag cag agc cac      170
               Met Gly Leu Gly Phe Lys Met Gln Gln Ser His
                   1             5             10
cac cct gtg gct ttg aag tta cat gag caa gct ggg agc cag aga aag      218
20 His Pro Val Ala Leu Lys Leu His Glu Gln Ala Gly Ser Gln Arg Lys
                   15             20             25
   ttc act ttc att cag aac ttc aga aac tgg ttt cta ccc ctt ctg ttt      266
   Phe Thr Phe Ile Gln Asn Phe Arg Asn Trp Phe Leu Pro Leu Leu Phe
                   30             35             40
25 gta tgg ttc att gtt atg gct gca ttt ggt gcc tgc atc tac cat aaa      314
   Val Trp Phe Ile Val Met Ala Ala Phe Gly Ala Cys Ile Tyr His Lys
                   45             50             55
   atg gat gct gaa act aaa gtc aga agg aaa gag gtg ctg ggt agc ctc      362
   Met Asp Ala Glu Thr Lys Val Arg Arg Lys Glu Val Leu Gly Ser Leu
30 60             65             70             75
   tgt gat caa agg gct aga atg cta caa gac caa ttc agt gtc agt gtc      410
   Cys Asp Gln Arg Ala Arg Met Leu Gln Asp Gln Phe Ser Val Ser Val

```

		80	85	90	
		aac cat gtc cat gcc ctt gcc atc ctt gtt tca acc ttc cat tac tac	458		
		Asn His Val His Ala Leu Ala Ile Leu Val Ser Thr Phe His Tyr Tyr			
		95	100	105	
5		aga aat act tca gcc att gac cag gaa acc ttt gca gaa tac acg gcc	506		
		Arg Asn Thr Ser Ala Ile Asp Gln Glu Thr Phe Ala Glu Tyr Thr Ala			
		110	115	120	
		agg aca gca ttt gaa cgg cca tta atg agt ggg gtg gcc tat gca cag	554		
		Arg Thr Ala Phe Glu Arg Pro Leu Met Ser Gly Val Ala Tyr Ala Gln			
10		125	130	135	
		aga gtg gtt cac tca gag aga gaa aga ttt gag aag caa cat ggg tgg	602		
		Arg Val Val His Ser Glu Arg Glu Arg Phe Glu Lys Gln His Gly Trp			
		140	145	150	155
		gtt ata aag aca atg gaa aga gtg cct tca ggg gtt agg gat gag tat	650		
15		Val Ile Lys Thr Met Glu Arg Val Pro Ser Gly Val Arg Asp Glu Tyr			
		160	165	170	
		gca gca gtg ata ttt gca cag gaa act gtc tct tac ctt gaa tct att	698		
		Ala Ala Val Ile Phe Ala Gln Glu Thr Val Ser Tyr Leu Glu Ser Ile			
		175	180	185	
20		gat atg atg tct ggg gag gag gac cga gag aac att ttg agg gct aga	746		
		Asp Met Met Ser Gly Glu Glu Asp Arg Glu Asn Ile Leu Arg Ala Arg			
		190	195	200	
		gcc act ggg aaa gct gtt ctg act agc cct ttc aga ctg ctg gat tct	794		
		Ala Thr Gly Lys Ala Val Leu Thr Ser Pro Phe Arg Leu Leu Asp Ser			
25		205	210	215	
		cat cat ctt ggc gtg gtt cta aca ttt cct gtt tat aaa tct aag ctc	842		
		His His Leu Gly Val Val Leu Thr Phe Pro Val Tyr Lys Ser Lys Leu			
		220	225	230	235
		cct cca gag cca acg acg gaa gag gtc att aaa gcc ata gca gga tat	890		
30		Pro Pro Glu Pro Thr Thr Glu Glu Val Ile Lys Ala Ile Ala Gly Tyr			
		240	245	250	
		att gga gga tcc ttt gat gtt gag tcc ctt gtg gag aat tta ttt ggt	938		

	Ile Gly Gly Ser Phe Asp Val Glu Ser Leu Val Glu Asn Leu Phe Gly	
	255 260 265	
	caa ctt gct ggt aac caa gca att ttg gtg aag gta tat gat ata aca	986
	Gln Leu Ala Gly Asn Gln Ala Ile Leu Val Lys Val Tyr Asp Ile Thr	
5	270 275 280	
	aac tct agc gac ccc cta atc atg tat ggc agc caa tat gaa gag ggt	1034
	Asn Ser Ser Asp Pro Leu Ile Met Tyr Gly Ser Gln Tyr Glu Glu Gly	
	285 290 295	
	gat atg tct ctt gtc cat gaa agt aag ctt gat ttt gga gat cca tac	1082
10	Asp Met Ser Leu Val His Glu Ser Lys Leu Asp Phe Gly Asp Pro Tyr	
	300 305 310 315	
	agg aaa cat cac atg atc tgt aga tat cac caa cag gca cca aca aat	1130
	Arg Lys His His Met Ile Cys Arg Tyr His Gln Gln Ala Pro Thr Asn	
	320 325 330	
15	tgg ata gca tat acc acg gca ttc cta ttc ttt gtg att ctt tgt tta	1178
	Trp Ile Ala Tyr Thr Thr Ala Phe Leu Phe Phe Val Ile Leu Cys Leu	
	335 340 345	
	gtg ggt tac att tta tat gct gct gga act cac att gtc aag gta gaa	1226
	Val Gly Tyr Ile Leu Tyr Ala Ala Gly Thr His Ile Val Lys Val Glu	
20	350 355 360	
	gat gat tac aat gca atg cag gat tta aaa gtc aaa gca gaa gca gct	1274
	Asp Asp Tyr Asn Ala Met Gln Asp Leu Lys Val Lys Ala Glu Ala Ala	
	365 370 375	
	gat att gcc aag tca cag ttt cta gct acc gtc tct cat gaa att aga	1322
25	Asp Ile Ala Lys Ser Gln Phe Leu Ala Thr Val Ser His Glu Ile Arg	
	380 385 390 395	
	act ccc atg aat gga att tta gga atg ctt ggt ctg ctt tta cgc aca	1370
	Thr Pro Met Asn Gly Ile Leu Gly Met Leu Gly Leu Leu Leu Arg Thr	
	400 405 410	
30	gaa ttg agt tca aca caa aga gac tat gct cag act gct caa gca tgt	1418
	Glu Leu Ser Ser Thr Gln Arg Asp Tyr Ala Gln Thr Ala Gln Ala Cys	
	415 420 425	

	ggg aag gca cta ata gca tta ata aat gag gtg ctt gac cga gct aaa	1466
	Gly Lys Ala Leu Ile Ala Leu Ile Asn Glu Val Leu Asp Arg Ala Lys	
	430 435 440	
	att gaa gca ggc aaa tta gag cta gaa gca gtt cca ttt gac ctt cgt	1514
5	Ile Glu Ala Gly Lys Leu Glu Leu Glu Ala Val Pro Phe Asp Leu Arg	
	445 450 455	
	tcc ata ctt gac gat gtc ctt tct ctt ttt tct gag aag tca aga cac	1562
	Ser Ile Leu Asp Asp Val Leu Ser Leu Phe Ser Glu Lys Ser Arg His	
	460 465 470 475	
10	aaa ggc tta gag ctg gca gtg ttt gtt tct gac aaa gtt ccg gat ata	1610
	Lys Gly Leu Glu Leu Ala Val Phe Val Ser Asp Lys Val Pro Asp Ile	
	480 485 490	
	gtt atg ggc gat cct ggg cga ttc aga caa ata gtg aca aat ctt gtt	1658
	Val Met Gly Asp Pro Gly Arg Phe Arg Gln Ile Val Thr Asn Leu Val	
15	495 500 505	
	gga aac tct gtt aag ttc act gag cga ggt cat ata ttt gtt aaa gtc	1706
	Gly Asn Ser Val Lys Phe Thr Glu Arg Gly His Ile Phe Val Lys Val	
	510 515 520	
	cat tta gct gaa aaa aga cag tgc aca atg aat gga aaa tgt gag act	1754
20	His Leu Ala Glu Lys Arg Gln Cys Thr Met Asn Gly Lys Cys Glu Thr	
	525 530 535	
	ttt cta aat gga ggc tgt gat gat gtt ttg cat gta tct ggc agt tat	1802
	Phe Leu Asn Gly Gly Cys Asp Asp Val Leu His Val Ser Gly Ser Tyr	
	540 545 550 555	
25	aat ttg aaa acc ctt agt gga tat gaa gcc gct gat gaa cgg aac agc	1850
	Asn Leu Lys Thr Leu Ser Gly Tyr Glu Ala Ala Asp Glu Arg Asn Ser	
	560 565 570	
	tgg gat aat ttt aag cat cat att gct gac gaa gaa ttt ttc ttt gat	1898
	Trp Asp Asn Phe Lys His His Ile Ala Asp Glu Glu Phe Phe Phe Asp	
30	575 580 585	
	gct tcg gtt aaa aag ttg gcc tct agt gaa tct tat gag caa gtc acc	1946
	Ala Ser Val Lys Lys Leu Ala Ser Ser Glu Ser Tyr Glu Gln Val Thr	

	590	595	600	
	ttg atg gtc agc gtg gag gac act gga att ggg att tct ttc tct gcc	1994		
	Leu Met Val Ser Val Glu Asp Thr Gly Ile Gly Ile Ser Phe Ser Ala			
	605	610	615	
5	caa gat agt att ttc atg cct ttt gtg cag gct gac agc tca acc tct	2042		
	Gln Asp Ser Ile Phe Met Pro Phe Val Gln Ala Asp Ser Ser Thr Ser			
	620	625	630	635
	cga aac tat ggg ggt acc ggg atc ggc ttg agt atc agt aag tgc ttg	2090		
	Arg Asn Tyr Gly Gly Thr Gly Ile Gly Leu Ser Ile Ser Lys Cys Leu			
10	640	645	650	
	gtt gaa ctg atg ggc ggt cag ata aac ttc ata agc cga ccc cag gtc	2138		
	Val Glu Leu Met Gly Gly Gln Ile Asn Phe Ile Ser Arg Pro Gln Val			
	655	660	665	
	ggg agc acg ttt tca ttc act gca gat ttc gga aca ttt aag aaa aac	2186		
15	Gly Ser Thr Phe Ser Phe Thr Ala Asp Phe Gly Thr Phe Lys Lys Asn			
	670	675	680	
	tca aca act gac atg aag aaa ctt aac ttt gaa gat cta cct tct agt	2234		
	Ser Thr Thr Asp Met Lys Lys Leu Asn Phe Glu Asp Leu Pro Ser Ser			
	685	690	695	
20	ttt aga ggt ctt aaa gcc att gtg gtt gat gga aaa cct gtt aga gct	2282		
	Phe Arg Gly Leu Lys Ala Ile Val Val Asp Gly Lys Pro Val Arg Ala			
	700	705	710	715
	gca gtg act aga tac cat ttg aag aga cta ggg ata caa gct aaa gtt	2330		
	Ala Val Thr Arg Tyr His Leu Lys Arg Leu Gly Ile Gln Ala Lys Val			
25	720	725	730	
	gca att agc atc aat aag gct gtt tct tta tgt ggg aaa aat ggt tct	2378		
	Ala Ile Ser Ile Asn Lys Ala Val Ser Leu Cys Gly Lys Asn Gly Ser			
	735	740	745	
	ttg acc tcg gca tta ttt cag cct gat att att ttt gtt gag aag gac	2426		
30	Leu Thr Ser Ala Leu Phe Gln Pro Asp Ile Ile Phe Val Glu Lys Asp			
	750	755	760	
	tct tgg gtt tct gga gag gat ggt ggt atc ttc aat gcg ttt aag atg	2474		

	Ser Trp Val Ser Gly Glu Asp Gly Gly Ile Phe Asn Ala Phe Lys Met	
	765 770 775	
	cct caa atg atc ctt ctt gca acc aat atc tgt aac gct gaa ttt gat	2522
	Pro Gln Met Ile Leu Leu Ala Thr Asn Ile Cys Asn Ala Glu Phe Asp	
5	780 785 790 795	
	aaa gcc aaa gct gca ggt ttc agt gat aca gtg atc atg aag cca ctg	2570
	Lys Ala Lys Ala Ala Gly Phe Ser Asp Thr Val Ile Met Lys Pro Leu	
	800 805 810	
	aga gct agt atg ctg gct gct tgt ctt cag caa gtt ttc ggg act ggc	2618
10	Arg Ala Ser Met Leu Ala Ala Cys Leu Gln Gln Val Phe Gly Thr Gly	
	815 820 825	
	aag acg agg cag ttt ggg aaa gac atg tcg aat ggt tct tca gta cga	2666
	Lys Thr Arg Gln Phe Gly Lys Asp Met Ser Asn Gly Ser Ser Val Arg	
	830 835 840	
15	agc ctt ctt tgc gga aag aaa atc tta gtg gtt gat gat aat ttg gtg	2714
	Ser Leu Leu Cys Gly Lys Lys Ile Leu Val Val Asp Asp Asn Leu Val	
	845 850 855	
	aac cga agg gtc gcc gcc ggc gcg ttg aaa aac ttt gga gct gat gtc	2762
	Asn Arg Arg Val Ala Ala Gly Ala Leu Lys Asn Phe Gly Ala Asp Val	
20	860 865 870 875	
	aaa tgt gca gca agt ggc aaa gct gct ctt gaa atg ctt caa tat cct	2810
	Lys Cys Ala Ala Ser Gly Lys Ala Ala Leu Glu Met Leu Gln Tyr Pro	
	880 885 890	
	cac gat ttc gat gct tgc ttc atg gat att caa atg cca gaa atg gat	2858
25	His Asp Phe Asp Ala Cys Phe Met Asp Ile Gln Met Pro Glu Met Asp	
	895 900 905	
	ggg ttt gag gcc act cgt cga att cgg atg atg gaa aga gag gca agt	2906
	Gly Phe Glu Ala Thr Arg Arg Ile Arg Met Met Glu Arg Glu Ala Ser	
	910 915 920	
30	gag cag ctg aaa agt gaa tct ggt gaa gaa aat ggt aag aaa agt gag	2954
	Glu Gln Leu Lys Ser Glu Ser Gly Glu Glu Asn Gly Lys Lys Ser Glu	
	925 930 935	

	ttc cac atg cct ata ttg gcc atg aca gct gat gta atc cat gct aca	3002
	Phe His Met Pro Ile Leu Ala Met Thr Ala Asp Val Ile His Ala Thr	
	940 945 950 955	
	tat gat aag tgc tta aat tgt ggg atg gat gga tac gtc tca aag cct	3050
5	Tyr Asp Lys Cys Leu Asn Cys Gly Met Asp Gly Tyr Val Ser Lys Pro	
	960 965 970	
	ttt gaa gaa gag aat ctc tat caa gca gtt gca aag ttc ttc aag tcc	3098
	Phe Glu Glu Glu Asn Leu Tyr Gln Ala Val Ala Lys Phe Phe Lys Ser	
	975 980 985	
10	aag cct gct tca gac tca tga cactgcttat tctgcagaac aggtcaacca	3149
	Lys Pro Ala Ser Asp Ser	
	990	
	acttttgatt gagaaacatt tagtgtagc atgtttggat caacttctcc cagcatcaat	3209
	tctgaaactc agaagctact catagaagct tctagccaga attgattttg gcttcagaag	3269
15	ctactaatgg ttttatgtag agagcaaaag ttggtttcca agatatgcga ggatccatga	3329
	tgatacacac gtctgaagtg atcatttcta accagttgaa gtttactcg acgtgatttg	3389
	aatccaagta aatgcatacc acataattat ccatcccat cttgtgtaca gattctccca	3449
	agggataaga aaatttatgt aaattcaatt tttttctttt gcatctcaat acttcctgt	3509
	tagaactttt tccctatgat tatttccact tttcattttc aattatatatt tttgtaaaa	3568
20	<210> 6	
	<211> 993	
	<212> PRT	
	<213> Lotus japonicus	
	<400> 6	
25	Met Gly Leu Gly Phe Lys Met Gln Gln Ser His His Pro Val Ala Leu	
	1 5 10 15	
	Lys Leu His Glu Gln Ala Gly Ser Gln Arg Lys Phe Thr Phe Ile Gln	
	20 25 30	
	Asn Phe Arg Asn Trp Phe Leu Pro Leu Leu Phe Val Trp Phe Ile Val	
30	35 40 45	
	Met Ala Ala Phe Gly Ala Cys Ile Tyr His Lys Met Asp Ala Glu Thr	
	50 55 60	

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

Ile Cys Arg Tyr His Gln Gln Ala Pro Thr Asn Trp Ile Ala Tyr Thr
 325 330 335
 Thr Ala Phe Leu Phe Phe Val Ile Leu Cys Leu Val Gly Tyr Ile Leu
 340 345 350
 5 Tyr Ala Ala Gly Thr His Ile Val Lys Val Glu Asp Asp Tyr Asn Ala
 355 360 365
 Met Gln Asp Leu Lys Val Lys Ala Glu Ala Ala Asp Ile Ala Lys Ser
 370 375 380
 Gln Phe Leu Ala Thr Val Ser His Glu Ile Arg Thr Pro Met Asn Gly
 10 385 390 395 400
 Ile Leu Gly Met Leu Gly Leu Leu Leu Arg Thr Glu Leu Ser Ser Thr
 405 410 415
 Gln Arg Asp Tyr Ala Gln Thr Ala Gln Ala Cys Gly Lys Ala Leu Ile
 420 425 430
 15 Ala Leu Ile Asn Glu Val Leu Asp Arg Ala Lys Ile Glu Ala Gly Lys
 435 440 445
 Leu Glu Leu Glu Ala Val Pro Phe Asp Leu Arg Ser Ile Leu Asp Asp
 450 455 460
 Val Leu Ser Leu Phe Ser Glu Lys Ser Arg His Lys Gly Leu Glu Leu
 20 465 470 475 480
 Ala Val Phe Val Ser Asp Lys Val Pro Asp Ile Val Met Gly Asp Pro
 485 490 495
 Gly Arg Phe Arg Gln Ile Val Thr Asn Leu Val Gly Asn Ser Val Lys
 500 505 510
 25 Phe Thr Glu Arg Gly His Ile Phe Val Lys Val His Leu Ala Glu Lys
 515 520 525
 Arg Gln Cys Thr Met Asn Gly Lys Cys Glu Thr Phe Leu Asn Gly Gly
 530 535 540
 Cys Asp Asp Val Leu His Val Ser Gly Ser Tyr Asn Leu Lys Thr Leu
 30 545 550 555 560
 Ser Gly Tyr Glu Ala Ala Asp Glu Arg Asn Ser Trp Asp Asn Phe Lys
 565 570 575

His His Ile Ala Asp Glu Glu Phe Phe Phe Asp Ala Ser Val Lys Lys
 580 585 590
 Leu Ala Ser Ser Glu Ser Tyr Glu Gln Val Thr Leu Met Val Ser Val
 595 600 605
 5 Glu Asp Thr Gly Ile Gly Ile Ser Phe Ser Ala Gln Asp Ser Ile Phe
 610 615 620
 Met Pro Phe Val Gln Ala Asp Ser Ser Thr Ser Arg Asn Tyr Gly Gly
 625 630 635 640
 Thr Gly Ile Gly Leu Ser Ile Ser Lys Cys Leu Val Glu Leu Met Gly
 10 645 650 655
 Gly Gln Ile Asn Phe Ile Ser Arg Pro Gln Val Gly Ser Thr Phe Ser
 660 665 670
 Phe Thr Ala Asp Phe Gly Thr Phe Lys Lys Asn Ser Thr Thr Asp Met
 675 680 685
 15 Lys Lys Leu Asn Phe Glu Asp Leu Pro Ser Ser Phe Arg Gly Leu Lys
 690 695 700
 Ala Ile Val Val Asp Gly Lys Pro Val Arg Ala Ala Val Thr Arg Tyr
 705 710 715 720
 His Leu Lys Arg Leu Gly Ile Gln Ala Lys Val Ala Ile Ser Ile Asn
 20 725 730 735
 Lys Ala Val Ser Leu Cys Gly Lys Asn Gly Ser Leu Thr Ser Ala Leu
 740 745 750
 Phe Gln Pro Asp Ile Ile Phe Val Glu Lys Asp Ser Trp Val Ser Gly
 755 760 765
 25 Glu Asp Gly Gly Ile Phe Asn Ala Phe Lys Met Pro Gln Met Ile Leu
 770 775 780
 Leu Ala Thr Asn Ile Cys Asn Ala Glu Phe Asp Lys Ala Lys Ala Ala
 785 790 795 800
 Gly Phe Ser Asp Thr Val Ile Met Lys Pro Leu Arg Ala Ser Met Leu
 30 805 810 815
 Ala Ala Cys Leu Gln Gln Val Phe Gly Thr Gly Lys Thr Arg Gln Phe
 820 825 830

Gly Lys Asp Met Ser Asn Gly Ser Ser Val Arg Ser Leu Leu Cys Gly
 835 840 845
 Lys Lys Ile Leu Val Val Asp Asp Asn Leu Val Asn Arg Arg Val Ala
 850 855 860
 5 Ala Gly Ala Leu Lys Asn Phe Gly Ala Asp Val Lys Cys Ala Ala Ser
 865 870 875 880
 Gly Lys Ala Ala Leu Glu Met Leu Gln Tyr Pro His Asp Phe Asp Ala
 885 890 895
 Cys Phe Met Asp Ile Gln Met Pro Glu Met Asp Gly Phe Glu Ala Thr
 10 900 905 910
 Arg Arg Ile Arg Met Met Glu Arg Glu Ala Ser Glu Gln Leu Lys Ser
 915 920 925
 Glu Ser Gly Glu Glu Asn Gly Lys Lys Ser Glu Phe His Met Pro Ile
 930 935 940
 15 Leu Ala Met Thr Ala Asp Val Ile His Ala Thr Tyr Asp Lys Cys Leu
 945 950 955 960
 Asn Cys Gly Met Asp Gly Tyr Val Ser Lys Pro Phe Glu Glu Glu Asn
 965 970 975
 Leu Tyr Gln Ala Val Ala Lys Phe Phe Lys Ser Lys Pro Ala Ser Asp
 20 980 985 990
 Ser

<210> 7
 <211> 4743
 25 <212> DNA
 <213> Medicago truncatula
 <220>
 <221> exon
 <222> (1)..(348)
 30 <220>
 <221> Intron
 <222> (349)..(462)

<220>
 <221> exon
 <222> (463)..(696)
 <220>
 5 <221> Intron
 <222> (697)..(799)
 <220>
 <221> exon
 <222> (800)..(969)
 10 <220>
 <221> Intron
 <222> (970)..(1261)
 <220>
 <221> exon
 15 <222> (1262)..(1480)
 <223> mutated codon encoding L/F substitution
 <220>
 <221> Intron
 <222> (1481)..(1574)
 20 <220>
 <221> exon
 <222> (1575)..(1761)
 <220>
 <221> Intron
 25 <222> (1762)..(1857)
 <220>
 <221> exon
 <222> (1858)..(1909)
 <220>
 30 <221> Intron
 <222> (1910)..(2018)
 <220>

	<221>	exon		
	<222>	(2019)..(2248)		
	<220>			
	<221>	Intron		
5	<222>	(2249)..(2383)		
	<220>			
	<221>	exon		
	<222>	(2384)..(2482)		
	<220>			
10	<221>	Intron		
	<222>	(2483)..(2624)		
	<220>			
	<221>	exon		
	<222>	(2625)..(3340)		
15	<220>			
	<221>	Intron		
	<222>	(3341)..(3597)		
	<220>			
	<221>	exon		
20	<222>	(3598)..(4101)		
	<220>			
	<221>	Intron		
	<222>	(4102)..(4490)		
	<220>			
25	<221>	exon		
	<222>	(4491)..(4743)		
	<400>	7		
	atg ggt ctt ctc ttg aag atg aag atg cag aat cag cac cac cct ttg			48
	Met Gly Leu Leu Leu Lys Met Lys Met Gln Asn Gln His His Pro Leu			
30	1 5 10 15			
	gct tct aag tta caa gaa caa acg ggg aac aaa aga tac aca ttc att			96
	Ala Ser Lys Leu Gln Glu Gln Thr Gly Asn Lys Arg Tyr Thr Phe Ile			

	20	25	30	
	caa gca cat aga gct tgg ctt ctc aaa tta atg ttt cta tgg att ctt	144		
	Gln Ala His Arg Ala Trp Leu Leu Lys Leu Met Phe Leu Trp Ile Leu			
	35	40	45	
5	ctg atg gct ctg att agt cgt atc atc tac agc aaa atg gat gtg ggt	192		
	Leu Met Ala Leu Ile Ser Arg Ile Ile Tyr Ser Lys Met Asp Val Gly			
	50	55	60	
	act aaa gtg aga agg aaa gag gtt ttg ggt agt ctt tgt gat caa agg	240		
	Thr Lys Val Arg Arg Lys Glu Val Leu Gly Ser Leu Cys Asp Gln Arg			
10	65	70	75	80
	gct aga atg ttg caa gac caa ttc agt gtt agt gtc aac cat gtt cat	288		
	Ala Arg Met Leu Gln Asp Gln Phe Ser Val Ser Val Asn His Val His			
	85	90	95	
	gct ctt gcc atc ctt gtt tca act ttc cat tat tac aga aac cct tct	336		
15	Ala Leu Ala Ile Leu Val Ser Thr Phe His Tyr Tyr Arg Asn Pro Ser			
	100	105	110	
	gcc att gac cag gtttgtgctt gaaagttttg atcattctgt gttggaaatg	388		
	Ala Ile Asp Gln			
	115			
20	aaaaatactc aatctttggtt gtgtttttga acctttatgt atcctctgat gattatttga	448		
	tgatttccct tcag gaa act ttt gca gaa tat acg gct agg acc gct ttc	498		
	Glu Thr Phe Ala Glu Tyr Thr Ala Arg Thr Ala Phe			
	120	125		
	gaa agg ccg cta ctt agt gga gtg gcc tat gca caa aga gtt gtt aac	546		
25	Glu Arg Pro Leu Leu Ser Gly Val Ala Tyr Ala Gln Arg Val Val Asn			
	130	135	140	
	tcg gaa aga gag cag ttt gag aag cag cat gga gtg gtt ata aag aca	594		
	Ser Glu Arg Glu Gln Phe Glu Lys Gln His Gly Val Val Ile Lys Thr			
	145	150	155	160
30	atg gaa aga gag gct tca ccg gtt agg gat gag tat gca ccg gtc ata	642		
	Met Glu Arg Glu Ala Ser Pro Val Arg Asp Glu Tyr Ala Pro Val Ile			
	165	170	175	

	ttt gct cag gaa act gtc tct tac ctt gag tct att gat atg atg tct	690
	Phe Ala Gln Glu Thr Val Ser Tyr Leu Glu Ser Ile Asp Met Met Ser	
	180 185 190	
	gga gag gtaaagaacg acacttgtga accattccac cgcttgTTTT ttttttttt	746
5	Gly Glu	
	tttttggtgg tggattactt ttattactac aatttctcat ttttgcaatg cag gag	802
	Glu	
	195	
10	gat cga gag aac ata atg aga gct aga gct act ggg aaa gct gtt ctg	850
	Asp Arg Glu Asn Ile Met Arg Ala Arg Ala Thr Gly Lys Ala Val Leu	
	200 205 210	
	act agc cct ttt agg ttg ttg ggt tct cat cat ctc ggt gtg gtt tta	898
	Thr Ser Pro Phe Arg Leu Leu Gly Ser His His Leu Gly Val Val Leu	
15	215 220 225	
	aca ttt cct gtt tac aaa tct aag ctc cct ccc aac cca aca aca gaa	946
	Thr Phe Pro Val Tyr Lys Ser Lys Leu Pro Pro Asn Pro Thr Thr Glu	
	230 235 240	
	gag ctc att aaa gcg acc gca gg gtatatgctt atttcactaa ttgtcagctt	999
20	Glu Leu Ile Lys Ala Thr Ala Gly	
	245 250	
	tttgTTTT gaatctTTTT tcttatgttt catattagta attagaggat gaaggttgtg	1059
	tttagtataa aaatttttagg tacattatta gcattttatt aagcagcagt taatctgtca	1119
	atgccactat tagtgttatc gacacaactt ctatcattga gttgaagaaa actgtttctg	1179
25	ttacaaggct tgtatttggt tatattttgt tgtaagcatt tcaatagaac atgtgattaa	1239
	tgaactTTTT aattggaggc ag a tat gtt gga gga tcc ttt gat gtg gag	1289
	Tyr Val Gly Gly Ser Phe Asp Val Glu	
	255 260	
	tca ctt gtg gaa aat tta ttt ggt caa ctt gct ggt cat caa gca att	1337
30	Ser Leu Val Glu Asn Leu Phe Gly Gln Leu Ala Gly His Gln Ala Ile	
	265 270 275	
	ttg gtc aat gta tat gat gtc acg aac tct tct gat ccc cta atc atg	1385

Leu Val Asn Val Tyr Asp Val Thr Asn Ser Ser Asp Pro Leu Ile Met
 280 285 290
 tat ggc aac caa tac gaa gaa ggt gat gtt tct ctt gtc cat gaa agt 1433
 Tyr Gly Asn Gln Tyr Glu Glu Gly Asp Val Ser Leu Val His Glu Ser
 5 295 300 305
 aag ctt gat ttt gga gat cca tac agg aaa cat caa atg ata tgt ag 1480
 Lys Leu Asp Phe Gly Asp Pro Tyr Arg Lys His Gln Met Ile Cys Arg
 310 315 320
 gtaggttctt ctggctactg tacattacat tttctggaac tgactttcgg cattttgtca 1540
 10 gcagaaagat gcttatgatc aaacttttgt acag g tat cac cag aag gca ccg 1593
 Tyr His Gln Lys Ala Pro
 325 330
 cca aat tgg aca gca ctt tct act gca atc cta ttc ttt gtg att ctt 1641
 Pro Asn Trp Thr Ala Leu Ser Thr Ala Ile Leu Phe Phe Val Ile Leu
 15 335 340 345
 ctt tta atc ggt tac att tta tat ggt gct ggg aat cat att gtc aaa 1689
 Leu Leu Ile Gly Tyr Ile Leu Tyr Gly Ala Gly Asn His Ile Val Lys
 350 355 360
 gta gag gat gat ttc cac gaa atg cag gag cta aaa gtt cga gca gag 1737
 20 Val Glu Asp Asp Phe His Glu Met Gln Glu Leu Lys Val Arg Ala Glu
 365 370 375
 gca gcc gat gtt gcc aag tca cag gtacttttcc ttgacatgtc gttggcactt 1791
 Ala Ala Asp Val Ala Lys Ser Gln
 380 385
 25 ggcatatttc gatatttcct ttgattacgc aacaataact aatactgaaa tttttgtttt 1851
 tgtcag ttt cta gct act gtc tct cat gaa att aga aca cct atg aat 1899
 Phe Leu Ala Thr Val Ser His Glu Ile Arg Thr Pro Met Asn
 390 395 400
 ggc att tta g gtaactgcaa aattctctct ctctttgcct aacatgtaac 1949
 30 Gly Ile Leu
 atatgctttc atagctttca ctgctcatt aataatgtta tcacaaaaca aaactctgta 2009

	tgattgtag ga atg ctt ggt ctg ctt cta cgc acg gaa ttg aac tca act	2059
	Gly Met Leu Gly Leu Leu Leu Arg Thr Glu Leu Asn Ser Thr	
	405 410 415	
	caa cgg gac tat gct cag act gct caa gca tgt ggg aag gct ctg ata	2107
5	Gln Arg Asp Tyr Ala Gln Thr Ala Gln Ala Cys Gly Lys Ala Leu Ile	
	420 425 430	
	gca tta ata aat gaa gtg ctt gac cga gca aaa att gaa gct ggc aaa	2155
	Ala Leu Ile Asn Glu Val Leu Asp Arg Ala Lys Ile Glu Ala Gly Lys	
	435 440 445	
10	tta gag ctg gaa gcg gtt cca ttt gac ctt cgt tcc ata ctc gat gat	2203
	Leu Glu Leu Glu Ala Val Pro Phe Asp Leu Arg Ser Ile Leu Asp Asp	
	450 455 460 465	
	gtc ctc tct ctt ttt tca gag aag tct aga cac aaa ggt tta gag	2248
	Val Leu Ser Leu Phe Ser Glu Lys Ser Arg His Lys Gly Leu Glu	
15	470 475 480	
	gtatgctcaa tcgttactaa acctattgtg aaggccaata cgagtttcat aataatatct	2308
	taattcgggtt tcctctcttt atttgatgat ctaacttcat tacccttttt ttttttttca	2368
	ttttacccttt catag ctg gca gtg ttt gtt tct gat aaa gtt cca gat att	2419
	Leu Ala Val Phe Val Ser Asp Lys Val Pro Asp Ile	
20	485 490	
	gtt atg gga gat cct ggg aga ttc aga caa att gta aca aat ctt gtt	2467
	Val Met Gly Asp Pro Gly Arg Phe Arg Gln Ile Val Thr Asn Leu Val	
	495 500 505	
	ggc aac tct gtt aaa gtaagtggaa ttttcaaatt ttatttgcct aaagttattt	2522
25	Gly Asn Ser Val Lys	
	510	
	gcattactaa tattcatctt atgttggaat tgtttgaaat tttgctgttg tatagcaaga	2582
	tacttggtga ttacatttct gattaacttc aatatattgc ag ttc act gag cga	2636
	Phe Thr Glu Arg	
30	515	
	ggc cat ata ttt gtt aaa gtt cat tta tct gaa aac aga aag ccc gta	2684
	Gly His Ile Phe Val Lys Val His Leu Ser Glu Asn Arg Lys Pro Val	

	520	525	530	
	aca aat gga aag cat gag act tat cga aat gga ggg tct gaa gaa gtt	2732		
	Thr Asn Gly Lys His Glu Thr Tyr Arg Asn Gly Gly Ser Glu Glu Val			
	535	540	545	
5	gtg cat gca tct ggc ggt tat aat ctc aaa acg cta agt gga tat gaa	2780		
	Val His Ala Ser Gly Gly Tyr Asn Leu Lys Thr Leu Ser Gly Tyr Glu			
	550	555	560	565
	gct gct gat gaa cgc aac aac tgg gat aat ttt aac cac tta att gct	2828		
	Ala Ala Asp Glu Arg Asn Asn Trp Asp Asn Phe Asn His Leu Ile Ala			
10	570	575	580	
	gat gaa gag ttt ttc tgc gat gct tca act aaa aaa gtg gcc tcg aat	2876		
	Asp Glu Glu Phe Phe Cys Asp Ala Ser Thr Lys Lys Val Ala Ser Asn			
	585	590	595	
	gaa ttt tat gaa caa gtc acc ttg atg gtc tgt gtc gaa gac act gga	2924		
15	Glu Phe Tyr Glu Gln Val Thr Leu Met Val Cys Val Glu Asp Thr Gly			
	600	605	610	
	att gga att cct ttc tcg gcc caa gat agg att ttc atg cct ttt gtt	2972		
	Ile Gly Ile Pro Phe Ser Ala Gln Asp Arg Ile Phe Met Pro Phe Val			
	615	620	625	
20	cag gca gat agc tcg act tct aga aat tat ggt ggt acc ggc att ggt	3020		
	Gln Ala Asp Ser Ser Thr Ser Arg Asn Tyr Gly Gly Thr Gly Ile Gly			
	630	635	640	645
	ttg agt atc agt aag tgc cta gtt gaa cta atg ggt ggt caa ata aac	3068		
	Leu Ser Ile Ser Lys Cys Leu Val Glu Leu Met Gly Gly Gln Ile Asn			
25	650	655	660	
	ttt ata agc cgg ccg cag gtt gga agc acc ttt tca ttt act gcg gat	3116		
	Phe Ile Ser Arg Pro Gln Val Gly Ser Thr Phe Ser Phe Thr Ala Asp			
	665	670	675	
	ttt gga ata ttt aag aag aat cca ata act gag gtg aag aag gtt aac	3164		
30	Phe Gly Ile Phe Lys Lys Asn Pro Ile Thr Glu Val Lys Lys Val Asn			
	680	685	690	
	tat gaa gat cta cca tcc agt ttt aga ggg ctt aaa gcc gtt gtg gtt	3212		

	Tyr	Glu	Asp	Leu	Pro	Ser	Ser	Phe	Arg	Gly	Leu	Lys	Ala	Val	Val	Val	
	695							700					705				
	gat	ggg	aaa	cct	gtt	aga	gct	gct	gtg	act	aga	tac	cat	ttg	aag	aga	3260
	Asp	Gly	Lys	Pro	Val	Arg	Ala	Ala	Val	Thr	Arg	Tyr	His	Leu	Lys	Arg	
5	710					715				720				725			
	ctt	ggg	ata	caa	gtt	aaa	gtc	gca	aat	gcc	atc	aat	aag	gct	gtt	tcc	3308
	Leu	Gly	Ile	Gln	Val	Lys	Val	Ala	Asn	Ala	Ile	Asn	Lys	Ala	Val	Ser	
					730					735				740			
	ttg	tgt	ggg	aaa	aat	ggg	gct	tcc	agc	aca	gg	gtaagttttt	aatttttcctt				3360
10	Leu	Cys	Gly	Lys	Asn	Gly	Ala	Ser	Ser	Thr	Gly						
					745					750							
	tttgtaatta	atgcattgct	atttttaatg	aattaatggt	aggactatgt	gtgtttaatc											3420
	ataattccat	ctaaagtcac	acatgatagg	ctgtgacttg	ctgacagtgt	aaaacgtata											3480
	aaacattata	cactaacagt	gtatcaaaat	taaactcttt	gaagaattgc	atttgttctt											3540
15	gttgtgtggt	tcaatacttg	agttcattaa	ctgagttaaa	atatttggtt	gtggcag g											3598
	tta	ttc	cag	cct	gat	att	att	ttt	gtt	gag	aaa	gat	tca	tgg	gtt	tgt	3646
	Leu	Phe	Gln	Pro	Asp	Ile	Ile	Phe	Val	Glu	Lys	Asp	Ser	Trp	Val	Cys	
					755					760				765			
	gga	gag	gac	ggg	atc	ttc	agt	gtg	cgc	caa	ctg	gac	tgg	aaa	cag	aat	3694
20	Gly	Glu	Asp	Gly	Ile	Phe	Ser	Val	Arg	Gln	Leu	Asp	Trp	Lys	Gln	Asn	
					770					775				780			
	gga	cat	ata	ttt	aag	atg	cct	caa	atg	atc	ctt	cta	gct	aca	aat	att	3742
	Gly	His	Ile	Phe	Lys	Met	Pro	Gln	Met	Ile	Leu	Leu	Ala	Thr	Asn	Ile	
					785					790				795		800	
25	agt	aat	gac	gaa	ttt	gat	aaa	gct	aaa	tcc	gca	ggg	ttt	agt	gat	acg	3790
	Ser	Asn	Asp	Glu	Phe	Asp	Lys	Ala	Lys	Ser	Ala	Gly	Phe	Ser	Asp	Thr	
					805					810				815			
	gtg	atc	atg	aag	cca	ctg	aga	gct	agt	atg	gtg	gga	gct	tgc	ctt	cag	3838
	Val	Ile	Met	Lys	Pro	Leu	Arg	Ala	Ser	Met	Val	Gly	Ala	Cys	Leu	Gln	
30					820					825				830			
	caa	gtt	ttg	gga	aca	ggc	aag	aag	aga	cag	ctg	gga	aaa	gag	atg	cct	3886
	Gln	Val	Leu	Gly	Thr	Gly	Lys	Lys	Arg	Gln	Leu	Gly	Lys	Glu	Met	Pro	

	835	840	845	
	aat ggt tca act tca gtt cga agc ctt ctg ttc gga aag aaa att tta	3934		
	Asn Gly Ser Thr Ser Val Arg Ser Leu Leu Phe Gly Lys Lys Ile Leu			
	850	855	860	
5	gtg gtt gat gac aat gta gta aac cgg agg gtg gct gca ggt gcc ttg	3982		
	Val Val Asp Asp Asn Val Val Asn Arg Arg Val Ala Ala Gly Ala Leu			
	865	870	875	880
	aaa aac ttt gga gcg gat gtg aag tgt gca gat agc ggc aaa gct gct	4030		
	Lys Asn Phe Gly Ala Asp Val Lys Cys Ala Asp Ser Gly Lys Ala Ala			
10	885	890	895	
	ctt gaa atg ctt caa ttc cct cac aag ttt gat gct tgc ttc atg gat	4078		
	Leu Glu Met Leu Gln Phe Pro His Lys Phe Asp Ala Cys Phe Met Asp			
	900	905	910	
	att caa atg cca gaa atg gac gg gtatgtttgt ttgcatcgat tattaatttt	4131		
15	Ile Gln Met Pro Glu Met Asp Gly			
	915			
	ttttttggtta ttatttgaac ataacatggt aacatgatag ttgcattgct gataataaaa	4191		
	aaaaaaaaactt gcataagttg aatgcaagtt acgtgatagt tgccctgctg ataaattcaa	4251		
	tgatatgcac gtaagttgaa tgcacgtaac ttgcatgttg gctatactaa ttcacattca	4311		
20	aattatTTTTg tttgtatata aactctaatt acacatgtta tgtatgttct aattcacatc	4371		
	catactatat ttgtttgtat ataaacctta tcaatagatg ttatgtgtgt tctaattcac	4431		
	atgaatatta tttttacgca tattgttttt tataaatcat gtacttgcat tattttcag	4490		
	g ttt gaa gca act cgt cga att cgg gag atg gag agg aca gcg aat gag	4539		
	Phe Glu Ala Thr Arg Arg Ile Arg Glu Met Glu Arg Thr Ala Asn Glu			
25	925	930	935	
	gag acg aat agt gaa tgt ggt gaa agg aaa agt gaa ttc cat tta cct	4587		
	Glu Thr Asn Ser Glu Cys Gly Glu Arg Lys Ser Glu Phe His Leu Pro			
	940	945	950	
	ata ttg gcc atg aca gca gat gta atc cat gct aca tat gaa gag tgt	4635		
30	Ile Leu Ala Met Thr Ala Asp Val Ile His Ala Thr Tyr Glu Glu Cys			
	955	960	965	
	ttg aaa tgt ggg atg gat ggt tat gtt tca aaa cct ttt gag gaa gag	4683		

```

Leu Lys Cys Gly Met Asp Gly Tyr Val Ser Lys Pro Phe Glu Glu Glu
      970              975              980
aat ctt tat caa gca gtt gca aag ttt ttc cag aca aaa cct act tca      4731
Asn Leu Tyr Gln Ala Val Ala Lys Phe Phe Gln Thr Lys Pro Thr Ser
5   985              990              995              1000
gta gat tca tga      4743
Val Asp Ser

<210> 8
10 <211> 3694
    <212> DNA
    <213> Arabidopsis thaliana
    <220>
    <221> 5'UTR
15 <222> (1)..(68)
    <220>
    <221> CDS
    <222> (69)..(3311)
    <220>
20 <221> 3'UTR
    <222> (3312)..(3694)
    <220>
    <221> misc_feature
    <222> (3594)..(3594)
25 <223> n is a, c, g, or t
    <220>
    <221> misc_feature
    <222> (3631)..(3631)
    <223> n is a, c, g, or t
30 <400> 8
aaaaaatctc actaaaacaa aagaagaaga aagaagaaag aaaatggaat acctacattt      60
ttgaagtg atg aga aga gat ttt gtg tat aat aat aat gca atg ttc aat      110

```

	Met Arg Arg Asp Phe Val Tyr Asn Asn Asn Ala Met Phe Asn																			
	1					5					10									
	cct	ctc	aca	act	cat	tac	agc	tca	gat	atg	aac	tgg	gca	ctc	aac	aat	158			
	Pro	Leu	Thr	Thr	His	Tyr	Ser	Ser	Asp	Met	Asn	Trp	Ala	Leu	Asn	Asn				
5	15					20					25					30				
	cat	caa	gaa	gaa	gaa	gaa	gag	cca	cga	aga	att	gaa	att	tct	gat	tcc	206			
	His	Gln	Glu	Glu	Glu	Glu	Glu	Pro	Arg	Arg	Ile	Glu	Ile	Ser	Asp	Ser				
	35					40					45									
	gag	tca	cta	gaa	aac	ttg	aaa	agc	agc	gat	ttt	tat	caa	ctg	ggg	ggg	254			
10	Glu	Ser	Leu	Glu	Asn	Leu	Lys	Ser	Ser	Asp	Phe	Tyr	Gln	Leu	Gly	Gly				
	50					55					60									
	ggg	ggg	gct	ctg	aat	tcg	tca	gaa	aag	ccg	aga	aag	atc	gat	ttt	tgg	302			
	Gly	Gly	Ala	Leu	Asn	Ser	Ser	Glu	Lys	Pro	Arg	Lys	Ile	Asp	Phe	Trp				
	65					70					75									
15	cgt	tcg	ggg	ttg	atg	ggg	ttt	gcg	aag	atg	cag	cag	cag	caa	cag	ctt	350			
	Arg	Ser	Gly	Leu	Met	Gly	Phe	Ala	Lys	Met	Gln	Gln	Gln	Gln	Gln	Leu				
	80					85					90									
	cag	cat	tca	gtg	gcg	gtg	aag	atg	aac	aat	aat	aat	aat	aac	gat	cta	398			
	Gln	His	Ser	Val	Ala	Val	Lys	Met	Asn	Asn	Asn	Asn	Asn	Asn	Asp	Leu				
20	95					100					105					110				
	atg	ggg	aat	aaa	aaa	ggg	tca	act	ttc	ata	caa	gaa	cat	cga	gca	ttg	446			
	Met	Gly	Asn	Lys	Lys	Gly	Ser	Thr	Phe	Ile	Gln	Glu	His	Arg	Ala	Leu				
	115					120					125									
	tta	cca	aaa	gct	ttg	att	ctg	tgg	atc	atc	att	gtt	ggg	ttt	ata	agc	494			
25	Leu	Pro	Lys	Ala	Leu	Ile	Leu	Trp	Ile	Ile	Ile	Val	Gly	Phe	Ile	Ser				
	130					135					140									
	agt	ggg	att	tat	cag	tgg	atg	gat	gat	gct	aat	aag	att	aga	agg	gaa	542			
	Ser	Gly	Ile	Tyr	Gln	Trp	Met	Asp	Asp	Ala	Asn	Lys	Ile	Arg	Arg	Glu				
	145					150					155									
30	gag	gtt	ttg	gtc	agc	atg	tgt	gat	caa	aga	gct	aga	atg	ttg	cag	gat	590			
	Glu	Val	Leu	Val	Ser	Met	Cys	Asp	Gln	Arg	Ala	Arg	Met	Leu	Gln	Asp				
	160					165					170									

	caa ttt agt gtt agt gtt aat cat gtt cat gct ttg gct att ctc gtc	638
	Gln Phe Ser Val Ser Val Asn His Val His Ala Leu Ala Ile Leu Val	
	175 180 185 190	
	tcc act ttt cat tac cac aag aac cct tct gca att gat cag gag aca	686
5	Ser Thr Phe His Tyr His Lys Asn Pro Ser Ala Ile Asp Gln Glu Thr	
	195 200 205	
	ttt gcg gag tac acg gca aga aca gca ttt gag aga ccg ttg cta agt	734
	Phe Ala Glu Tyr Thr Ala Arg Thr Ala Phe Glu Arg Pro Leu Leu Ser	
	210 215 220	
10	gga gtg gct tat gct gaa aaa gtt gtg aat ttt gag agg gag atg ttt	782
	Gly Val Ala Tyr Ala Glu Lys Val Val Asn Phe Glu Arg Glu Met Phe	
	225 230 235	
	gag cgg cag cac aat tgg gtt ata aag aca atg gat aga gga gag cct	830
	Glu Arg Gln His Asn Trp Val Ile Lys Thr Met Asp Arg Gly Glu Pro	
15	240 245 250	
	tca ccg gtt agg gat gag tat gct cct gtt ata ttc tct caa gat agt	878
	Ser Pro Val Arg Asp Glu Tyr Ala Pro Val Ile Phe Ser Gln Asp Ser	
	255 260 265 270	
	gtc tct tac ctt gag tca ctc gat atg atg tca ggc gag gag gat cgt	926
20	Val Ser Tyr Leu Glu Ser Leu Asp Met Met Ser Gly Glu Glu Asp Arg	
	275 280 285	
	gag aat att ttg cga gct aga gaa acc gga aaa gct gtc ttg act agc	974
	Glu Asn Ile Leu Arg Ala Arg Glu Thr Gly Lys Ala Val Leu Thr Ser	
	290 295 300	
25	cct ttt agg ttg ttg gaa act cac cat ctc gga gtt gtg ttg aca ttc	1022
	Pro Phe Arg Leu Leu Glu Thr His His Leu Gly Val Val Leu Thr Phe	
	305 310 315	
	cct gtc tac aag tct tct ctt cct gaa aat ccg act gtc gaa gag cgt	1070
	Pro Val Tyr Lys Ser Ser Leu Pro Glu Asn Pro Thr Val Glu Glu Arg	
30	320 325 330	
	att gca gcc act gca ggg tac ctt ggt ggt gcg ttt gat gtg gag tct	1118
	Ile Ala Ala Thr Ala Gly Tyr Leu Gly Gly Ala Phe Asp Val Glu Ser	

	335	340	345	350	
	cta gtc gag aat tta ttt ggt cag ctt gct ggt aac caa gca ata gtt	1166			
	Leu Val Glu Asn Leu Phe Gly Gln Leu Ala Gly Asn Gln Ala Ile Val				
	355	360	365		
5	gtg cat gtg tat gat atc acc aat gca tca gat cca ctt gtc atg tat	1214			
	Val His Val Tyr Asp Ile Thr Asn Ala Ser Asp Pro Leu Val Met Tyr				
	370	375	380		
	ggg aat caa gat gaa gaa gcc gac aga tct ctc tct cat gag agc aag	1262			
	Gly Asn Gln Asp Glu Glu Ala Asp Arg Ser Leu Ser His Glu Ser Lys				
10	385	390	395		
	ctc gat ttt gga gac ccc ttc agg aaa cat aag atg ata tgc agg tac	1310			
	Leu Asp Phe Gly Asp Pro Phe Arg Lys His Lys Met Ile Cys Arg Tyr				
	400	405	410		
	cac caa aag gca cca ata cca ttg aat gtg ctc aca act gtg cca ttg	1358			
15	His Gln Lys Ala Pro Ile Pro Leu Asn Val Leu Thr Thr Val Pro Leu				
	415	420	425	430	
	ttc ttt gcg att ggt ttc ttg gtg ggt tat ata ctg tat ggt gca gct	1406			
	Phe Phe Ala Ile Gly Phe Leu Val Gly Tyr Ile Leu Tyr Gly Ala Ala				
	435	440	445		
20	atg cac ata gta aaa gtc gaa gat gat ttc cat gaa atg caa gag ctt	1454			
	Met His Ile Val Lys Val Glu Asp Asp Phe His Glu Met Gln Glu Leu				
	450	455	460		
	aaa gtg cga gca gaa gct gct gat gtc gct aaa tcg cag ttt ctt gct	1502			
	Lys Val Arg Ala Glu Ala Ala Asp Val Ala Lys Ser Gln Phe Leu Ala				
25	465	470	475		
	acc gtg tct cac gag atc agg aca cca atg aat ggc att ctc gga atg	1550			
	Thr Val Ser His Glu Ile Arg Thr Pro Met Asn Gly Ile Leu Gly Met				
	480	485	490		
	ctt gct atg ctc cta gat aca gaa cta agc tcg aca cag aga gat tac	1598			
30	Leu Ala Met Leu Leu Asp Thr Glu Leu Ser Ser Thr Gln Arg Asp Tyr				
	495	500	505	510	
	gct caa acc gct caa gta tgt ggt aaa gct ttg att gca ttg ata aat	1646			

	Ala Gln Thr Ala Gln Val Cys Gly Lys Ala Leu Ile Ala Leu Ile Asn	
	515 520 525	
	gag gtt ctt gat cgc gcc aag att gaa gct gga aag ctg gag ttg gaa	1694
	Glu Val Leu Asp Arg Ala Lys Ile Glu Ala Gly Lys Leu Glu Leu Glu	
5	530 535 540	
	tca gta cca ttt gat atc cgt tca ata ttg gat gat gtc ctt tct cta	1742
	Ser Val Pro Phe Asp Ile Arg Ser Ile Leu Asp Asp Val Leu Ser Leu	
	545 550 555	
	ttc tct gag gag tca agg aac aaa ggc att gag ctc gcg gtt ttc gtt	1790
10	Phe Ser Glu Glu Ser Arg Asn Lys Gly Ile Glu Leu Ala Val Phe Val	
	560 565 570	
	tca gac aaa gta cca gag ata gtc aaa gga gat tca ggg aga ttt aga	1838
	Ser Asp Lys Val Pro Glu Ile Val Lys Gly Asp Ser Gly Arg Phe Arg	
	575 580 585 590	
15	cag ata atc ata aac ctt gtt gga aat tcg gtt aaa ttc aca gag aaa	1886
	Gln Ile Ile Ile Asn Leu Val Gly Asn Ser Val Lys Phe Thr Glu Lys	
	595 600 605	
	gga cat atc ttt gtt aaa gtc cat ctt gcg gaa caa tca aaa gat gaa	1934
	Gly His Ile Phe Val Lys Val His Leu Ala Glu Gln Ser Lys Asp Glu	
20	610 615 620	
	tct gaa ccg aaa aat gca ttg aat ggt gga gtg tct gaa gaa atg atc	1982
	Ser Glu Pro Lys Asn Ala Leu Asn Gly Gly Val Ser Glu Glu Met Ile	
	625 630 635	
	gtt gtt tcc aaa cag tca agt tac aac aca ttg agc ggt tac gaa gct	2030
25	Val Val Ser Lys Gln Ser Ser Tyr Asn Thr Leu Ser Gly Tyr Glu Ala	
	640 645 650	
	gct gat ggt cgg aat agc tgg gat tca ttc aag cat ttg gtc tct gag	2078
	Ala Asp Gly Arg Asn Ser Trp Asp Ser Phe Lys His Leu Val Ser Glu	
	655 660 665 670	
30	gag cag tca tta tcg gag ttt gat att tct agc aat gtt agg ctt atg	2126
	Glu Gln Ser Leu Ser Glu Phe Asp Ile Ser Ser Asn Val Arg Leu Met	
	675 680 685	

	gtt tca atc gaa gac acg ggt att gga atc cct tta gtt gca caa ggc	2174
	Val Ser Ile Glu Asp Thr Gly Ile Gly Ile Pro Leu Val Ala Gln Gly	
	690 695 700	
	cgt gtg ttt atg ccg ttt atg caa gca gat agc tcg act tca aga aac	2222
5	Arg Val Phe Met Pro Phe Met Gln Ala Asp Ser Ser Thr Ser Arg Asn	
	705 710 715	
	tat gga ggt act ggt att ggt ttg agt ata agc aag tgt ctt gtt gaa	2270
	Tyr Gly Gly Thr Gly Ile Gly Leu Ser Ile Ser Lys Cys Leu Val Glu	
	720 725 730	
10	ctt atg cgt ggt cag ata aat ttc ata agc cgg cct cat att gga agc	2318
	Leu Met Arg Gly Gln Ile Asn Phe Ile Ser Arg Pro His Ile Gly Ser	
	735 740 745 750	
	acg ttc tgg ttc acg gct gtt tta gag aaa tgc gat aaa tgc agt gcg	2366
	Thr Phe Trp Phe Thr Ala Val Leu Glu Lys Cys Asp Lys Cys Ser Ala	
15	755 760 765	
	att aac cat atg aag aaa cct aat gtg gaa cac ttg cct tct act ttt	2414
	Ile Asn His Met Lys Lys Pro Asn Val Glu His Leu Pro Ser Thr Phe	
	770 775 780	
	aaa gga atg aaa gct ata gtt gtt gat gct aag cct gtt aga gct gct	2462
20	Lys Gly Met Lys Ala Ile Val Val Asp Ala Lys Pro Val Arg Ala Ala	
	785 790 795	
	gtg act aga tac cat atg aaa aga ctc gga atc aat gtt gat gtc gtg	2510
	Val Thr Arg Tyr His Met Lys Arg Leu Gly Ile Asn Val Asp Val Val	
	800 805 810	
25	aca agt ctc aaa acc gct gtt gtt gca gct gct gcg ttt gaa aga aac	2558
	Thr Ser Leu Lys Thr Ala Val Val Ala Ala Ala Ala Phe Glu Arg Asn	
	815 820 825 830	
	ggg tct cct ctc cca aca aaa ccg caa ctt gat atg atc tta gta gag	2606
	Gly Ser Pro Leu Pro Thr Lys Pro Gln Leu Asp Met Ile Leu Val Glu	
30	835 840 845	
	aaa gat tca tgg att tca act gaa gat aat gac tca gag att cgt tta	2654
	Lys Asp Ser Trp Ile Ser Thr Glu Asp Asn Asp Ser Glu Ile Arg Leu	

	850	855	860	
	ttg aat tca aga acc aac gga aac gtt cat cac aag tct ccg aaa cta	2702		
	Leu Asn Ser Arg Thr Asn Gly Asn Val His His Lys Ser Pro Lys Leu			
	865	870	875	
5	gct cta ttc gca aca aac atc aca aat tcg gag ttc gac aga gct aaa	2750		
	Ala Leu Phe Ala Thr Asn Ile Thr Asn Ser Glu Phe Asp Arg Ala Lys			
	880	885	890	
	tcc gca gga ttt gca gat acg gta ata atg aaa ccg tta aga gca agc	2798		
	Ser Ala Gly Phe Ala Asp Thr Val Ile Met Lys Pro Leu Arg Ala Ser			
10	895	900	905	910
	atg att ggg gcg tgt ctg caa caa gtt ctc gag ctg aga aaa aca aga	2846		
	Met Ile Gly Ala Cys Leu Gln Gln Val Leu Glu Leu Arg Lys Thr Arg			
	915	920	925	
	caa caa cat cca gaa gga tca tca ccc gca act ctc aag agc ttg ctt	2894		
15	Gln Gln His Pro Glu Gly Ser Ser Pro Ala Thr Leu Lys Ser Leu Leu			
	930	935	940	
	aca ggg aag aag att ctt gtg gtt gat gat aat ata gtt aac agg aga	2942		
	Thr Gly Lys Lys Ile Leu Val Val Asp Asp Asn Ile Val Asn Arg Arg			
	945	950	955	
20	gta gct gca gga gct ctc aag aaa ttt gga gca gaa gtg gtt tgt gca	2990		
	Val Ala Ala Gly Ala Leu Lys Lys Phe Gly Ala Glu Val Val Cys Ala			
	960	965	970	
	gag agt ggt caa gtt gct ttg ggt ttg ctt cag att cca cac act ttc	3038		
	Glu Ser Gly Gln Val Ala Leu Gly Leu Leu Gln Ile Pro His Thr Phe			
25	975	980	985	990
	gat gct tgc ttc atg gat att caa atg cca cag atg gac gga ttt gaa	3086		
	Asp Ala Cys Phe Met Asp Ile Gln Met Pro Gln Met Asp Gly Phe Glu			
	995	1000	1005	
	gca act cgt cag ata aga atg atg gag aag gaa gct aaa gag aag	3131		
30	Ala Thr Arg Gln Ile Arg Met Met Glu Lys Glu Ala Lys Glu Lys			
	1010	1015	1020	
	acg aat ctc gaa tgg cat tta ccg att cta gcg atg act gcg gat	3176		

Thr Asn Leu Glu Trp His Leu Pro Ile Leu Ala Met Thr Ala Asp
 1025 1030 1035
 gtg ata cac gcg acc tac gag gaa tgt ctg aaa agt ggg atg gat 3221
 Val Ile His Ala Thr Tyr Glu Glu Cys Leu Lys Ser Gly Met Asp
 5 1040 1045 1050
 ggt tac gtc tcc aaa cct ttt gaa gaa gag aat ctc tat aaa tcc 3266
 Gly Tyr Val Ser Lys Pro Phe Glu Glu Glu Asn Leu Tyr Lys Ser
 1055 1060 1065
 gtt gcc aaa tca ttc aaa cct aat cct atc tca cct tcg tcg taa 3311
 10 Val Ala Lys Ser Phe Lys Pro Asn Pro Ile Ser Pro Ser Ser
 1070 1075 1080
 tccaatcttc cggcgagttt ttttctctcc gcagccggaa gaatggactg cttctgctga 3371
 ttgattggga taaatatgca ttttggtttc tgtacatata gtaggttcac aatctagaga 3431
 ttttgaaggt tttttttaat ttctcactta agtaatgtag cttgccatga ctagtgtatg 3491
 15 ttgttaaacg acgacgtcta agatggttca gtgttgatct tagcgtaagt attaatccca 3551
 tgggaatcgg ttgtactgta tcagatttgg ttagtcgttt aancattgta atgttctaata 3611
 aatcactttt tcatatatan cctcatgtta taacatgaga cgagaccatt ttgattaaaa 3671
 aaaaaaaaaa aaaaaaaaaa aaa 3694
 <210> 9
 20 <211> 1080
 <212> PRT
 <213> Arabidopsis thaliana
 <400> 9
 Met Arg Arg Asp Phe Val Tyr Asn Asn Asn Ala Met Phe Asn Pro Leu
 25 1 5 10 15
 Thr Thr His Tyr Ser Ser Asp Met Asn Trp Ala Leu Asn Asn His Gln
 20 25 30
 Glu Glu Glu Glu Glu Pro Arg Arg Ile Glu Ile Ser Asp Ser Glu Ser
 35 40 45
 30 Leu Glu Asn Leu Lys Ser Ser Asp Phe Tyr Gln Leu Gly Gly Gly Gly
 50 55 60
 Ala Leu Asn Ser Ser Glu Lys Pro Arg Lys Ile Asp Phe Trp Arg Ser

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

	325	330	335
	Ala Thr Ala Gly Tyr Leu Gly Gly Ala Phe Asp Val Glu Ser Leu Val		
	340	345	350
	Glu Asn Leu Phe Gly Gln Leu Ala Gly Asn Gln Ala Ile Val Val His		
5	355	360	365
	Val Tyr Asp Ile Thr Asn Ala Ser Asp Pro Leu Val Met Tyr Gly Asn		
	370	375	380
	Gln Asp Glu Glu Ala Asp Arg Ser Leu Ser His Glu Ser Lys Leu Asp		
	385	390	395
10	Phe Gly Asp Pro Phe Arg Lys His Lys Met Ile Cys Arg Tyr His Gln		
	405	410	415
	Lys Ala Pro Ile Pro Leu Asn Val Leu Thr Thr Val Pro Leu Phe Phe		
	420	425	430
	Ala Ile Gly Phe Leu Val Gly Tyr Ile Leu Tyr Gly Ala Ala Met His		
15	435	440	445
	Ile Val Lys Val Glu Asp Asp Phe His Glu Met Gln Glu Leu Lys Val		
	450	455	460
	Arg Ala Glu Ala Ala Asp Val Ala Lys Ser Gln Phe Leu Ala Thr Val		
	465	470	475
20	Ser His Glu Ile Arg Thr Pro Met Asn Gly Ile Leu Gly Met Leu Ala		
	485	490	495
	Met Leu Leu Asp Thr Glu Leu Ser Ser Thr Gln Arg Asp Tyr Ala Gln		
	500	505	510
	Thr Ala Gln Val Cys Gly Lys Ala Leu Ile Ala Leu Ile Asn Glu Val		
25	515	520	525
	Leu Asp Arg Ala Lys Ile Glu Ala Gly Lys Leu Glu Leu Glu Ser Val		
	530	535	540
	Pro Phe Asp Ile Arg Ser Ile Leu Asp Asp Val Leu Ser Leu Phe Ser		
	545	550	555
30	Glu Glu Ser Arg Asn Lys Gly Ile Glu Leu Ala Val Phe Val Ser Asp		
	565	570	575
	Lys Val Pro Glu Ile Val Lys Gly Asp Ser Gly Arg Phe Arg Gln Ile		

	580	585	590
	Ile Ile Asn Leu Val Gly Asn Ser Val Lys Phe Thr Glu Lys Gly His		
	595	600	605
	Ile Phe Val Lys Val His Leu Ala Glu Gln Ser Lys Asp Glu Ser Glu		
5	610	615	620
	Pro Lys Asn Ala Leu Asn Gly Gly Val Ser Glu Glu Met Ile Val Val		
	625	630	635
	Ser Lys Gln Ser Ser Tyr Asn Thr Leu Ser Gly Tyr Glu Ala Ala Asp		
	645	650	655
10	Gly Arg Asn Ser Trp Asp Ser Phe Lys His Leu Val Ser Glu Glu Gln		
	660	665	670
	Ser Leu Ser Glu Phe Asp Ile Ser Ser Asn Val Arg Leu Met Val Ser		
	675	680	685
	Ile Glu Asp Thr Gly Ile Gly Ile Pro Leu Val Ala Gln Gly Arg Val		
15	690	695	700
	Phe Met Pro Phe Met Gln Ala Asp Ser Ser Thr Ser Arg Asn Tyr Gly		
	705	710	715
	Gly Thr Gly Ile Gly Leu Ser Ile Ser Lys Cys Leu Val Glu Leu Met		
	725	730	735
20	Arg Gly Gln Ile Asn Phe Ile Ser Arg Pro His Ile Gly Ser Thr Phe		
	740	745	750
	Trp Phe Thr Ala Val Leu Glu Lys Cys Asp Lys Cys Ser Ala Ile Asn		
	755	760	765
	His Met Lys Lys Pro Asn Val Glu His Leu Pro Ser Thr Phe Lys Gly		
25	770	775	780
	Met Lys Ala Ile Val Val Asp Ala Lys Pro Val Arg Ala Ala Val Thr		
	785	790	795
	Arg Tyr His Met Lys Arg Leu Gly Ile Asn Val Asp Val Val Thr Ser		
	805	810	815
30	Leu Lys Thr Ala Val Val Ala Ala Ala Ala Phe Glu Arg Asn Gly Ser		
	820	825	830
	Pro Leu Pro Thr Lys Pro Gln Leu Asp Met Ile Leu Val Glu Lys Asp		

	835	840	845
	Ser Trp Ile Ser Thr Glu Asp Asn Asp Ser Glu Ile Arg Leu Leu Asn		
	850	855	860
	Ser Arg Thr Asn Gly Asn Val His His Lys Ser Pro Lys Leu Ala Leu		
5	865	870	875
	Phe Ala Thr Asn Ile Thr Asn Ser Glu Phe Asp Arg Ala Lys Ser Ala		880
		885	890
	Gly Phe Ala Asp Thr Val Ile Met Lys Pro Leu Arg Ala Ser Met Ile		895
		900	905
10	Gly Ala Cys Leu Gln Gln Val Leu Glu Leu Arg Lys Thr Arg Gln Gln		910
	915	920	925
	His Pro Glu Gly Ser Ser Pro Ala Thr Leu Lys Ser Leu Leu Thr Gly		
	930	935	940
	Lys Lys Ile Leu Val Val Asp Asp Asn Ile Val Asn Arg Arg Val Ala		
15	945	950	955
	Ala Gly Ala Leu Lys Lys Phe Gly Ala Glu Val Val Cys Ala Glu Ser		960
		965	970
	Gly Gln Val Ala Leu Gly Leu Leu Gln Ile Pro His Thr Phe Asp Ala		975
		980	985
20	Cys Phe Met Asp Ile Gln Met Pro Gln Met Asp Gly Phe Glu Ala Thr		990
	995	1000	1005
	Arg Gln Ile Arg Met Met Glu Lys Glu Ala Lys Glu Lys Thr Asn		
	1010	1015	1020
	Leu Glu Trp His Leu Pro Ile Leu Ala Met Thr Ala Asp Val Ile		
25	1025	1030	1035
	His Ala Thr Tyr Glu Glu Cys Leu Lys Ser Gly Met Asp Gly Tyr		
	1040	1045	1050
	Val Ser Lys Pro Phe Glu Glu Glu Asn Leu Tyr Lys Ser Val Ala		
	1055	1060	1065
30	Lys Ser Phe Lys Pro Asn Pro Ile Ser Pro Ser Ser		
	1070	1075	1080

<210> 10

<211> 5474
 <212> DNA
 <213> Oryza (rice)
 <220>
 5 <221> exon
 <222> (1)..(348)
 <220>
 <221> Intron
 <222> (349)..(711)
 10 <220>
 <221> exon
 <222> (712)..(945)
 <220>
 <221> Intron
 15 <222> (946)..(1040)
 <220>
 <221> exon
 <222> (1041)..(1207)
 <220>
 20 <221> Intron
 <222> (1208)..(1508)
 <220>
 <221> exon
 <222> (1509)..(1727)
 25 <220>
 <221> Intron
 <222> (1728)..(1818)
 <220>
 <221> exon
 30 <222> (1819)..(2005)
 <220>
 <221> Intron

<222> (2006)..(2087)
 <220>
 <221> exon
 <222> (2088)..(2139)
 5 <220>
 <221> Intron
 <222> (2140)..(2233)
 <220>
 <221> exon
 10 <222> (2234)..(2463)
 <220>
 <221> Intron
 <222> (2464)..(2694)
 <220>
 15 <221> exon
 <222> (2695)..(2793)
 <220>
 <221> Intron
 <222> (2794)..(2910)
 20 <220>
 <221> exon
 <222> (2911)..(3641)
 <220>
 <221> Intron
 25 <222> (3642)..(4079)
 <220>
 <221> exon
 <222> (4080)..(4595)
 <220>
 30 <221> Intron
 <222> (4596)..(5215)
 <220>

<221> exon

<222> (5216) .. (5474)

<400> 10

	atg ggt gtg gga gga ggc gga gga gga gga gga gga gga ggc ggc ggc	48
5	met Gly Val Gly Gly Gly Gly Gly Gly Gly Gly Gly Glu Ala Ala Ala	
	1 5 10 15	
	gcg gtg gcg gtg gag ggg gat gag gcg ggg aag ggg agg agg tgg tgg	96
	Ala Val Ala Val Glu Gly Asp Glu Ala Gly Lys Gly Arg Arg Trp Trp	
	20 25 30	
10	agg gtg aag gtg aag ctg agc acg gtg gcg gtg gtg gcg tgg gtg ctg	144
	Arg Val Lys Val Lys Leu Ser Thr Val Ala Val Val Ala Trp Val Leu	
	35 40 45	
	gcg tcg gcg gcg ctc tgg gcg ggg ctg cac tgg cgc ttc cgc cgc gcg	192
	Ala Ser Ala Ala Leu Trp Ala Gly Leu His Trp Arg Phe Arg Arg Ala	
15	50 55 60	
	gcg ctg cac aag gcc gag gag gcc ctc gtc tgc atg tgc gag gag cgc	240
	Ala Leu His Lys Ala Glu Glu Ala Leu Val Cys Met Cys Glu Glu Arg	
	65 70 75 80	
	gcc cgc atg ctg cag gac cag ttc gcc gtc tcc gtc aac cac gtc cac	288
20	Ala Arg Met Leu Gln Asp Gln Phe Ala Val Ser Val Asn His Val His	
	85 90 95	
	gcc ctc gcc atc ctc gtc gcc acc ttc cac tac gac aag cac cct ccc	336
	Ala Leu Ala Ile Leu Val Ala Thr Phe His Tyr Asp Lys His Pro Pro	
	100 105 110	
25	gcc ctc gac cag gtcggcccga actccgacga gctcttccgc cgccgccgcg	388
	Ala Leu Asp Gln	
	115	
	atgatactgt tgcatactgtt gtttttttgc cccccgcggt taattgcat aatgcctcga	448
	tttttactcc acatacttgcc cgtgtacttc gctctgctgc ttcttcggct tcatttaatt	508
30	ctaccgtgac cttccgtgtc agccatggaa gccatggatt actgttgctg tctcttgcta	568
	ttatatggag cgcacttttt gttgggaggg agtgaattga ttgtgcttgc ttgcttctgt	628
	tggtagtagt actagtatt tctttggctg tgtggctgat gaatcttctt ccatgtgttg	688

4	tcgctgctg cgcgtgtgtg cag gac acg ttc gcc gtg tac gcc gcg agg acg	741
	Asp Thr Phe Ala Val Tyr Ala Ala Arg Thr	
	120 125	
	tcc ttc gag cgg ccg ctg ctg agc ggc gtg gcg tac gcg cag cgg gtg	789
5	Ser Phe Glu Arg Pro Leu Leu Ser Gly Val Ala Tyr Ala Gln Arg Val	
	130 135 140	
	gtg cac gcc gac agg gag agc ttc gag cgg cag cag ggg tgg atc atc	837
	Val His Ala Asp Arg Glu Ser Phe Glu Arg Gln Gln Gly Trp Ile Ile	
	145 150 155	
10	aag acc atg aag cac gag ccg tcc ccg gcg cag gac gag tac gcc ccg	885
	Lys Thr Met Lys His Glu Pro Ser Pro Ala Gln Asp Glu Tyr Ala Pro	
	160 165 170	
	gtg atc tac tcc cag gag acc atc tcc tac atc gag ggc ctc gac gtc	933
	Val Ile Tyr Ser Gln Glu Thr Ile Ser Tyr Ile Glu Gly Leu Asp Val	
15	175 180 185 190	
	atg tcc ggc gag gtgcgtttct tgggttacag cttcgcagct gctgctgcgg	985
	Met Ser Gly Glu	
	ttatcgccat gtccgtgct ctgaactgtg ctgctgggtg tgcttcggcc tgcag gag	1043
20	Glu	
	195	
	gac agg gag aac atc ttg agg gcg agg gcg aca ggg aag gcc gtc ctc	1091
	Asp Arg Glu Asn Ile Leu Arg Ala Arg Ala Thr Gly Lys Ala Val Leu	
	200 205 210	
25	acg agg ccg ttc cgg ctg atg tcg aat cac ttg ggt gtt gtc ttg acg	1139
	Thr Arg Pro Phe Arg Leu Met Ser Asn His Leu Gly Val Val Leu Thr	
	215 220 225	
	ttt cct gtc tac ctc gtc gat ctc cca aat gat acc gcg gtg gag gat	1187
	Phe Pro Val Tyr Leu Val Asp Leu Pro Asn Asp Thr Ala Val Glu Asp	
30	230 235 240	
	cgt gtt gct gct act gca gg gtgagggatt actttacttt tctgaatgaa	1237
	Arg Val Ala Ala Thr Ala Gly	

	245		
		gattattctc tccaactgat tcctcttctg tctggaatcc actgccttca gctcttcggt	1297
		ttgttgcaagt cgttgttgga tgcttttagt agtggaaatg tgtgcgtttc agggatattt	1357
		gatcacatgc aacattttca ctcataactg gctgaaaaag ttttgcatta atagagctga	1417
5		aatgtctaga tggataagca attgcagtgg tattttaagt acaacatgtg caacgaatgg	1477
		ctccatttaa ctttttcttt ttgttcggca g a tac ctt gga gga gca ttt gat	1530
		Tyr Leu Gly Gly Ala Phe Asp	
		255	
		gtg gag tca cta gta gaa aat ttg ttt aga cag tta gct ggt aac cag	1578
10		Val Glu Ser Leu Val Glu Asn Leu Phe Arg Gln Leu Ala Gly Asn Gln	
		260 265 270	
		gaa ttg gtg gtc aat gtt tat gat gtc aca aac cac tca aac cct ctt	1626
		Glu Leu Val Val Asn Val Tyr Asp Val Thr Asn His Ser Asn Pro Leu	
		275 280 285	
15		gtg atg tat gga tct gag gtt cct ctt ggt att ccc tca cca tca cac	1674
		Val Met Tyr Gly Ser Glu Val Pro Leu Gly Ile Pro Ser Pro Ser His	
		290 295 300 305	
		acc tat acg ttg gat ttt ggt gat cca ttg aga aag cat cag atg gtt	1722
		Thr Tyr Thr Leu Asp Phe Gly Asp Pro Leu Arg Lys His Gln Met Val	
20		310 315 320	
		tgc ag gtaaatttgt gtgaattgat cgttggtttt cccattttat attatagaac	1777
		Cys Arg	
		gatcgggtttt tttaacatcc attggccata aatctgagca g a tac aga aac aag	1831
25		Tyr Arg Asn Lys	
		325	
		ctt cat gtt tca tgg tct gca att act aca cca tca ggg gtc ttt gtt	1879
		Leu His Val Ser Trp Ser Ala Ile Thr Thr Pro Ser Gly Val Phe Val	
		330 335 340	
30		ata tgt atg ctg gtg ggc tac ata ata tat gct gct tgg agt cgc tac	1927
		Ile Cys Met Leu Val Gly Tyr Ile Ile Tyr Ala Ala Trp Ser Arg Tyr	
		345 350 355	

	gat aat gtt aag gaa gat tgc cgg aaa atg gaa gcg ctg aaa aaa cgg	1975
	Asp Asn Val Lys Glu Asp Cys Arg Lys Met Glu Ala Leu Lys Lys Arg	
	360 365 370 375	
	gca gaa gcg gct gat att gct aaa tct cag gtatagttgg atgttgtttg	2025
5	Ala Glu Ala Ala Asp Ile Ala Lys Ser Gln	
	380 385	
	cttctctatt tctattgcaa gcttattggt atatctaaaa gggtcttatt catttatgac	2085
	ag ttc ctt gca act gtt tct cat gag atc aga aca ccc atg aat ggc	2132
	Phe Leu Ala Thr Val Ser His Glu Ile Arg Thr Pro Met Asn Gly	
10	390 395 400	
	gtg ctg g gtattttctt tgatcttaca acacattcag tttaatgtta tgcaactcat	2189
	Val Leu	
	ttcttttgaa aaaatggaat catctctttg tttcttttcc ctag ga atg ctt gat	2244
15	Gly Met Leu Asp	
	405	
	atg tta tta gac aca gag ctg aag tca acc cag agg gat tat gca caa	2292
	Met Leu Leu Asp Thr Glu Leu Lys Ser Thr Gln Arg Asp Tyr Ala Gln	
	410 415 420	
20	aca gcc caa gtc tgt gga aag gca tta ata tcc ctg att aac gaa gtg	2340
	Thr Ala Gln Val Cys Gly Lys Ala Leu Ile Ser Leu Ile Asn Glu Val	
	425 430 435	
	ctt gac agg gcc aaa atc gag gct ggc aag ata gat ctc gag tca gta	2388
	Leu Asp Arg Ala Lys Ile Glu Ala Gly Lys Ile Asp Leu Glu Ser Val	
25	440 445 450	
	cca ttt gac ctg agg tcc atc ctt gat gat gtc atc tcg tta ttt tct	2436
	Pro Phe Asp Leu Arg Ser Ile Leu Asp Asp Val Ile Ser Leu Phe Ser	
	455 460 465 470	
	tca aaa tca aga gag aaa gga att gag gttagttaaa ctgatttcgg	2483
30	Ser Lys Ser Arg Glu Lys Gly Ile Glu	
	475	
	tcatggttg acaaagatca ctaaactgat taagtttctg ccagccatca attatttctt	2543

	ttaggaaaat atcatgcact agttccaccg acatctttta gtctcttagc ttgatactct	2603
	ttccatgaac ttctctgcat taccgtcatg caccatgcac gtttaacttt gtttaatccc	2663
	agttgatttt cttctatggt gtaacttcca g ctt gct gta tat gtt tct gaa	2715
	Leu Ala Val Tyr Val Ser Glu	
5	480 485	
	aga gtt cct gaa atc ctg ttg ggc gac cct gga agg ttt cgt cag ata	2763
	Arg Val Pro Glu Ile Leu Leu Gly Asp Pro Gly Arg Phe Arg Gln Ile	
	490 495 500	
	att aca aac ttg gtg gga aac tcg atc aag gtaaattgcgc ataactttg	2813
10	Ile Thr Asn Leu Val Gly Asn Ser Ile Lys	
	505 510	
	tatccattca tgattttctt taacgatacc aatagttctc accaatgaca tcaggcaact	2873
	tgtttcttag tatactattg ttcaatgtga acacaag ata aca ata ttt acc ttg	2928
	Ile Thr Ile Phe Thr Leu	
15	515	
	tcg cag ttc aca gaa cgg ggg cac att ttt gta caa gtt cac ctg gca	2976
	Ser Gln Phe Thr Glu Arg Gly His Ile Phe Val Gln Val His Leu Ala	
	520 525 530	
	gat cac tca aat ctt gca aca gaa gca aaa att gaa cca gta gtc aat	3024
20	Asp His Ser Asn Leu Ala Thr Glu Ala Lys Ile Glu Pro Val Val Asn	
	535 540 545 550	
	ggg atg aat gga cat aaa gac gag gct att gct ata ccc acc agt ggg	3072
	Gly Met Asn Gly His Lys Asp Glu Ala Ile Ala Ile Pro Thr Ser Gly	
	555 560 565	
25	tct cat aac act tta agt ggt ttt gaa gca gct gat agc cga aat aac	3120
	Ser His Asn Thr Leu Ser Gly Phe Glu Ala Ala Asp Ser Arg Asn Asn	
	570 575 580	
	tgg gaa aac ttc aag ctt ttg ctc tct tac gag aaa aat gaa atg cca	3168
	Trp Glu Asn Phe Lys Leu Leu Leu Ser Tyr Glu Lys Asn Glu Met Pro	
30	585 590 595	
	tat gaa agt gat tct gat aaa gta act ctt gtt gtt agt gtg gaa gat	3216
	Tyr Glu Ser Asp Ser Asp Lys Val Thr Leu Val Val Ser Val Glu Asp	

	600	605	610	
	act ggg ata ggc ata cca ctg cat gcc caa ggc cgg gtc ttc acg cct			3264
	Thr Gly Ile Gly Ile Pro Leu His Ala Gln Gly Arg Val Phe Thr Pro			
	615	620	625	630
5	ttc atg caa gct gac agc tca act tct agg aac tat ggt gga act ggc			3312
	Phe Met Gln Ala Asp Ser Ser Thr Ser Arg Asn Tyr Gly Gly Thr Gly			
	635	640	645	
	att gga ttg agc atc agc aaa tgt ctt gtt gaa ata atg ggt ggt cag			3360
	Ile Gly Leu Ser Ile Ser Lys Cys Leu Val Glu Ile Met Gly Gly Gln			
10	650	655	660	
	ata aac ttt gtc agc cga cct ctt gtt ggg agc aca ttc aca ttc act			3408
	Ile Asn Phe Val Ser Arg Pro Leu Val Gly Ser Thr Phe Thr Phe Thr			
	665	670	675	
	gct gtt ctg aga agg tgt gac aaa aat gct att agt gac agt aag act			3456
15	Ala Val Leu Arg Arg Cys Asp Lys Asn Ala Ile Ser Asp Ser Lys Thr			
	680	685	690	
	gtt gct ttg cac cca tta ccg tcc agt ttt aaa ggc tta tct gcg cta			3504
	Val Ala Leu His Pro Leu Pro Ser Ser Phe Lys Gly Leu Ser Ala Leu			
	695	700	705	710
20	ttg gtt gat aaa aga cct gta aga gca act gtg act aag tat cat ttg			3552
	Leu Val Asp Lys Arg Pro Val Arg Ala Thr Val Thr Lys Tyr His Leu			
	715	720	725	
	caa agg ctt gga atc act tct gaa gtt gtt ggt acc att gat ccg aca			3600
	Gln Arg Leu Gly Ile Thr Ser Glu Val Val Gly Thr Ile Asp Pro Thr			
25	730	735	740	
	ttt ggt gtg ttg tct ggg aga aat ggc agt tct cta acc ag			3641
	Phe Gly Val Leu Ser Gly Arg Asn Gly Ser Ser Leu Thr Ser			
	745	750	755	
	gtacttctat cttctacatt ctttcaaaa aattgaaatc ctggagttaa taggctactt			3701
30	ttctctggaa attagaataa acggagcatg cttgcatact aacttcttat gcaaatatca			3761
	ctgttctatg tataaatgat tacacgataa tgattgtttt ttggtaagta ttactgggta			3821
	atacttggcc atcatagttc ttgcctttat tttgtatcac ttcggtatat tgctacttct			3881

	gcggcagttc tctttacgct accgaatgtg atatatttaa ttgaaaattg attttatttt	3941
	aatctgtgaa aagaacattt ttttaaggcc ccacaattct caaataacta gtaagctgat	4001
	tgcagatgat atttgaacta tcttgacacc cagttttttt ttttaacta actcgttttg	4061
	ttccaaattt gttgtcag c att ggt aag aag cag cca tgc atg ttg cta atc	4113
5	Ile Gly Lys Lys Gln Pro Cys Met Leu Leu Ile	
	760 765	
	gag agt gat tcc tgg gga cca cag atg gat gtc tcc tta cat gct aga	4161
	Glu Ser Asp Ser Trp Gly Pro Gln Met Asp Val Ser Leu His Ala Arg	
	770 775 780	
10	ctt cag gag atg aaa cag agt gat cgc ata cat gta ttg ccc aag gtt	4209
	Leu Gln Glu Met Lys Gln Ser Asp Arg Ile His Val Leu Pro Lys Val	
	785 790 795	
	ttc ctt ctt tct gct gca gaa tca gac aaa gta aag aag ata cat gca	4257
	Phe Leu Leu Ser Ala Ala Glu Ser Asp Lys Val Lys Lys Ile His Ala	
15	800 805 810 815	
	gtt gat tct gtg ata cca aag cct ctg aaa gca agt gca ctt gcg gcc	4305
	Val Asp Ser Val Ile Pro Lys Pro Leu Lys Ala Ser Ala Leu Ala Ala	
	820 825 830	
	tgt ctg ttc caa gca ctt ggt atc aca cag ccg agc cat gag aaa cgt	4353
20	Cys Leu Phe Gln Ala Leu Gly Ile Thr Gln Pro Ser His Glu Lys Arg	
	835 840 845	
	gac gat tca ggt tct ctt cat ggg cgt gat ggt tca ggt tct ctt cat	4401
	Asp Asp Ser Gly Ser Leu His Gly Arg Asp Gly Ser Gly Ser Leu His	
	850 855 860	
25	ggg ttg ctt ctt ggc aag aac ata ttg gta gtt gat gac aac aag gta	4449
	Gly Leu Leu Leu Gly Lys Asn Ile Leu Val Val Asp Asp Asn Lys Val	
	865 870 875	
	aac ctc aga gtg gcc gct ggt aca ttg aag aaa tat ggg gca aag gtg	4497
	Asn Leu Arg Val Ala Ala Gly Thr Leu Lys Lys Tyr Gly Ala Lys Val	
30	880 885 890 895	
	gag tgt gtg gag agt gga aaa gat gct ctt tcc ctt cta caa gtg ccg	4545
	Glu Cys Val Glu Ser Gly Lys Asp Ala Leu Ser Leu Leu Gln Val Pro	

	900	905	910	
	cac aag ttt gat ctg tgt ctc atg gac att cag atg ccg gag atg gat			4593
	His Lys Phe Asp Leu Cys Leu Met Asp Ile Gln Met Pro Glu Met Asp			
	915	920	925	
5	gg gtaagcttat gtcccggttca agatttattt cttttgaatt tgagcctttt			4645
	Gly			
	atttcatatg atccaaagcg tgacatgtca tctggataag tgccacctcg ttgactatca			4705
	atttattcac gcaatgaatc aggctttctg cctctttttg aagaaaaaaa aaatgttatt			4765
10	cacacttcgt ggttaatgga aacaagcata tacattgctg acgccaaata ctgattataa			4825
	ctagtaaaaa tgtcagtctt gtgttagttt tttatgtgcg aaaaatgccg tgtctaattc			4885
	attatgtgaa aattttcaaa gacattagcc ctgatatcgt cctgcattat gccaaacagt			4945
	actttaaata ttaatatcag aaggtctaaa atccgtgcag ttaccatttg ttcattgtag			5005
	ttttaaccat catactgcaa tatgcagttc ttgggctaata gaacataatg gtgtgctaac			5065
15	accctcatga ccaattaact tatgtctctt gaacacttgc tgcattattga catctctgtt			5125
	cctatatattt ctgaatagta accaagtaat gtcaaaccat gtcattattt tgtcttcgtt			5185
	taatcaacat agttttgctt catgtgctag a ttt gag gca act cgg caa ata			5237
		Phe Glu Ala Thr Arg Gln Ile		
		930	935	
20	cga gca atg gaa ggg aag gca aat gag cag gca gac gac agc gaa tcg			5285
	Arg Ala Met Glu Gly Lys Ala Asn Glu Gln Ala Asp Asp Ser Glu Ser			
	940	945	950	
	ggc tca gaa atc gca gca aag acg gcc aaa tgg cac ttg cca atc ctg			5333
	Gly Ser Glu Ile Ala Ala Lys Thr Ala Lys Trp His Leu Pro Ile Leu			
25	955	960	965	
	gca atg acc gct gat gtc atc cag gcc acc cac gag gaa tgc aca aag			5381
	Ala Met Thr Ala Asp Val Ile Gln Ala Thr His Glu Glu Cys Thr Lys			
	970	975	980	
	tgc ggg atg gat ggc tac gtc tcg aag ccc ttt gag gag aag cag ctc			5429
30	cys Gly Met Asp Gly Tyr Val Ser Lys Pro Phe Glu Glu Lys Gln Leu			
	985	990	995	
	ttc cag gca gta cag aag ttc ttg ggc cca tgc gtt tcc agc tga			5474

Phe Gln Ala Val Gln Lys Phe Leu Gly Pro Cys Val Ser Ser
 1000 1005 1010
 <210> 11
 <211> 3046
 5 <212> DNA
 <213> Zea mays
 <220>
 <221> 5'UTR
 <222> (1)..(15)
 10 <220>
 <221> CDS
 <222> (16)..(3003)
 <220>
 <221> 3'UTR
 15 <222> (3004)..(3046)
 <400> 11
 ggggtggagc ctggg atg ggg gtg gga ggc gga ggc gga ggg gag gcc gct 51
 Met Gly Val Gly Gly Gly Gly Gly Glu Ala Ala
 1 5 10
 20 gcg gtg tcg gcg ccg gcg ccg gcg gag gag gcg ggg aag gac gcg gag 99
 Ala Val Ser Ala Pro Ala Pro Ala Glu Glu Ala Gly Lys Asp Ala Glu
 15 20 25
 gat ggc ggc ggc tgg acc ttg aag gcg aag ctg atc gcc gta gcg gtg 147
 Asp Gly Gly Gly Trp Thr Leu Lys Ala Lys Leu Ile Ala Val Ala Val
 25 30 35 40
 ctg gtg tgg gtg ctg ggg gcc ttg gcg ctc ggg gtg ttc ctg cac tcc 195
 Leu Val Trp Val Leu Gly Ala Leu Ala Leu Gly Val Phe Leu His Ser
 45 50 55 60
 tac ttc cgc cac gcg gcg ctg cgc aag gcg gag gaa ggg ctc gtc agc 243
 30 Tyr Phe Arg His Ala Ala Leu Arg Lys Ala Glu Glu Gly Leu Val Ser
 65 70 75
 atg tgc gag gag cgc gcg cgc atg ctg cag gac cag ttc gcc gtc tcc 291

	Met Cys Glu Glu Arg Ala Arg Met Leu Gln Asp Gln Phe Ala Val Ser	
	80 85 90	
	gtc aac cac gtc cac gcc ctc gcc atc ctc gtc gcc acc ttc cac tac	339
	Val Asn His Val His Ala Leu Ala Ile Leu Val Ala Thr Phe His Tyr	
5	95 100 105	
	gag aag cgc ccg ccc gcg ctc gac cag aac acg ttc gcg gac tac acc	387
	Glu Lys Arg Pro Pro Ala Leu Asp Gln Asn Thr Phe Ala Asp Tyr Thr	
	110 115 120	
	gcc agg acg tcg ttc gag cgg ccg ctg ctg agc ggg gtg gcg tac gcg	435
10	Ala Arg Thr Ser Phe Glu Arg Pro Leu Leu Ser Gly Val Ala Tyr Ala	
	125 130 135 140	
	cag cgg gtg gtg cac ggc gac agg gag agc ttc gag cgc cag cag ggc	483
	Gln Arg Val Val His Gly Asp Arg Glu Ser Phe Glu Arg Gln Gln Gly	
	145 150 155	
15	tgg atc atc aag acc atg aag cac gag ccg tct ccg gtg cag gat gag	531
	Trp Ile Ile Lys Thr Met Lys His Glu Pro Ser Pro Val Gln Asp Glu	
	160 165 170	
	tac gcg ccg gtc gtt tac tcg cag gag acc gtc tcc tac att gag ggg	579
	Tyr Ala Pro Val Val Tyr Ser Gln Glu Thr Val Ser Tyr Ile Glu Gly	
20	175 180 185	
	ctc gac atg atg tcc ggc gag gag gac cgt gag aac att ttg agg tca	627
	Leu Asp Met Met Ser Gly Glu Glu Asp Arg Glu Asn Ile Leu Arg Ser	
	190 195 200	
	aga gca tcg ggg aag gca gtt ctt act aga ccg ttc ccg ctc atg tcg	675
25	Arg Ala Ser Gly Lys Ala Val Leu Thr Arg Pro Phe Arg Leu Met Ser	
	205 210 215 220	
	aat cac ctg ggt gtc gta ttg act ttt cct gtc tac cat gtc gat ctt	723
	Asn His Leu Gly Val Val Leu Thr Phe Pro Val Tyr His Val Asp Leu	
	225 230 235	
30	tct tct gat gcc aag gag gag gat cgt gtt gct gcc acc gca gga tac	771
	Ser Ser Asp Ala Lys Glu Glu Asp Arg Val Ala Ala Thr Ala Gly Tyr	
	240 245 250	

	ctt ggg gga tca ttt gat gta gaa tca tta gtg gaa aat ttg ttt agg	819
	Leu Gly Gly Ser Phe Asp Val Glu Ser Leu Val Glu Asn Leu Phe Arg	
	255 260 265	
	cag cta gct ggc aat cag gaa ttg gtg gta aat gtt tat gat gtc aca	867
5	Gln Leu Ala Gly Asn Gln Glu Leu Val Val Asn Val Tyr Asp Val Thr	
	270 275 280	
	aac agt tcg aac cct ctt gtc atg tat gga tcg gaa gtt tct ctt ggc	915
	Asn Ser Ser Asn Pro Leu Val Met Tyr Gly Ser Glu Val Ser Leu Gly	
	285 290 295 300	
10	aac ccc tca cca tcg cac atc tgc atg cta gat ttt ggc gat cca ttc	963
	Asn Pro Ser Pro Ser His Ile Cys Met Leu Asp Phe Gly Asp Pro Phe	
	305 310 315	
	aga aag cat cat atg gtt tgc aga tac aga aac aag cct cag ctc cca	1011
	Arg Lys His His Met Val Cys Arg Tyr Arg Asn Lys Pro Gln Leu Pro	
15	320 325 330	
	tgg tct gca ata tct tcg tca tct ggt gta ttt gtc ata tgt atg ctt	1059
	Trp Ser Ala Ile Ser Ser Ser Ser Gly Val Phe Val Ile Cys Met Leu	
	335 340 345	
	gtg ggg tac atc gtg ggt gcc gct tgg agt cgt tat gat aat gtt aag	1107
20	Val Gly Tyr Ile Val Gly Ala Ala Trp Ser Arg Tyr Asp Asn Val Lys	
	350 355 360	
	gaa gat tgc cgg aaa atg gag gag ctg aaa aaa cag gca gaa gca gcc	1155
	Glu Asp Cys Arg Lys Met Glu Glu Leu Lys Lys Gln Ala Glu Ala Ala	
	365 370 375 380	
25	gat gtt gct aaa tct cag ttc ctt gca act gtt tct cat gag atc aga	1203
	Asp Val Ala Lys Ser Gln Phe Leu Ala Thr Val Ser His Glu Ile Arg	
	385 390 395	
	acg ccc atg aat gga gtt cta ggg atg ctt gat atg ctg tta gac act	1251
	Thr Pro Met Asn Gly Val Leu Gly Met Leu Asp Met Leu Leu Asp Thr	
30	400 405 410	
	gac cta acg tcg acc cag agg gat ttt gca caa aca gct caa gtc tgt	1299
	Asp Leu Thr Ser Thr Gln Arg Asp Phe Ala Gln Thr Ala Gln Val Cys	

	415	420	425	
	gga aag gct tta ata tca cta atc aat gaa gtg ctt gac aga gcg aaa			1347
	Gly Lys Ala Leu Ile Ser Leu Ile Asn Glu Val Leu Asp Arg Ala Lys			
	430	435	440	
5	att gaa gcc gga aag ttg gat ctt gag tct gta cca ttt gac ctg aga			1395
	Ile Glu Ala Gly Lys Leu Asp Leu Glu Ser Val Pro Phe Asp Leu Arg			
	445	450	455	460
	tcc atc ctt gat gat gtc atc tca tta ttt tct tca aag tca aga gag			1443
	Ser Ile Leu Asp Asp Val Ile Ser Leu Phe Ser Ser Lys Ser Arg Glu			
10	465	470	475	
	aag gga att gag ctt gct gta tat gtc tct gaa aga gtt cct gaa ctc			1491
	Lys Gly Ile Glu Leu Ala Val Tyr Val Ser Glu Arg Val Pro Glu Leu			
	480	485	490	
	ttg ttg ggt gat cct gga agg ttt cgg cag ata att aca aat tta gtg			1539
15	Leu Leu Gly Asp Pro Gly Arg Phe Arg Gln Ile Ile Thr Asn Leu Val			
	495	500	505	
	ggc aac tca att aag ttc aca gaa cgg gga cat att ttt gta caa gtt			1587
	Gly Asn Ser Ile Lys Phe Thr Glu Arg Gly His Ile Phe Val Gln Val			
	510	515	520	
20	cat ctg gca gat cac tca aat ctt gca aca gaa tcc aaa gtt gag tca			1635
	His Leu Ala Asp His Ser Asn Leu Ala Thr Glu Ser Lys Val Glu Ser			
	525	530	535	540
	gtg gct aac ggg atg aat gga cat aaa gat gag aaa act gct gta gca			1683
	Val Ala Asn Gly Met Asn Gly His Lys Asp Glu Lys Thr Ala Val Ala			
25	545	550	555	
	acc agt gtt tct ctc aac aca cta agt ggt ttt gaa gct gct gat agc			1731
	Thr Ser Val Ser Leu Asn Thr Leu Ser Gly Phe Glu Ala Ala Asp Ser			
	560	565	570	
	cga aat agt tgg gaa aac ttc aag ctt ttg ctt tct tat gag aaa aat			1779
30	Arg Asn Ser Trp Glu Asn Phe Lys Leu Leu Leu Ser Tyr Glu Lys Asn			
	575	580	585	
	gag atg ccc tat gaa agt gta tct gat aaa gtt act ctt gta gta agt			1827

	Glu Met Pro Tyr Glu Ser Val Ser Asp Lys Val Thr Leu Val Val Ser	
	590 595 600	
	gtg gaa gat aca ggg ata ggt ata cca ttg gat gcc caa gcc aag gtg	1875
	Val Glu Asp Thr Gly Ile Gly Ile Pro Leu Asp Ala Gln Ala Lys Val	
5	605 610 615 620	
	ttc act cct ttc atg cag gcc gat agt tcg act tcc agg aca tat ggt	1923
	Phe Thr Pro Phe Met Gln Ala Asp Ser Ser Thr Ser Arg Thr Tyr Gly	
	625 630 635	
	ggg act gga att gga ttg agc atc agc aaa tgt ctt gtt gaa cta atg	1971
10	Gly Thr Gly Ile Gly Leu Ser Ile Ser Lys Cys Leu Val Glu Leu Met	
	640 645 650	
	ggg ggt cag ata aac ttt gtt agc cga cca cat gtt ggg agt aca ttc	2019
	Gly Gly Gln Ile Asn Phe Val Ser Arg Pro His Val Gly Ser Thr Phe	
	655 660 665	
15	aca ttt act gca gct ctc caa aga tgt gac aga agc gct att ggt gac	2067
	Thr Phe Thr Ala Ala Leu Gln Arg Cys Asp Arg Ser Ala Ile Gly Asp	
	670 675 680	
	agt aag cct gtt atg ttg cac cct ctt cca tcc agt ttc aaa ggt tta	2115
	Ser Lys Pro Val Met Leu His Pro Leu Pro Ser Ser Phe Lys Gly Leu	
20	685 690 695 700	
	tct gca tta ttg gtt gat aga aga cca gta aga gct act gta act aag	2163
	Ser Ala Leu Leu Val Asp Arg Arg Pro Val Arg Ala Thr Val Thr Lys	
	705 710 715	
	tat cac ttg caa agg ctg gga att gcc tgc gat gtt gtt gct acc att	2211
25	Tyr His Leu Gln Arg Leu Gly Ile Ala Cys Asp Val Val Ala Thr Ile	
	720 725 730	
	gaa ttg gct ctt ggt gtg ctg tcc ggg aga aat ggc agt tct cta acc	2259
	Glu Leu Ala Leu Gly Val Leu Ser Gly Arg Asn Gly Ser Ser Leu Thr	
	735 740 745	
30	agc acg aag caa cca tgc atg tta ttg att gag agt gat tca tgg ggc	2307
	Ser Thr Lys Gln Pro Cys Met Leu Leu Ile Glu Ser Asp Ser Trp Gly	
	750 755 760	

	ttc aag att gat gta cct tta cga tct cga ctc ctg gag atg aag cag	2355
	Phe Lys Ile Asp Val Pro Leu Arg Ser Arg Leu Leu Glu Met Lys Gln	
	765 770 775 780	
5	aat ggt cca cct gga ttg ccc aaa act atc ctt ctc gca gct gca gaa	2403
	Asn Gly Pro Pro Gly Leu Pro Lys Thr Ile Leu Leu Ala Ala Ala Glu	
	785 790 795	
	tcg ggc aaa ctc aaa gca cac tat gca gtt gat tct gtg atc acg aag	2451
	Ser Gly Lys Leu Lys Ala His Tyr Ala Val Asp Ser Val Ile Thr Lys	
	800 805 810	
10	cct ctg aaa gca agc gga ctt gcc gct tgt cta ttc caa aca ctt ggc	2499
	Pro Leu Lys Ala Ser Gly Leu Ala Ala Cys Leu Phe Gln Thr Leu Gly	
	815 820 825	
	atc aca cag tca agc aac gag aga cgc gac aac tca ggt tcc ctt cat	2547
	Ile Thr Gln Ser Ser Asn Glu Arg Arg Asp Asn Ser Gly Ser Leu His	
15	830 835 840	
	ggg ttg ctc ctt ggc aag aac ata ttg gtg gtt gat gac aac aag gta	2595
	Gly Leu Leu Leu Gly Lys Asn Ile Leu Val Val Asp Asp Asn Lys Val	
	845 850 855 860	
	aat ctt aga gtg gct gct ggc aca tta aag aaa ttc gga gcg aag gtg	2643
	Asn Leu Arg Val Ala Ala Gly Thr Leu Lys Lys Phe Gly Ala Lys Val	
	865 870 875	
	gag tgc gtg gag agt gga aaa gat gct ctc gcc agc cta caa gtt cca	2691
	Glu Cys Val Glu Ser Gly Lys Asp Ala Leu Ala Ser Leu Gln Val Pro	
	880 885 890	
25	cat aag ttc cat ctt tgt ctc atg gac att cag atg ccc gaa atg gat	2739
	His Lys Phe His Leu Cys Leu Met Asp Ile Gln Met Pro Glu Met Asp	
	895 900 905	
	ggg ttc gag gcc acc aag caa ata agg gca atg gaa gcg aag gca aat	2787
	Gly Phe Glu Ala Thr Lys Gln Ile Arg Ala Met Glu Ala Lys Ala Asn	
30	910 915 920	
	gag cag gca gtc gcc tgt gac gat tca gat acg gat ggc gcg aca agg	2835
	Glu Gln Ala Val Ala Cys Asp Asp Ser Asp Thr Asp Gly Ala Thr Arg	

	925	930	935	940	
	gcg gca aga tgg cac ctg cct gtc ctt gca atg acc gcc gat gtc atc				2883
	Ala Ala Arg Trp His Leu Pro Val Leu Ala Met Thr Ala Asp Val Ile				
		945	950	955	
5	cag gcc acc cat gag gag tgc aca aag tac ggg atg gat ggg tac gtc				2931
	Gln Ala Thr His Glu Glu Cys Thr Lys Tyr Gly Met Asp Gly Tyr Val				
		960	965	970	
	acg aag ccc ttc gag gag aag cag ctc ttc cag gcg ctg cag aag ttc				2979
	Thr Lys Pro Phe Glu Glu Lys Gln Leu Phe Gln Ala Leu Gln Lys Phe				
10		975	980	985	
	ttg gac cct ggc atg tcc agc taa cacccaagtg ctgcgttcgt tgcaagtgag				3033
	Leu Asp Pro Gly Met Ser Ser				
		990	995		
	gcaccattct cct				3046
15	<210> 12				
	<211> 995				
	<212> PRT				
	<213> Zea mays				
	<400> 12				
20	Met Gly Val Gly Gly Gly Gly Gly Gly Glu Ala Ala Ala Val Ser Ala				
	1 5 10 15				
	Pro Ala Pro Ala Glu Glu Ala Gly Lys Asp Ala Glu Asp Gly Gly Gly				
	20 25 30				
	Trp Thr Leu Lys Ala Lys Leu Ile Ala Val Ala Val Leu Val Trp Val				
25		35	40	45	
	Leu Gly Ala Leu Ala Leu Gly Val Phe Leu His Ser Tyr Phe Arg His				
	50 55 60				
	Ala Ala Leu Arg Lys Ala Glu Glu Gly Leu Val Ser Met Cys Glu Glu				
	65 70 75 80				
30	Arg Ala Arg Met Leu Gln Asp Gln Phe Ala Val Ser Val Asn His Val				
		85	90	95	
	His Ala Leu Ala Ile Leu Val Ala Thr Phe His Tyr Glu Lys Arg Pro				

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

	355		360		365
	Lys Met Glu Glu Leu Lys Lys Gln Ala Glu Ala Ala Asp Val Ala Lys				
	370		375		380
	Ser Gln Phe Leu Ala Thr Val Ser His Glu Ile Arg Thr Pro Met Asn				
5	385		390		395
	Gly Val Leu Gly Met Leu Asp Met Leu Leu Asp Thr Asp Leu Thr Ser				
		405		410	415
	Thr Gln Arg Asp Phe Ala Gln Thr Ala Gln Val Cys Gly Lys Ala Leu				
		420		425	430
10	Ile Ser Leu Ile Asn Glu Val Leu Asp Arg Ala Lys Ile Glu Ala Gly				
		435		440	445
	Lys Leu Asp Leu Glu Ser Val Pro Phe Asp Leu Arg Ser Ile Leu Asp				
		450		455	460
	Asp Val Ile Ser Leu Phe Ser Ser Lys Ser Arg Glu Lys Gly Ile Glu				
15	465		470		475
	Leu Ala Val Tyr Val Ser Glu Arg Val Pro Glu Leu Leu Leu Gly Asp				
		485		490	495
	Pro Gly Arg Phe Arg Gln Ile Ile Thr Asn Leu Val Gly Asn Ser Ile				
		500		505	510
20	Lys Phe Thr Glu Arg Gly His Ile Phe Val Gln Val His Leu Ala Asp				
		515		520	525
	His Ser Asn Leu Ala Thr Glu Ser Lys Val Glu Ser Val Ala Asn Gly				
		530		535	540
	Met Asn Gly His Lys Asp Glu Lys Thr Ala Val Ala Thr Ser Val Ser				
25	545		550		555
	Leu Asn Thr Leu Ser Gly Phe Glu Ala Ala Asp Ser Arg Asn Ser Trp				
		565		570	575
	Glu Asn Phe Lys Leu Leu Leu Ser Tyr Glu Lys Asn Glu Met Pro Tyr				
		580		585	590
30	Glu Ser Val Ser Asp Lys Val Thr Leu Val Val Ser Val Glu Asp Thr				
		595		600	605
	Gly Ile Gly Ile Pro Leu Asp Ala Gln Ala Lys Val Phe Thr Pro Phe				

	610	615	620
	Met Gln Ala Asp Ser Ser Thr Ser Arg Thr Tyr Gly Gly Thr Gly Ile		
	625	630	635 640
	Gly Leu Ser Ile Ser Lys Cys Leu Val Glu Leu Met Gly Gly Gln Ile		
5	645	650	655
	Asn Phe Val Ser Arg Pro His Val Gly Ser Thr Phe Thr Phe Thr Ala		
	660	665	670
	Ala Leu Gln Arg Cys Asp Arg Ser Ala Ile Gly Asp Ser Lys Pro Val		
	675	680	685
10	Met Leu His Pro Leu Pro Ser Ser Phe Lys Gly Leu Ser Ala Leu Leu		
	690	695	700
	Val Asp Arg Arg Pro Val Arg Ala Thr Val Thr Lys Tyr His Leu Gln		
	705	710	715 720
	Arg Leu Gly Ile Ala Cys Asp Val Val Ala Thr Ile Glu Leu Ala Leu		
15	725	730	735
	Gly Val Leu Ser Gly Arg Asn Gly Ser Ser Leu Thr Ser Thr Lys Gln		
	740	745	750
	Pro Cys Met Leu Leu Ile Glu Ser Asp Ser Trp Gly Phe Lys Ile Asp		
	755	760	765
20	Val Pro Leu Arg Ser Arg Leu Leu Glu Met Lys Gln Asn Gly Pro Pro		
	770	775	780
	Gly Leu Pro Lys Thr Ile Leu Leu Ala Ala Ala Glu Ser Gly Lys Leu		
	785	790	795 800
	Lys Ala His Tyr Ala Val Asp Ser Val Ile Thr Lys Pro Leu Lys Ala		
25	805	810	815
	Ser Gly Leu Ala Ala Cys Leu Phe Gln Thr Leu Gly Ile Thr Gln Ser		
	820	825	830
	Ser Asn Glu Arg Arg Asp Asn Ser Gly Ser Leu His Gly Leu Leu Leu		
	835	840	845
30	Gly Lys Asn Ile Leu Val Val Asp Asp Asn Lys Val Asn Leu Arg Val		
	850	855	860
	Ala Ala Gly Thr Leu Lys Lys Phe Gly Ala Lys Val Glu Cys Val Glu		

865 870 875 880
 Ser Gly Lys Asp Ala Leu Ala Ser Leu Gln Val Pro His Lys Phe His
 885 890 895
 Leu Cys Leu Met Asp Ile Gln Met Pro Glu Met Asp Gly Phe Glu Ala
 5 900 905 910
 Thr Lys Gln Ile Arg Ala Met Glu Ala Lys Ala Asn Glu Gln Ala Val
 915 920 925
 Ala Cys Asp Asp Ser Asp Thr Asp Gly Ala Thr Arg Ala Ala Arg Trp
 930 935 940
 10 His Leu Pro Val Leu Ala Met Thr Ala Asp Val Ile Gln Ala Thr His
 945 950 955 960
 Glu Glu Cys Thr Lys Tyr Gly Met Asp Gly Tyr Val Thr Lys Pro Phe
 965 970 975
 Glu Glu Lys Gln Leu Phe Gln Ala Leu Gln Lys Phe Leu Asp Pro Gly
 15 980 985 990
 Met Ser Ser
 995
 <210> 13
 <211> 3318
 20 <212> DNA
 <213> Cucurbita maxima
 <220>
 <221> CDS
 <222> (1)..(2946)
 25 <220>
 <221> 3'UTR
 <222> (2947)..(3318)
 <400> 13
 atg cag gtg agc gat aac tct gtg ggt ttg aag tgg aat gag caa atg 48
 30 Met Gln Val Ser Asp Asn Ser Val Gly Leu Lys Trp Asn Glu Gln Met
 1 5 10 15
 gga aca aca aag aag ggt tac aca ttt gtt caa gct aac agg gct tgg 96

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

	aat att ttg agg gct aga gca aca gga aag gct gtc tta aca aga ccc	624
	Asn Ile Leu Arg Ala Arg Ala Thr Gly Lys Ala Val Leu Thr Arg Pro	
	195 200 205	
	ttc agg ctg ctg ggt tcc cat cat ctt gga gtt gtt ttg aca ttt cct	672
5	Phe Arg Leu Leu Gly Ser His His Leu Gly Val Val Leu Thr Phe Pro	
	210 215 220	
	gtt tac aaa ttc aaa ttg cca tcc ata ccg act gaa gaa gaa cgg ata	720
	Val Tyr Lys Phe Lys Leu Pro Ser Ile Pro Thr Glu Glu Glu Arg Ile	
	225 230 235 240	
10	gaa gca aca gca ggc tac gtt ggc gga gcc ttt gat gtt gag tca ctc	768
	Glu Ala Thr Ala Gly Tyr Val Gly Gly Ala Phe Asp Val Glu Ser Leu	
	245 250 255	
	gtg gag aac ttg ttt ggg cag ctt gca ggg aat cag gcc att ttg gta	816
	Val Glu Asn Leu Phe Gly Gln Leu Ala Gly Asn Gln Ala Ile Leu Val	
15	260 265 270	
	aat gta tat gat gtc acg aac tct tct gat ctt ctc gtg atg tat ggt	864
	Asn Val Tyr Asp Val Thr Asn Ser Ser Asp Leu Leu Val Met Tyr Gly	
	275 280 285	
	cat caa tat caa gat ggt gac ttg tcg ctt tca cat gag agc agc ctt	912
20	His Gln Tyr Gln Asp Gly Asp Leu Ser Leu Ser His Glu Ser Ser Leu	
	290 295 300	
	gat ttc gga gat cca ttc agg aag cat ttg atg att tgt aga tat cag	960
	Asp Phe Gly Asp Pro Phe Arg Lys His Leu Met Ile Cys Arg Tyr Gln	
	305 310 315 320	
25	cag agg gct ccc aca tcc tgg act gcc cta act act gca ttc tta ttc	1008
	Gln Arg Ala Pro Thr Ser Trp Thr Ala Leu Thr Thr Ala Phe Leu Phe	
	325 330 335	
	ttc gtg atc ggt ttg tta gtt gga tat att ttg tat ggt gca gca act	1056
	Phe Val Ile Gly Leu Leu Val Gly Tyr Ile Leu Tyr Gly Ala Ala Thr	
30	340 345 350	
	cac att gtg aag gtt gaa gat gat ttt cat gaa atg caa gta ctg aaa	1104
	His Ile Val Lys Val Glu Asp Asp Phe His Glu Met Gln Val Leu Lys	

	355	360	365	
	ggt cga gcg gag gct gcc gat gta gca aaa tcc cag ttt ctt gca act	1152		
	Val Arg Ala Glu Ala Ala Asp Val Ala Lys Ser Gln Phe Leu Ala Thr			
	370	375	380	
5	ggt tct cat gaa att agg aca cca atg aat ggc atc ctc gga atg ctt	1200		
	Val Ser His Glu Ile Arg Thr Pro Met Asn Gly Ile Leu Gly Met Leu			
	385	390	395	400
	gct ctg ctt ctg gat aca gat cta agt tcc aca cag aag gat tat gct	1248		
	Ala Leu Leu Leu Asp Thr Asp Leu Ser Ser Thr Gln Lys Asp Tyr Ala			
10	405	410	415	
	caa act gcc cag gct tgt gga aag gca ttg ata gca tta ata aat gag	1296		
	Gln Thr Ala Gln Ala Cys Gly Lys Ala Leu Ile Ala Leu Ile Asn Glu			
	420	425	430	
	ggt ctt gac cgg gca aaa att gaa gct gga aag tta gaa ctg gaa gca	1344		
15	Val Leu Asp Arg Ala Lys Ile Glu Ala Gly Lys Leu Glu Leu Glu Ala			
	435	440	445	
	ggt cca ttc gac att cga tca ata ctt gat gac gtg cta tct tta ttt	1392		
	Val Pro Phe Asp Ile Arg Ser Ile Leu Asp Asp Val Leu Ser Leu Phe			
	450	455	460	
20	tcc gag aag tcc aga caa aag ggt ctg gag ctg gca gtt ttt gtt tct	1440		
	Ser Glu Lys Ser Arg Gln Lys Gly Leu Glu Leu Ala Val Phe Val Ser			
	465	470	475	480
	gat aaa gtt cca gaa att gta att gga gat cct gga aga ttc aga caa	1488		
	Asp Lys Val Pro Glu Ile Val Ile Gly Asp Pro Gly Arg Phe Arg Gln			
25	485	490	495	
	att ata aca aat ctt gtg ggt aac tct gtt aag ttt act gaa aga gga	1536		
	Ile Ile Thr Asn Leu Val Gly Asn Ser Val Lys Phe Thr Glu Arg Gly			
	500	505	510	
	cat ata ttt gtt aaa gta cac cta gct gag aat tca aaa gtc tcc atg	1584		
30	His Ile Phe Val Lys Val His Leu Ala Glu Asn Ser Lys Val Ser Met			
	515	520	525	
	gac tcg gaa tac gtc aac gga ata tcc gac agt ggc tta ttc gta ttg	1632		

	Asp Ser Glu Tyr Val Asn Gly Ile Ser Asp Ser Gly Leu Phe Val Leu	
	530 535 540	
	gat ggt cgt gaa ttt caa act ttg agt gga cgc gag gca gcc gat gat	1680
	Asp Gly Arg Glu Phe Gln Thr Leu Ser Gly Arg Glu Ala Ala Asp Asp	
5	545 550 555 560	
	cag aac agt tgg gat aac ttc aag cat cta atc gct gac gac aac ttc	1728
	Gln Asn Ser Trp Asp Asn Phe Lys His Leu Ile Ala Asp Asp Asn Phe	
	565 570 575	
	cag tcg aat gcc gct tca aac aac tca gca gtt acc aac aag ggt tgt	1776
10	Gln Ser Asn Ala Ala Ser Asn Asn Ser Ala Val Thr Asn Lys Gly Cys	
	580 585 590	
	gat cat gtt act ttg atg gta agt gtg gag gat act gga att ggg atc	1824
	Asp His Val Thr Leu Met Val Ser Val Glu Asp Thr Gly Ile Gly Ile	
	595 600 605	
15	ctt tta cat gcc caa aat cga gtt ttc aca ccc ttc atg caa gca gat	1872
	Leu Leu His Ala Gln Asn Arg Val Phe Thr Pro Phe Met Gln Ala Asp	
	610 615 620	
	agc tcg acc tcc cga aat tat gga ggg act ggt att ggt ttg agt atc	1920
	Ser Ser Thr Ser Arg Asn Tyr Gly Gly Thr Gly Ile Gly Leu Ser Ile	
20	625 630 635 640	
	agc aaa tgt tta gtt gag tta atg ggt ggt cag atc aac ttc ata agc	1968
	Ser Lys Cys Leu Val Glu Leu Met Gly Gly Gln Ile Asn Phe Ile Ser	
	645 650 655	
	cgg cct cag att gga agc acg ttt tcc ttc act gct gta ttt gga aaa	2016
25	Arg Pro Gln Ile Gly Ser Thr Phe Ser Phe Thr Ala Val Phe Gly Lys	
	660 665 670	
	tgt aag aaa aac tcg atg aat gat atg aaa aag ccc aac tct gaa gaa	2064
	Cys Lys Lys Asn Ser Met Asn Asp Met Lys Lys Pro Asn Ser Glu Glu	
	675 680 685	
30	ctt ccc ccc agt ttt aga gga atg aaa gca ata gta gtt gat agc aaa	2112
	Leu Pro Pro Ser Phe Arg Gly Met Lys Ala Ile Val Val Asp Ser Lys	
	690 695 700	

	cat gta cga gct tct gta acc agg tat cat ttg aag aga ctt ggt atc	2160
	His Val Arg Ala Ser Val Thr Arg Tyr His Leu Lys Arg Leu Gly Ile	
	705 710 715 720	
	ata gtt gaa gtc acc aat agc atc aac atg gca gct tct tta ttc aga	2208
5	Ile Val Glu Val Thr Asn Ser Ile Asn Met Ala Ala Ser Leu Phe Arg	
	725 730 735	
	gaa aat gga tcc aca ctg cca aga aac aca atc ctt cca gat atg atc	2256
	Glu Asn Gly Ser Thr Leu Pro Arg Asn Thr Ile Leu Pro Asp Met Ile	
	740 745 750	
10	tta gtt gaa aag gac ata cta aat tct gat gag gaa tgt ggg atc att	2304
	Leu Val Glu Lys Asp Ile Leu Asn Ser Asp Glu Glu Cys Gly Ile Ile	
	755 760 765	
	cat cat ctg aac tgg aaa ccg aac ggt agt tcg gtt aag ttt cca aag	2352
	His His Leu Asn Trp Lys Pro Asn Gly Ser Ser Val Lys Phe Pro Lys	
15	770 775 780	
	ctg atc ctt ctc gct acc aat att gcc act gct gaa cta gac aag gca	2400
	Leu Ile Leu Leu Ala Thr Asn Ile Ala Thr Ala Glu Leu Asp Lys Ala	
	785 790 795 800	
	aga gca gca ggt ttt gca gac acc gtg atc atg aag ccg ttg agg gcg	2448
20	Arg Ala Ala Gly Phe Ala Asp Thr Val Ile Met Lys Pro Leu Arg Ala	
	805 810 815	
	act atg gtg gct gcc tgt ctt caa caa gta ctc ggg gtt aag aat cag	2496
	Thr Met Val Ala Ala Cys Leu Gln Gln Val Leu Gly Val Lys Asn Gln	
	820 825 830	
25	aga cgg ccg aat ggt tct gct ttc ctc cag agc ctt ctc tgt ggc aag	2544
	Arg Arg Pro Asn Gly Ser Ala Phe Leu Gln Ser Leu Leu Cys Gly Lys	
	835 840 845	
	aga atc tta att gtt gat gac aac cga gta aac cgt cgg gtc gct gca	2592
	Arg Ile Leu Ile Val Asp Asp Asn Arg Val Asn Arg Arg Val Ala Ala	
30	850 855 860	
	ggc gct ctg aag aaa ttt ggt gca gat gtt gag tgt gca gat agc ggg	2640
	Gly Ala Leu Lys Lys Phe Gly Ala Asp Val Glu Cys Ala Asp Ser Gly	

	865	870	875	880	
	aaa tct gca ctg aag ttg ctt cag cta ccg cat aat ttt gat gct tgc				2688
	Lys Ser Ala Leu Lys Leu Leu Gln Leu Pro His Asn Phe Asp Ala Cys				
		885	890	895	
5	ttc atg gat att caa atg cct gaa atg gat ggg ttt gag gcg act cgt				2736
	Phe Met Asp Ile Gln Met Pro Glu Met Asp Gly Phe Glu Ala Thr Arg				
		900	905	910	
	cgt atc agg aca atg gag gtc gag gca aac aaa gga gga ttg tct gca				2784
	Arg Ile Arg Thr Met Glu Val Glu Ala Asn Lys Gly Gly Leu Ser Ala				
10		915	920	925	
	aca gaa ggc aaa cgg cct ata cca ata tta gca atg act gca gac gtg				2832
	Thr Glu Gly Lys Arg Pro Ile Pro Ile Leu Ala Met Thr Ala Asp Val				
		930	935	940	
	att cat gct acg tac gaa gaa tgc ctg aaa tgc ggt atg aat ggt tac				2880
15	Ile His Ala Thr Tyr Glu Glu Cys Leu Lys Cys Gly Met Asn Gly Tyr				
		945	950	955	960
	gtc tcg aaa ccc ttt gaa gaa gaa aat cta tac aag gaa gtt gcc cga				2928
	Val Ser Lys Pro Phe Glu Glu Glu Asn Leu Tyr Lys Glu Val Ala Arg				
		965	970	975	
20	ttt ttc aaa aaa cca tag tccatcaaaa gcttcatgaa tgacaagagg				2976
	Phe Phe Lys Lys Pro				
		980			
	tcatcagctg tagagctcct ttggtgggt tggaaagtac cagcaagttt tgacaccatt				3036
	gctggtgcta actgtcctgt tgctggtgac cgaggaaccg agttcaacga gggcgctcggg				3096
25	aattctcgac tgccattgag actcgggttt gagctgctac catttctaac caactaaata				3156
	ttttattttg actagaatgt gactacctac ctgtatacta caccagaaat atccatcccc				3216
	aaatggatgt ataattatgg ttgcaagggg aaggagctaa attgtaaatg ctcatatattc				3276
	taagacctct tctaaacatc tttatagttg gttggcttag gc				3318
	<210> 14				
30	<211> 981				
	<212> PRT				
	<213> Cucurbita maxima				

<400> 14

Met Gln Val Ser Asp Asn Ser Val Gly Leu Lys Trp Asn Glu Gln Met
 1 5 10 15

Gly Thr Thr Lys Lys Gly Tyr Thr Phe Val Gln Ala Asn Arg Ala Trp
 5 20 25 30

Leu Arg Lys Tyr Leu Leu Phe Trp Ile Met Gly Met Ala Phe Ile Ser
 35 40 45

Met Leu Ile Tyr Asn Gly Met Asp Ala Asp Ile Lys Val Arg Arg Asn
 50 55 60

10 Glu Val Leu Gly Ser Met Cys Glu Gln Arg Ala Arg Met Leu Gln Asp
 65 70 75 80

Gln Phe Asn Val Ser Val Asn His Val His Ala Leu Ala Val Leu Val
 85 90 95

Ser Thr Phe His Tyr Phe Lys Asn Pro Ser Ala Ile Asp Gln Glu Thr
 15 100 105 110

Phe Ala Glu Tyr Thr Ala Arg Thr Ala Phe Glu Arg Pro Leu Leu Ser
 115 120 125

Gly Val Ala Tyr Ala Gln Arg Val Ile His Ser Glu Arg Asp Ile Phe
 130 135 140

20 Glu Lys Gln His Gly Trp Met Ile Arg Thr Met Glu Lys Glu Pro Ser
 145 150 155 160

Pro Asp Arg Asp Glu Tyr Ala Pro Val Ile Phe Ser Gln Glu Thr Val
 165 170 175

Ser Tyr Ile Glu Ser Leu Asp Met Met Ser Gly Glu Glu Asp Arg Glu
 25 180 185 190

Asn Ile Leu Arg Ala Arg Ala Thr Gly Lys Ala Val Leu Thr Arg Pro
 195 200 205

Phe Arg Leu Leu Gly Ser His His Leu Gly Val Val Leu Thr Phe Pro
 210 215 220

30 Val Tyr Lys Phe Lys Leu Pro Ser Ile Pro Thr Glu Glu Glu Arg Ile
 225 230 235 240

Glu Ala Thr Ala Gly Tyr Val Gly Gly Ala Phe Asp Val Glu Ser Leu

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

	500	505	510	
	His Ile Phe Val Lys Val	His Leu Ala Glu Asn Ser Lys Val	Ser Met	
	515	520	525	
	Asp Ser Glu Tyr Val Asn Gly Ile Ser Asp Ser Gly Leu Phe Val	Leu		
5	530	535	540	
	Asp Gly Arg Glu Phe Gln Thr Leu Ser Gly Arg Glu Ala Ala Asp Asp			
	545	550	555	560
	Gln Asn Ser Trp Asp Asn Phe Lys His Leu Ile Ala Asp Asp Asn Phe			
	565	570	575	
10	Gln Ser Asn Ala Ala Ser Asn Asn Ser Ala Val Thr Asn Lys Gly Cys			
	580	585	590	
	Asp His Val Thr Leu Met Val Ser Val Glu Asp Thr Gly Ile Gly Ile			
	595	600	605	
	Leu Leu His Ala Gln Asn Arg Val Phe Thr Pro Phe Met Gln Ala Asp			
15	610	615	620	
	Ser Ser Thr Ser Arg Asn Tyr Gly Gly Thr Gly Ile Gly Leu Ser Ile			
	625	630	635	640
	Ser Lys Cys Leu Val Glu Leu Met Gly Gly Gln Ile Asn Phe Ile Ser			
	645	650	655	
20	Arg Pro Gln Ile Gly Ser Thr Phe Ser Phe Thr Ala Val Phe Gly Lys			
	660	665	670	
	Cys Lys Lys Asn Ser Met Asn Asp Met Lys Lys Pro Asn Ser Glu Glu			
	675	680	685	
	Leu Pro Pro Ser Phe Arg Gly Met Lys Ala Ile Val Val Asp Ser Lys			
25	690	695	700	
	His Val Arg Ala Ser Val Thr Arg Tyr His Leu Lys Arg Leu Gly Ile			
	705	710	715	720
	Ile Val Glu Val Thr Asn Ser Ile Asn Met Ala Ala Ser Leu Phe Arg			
	725	730	735	
30	Glu Asn Gly Ser Thr Leu Pro Arg Asn Thr Ile Leu Pro Asp Met Ile			
	740	745	750	
	Leu Val Glu Lys Asp Ile Leu Asn Ser Asp Glu Glu Cys Gly Ile Ile			

	755	760	765
	His His Leu Asn Trp Lys Pro Asn Gly Ser Ser Val Lys Phe Pro Lys		
	770	775	780
	Leu Ile Leu Leu Ala Thr Asn Ile Ala Thr Ala Glu Leu Asp Lys Ala		
5	785	790	795
	Arg Ala Ala Gly Phe Ala Asp Thr Val Ile Met Lys Pro Leu Arg Ala		800
		805	810
	Thr Met Val Ala Ala Cys Leu Gln Gln Val Leu Gly Val Lys Asn Gln		815
		820	825
10	Arg Arg Pro Asn Gly Ser Ala Phe Leu Gln Ser Leu Leu Cys Gly Lys		830
	835	840	845
	Arg Ile Leu Ile Val Asp Asp Asn Arg Val Asn Arg Arg Val Ala Ala		
	850	855	860
	Gly Ala Leu Lys Lys Phe Gly Ala Asp Val Glu Cys Ala Asp Ser Gly		
15	865	870	875
	Lys Ser Ala Leu Lys Leu Leu Gln Leu Pro His Asn Phe Asp Ala Cys		880
		885	890
	Phe Met Asp Ile Gln Met Pro Glu Met Asp Gly Phe Glu Ala Thr Arg		895
		900	905
20	Arg Ile Arg Thr Met Glu Val Glu Ala Asn Lys Gly Gly Leu Ser Ala		910
	915	920	925
	Thr Glu Gly Lys Arg Pro Ile Pro Ile Leu Ala Met Thr Ala Asp Val		
	930	935	940
	Ile His Ala Thr Tyr Glu Glu Cys Leu Lys Cys Gly Met Asn Gly Tyr		
25	945	950	955
	Val Ser Lys Pro Phe Glu Glu Glu Asn Leu Tyr Lys Glu Val Ala Arg		960
		965	970
	Phe Phe Lys Lys Pro		975
		980	
30	<210> 15		
	<211> 993		
	<212> PRT		

<213> Lotus corniculatus

<220>

<221> Mutation

<222> (266)..(266)

5 <223> X: is Xaa (any amino acid other than Leu)

<220>

<221> misc_feature

<222> (266)..(266)

<223> xaa can be any naturally occurring amino acid

10 <400> 15

Met Gly Leu Gly Phe Lys Met Gln Gln Ser His His Pro Val Ala Leu

1 5 10 15

Lys Leu His Glu Gln Ala Gly Ser Gln Arg Lys Phe Thr Phe Ile Gln

20 25 30

15 Asn Phe Arg Asn Trp Phe Leu Pro Leu Leu Phe Val Trp Phe Ile Val

35 40 45

Met Ala Ala Phe Gly Ala Cys Ile Tyr His Lys Met Asp Ala Glu Thr

50 55 60

Lys Val Arg Arg Lys Glu Val Leu Gly Ser Leu Cys Asp Gln Arg Ala

20 65 70 75 80

Arg Met Leu Gln Asp Gln Phe Ser Val Ser Val Asn His Val His Ala

85 90 95

Leu Ala Ile Leu Val Ser Thr Phe His Tyr Tyr Arg Asn Thr Ser Ala

100 105 110

25 Ile Asp Gln Glu Thr Phe Ala Glu Tyr Thr Ala Arg Thr Ala Phe Glu

115 120 125

Arg Pro Leu Met Ser Gly Val Ala Tyr Ala Gln Arg Val Val His Ser

130 135 140

Glu Arg Glu Arg Phe Glu Lys Gln His Gly Trp Val Ile Lys Thr Met

30 145 150 155 160

Glu Arg Val Pro Ser Gly Val Arg Asp Glu Tyr Ala Ala Val Ile Phe

165 170 175

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

Ala Leu Ile Asn Glu Val Leu Asp Arg Ala Lys Ile Glu Ala Gly Lys
 435 440 445
 Leu Glu Leu Glu Ala Val Pro Phe Asp Leu Arg Ser Ile Leu Asp Asp
 450 455 460
 5 Val Leu Ser Leu Phe Ser Glu Lys Ser Arg His Lys Gly Leu Glu Leu
 465 470 475 480
 Ala Val Phe Val Ser Asp Lys Val Pro Asp Ile Val Met Gly Asp Pro
 485 490 495
 Gly Arg Phe Arg Gln Ile Val Thr Asn Leu Val Gly Asn Ser Val Lys
 10 500 505 510
 Phe Thr Glu Arg Gly His Ile Phe Val Lys Val His Leu Ala Glu Lys
 515 520 525
 Arg Gln Cys Thr Met Asn Gly Lys Cys Glu Thr Phe Leu Asn Gly Gly
 530 535 540
 15 Cys Asp Asp Val Leu His Val Ser Gly Ser Tyr Asn Leu Lys Thr Leu
 545 550 555 560
 Ser Gly Tyr Glu Ala Ala Asp Glu Arg Asn Ser Trp Asp Asn Phe Lys
 565 570 575
 His His Ile Ala Asp Glu Glu Phe Phe Phe Asp Ala Ser Val Lys Lys
 20 580 585 590
 Leu Ala Ser Ser Glu Ser Tyr Glu Gln Val Thr Leu Met Val Ser Val
 595 600 605
 Glu Asp Thr Gly Ile Gly Ile Ser Phe Ser Ala Gln Asp Ser Ile Phe
 610 615 620
 25 Met Pro Phe Val Gln Ala Asp Ser Ser Thr Ser Arg Asn Tyr Gly Gly
 625 630 635 640
 Thr Gly Ile Gly Leu Ser Ile Ser Lys Cys Leu Val Glu Leu Met Gly
 645 650 655
 Gly Gln Ile Asn Phe Ile Ser Arg Pro Gln Val Gly Ser Thr Phe Ser
 30 660 665 670
 Phe Thr Ala Asp Phe Gly Thr Phe Lys Lys Asn Ser Thr Thr Asp Met
 675 680 685

Lys Lys Leu Asn Phe Glu Asp Leu Pro Ser Ser Phe Arg Gly Leu Lys
 690 695 700
 Ala Ile Val Val Asp Gly Lys Pro Val Arg Ala Ala Val Thr Arg Tyr
 705 710 715 720
 5 His Leu Lys Arg Leu Gly Ile Gln Ala Lys Val Ala Ile Ser Ile Asn
 725 730 735
 Lys Ala Val Ser Leu Cys Gly Lys Asn Gly Ser Leu Thr Ser Ala Leu
 740 745 750
 Phe Gln Pro Asp Ile Ile Phe Val Glu Lys Asp Ser Trp Val Ser Gly
 10 755 760 765
 Glu Asp Gly Gly Ile Phe Asn Ala Phe Lys Met Pro Gln Met Ile Leu
 770 775 780
 Leu Ala Thr Asn Ile Cys Asn Ala Glu Phe Asp Lys Ala Lys Ala Ala
 785 790 795 800
 15 Gly Phe Ser Asp Thr Val Ile Met Lys Pro Leu Arg Ala Ser Met Leu
 805 810 815
 Ala Ala Cys Leu Gln Gln Val Phe Gly Thr Gly Lys Thr Arg Gln Phe
 820 825 830
 Gly Lys Asp Met Ser Asn Gly Ser Ser Val Arg Ser Leu Leu Cys Gly
 20 835 840 845
 Lys Lys Ile Leu Val Val Asp Asp Asn Leu Val Asn Arg Arg Val Ala
 850 855 860
 Ala Gly Ala Leu Lys Asn Phe Gly Ala Asp Val Lys Cys Ala Ala Ser
 865 870 875 880
 25 Gly Lys Ala Ala Leu Glu Met Leu Gln Tyr Pro His Asp Phe Asp Ala
 885 890 895
 Cys Phe Met Asp Ile Gln Met Pro Glu Met Asp Gly Phe Glu Ala Thr
 900 905 910
 Arg Arg Ile Arg Met Met Glu Arg Glu Ala Ser Glu Gln Leu Lys Ser
 30 915 920 925
 Glu Ser Gly Glu Glu Asn Gly Lys Lys Ser Glu Phe His Met Pro Ile
 930 935 940

Leu Ala Met Thr Ala Asp Val Ile His Ala Thr Tyr Asp Lys Cys Leu
 945 950 955 960
 Asn Cys Gly Met Asp Gly Tyr Val Ser Lys Pro Phe Glu Glu Glu Asn
 965 970 975
 5 Leu Tyr Gln Ala Val Ala Lys Phe Phe Lys Ser Lys Pro Ala Ser Asp
 980 985 990
 Ser

<210> 16

10 <211> 1003

<212> PRT

<213> Medicago truncatula

<220>

<221> Mutation

15 <222> (267)..(267)

<223> X = Xaa (any amino acid other than Leu)

<220>

<221> misc_feature

<222> (267)..(267)

20 <223> Xaa can be any naturally occurring amino acid

<400> 16

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

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

	340	345	350
	Leu Tyr Gly Ala Gly Asn His Ile Val Lys Val Glu Asp Asp Phe His		
	355	360	365
	Glu Met Gln Glu Leu Lys Val Arg Ala Glu Ala Ala Asp Val Ala Lys		
5	370	375	380
	Ser Gln Phe Leu Ala Thr Val Ser His Glu Ile Arg Thr Pro Met Asn		
	385	390	395
	Gly Ile Leu Gly Met Leu Gly Leu Leu Leu Arg Thr Glu Leu Asn Ser		
	405	410	415
10	Thr Gln Arg Asp Tyr Ala Gln Thr Ala Gln Ala Cys Gly Lys Ala Leu		
	420	425	430
	Ile Ala Leu Ile Asn Glu Val Leu Asp Arg Ala Lys Ile Glu Ala Gly		
	435	440	445
	Lys Leu Glu Leu Glu Ala Val Pro Phe Asp Leu Arg Ser Ile Leu Asp		
15	450	455	460
	Asp Val Leu Ser Leu Phe Ser Glu Lys Ser Arg His Lys Gly Leu Glu		
	465	470	475
	Leu Ala Val Phe Val Ser Asp Lys Val Pro Asp Ile Val Met Gly Asp		
	485	490	495
20	Pro Gly Arg Phe Arg Gln Ile Val Thr Asn Leu Val Gly Asn Ser Val		
	500	505	510
	Lys Phe Thr Glu Arg Gly His Ile Phe Val Lys Val His Leu Ser Glu		
	515	520	525
	Asn Arg Lys Pro Val Thr Asn Gly Lys His Glu Thr Tyr Arg Asn Gly		
25	530	535	540
	Gly Ser Glu Glu Val Val His Ala Ser Gly Gly Tyr Asn Leu Lys Thr		
	545	550	555
	Leu Ser Gly Tyr Glu Ala Ala Asp Glu Arg Asn Asn Trp Asp Asn Phe		
	565	570	575
30	Asn His Leu Ile Ala Asp Glu Glu Phe Phe Cys Asp Ala Ser Thr Lys		
	580	585	590
	Lys Val Ala Ser Asn Glu Phe Tyr Glu Gln Val Thr Leu Met Val Cys		

	595	600	605
	Val Glu Asp Thr Gly Ile Gly Ile Pro Phe Ser Ala Gln Asp Arg Ile		
	610	615	620
	Phe Met Pro Phe Val Gln Ala Asp Ser Ser Thr Ser Arg Asn Tyr Gly		
5	625	630	635
	Gly Thr Gly Ile Gly Leu Ser Ile Ser Lys Cys Leu Val Glu Leu Met		
	645	650	655
	Gly Gly Gln Ile Asn Phe Ile Ser Arg Pro Gln Val Gly Ser Thr Phe		
	660	665	670
10	Ser Phe Thr Ala Asp Phe Gly Ile Phe Lys Lys Asn Pro Ile Thr Glu		
	675	680	685
	Val Lys Lys Val Asn Tyr Glu Asp Leu Pro Ser Ser Phe Arg Gly Leu		
	690	695	700
	Lys Ala Val Val Val Asp Gly Lys Pro Val Arg Ala Ala Val Thr Arg		
15	705	710	715
	Tyr His Leu Lys Arg Leu Gly Ile Gln Val Lys Val Ala Asn Ala Ile		
	725	730	735
	Asn Lys Ala Val Ser Leu Cys Gly Lys Asn Gly Ala Ser Ser Thr Gly		
	740	745	750
20	Leu Phe Gln Pro Asp Ile Ile Phe Val Glu Lys Asp Ser Trp Val Cys		
	755	760	765
	Gly Glu Asp Gly Ile Phe Ser Val Arg Gln Leu Asp Trp Lys Gln Asn		
	770	775	780
	Gly His Ile Phe Lys Met Pro Gln Met Ile Leu Leu Ala Thr Asn Ile		
25	785	790	795
	Ser Asn Asp Glu Phe Asp Lys Ala Lys Ser Ala Gly Phe Ser Asp Thr		
	805	810	815
	Val Ile Met Lys Pro Leu Arg Ala Ser Met Val Gly Ala Cys Leu Gln		
	820	825	830
30	Gln Val Leu Gly Thr Gly Lys Lys Arg Gln Leu Gly Lys Glu Met Pro		
	835	840	845
	Asn Gly Ser Thr Ser Val Arg Ser Leu Leu Phe Gly Lys Lys Ile Leu		

850 855 860
 Val Val Asp Asp Asn Val Val Asn Arg Arg Val Ala Ala Gly Ala Leu
 865 870 875 880
 Lys Asn Phe Gly Ala Asp Val Lys Cys Ala Asp Ser Gly Lys Ala Ala
 5 885 890 895
 Leu Glu Met Leu Gln Phe Pro His Lys Phe Asp Ala Cys Phe Met Asp
 900 905 910
 Ile Gln Met Pro Glu Met Asp Gly Phe Glu Ala Thr Arg Arg Ile Arg
 915 920 925
 10 Glu Met Glu Arg Thr Ala Asn Glu Glu Thr Asn Ser Glu Cys Gly Glu
 930 935 940
 Arg Lys Ser Glu Phe His Leu Pro Ile Leu Ala Met Thr Ala Asp Val
 945 950 955 960
 Ile His Ala Thr Tyr Glu Glu Cys Leu Lys Cys Gly Met Asp Gly Tyr
 15 965 970 975
 Val Ser Lys Pro Phe Glu Glu Glu Asn Leu Tyr Gln Ala Val Ala Lys
 980 985 990
 Phe Phe Gln Thr Lys Pro Thr Ser Val Asp Ser
 995 1000
 20 <210> 17
 <211> 1080
 <212> PRT
 <213> Arabidopsis thaliana
 <220>
 25 <221> Mutation
 <222> (356)..(356)
 <223> X= Xaa (any amino acid other than Leucine)
 <220>
 <221> misc_feature
 30 <222> (356)..(356)
 <223> Xaa can be any naturally occurring amino acid
 <400> 17

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

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

Thr Ala Gln Val Cys Gly Lys Ala Leu Ile Ala Leu Ile Asn Glu Val
 515 520 525
 Leu Asp Arg Ala Lys Ile Glu Ala Gly Lys Leu Glu Leu Glu Ser Val
 530 535 540
 5 Pro Phe Asp Ile Arg Ser Ile Leu Asp Asp Val Leu Ser Leu Phe Ser
 545 550 555 560
 Glu Glu Ser Arg Asn Lys Gly Ile Glu Leu Ala Val Phe Val Ser Asp
 565 570 575
 Lys Val Pro Glu Ile Val Lys Gly Asp Ser Gly Arg Phe Arg Gln Ile
 10 580 585 590
 Ile Ile Asn Leu Val Gly Asn Ser Val Lys Phe Thr Glu Lys Gly His
 595 600 605
 Ile Phe Val Lys Val His Leu Ala Glu Gln Ser Lys Asp Glu Ser Glu
 610 615 620
 15 Pro Lys Asn Ala Leu Asn Gly Gly Val Ser Glu Glu Met Ile Val Val
 625 630 635 640
 Ser Lys Gln Ser Ser Tyr Asn Thr Leu Ser Gly Tyr Glu Ala Ala Asp
 645 650 655
 Gly Arg Asn Ser Trp Asp Ser Phe Lys His Leu Val Ser Glu Glu Gln
 20 660 665 670
 Ser Leu Ser Glu Phe Asp Ile Ser Ser Asn Val Arg Leu Met Val Ser
 675 680 685
 Ile Glu Asp Thr Gly Ile Gly Ile Pro Leu Val Ala Gln Gly Arg Val
 690 695 700
 25 Phe Met Pro Phe Met Gln Ala Asp Ser Ser Thr Ser Arg Asn Tyr Gly
 705 710 715 720
 Gly Thr Gly Ile Gly Leu Ser Ile Ser Lys Cys Leu Val Glu Leu Met
 725 730 735
 Arg Gly Gln Ile Asn Phe Ile Ser Arg Pro His Ile Gly Ser Thr Phe
 30 740 745 750
 Trp Phe Thr Ala Val Leu Glu Lys Cys Asp Lys Cys Ser Ala Ile Asn
 755 760 765

His Met Lys Lys Pro Asn Val Glu His Leu Pro Ser Thr Phe Lys Gly
 770 775 780
 Met Lys Ala Ile Val Val Asp Ala Lys Pro Val Arg Ala Ala Val Thr
 785 790 795 800
 5 Arg Tyr His Met Lys Arg Leu Gly Ile Asn Val Asp Val Val Thr Ser
 805 810 815
 Leu Lys Thr Ala Val Val Ala Ala Ala Ala Phe Glu Arg Asn Gly Ser
 820 825 830
 Pro Leu Pro Thr Lys Pro Gln Leu Asp Met Ile Leu Val Glu Lys Asp
 10 835 840 845
 Ser Trp Ile Ser Thr Glu Asp Asn Asp Ser Glu Ile Arg Leu Leu Asn
 850 855 860
 Ser Arg Thr Asn Gly Asn Val His His Lys Ser Pro Lys Leu Ala Leu
 865 870 875 880
 15 Phe Ala Thr Asn Ile Thr Asn Ser Glu Phe Asp Arg Ala Lys Ser Ala
 885 890 895
 Gly Phe Ala Asp Thr Val Ile Met Lys Pro Leu Arg Ala Ser Met Ile
 900 905 910
 Gly Ala Cys Leu Gln Gln Val Leu Glu Leu Arg Lys Thr Arg Gln Gln
 20 915 920 925
 His Pro Glu Gly Ser Ser Pro Ala Thr Leu Lys Ser Leu Leu Thr Gly
 930 935 940
 Lys Lys Ile Leu Val Val Asp Asp Asn Ile Val Asn Arg Arg Val Ala
 945 950 955 960
 25 Ala Gly Ala Leu Lys Lys Phe Gly Ala Glu Val Val Cys Ala Glu Ser
 965 970 975
 Gly Gln Val Ala Leu Gly Leu Leu Gln Ile Pro His Thr Phe Asp Ala
 980 985 990
 Cys Phe Met Asp Ile Gln Met Pro Gln Met Asp Gly Phe Glu Ala Thr
 30 995 1000 1005
 Arg Gln Ile Arg Met Met Glu Lys Glu Ala Lys Glu Lys Thr Asn
 1010 1015 1020

Leu Glu Trp His Leu Pro Ile Leu Ala Met Thr Ala Asp Val Ile
 1025 1030 1035
 His Ala Thr Tyr Glu Glu Cys Leu Lys Ser Gly Met Asp Gly Tyr
 1040 1045 1050
 5 Val Ser Lys Pro Phe Glu Glu Glu Asn Leu Tyr Lys Ser Val Ala
 1055 1060 1065
 Lys Ser Phe Lys Pro Asn Pro Ile Ser Pro Ser Ser
 1070 1075 1080
 <210> 18
 10 <211> 1013
 <212> PRT
 <213> Oryza
 <220>
 <221> Mutation
 15 <222> (266)..(266)
 <223> X=Xaa (any amino acid other than Leu)
 <220>
 <221> misc_feature
 <222> (266)..(266)
 20 <223> Xaa can be any naturally occurring amino acid
 <400> 18
 Met Gly Val Gly Gly Gly Gly Gly Gly Gly Gly Glu Ala Ala Ala
 1 5 10 15
 Ala Val Ala Val Glu Gly Asp Glu Ala Gly Lys Gly Arg Arg Trp Trp
 25 20 25 30
 Arg Val Lys Val Lys Leu Ser Thr Val Ala Val Val Ala Trp Val Leu
 35 40 45
 Ala Ser Ala Ala Leu Trp Ala Gly Leu His Trp Arg Phe Arg Arg Ala
 50 55 60
 30 Ala Leu His Lys Ala Glu Glu Ala Leu Val Cys Met Cys Glu Glu Arg
 65 70 75 80
 Ala Arg Met Leu Gln Asp Gln Phe Ala Val Ser Val Asn His Val His

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

	340	345	350
	Tyr Ala Ala Trp Ser Arg Tyr Asp Asn Val Lys Glu Asp Cys Arg Lys		
	355	360	365
	Met Glu Ala Leu Lys Lys Arg Ala Glu Ala Ala Asp Ile Ala Lys Ser		
5	370	375	380
	Gln Phe Leu Ala Thr Val Ser His Glu Ile Arg Thr Pro Met Asn Gly		
	385	390	395
	Val Leu Gly Met Leu Asp Met Leu Leu Asp Thr Glu Leu Lys Ser Thr		
	405	410	415
10	Gln Arg Asp Tyr Ala Gln Thr Ala Gln Val Cys Gly Lys Ala Leu Ile		
	420	425	430
	Ser Leu Ile Asn Glu Val Leu Asp Arg Ala Lys Ile Glu Ala Gly Lys		
	435	440	445
	Ile Asp Leu Glu Ser Val Pro Phe Asp Leu Arg Ser Ile Leu Asp Asp		
15	450	455	460
	Val Ile Ser Leu Phe Ser Ser Lys Ser Arg Glu Lys Gly Ile Glu Leu		
	465	470	475
	Ala Val Tyr Val Ser Glu Arg Val Pro Glu Ile Leu Leu Gly Asp Pro		
	485	490	495
20	Gly Arg Phe Arg Gln Ile Ile Thr Asn Leu Val Gly Asn Ser Ile Lys		
	500	505	510
	Ile Thr Ile Phe Thr Leu Ser Gln Phe Thr Glu Arg Gly His Ile Phe		
	515	520	525
	Val Gln Val His Leu Ala Asp His Ser Asn Leu Ala Thr Glu Ala Lys		
25	530	535	540
	Ile Glu Pro Val Val Asn Gly Met Asn Gly His Lys Asp Glu Ala Ile		
	545	550	555
	Ala Ile Pro Thr Ser Gly Ser His Asn Thr Leu Ser Gly Phe Glu Ala		
	565	570	575
30	Ala Asp Ser Arg Asn Asn Trp Glu Asn Phe Lys Leu Leu Leu Ser Tyr		
	580	585	590
	Glu Lys Asn Glu Met Pro Tyr Glu Ser Asp Ser Asp Lys Val Thr Leu		

	595	600	605
	val val ser val glu asp thr gly ile gly ile pro leu his ala gln		
	610	615	620
	gly arg val phe thr pro phe met gln ala asp ser ser thr ser arg		
5	625	630	635
	asn tyr gly gly thr gly ile gly leu ser ile ser lys cys leu val		
	645	650	655
	glu ile met gly gly gln ile asn phe val ser arg pro leu val gly		
	660	665	670
10	ser thr phe thr phe thr ala val leu arg arg cys asp lys asn ala		
	675	680	685
	ile ser asp ser lys thr val ala leu his pro leu pro ser ser phe		
	690	695	700
	lys gly leu ser ala leu leu val asp lys arg pro val arg ala thr		
15	705	710	715
	val thr lys tyr his leu gln arg leu gly ile thr ser glu val val		
	725	730	735
	gly thr ile asp pro thr phe gly val leu ser gly arg asn gly ser		
	740	745	750
20	ser leu thr ser ile gly lys lys gln pro cys met leu leu ile glu		
	755	760	765
	ser asp ser trp gly pro gln met asp val ser leu his ala arg leu		
	770	775	780
	gln glu met lys gln ser asp arg ile his val leu pro lys val phe		
25	785	790	795
	leu leu ser ala ala glu ser asp lys val lys lys ile his ala val		
	805	810	815
	asp ser val ile pro lys pro leu lys ala ser ala leu ala ala cys		
	820	825	830
30	leu phe gln ala leu gly ile thr gln pro ser his glu lys arg asp		
	835	840	845
	asp ser gly ser leu his gly arg asp gly ser gly ser leu his gly		

850 855 860
 Leu Leu Leu Gly Lys Asn Ile Leu Val Val Asp Asp Asn Lys Val Asn
 865 870 875 880
 Leu Arg Val Ala Ala Gly Thr Leu Lys Lys Tyr Gly Ala Lys Val Glu
 5 885 890 895
 Cys Val Glu Ser Gly Lys Asp Ala Leu Ser Leu Leu Gln Val Pro His
 900 905 910
 Lys Phe Asp Leu Cys Leu Met Asp Ile Gln Met Pro Glu Met Asp Gly
 915 920 925
 10 Phe Glu Ala Thr Arg Gln Ile Arg Ala Met Glu Gly Lys Ala Asn Glu
 930 935 940
 Gln Ala Asp Asp Ser Glu Ser Gly Ser Glu Ile Ala Ala Lys Thr Ala
 945 950 955 960
 Lys Trp His Leu Pro Ile Leu Ala Met Thr Ala Asp Val Ile Gln Ala
 15 965 970 975
 Thr His Glu Glu Cys Thr Lys Cys Gly Met Asp Gly Tyr Val Ser Lys
 980 985 990
 Pro Phe Glu Glu Lys Gln Leu Phe Gln Ala Val Gln Lys Phe Leu Gly
 995 1000 1005
 20 Pro Cys Val Ser Ser
 1010
 <210> 19
 <211> 995
 <212> PRT
 25 <213> Zea mays
 <220>
 <221> Mutation
 <222> (267)..(267)
 <223> X= Xaa (any amino acid other than Leu)
 30 <220>
 <221> misc_feature
 <222> (267)..(267)

<223> Xaa can be any naturally occurring amino acid
 <400> 19
 Met Gly Val Gly Gly Gly Gly Gly Gly Glu Ala Ala Ala Val Ser Ala
 1 5 10 15
 5 Pro Ala Pro Ala Glu Glu Ala Gly Lys Asp Ala Glu Asp Gly Gly Gly
 20 25 30
 Trp Thr Leu Lys Ala Lys Leu Ile Ala Val Ala Val Leu Val Trp Val
 35 40 45
 Leu Gly Ala Leu Ala Leu Gly Val Phe Leu His Ser Tyr Phe Arg His
 10 50 55 60
 Ala Ala Leu Arg Lys Ala Glu Glu Gly Leu Val Ser Met Cys Glu Glu
 65 70 75 80
 Arg Ala Arg Met Leu Gln Asp Gln Phe Ala Val Ser Val Asn His Val
 85 90 95
 15 His Ala Leu Ala Ile Leu Val Ala Thr Phe His Tyr Glu Lys Arg Pro
 100 105 110
 Pro Ala Leu Asp Gln Asn Thr Phe Ala Asp Tyr Thr Ala Arg Thr Ser
 115 120 125
 Phe Glu Arg Pro Leu Leu Ser Gly Val Ala Tyr Ala Gln Arg Val Val
 130 135 140
 20 His Gly Asp Arg Glu Ser Phe Glu Arg Gln Gln Gly Trp Ile Ile Lys
 145 150 155 160
 Thr Met Lys His Glu Pro Ser Pro Val Gln Asp Glu Tyr Ala Pro Val
 165 170 175
 25 Val Tyr Ser Gln Glu Thr Val Ser Tyr Ile Glu Gly Leu Asp Met Met
 180 185 190
 Ser Gly Glu Glu Asp Arg Glu Asn Ile Leu Arg Ser Arg Ala Ser Gly
 195 200 205
 Lys Ala Val Leu Thr Arg Pro Phe Arg Leu Met Ser Asn His Leu Gly
 30 210 215 220
 Val Val Leu Thr Phe Pro Val Tyr His Val Asp Leu Ser Ser Asp Ala
 225 230 235 240

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

	Pro	Gly	Arg	Phe	Arg	Gln	Ile	Ile	Thr	Asn	Leu	Val	Gly	Asn	Ser	Ile	
				500					505					510			
	Lys	Phe	Thr	Glu	Arg	Gly	His	Ile	Phe	Val	Gln	Val	His	Leu	Ala	Asp	
				515				520					525				
5	His	Ser	Asn	Leu	Ala	Thr	Glu	Ser	Lys	Val	Glu	Ser	Val	Ala	Asn	Gly	
				530				535					540				
	Met	Asn	Gly	His	Lys	Asp	Glu	Lys	Thr	Ala	Val	Ala	Thr	Ser	Val	Ser	
	545					550					555					560	
	Leu	Asn	Thr	Leu	Ser	Gly	Phe	Glu	Ala	Ala	Asp	Ser	Arg	Asn	Ser	Trp	
10				565						570					575		
	Glu	Asn	Phe	Lys	Leu	Leu	Leu	Ser	Tyr	Glu	Lys	Asn	Glu	Met	Pro	Tyr	
				580					585					590			
	Glu	Ser	Val	Ser	Asp	Lys	Val	Thr	Leu	Val	Val	Ser	Val	Glu	Asp	Thr	
				595				600						605			
15	Gly	Ile	Gly	Ile	Pro	Leu	Asp	Ala	Gln	Ala	Lys	Val	Phe	Thr	Pro	Phe	
				610				615					620				
	Met	Gln	Ala	Asp	Ser	Ser	Thr	Ser	Arg	Thr	Tyr	Gly	Gly	Thr	Gly	Ile	
	625					630					635					640	
	Gly	Leu	Ser	Ile	Ser	Lys	Cys	Leu	Val	Glu	Leu	Met	Gly	Gly	Gln	Ile	
20				645						650					655		
	Asn	Phe	Val	Ser	Arg	Pro	His	Val	Gly	Ser	Thr	Phe	Thr	Phe	Thr	Ala	
				660					665					670			
	Ala	Leu	Gln	Arg	Cys	Asp	Arg	Ser	Ala	Ile	Gly	Asp	Ser	Lys	Pro	Val	
				675					680					685			
25	Met	Leu	His	Pro	Leu	Pro	Ser	Ser	Phe	Lys	Gly	Leu	Ser	Ala	Leu	Leu	
				690				695						700			
	Val	Asp	Arg	Arg	Pro	Val	Arg	Ala	Thr	Val	Thr	Lys	Tyr	His	Leu	Gln	
	705					710						715				720	
	Arg	Leu	Gly	Ile	Ala	Cys	Asp	Val	Val	Ala	Thr	Ile	Glu	Leu	Ala	Leu	
30				725						730					735		
	Gly	Val	Leu	Ser	Gly	Arg	Asn	Gly	Ser	Ser	Leu	Thr	Ser	Thr	Lys	Gln	
				740					745						750		

Pro Cys Met Leu Leu Ile Glu Ser Asp Ser Trp Gly Phe Lys Ile Asp
 755 760 765
 Val Pro Leu Arg Ser Arg Leu Leu Glu Met Lys Gln Asn Gly Pro Pro
 770 775 780
 5 Gly Leu Pro Lys Thr Ile Leu Leu Ala Ala Ala Glu Ser Gly Lys Leu
 785 790 795 800
 Lys Ala His Tyr Ala Val Asp Ser Val Ile Thr Lys Pro Leu Lys Ala
 805 810 815
 Ser Gly Leu Ala Ala Cys Leu Phe Gln Thr Leu Gly Ile Thr Gln Ser
 10 820 825 830
 Ser Asn Glu Arg Arg Asp Asn Ser Gly Ser Leu His Gly Leu Leu Leu
 835 840 845
 Gly Lys Asn Ile Leu Val Val Asp Asp Asn Lys Val Asn Leu Arg Val
 850 855 860
 15 Ala Ala Gly Thr Leu Lys Lys Phe Gly Ala Lys Val Glu Cys Val Glu
 865 870 875 880
 Ser Gly Lys Asp Ala Leu Ala Ser Leu Gln Val Pro His Lys Phe His
 885 890 895
 Leu Cys Leu Met Asp Ile Gln Met Pro Glu Met Asp Gly Phe Glu Ala
 20 900 905 910
 Thr Lys Gln Ile Arg Ala Met Glu Ala Lys Ala Asn Glu Gln Ala Val
 915 920 925
 Ala Cys Asp Asp Ser Asp Thr Asp Gly Ala Thr Arg Ala Ala Arg Trp
 930 935 940
 25 His Leu Pro Val Leu Ala Met Thr Ala Asp Val Ile Gln Ala Thr His
 945 950 955 960
 Glu Glu Cys Thr Lys Tyr Gly Met Asp Gly Tyr Val Thr Lys Pro Phe
 965 970 975
 Glu Glu Lys Gln Leu Phe Gln Ala Leu Gln Lys Phe Leu Asp Pro Gly
 30 980 985 990
 Met Ser Ser
 995

<210> 20
 <211> 981
 <212> PRT
 <213> Cucurbita maxima
 5 <220>
 <221> Mutation
 <222> (261)..(261)
 <223> X= Xaa (any amino acid other than Leu)
 <220>
 10 <221> misc_feature
 <222> (261)..(261)
 <223> Xaa can be any naturally occurring amino acid
 <400> 20
 Met Gln Val Ser Asp Asn Ser Val Gly Leu Lys Trp Asn Glu Gln Met
 15 1 5 10 15
 Gly Thr Thr Lys Lys Gly Tyr Thr Phe Val Gln Ala Asn Arg Ala Trp
 20 25 30
 Leu Arg Lys Tyr Leu Leu Phe Trp Ile Met Gly Met Ala Phe Ile Ser
 35 40 45
 20 Met Leu Ile Tyr Asn Gly Met Asp Ala Asp Ile Lys Val Arg Arg Asn
 50 55 60
 Glu Val Leu Gly Ser Met Cys Glu Gln Arg Ala Arg Met Leu Gln Asp
 65 70 75 80
 Gln Phe Asn Val Ser Val Asn His Val His Ala Leu Ala Val Leu Val
 25 85 90 95
 Ser Thr Phe His Tyr Phe Lys Asn Pro Ser Ala Ile Asp Gln Glu Thr
 100 105 110
 Phe Ala Glu Tyr Thr Ala Arg Thr Ala Phe Glu Arg Pro Leu Leu Ser
 115 120 125
 30 Gly Val Ala Tyr Ala Gln Arg Val Ile His Ser Glu Arg Asp Ile Phe
 130 135 140
 Glu Lys Gln His Gly Trp Met Ile Arg Thr Met Glu Lys Glu Pro Ser

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

	405	410	415
	Gln Thr Ala Gln Ala Cys Gly Lys Ala Leu Ile Ala Leu Ile Asn Glu		
	420	425	430
	Val Leu Asp Arg Ala Lys Ile Glu Ala Gly Lys Leu Glu Leu Glu Ala		
5	435	440	445
	Val Pro Phe Asp Ile Arg Ser Ile Leu Asp Asp Val Leu Ser Leu Phe		
	450	455	460
	Ser Glu Lys Ser Arg Gln Lys Gly Leu Glu Leu Ala Val Phe Val Ser		
	465	470	475
10	480	485	490
	Asp Lys Val Pro Glu Ile Val Ile Gly Asp Pro Gly Arg Phe Arg Gln		
	495	500	505
	Ile Ile Thr Asn Leu Val Gly Asn Ser Val Lys Phe Thr Glu Arg Gly		
	510	515	520
	His Ile Phe Val Lys Val His Leu Ala Glu Asn Ser Lys Val Ser Met		
15	525	530	535
	Asp Ser Glu Tyr Val Asn Gly Ile Ser Asp Ser Gly Leu Phe Val Leu		
	540	545	550
	Asp Gly Arg Glu Phe Gln Thr Leu Ser Gly Arg Glu Ala Ala Asp Asp		
	555	560	565
20	570	575	580
	Gln Asn Ser Trp Asp Asn Phe Lys His Leu Ile Ala Asp Asp Asn Phe		
	585	590	595
	Gln Ser Asn Ala Ala Ser Asn Asn Ser Ala Val Thr Asn Lys Gly Cys		
	600	605	610
	Asp His Val Thr Leu Met Val Ser Val Glu Asp Thr Gly Ile Gly Ile		
25	615	620	625
	Leu Leu His Ala Gln Asn Arg Val Phe Thr Pro Phe Met Gln Ala Asp		
	630	635	640
	Ser Ser Thr Ser Arg Asn Tyr Gly Gly Thr Gly Ile Gly Leu Ser Ile		
	645	650	655
30	660	665	670
	Ser Lys Cys Leu Val Glu Leu Met Gly Gly Gln Ile Asn Phe Ile Ser		
	675	680	685
	Arg Pro Gln Ile Gly Ser Thr Phe Ser Phe Thr Ala Val Phe Gly Lys		

	660	665	670
	Cys Lys Lys Asn Ser Met Asn Asp Met Lys Lys Pro Asn Ser Glu Glu		
	675	680	685
	Leu Pro Pro Ser Phe Arg Gly Met Lys Ala Ile Val Val Asp Ser Lys		
5	690	695	700
	His Val Arg Ala Ser Val Thr Arg Tyr His Leu Lys Arg Leu Gly Ile		
	705	710	715
	Ile Val Glu Val Thr Asn Ser Ile Asn Met Ala Ala Ser Leu Phe Arg		
	725	730	735
10	Glu Asn Gly Ser Thr Leu Pro Arg Asn Thr Ile Leu Pro Asp Met Ile		
	740	745	750
	Leu Val Glu Lys Asp Ile Leu Asn Ser Asp Glu Glu Cys Gly Ile Ile		
	755	760	765
	His His Leu Asn Trp Lys Pro Asn Gly Ser Ser Val Lys Phe Pro Lys		
15	770	775	780
	Leu Ile Leu Leu Ala Thr Asn Ile Ala Thr Ala Glu Leu Asp Lys Ala		
	785	790	795
	Arg Ala Ala Gly Phe Ala Asp Thr Val Ile Met Lys Pro Leu Arg Ala		
	805	810	815
20	Thr Met Val Ala Ala Cys Leu Gln Gln Val Leu Gly Val Lys Asn Gln		
	820	825	830
	Arg Arg Pro Asn Gly Ser Ala Phe Leu Gln Ser Leu Leu Cys Gly Lys		
	835	840	845
	Arg Ile Leu Ile Val Asp Asp Asn Arg Val Asn Arg Arg Val Ala Ala		
25	850	855	860
	Gly Ala Leu Lys Lys Phe Gly Ala Asp Val Glu Cys Ala Asp Ser Gly		
	865	870	875
	Lys Ser Ala Leu Lys Leu Leu Gln Leu Pro His Asn Phe Asp Ala Cys		
	885	890	895
30	Phe Met Asp Ile Gln Met Pro Glu Met Asp Gly Phe Glu Ala Thr Arg		
	900	905	910
	Arg Ile Arg Thr Met Glu Val Glu Ala Asn Lys Gly Gly Leu Ser Ala		

[illegible]

REIVINDICAÇÕES

1. Molécula de DNA que codifica um polipeptídeo de histidina quinase mutante, caracterizada pelo fato de que compreende uma sequência de aminoácidos selecionada entre:

5 a. SEQ ID NO: 6, 15 e 16, em que o resíduo de aminoácido 266 [Phe] na SEQ ID NO: 6 e 266 [Xaa] na SEQ ID NO: 15 e resíduo 267 [Xaa] na SEQ ID NO: 16 é selecionado entre isoleucina, serina, treonina, valina, metionina, alanina, fenilalanina, tirosina, triptofano, arginina, lisina, glicina, histidina, aspartato, asparagina, glutamato, glutamina, prolina e cisteína,

10 b. um truncamento de (a),
em que o referido peptídeo é capaz de induzir formação espontânea quando expressa em uma planta.

2. Polipeptídeo de histidina quinase mutante, caracterizado pelo fato de ser codificado pela molécula de DNA tal como definida na reivindicação 1.

15 3. Planta geneticamente modificada, caracterizada por ter uma sequência de nucleotídeo que codifica um polipeptídeo de histidina quinase mutante que compreende uma sequência de aminoácidos selecionada entre:

20 a. SEQ ID NO: 6, 15 e 16, em que o resíduo de aminoácido 266 [Phe]na SEQ ID NO: 6 e resíduo 266 [Xaa]na SEQ ID NO: 15 e resíduo 267 [Xaa] na SEQ ID NO: 16 é selecionado entre isoleucina, serina, treonina, valina, metionina, alanina, fenilalanina, tirosina, triptofano, arginina, lisina, glicina, histidina, aspartato, asparagina, glutamato, glutamina, prolina e cisteína,

25 b. um truncamento de (a),
em que a dita planta é capaz de formação espontânea de nódulo.

4. Planta geneticamente modificada de acordo com a reivindicação 3, caracterizada pelo fato de que o polipeptídeo de histidina quinase mutante é codificado por uma molécula de ácido nucleico tendo uma sequência de ácido nucleico selecionada do grupo: SEQ ID NO: 4, 5, e 7, ou a sequência codificante dessas.

5. Uso de uma molécula de DNA que codifica um polipeptídeo

de histidina quinase mutante, caracterizado pelo fato de ser capaz de induzir formação de nódulo espontânea quando expressa em uma planta consistindo em uma sequência de aminoácido selecionado entre:

a. SEQ ID NO: 6, 15 e 16, em que o resíduo de aminoácido 266 [Phe]na SEQ ID NO: 6 e resíduo 266 [Xaa]na SEQ ID NO: 15 e resíduo 267 [Xaa] na SEQ ID NO: 16 é selecionado entre isoleucina, serina, treonina, valina, metionina, alanina, fenilalanina, tirosina, triptofano, arginina, lisina, glicina, histidina, aspartato, asparagina, glutamato, glutamina, prolina e cisteína, e

b. um truncamento de (a),
como um transgene para produzir a planta geneticamente modificada como definida na reivindicação 5.

6. Planta geneticamente modificada como definida na reivindicação 3 ou 4, caracterizada pelo fato de que compreende adicionalmente uma sequência de nucleotídeos homóloga ou heteróloga do promotor operavelmente ligada à sequência de nucleotídeos que codifica o polipeptídeo.

7. Planta geneticamente modificada de acordo com a reivindicação 6, caracterizada pelo fato de que o dito promotor é um promotor regulado.

8. Planta geneticamente modificada de acordo com qualquer uma das reivindicações 3, 4, 6 ou 7, caracterizada pelo fato de que a dita planta é uma planta monocotiledônea ou dicotiledônea.

9. Planta geneticamente modificada de acordo com a reivindicação 8, caracterizada pelo fato de que a dita planta é selecionada entre arroz, cevada, milho, aveia, centeio, sorgo, trigo e grama *Poaceae*.

10. Planta geneticamente modificada de acordo com qualquer uma das reivindicações 3, 4, 6 a 9, obtível por um processo de mutagenese de DNA e seleção de uma planta capaz de formação espontânea de nódulo, caracterizada pelo fato de ter uma sequência de nucleotídeo que codifica um polipeptídeo de histidina quinase mutante, compreendendo uma

a. SEQ ID NO: 6, 15 e 16, em que o resíduo de aminoácido 266

[Phe]na SEQ ID NO: 6 e resíduo 266 [Xaa]na SEQ ID NO: 15 e resíduo 267 [Xaa] na SEQ ID NO: 16 é selecionado entre isoleucina, serina, treonina, valina, metionina, alanina, fenilalanina, tirosina, triptofano, arginina, lisina, glicina, histidina, aspartato, asparagina, glutamato, glutamina, prolina e cisteína, e

b. um truncamento de (a)

11. Método de produção de uma planta geneticamente modificada tal como definida em qualquer uma das reivindicações 3, 4, 6 a 9, caracterizado pelo fato de que a introdução de cassete de gene compreendendo a dita sequência de nucleotídeos que codifica o dito polipeptídeo e selecionando uma planta transgênica e a sua progênie que expressa o dito polipeptídeo.

12. Método de acordo com a reivindicação 11, caracterizado pelo fato de que o cassete de gene é introduzido em uma planta pela transformação ou é sujeito à "introgress" na planta por cruzamento sexual com a planta transgênica compreendendo o dito cassete de gene.

13. Planta geneticamente modificada produzida de acordo com o método como definido na reivindicação 11 ou 12, caracterizada pelo fato de ter uma sequência de nucleotídeo que codifica um polipeptídeo de histidina quinase mutante, compreendendo uma sequência de aminoácido selecionada entre:

a. SEQ ID NO: 6, 15 e 16, em que o resíduo de aminoácido 266 [Phe]na SEQ ID NO: 6 e resíduo 266 [Xaa]na SEQ ID NO: 15 e resíduo 267 [Xaa] na SEQ ID NO: 16 é selecionado entre isoleucina, serina, treonina, valina, metionina, alanina, fenilalanina, tirosina, triptofano, arginina, lisina, glicina, histidina, aspartato, asparagina, glutamato, glutamina, prolina e cisteína, e

b. um truncamento de (a)

em que o referido peptídeo é capaz de induzir formação espontânea quando expressa em uma planta.

14. Semente da planta geneticamente modificada tal como definida em qualquer uma das reivindicações 3, 4, 6 a 10 e 13, caracterizada

pelo fato de ter uma sequência de nucleotídeo que codifica um polipeptídeo de histidina quinase mutante, compreendendo uma sequência de aminoácido selecionada entre:

- a. SEQ ID NO: 6, 15 e 16, em que o resíduo de aminoácido 266 [Phe]na SEQ ID NO: 6 e resíduo 266 [Xaa]na SEQ ID NO: 15 e resíduo 267 [Xaa] na SEQ ID NO: 16 é selecionado entre isoleucina, serina, treonina, valina, metionina, alanina, fenilalanina, tirosina, triptofano, arginina, lisina, glicina, histidina, aspartato, asparagina, glutamato, glutamina, prolina e cisteína, e
 - b. um truncamento de (a)
- em que o referido polipeptídeo é capaz de induzir formação de nódulo espontânea quando expressa em uma planta.
15. Plantação, caracterizada pelo fato de que compreende uma planta geneticamente modificada tal como definida em qualquer uma das reivindicações 3, 4, 6 a 10 e 13.

RESUMO

Patente de Invenção: "MOLÉCULA DE DNA QUE CODIFICA UM POLI-
PEPTÍDEO DE HISTIDINA QUINASE MUTANTE E SEU USO RELACIO-
NADO, POLIPEPTÍDEO DE HISTIDINA QUINASE MUTANTE, BEM COMO
5 MÉTODO DE PRODUÇÃO DE UMA PLANTA GENETICAMENTE MODIFI-
CADA".

A presente invenção refere-se à formação de nódulos radícula-
res fixadores de nitrogênio em leguminosas que é induzida pela captação de
moléculas sinalizadoras de oligossacarídeo de lipoquitina secretadas por
10 bactérias *Rhizobium* compatíveis, que desencadeiam uma via simbiótica
comum. A presente invenção provê um mutante de *formação espontânea de
nódulo (snf2)*, no qual a formação de nódulos simbióticos é espontânea, le-
vando ao desenvolvimento nodular na ausência bem como na presença de
bactérias *Rhizobium* e/ou sinais rizobiais exógenos. A invenção ainda provê
15 uma sequência de DNA isolada que codifica uma histidina quinase mutante
independente de citocinina, cuja atividade resulta neste fenótipo dominante
de 'ganho de função' da formação espontânea de nódulo. Além disso, mos-
tra-se que o gene *snf2* confere um fenótipo caracterizado pela organogênese
regulada de nódulos espontâneos em plantas que têm um contexto genético
20 de nodulação deficiente. Um gene da invenção, que confere este fenótipo de
nodulação espontânea, tem utilidade para transferência e estabelecimento
da capacidade de fixação de nitrogênio em plantas não-noduladas, e por
meio disso reduz a dependência de fertilizantes nitrogenados das culturas de
plantas não-noduladas.