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

(21) **PI 0712449-0 A2**

* B R P I 0 7 1 2 4 4 9 A 2 *

(22) Data de Depósito: 30/05/2007
(43) Data da Publicação: 13/03/2012
(RPI 2149)

(51) Int.Cl.:
C12N 15/82
C07K 14/415
C07K 14/47
C12N 9/12

(54) Título: MÉTODO PARA AUMENTAR A RESISTÊNCIA À TENSÃO ABIÓTICA EM PLANTAS COM RELAÇÃO ÀS PLANTAS DE CONTROLE, E, USO DE UMA CONSTRUÇÃO

(30) Prioridade Unionista: 30/05/2006 EP 06114726.0, 30/05/2006 EP 06114735.1, 31/05/2006 EP 06114758.3, 08/06/2006 EP 06115142.9, 30/06/2006 EP 06116490.1, 12/07/2006 EP 06117060.1, 28/07/2006 EP 06118060.0, 02/06/2006 US 60/810572, 02/06/2006 US 60/810749, 02/06/2006 US 60/810759, 14/06/2006 US 60/813514, 13/07/2006 US 60/830551, 18/07/2006 US 60/831642, 03/08/2006 US 60/835303, 30/05/2006 EP 06114726.0, 30/05/2006 EP 06114735.1, 31/05/2006 EP 06114758.3, 08/06/2006 EP 06115142.9, 30/06/2006 EP 06116490.1, 12/07/2006 EP 06117060.1, 28/07/2006 EP 06118060.0, 02/06/2006 US 60/810572, 02/06/2006 US 60/810749, 02/06/2006 US 60/810759, 14/06/2006 US 60/813514, 13/07/2006 US 60/830551, 03/08/2006 US 60/835303, 30/05/2006 EP 06114726.0, 30/05/2006 EP 06114735.1, 31/05/2006 EP 06114758.3, 08/06/2006 EP 06115142.9, 30/06/2006 EP 06116490.1, 12/07/2006 EP 06117060.1, 28/07/2006 EP 06118060.0, 14/06/2006 US 60/813514, 13/07/2006 US 60/830551, 18/07/2006 US 60/831642

(73) Titular(es): Cropdesign N.V.

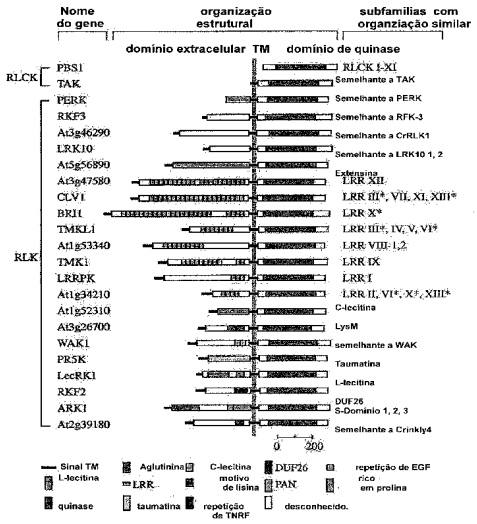
(72) Inventor(es): Christophe Reuzeau, Valerie Frankard, Vladimir Mironov

(74) Procurador(es): Momsen, Leonardos & CIA.

(86) Pedido Internacional: PCT EP2007055238 de 30/05/2007

(87) Publicação Internacional: WO 2007/138070de 06/12/2007

(57) Resumo: MÉTODO PARA AUMENTAR A RESISTÊNCIA A TENSÃO ABIÓTICA EM PLANTAS COM RELAÇÃO ÀS PLANTAS DE CONTROLE, E, USO DE UMA CONSTRUÇÃO. A presente invenção diz respeito no geral ao campo da biologia molecular e diz respeito a um método para melhorar várias características do crescimento vegetal pela modulação da expressão em uma planta de um ácido nucleico que codifica um GRP (Proteína Relacionada com o Crescimento). A presente invenção também diz respeito a plantas tendo expressão modulada de um ácido nucleico que codifica um GRP, cujas plantas tem características de crescimento melhoradas com relação às plantas do tipo selvagem correspondentes ou outras plantas de controle. A invenção também fornece construções úteis nos métodos da invenção. A GRP pode ser uma das seguintes: Quinase semelhante ao receptor de extensina (ERLK), polipeptídeo F-Box WD4O (FBXW), Proteína de ligação de RAN (RANBP), Fator de Transcrição semelhante a Golden2 (GLK), Polipeptídeo de domínio zíper de leucina de homeodomínio REV delta, proteína CLE e proteína reguladora do crescimento de semente (SYR).



"MÉTODO PARA AUMENTAR A RESISTÊNCIA À TENSÃO ABIÓTICA EM PLANTAS COM RELAÇÃO ÀS PLANTAS DE CONTROLE, E, USO DE UMA CONSTRUÇÃO"

No geral, a presente invenção diz respeito ao campo da biologia molecular e diz respeito a um método para melhorar várias características do crescimento vegetal pela modulação da expressão em uma planta de um ácido nucleico que codifica um GRP (Proteína Relacionada com o Crescimento). A presente invenção também diz respeito a plantas tendo expressão modulada de um ácido nucleico que codifica um GRP, cujas plantas tem características de crescimento melhoradas com relação às plantas do tipo selvagem correspondentes ou outras plantas de controle. A invenção também fornece construções úteis nos métodos da invenção.

A população mundial sempre crescente e o fornecimento em declínio de terra arável disponível para a pesquisa de combustíveis agrícolas com respeito ao aumento da eficiência da agricultura. Meios convencionais para as melhoras de safra e da horticultura utilizam técnicas de criação seletivas para a identificação de plantas tendo características desejáveis. Entretanto, tais técnicas de criação seletivas têm diversas desvantagens, isto é que estas técnicas são tipicamente trabalho intensivo e resulta nas plantas que freqüentemente contém componentes genéticos heterogêneos que nem sempre podem resultar na característica desejável sendo passadas das plantas precursoras. Avanças na biologia molecular permitiram ao homem e modificar o germoplasma de animais e plantas. A engenharia genética de plantas vincula o isolamento e a manipulação do material genético (tipicamente na forma de DNA ou RNA) e a introdução subseqüente daquele material genético em uma planta. Tal tecnologia tem a capacidade de liberar safras ou plantas tendo várias características econômicas, agronômicas ou de horticultura.

Uma característica de interesse econômico particular é

rendimento aumentado. O rendimento é normalmente definido como a produção mensurável de valor econômico de uma safra. Isto pode ser definido em termos de quantidade e/ou qualidade. O rendimento é diretamente dependente de diversos fatores, por exemplo, o número e o tamanho dos órgãos, arquitetura da planta (por exemplo, o número de ramificações), produção de semente, senescência foliar. O crescimento da raiz, absorção de nutriente, tolerância à tensão e vigor precoce também podem ser fatores importantes na determinação do rendimento. A otimização dos fatores mencionados acima pode, portanto contribuir para aumentar o rendimento da safra.

O rendimento da semente é uma característica particularmente importante, visto que as sementes de muitas plantas são importantes para a nutrição humana ou animal. Safras, tais como milho, arroz, trigo, canola e soja responde por mais da metade a absorção calórica humana total, através do consumo direto dos alimento por si só ou através dos produtos de carne unidos com sementes processadas. Estes também são uma fonte de açúcares, óleos e muitos tipos de metabólitos usados em processos industriais. As sementes contém um embrião (a fonte de novos brotos e raízes) e um endosperma (a fonte de nutrientes para o crescimento do embrião durante a germinação e durante o crescimento precoce de mudas). O crescimento de uma semente envolve muitos genes e requer a transferência de metabólitos das raízes, folhas e caules na semente em crescimento. O endosperma, em particular, assimila-se aos precursores metabólicos de carboidratos, óleos e proteínas e os sintetiza em macromoléculas de armazenamento para preencher o grão.

Uma outra característica importante para muitas safras é o vigor precoce. A melhora do vigor precoce é um objetivo importante de programas de geração de arroz moderno tanto em cultivares de arroz temperados quanto tropicais. Raízes longas são importantes para a ancoragem

apropriada no solo em arroz semeado em água. Quando o arroz é semeado diretamente em campos alagados e quando as plantas devem emergir rapidamente através da água, brotos mais longos são associados com vigor. Quando a semeadura em sulco é praticada, mesocotilas e coleóptilos mais longos são importantes para a boa emergência da semeadura. A capacidade de projetar vigor precoce em plantas deve ser de grande importância na agricultura. Por exemplo, vigor precoce deficiente foi uma limitação para a introdução de híbridos de milho (Zea mays L.) com base em germoplasma da região milheira no Atlântico europeu.

Uma característica importante adicional é aquela de tolerância à tensão biótica melhorada. A tensão biótica é uma causa primária da perda de safra no mundo todo, reduzindo-se os rendimentos médios para a maioria das plantas de safra principais por mais do que 50% (Wang et al., Planta (2003) 218: 1-14). As tensões abióticas As tensões abióticas podem ser causadas por seca, salinidade, extremos de temperatura, toxicidade química e tensões oxidativas. A capacidade de melhorar a tolerância da planta à tensão abiótica deve ser de vantagem econômica grande para fazendeiros do mundo todo e deve permitir o cultivo de safras durante condições adversas e em territórios onde o cultivo de safras, de outra maneira, pode não ser de outra maneira possível.

O rendimento da safra pode, portanto, ser aumentado pela otimização de um dos fatores mencionados acima.

Dependendo do uso final, a modificação de certas características de rendimento pode ser favorecida sobre outras. Por exemplo, para aplicações, tais como produção de forragem ou de madeira ou recurso de bio-combustível, um aumento nas partes vegetativas de uma planta pode ser desejável e para aplicações tais como produção de farinha, amido ou óleo, um aumento nos parâmetros de semente pode ser particularmente desejável. Ainda entre os parâmetros de semente, alguns podem ser favorecidos sobre

outros, dependendo da aplicação. Vários mecanismos podem contribuir podem contribuir para aumentar o rendimento de semente, quando aquele está na forma de tamanho de semente aumentado ou número de semente aumentado.

Um método para aumentar o rendimento (rendimento de semente e/ou biomassa) em plantas pode ocorrer através da modificação dos mecanismos de crescimento inerentes de uma planta, tal como o ciclo celular ou vários caminhos de sinalização envolvidos no crescimento da planta ou em mecanismos de defesa.

Foi observado agora que várias características de crescimento podem ser melhoradas em plantas pela modulação da expressão em uma planta de um ácido nucleico que codifica um GRP (Proteína Relacionada com o Crescimento de Interesse) em uma planta. O GRP pode ser um dos seguintes: Quinase Semelhante ao Receptor de Extensina (ERLK), polipeptídeo F-Box WD40 (FBXW), Proteína de Ligação RAN (RANBP), Fator de Transcrição semelhante a Golden2 (GLK), polipeptídeo de domínio zíper de leucina de homeodomínio REV delta, proteína CLE e proteína Reguladora do Domínio de Semente (SYR).

FUNDAMENTOS

Quinase semelhante ao Receptor de Extensina

As quinases semelhantes ao receptor (RLKs) estão envolvidas na transmissão de sinais extracelulares na célula. As proteínas RLK têm uma estrutura modular, começando do terminal N com um sinal de secreção que torna-se processado, um domínio extracelular, um domínio de transmembrana simples e um domínio de quinase citoplásmico. As quinases semelhantes ao receptor animal, na maioria das vezes tem atividade de tirosina quinase, visto que os RLKs vegetais usualmente tem especificidade de Ser/Thr quinase ou pode, algumas vezes, ter uma especificidade dupla. Em animais, a maioria dos RLKs atua como os receptores do fator de crescimento, visto que as quinases

semelhantes ao receptor vegetal pode funcionar em vários processos, incluindo crescimento, percepção hormonal ou resposta de patogênio. Uma visão geral das funções de crescimento de quinases semelhantes ao receptor vegetal, tal como crescimento de meristema, interações de pólen-pistilo, sinalização de hormônio, crescimento de gametófito, morfogênese e diferenciação celular, forma do órgão, abscisão do órgão e embriogênese somática é dada por Becraft (Annu. Rev. Cell Dev. Biol., 18, 163-192, 2002).

Alternativamente, as quinases semelhantes ao receptor podem ser agrupadas de acordo com a estrutura de seus domínios extracelulares ou intracelulares (Shiu and Bleecker, Proc. Natl. Acad. Sci. USA 98, 10763-10768, 2001). Uma visão geral é dada na Figura 1. PERK RLKs e ERLKs têm um domínio extracelular que é rico em prolina e que tem motivos típicos para proteínas de parede celular ricas em extensinas e hidroxiprolina. As extensinas são um grupo de glicoproteínas ricas em hidroxiprolina observadas em paredes celulares vegetais. Estes são, usualmente, ricos em hidroxiprolina (Hyp), serina e combinações de Val, Tyr, His e Lys. Os motivos típicos em proteínas de extensina são o motivo SP_x, em que x representa o número de repetições de (hidróxi)prolina, usualmente 2, 3, 4 ou mais. As Extensinas podem contar com até 20% do peso seco da parede celular. Estes são altamente glicosilados, refletindo, possivelmente suas interações com os carboidratos da parede celular. Entre as suas funções, está o reforço da parede celular em resposta à tensão mecânica (por exemplo, durante o ataque por pragas, curvando a planta no vento, etc.). os motivos de extensina também são observados no grupo pequeno de quinases semelhantes ao receptor de extensina, identificado pelo gene At5g56890 de *Arabidopsis*. Shiu and Bleeker (2001) identificaram 5 membros familiares em Arabidopsis; vários ortólogos foram observados no arroz (Shiu et al. Plant Cell 16, 1220-1234, 2004) e em outras espécies. Os RLKs de extensinas ainda não foram caracterizados.

Polipeptídeo F-box WD40 (FBXW)

As plantas devem ajustar seu metabolismo aos estímulos externos e internos para garantir um crescimento ótimo. Os níveis de proteínas reguladoras envolvidas nestes processos celulares são freqüentemente controlados por mecanismos proteolíticos. Entre os caminhos de degradação de proteína seletivos mais importantes neste assunto é o caminho proteolítico dependente de ubiquitina. Este caminho é conservado entre as plantas, animais e levedura e este controla a degradação de polipeptídeos maldobrados e proteínas de vida curta. Entre os últimos estão proteínas reguladoras importantes que regulam processos como respostas de defesa e de tensão, progressão do ciclo celular ou transdução de sinal. As proteínas destinadas serem degradadas são covalentemente rotuladas com diversas unidades de ubiquitina. A proteína ubiquitinada é subseqüentemente reconhecida pelo 26S proteassoma que degrada a proteína alvo mas recicla os monômeros de ubiquitina. A seleção e a ubiquitinação subseqüente de uma proteína alvo ocorre em etapas diferentes. Três classes de proteínas que estão envolvidas foram denominadas E1 a E3, com base em sua ação sequencial. A enzima E1, também conhecida como a enzima ativadora de ubiquitina, "ativa" uma molécula de ubiquitina livre no gasto de um ATP e complexa a ubiquitina em uma ligação de tioéster. A seguir, a molécula de ubiquitina ativada é transferida da enzima E1 a uma cisteína de uma enzima conjugadora de ubiquitina E2. Esta enzima E2 normalmente está associada com uma enzima E3 denominada ubiquitina proteína ligase. A ubiquitina ligase catalisa a transferência da ubiquitina a partir da ubiquitina que conjuga a enzima à proteína alvo, que ultimamente torna-se rotulada com uma cadeia de poli-ubiquitina. O complexo de enzimas E2 e E3 determina a especificidade para a proteína ser ubiquitinada. Visto que, em organismos diferentes uma ou apenas algumas enzimas E1 estão presentes, diversas espécies E2 que existem podem ser associadas com diversas enzimas E3. A ubiquitina proteína ligase por si só

também é um complexo de proteínas diferentes. Até hoje, cinco tipos diferentes de enzima E3 são conhecidos. Entre estes, o complexo de SCF desempenha um papel proeminente nos processos reguladores (Ciechanover (1998) EMBO J. 17: 7151-7160; Hershko and Ciechanover (1998) Annu. Rev. Biochem. 67: 425-479). O complexo consiste de quatro subunidades: cdc53/cullin, Skp1, Rbx1 e uma proteína F-box. A proteína Skp1, junto com cullin e Rbx1 forma a unidade de ligase do núcleo do complexo SCF. Dentro deste complexo, o Rbx1 é responsável pela ligação da enzima E2 carregada com uma ubiquitina ativada. As proteínas F-box contém, no terminal N, um motivo degenerado de 40 a 50 aminoácidos, conhecido com o F-box, denominado após a ciclina humana F. Esta proteína F-box é responsável pela associação com a subunidade Skp1 do complexo de SCF. No terminal C, um domínio de interação de proteína variável determina a ligação da proteína alvo. Um tal domínio de interação de proteína é um domínio de repetição WD40.

Em *Arabidopsis*, as proteínas F-box representa uma das superfamílias maiores observadas até agora nas plantas, em comparação com outros organismos (Gagne *et al.* (2002) Proc Natl Acad Science 99: 11519-11524; Kuroda *et al.* (2002) Plant Cell Physiol 43(10): 1073-85). entretanto, apenas duas proteínas F-box contém um domínio de repetição WD40, visto que muitas repetições WD40 são encontradas em proteínas F-box de outros organismos. As repetições WD40 (também são conhecidas como repetições de beta-transducina) são motivos de ~40 aminoácidos curtos, freqüentemente terminando em um dipeptídeo Trp-Asp (W-D). As repetições de WD40 contendo proteínas (ou proteínas WD40) tem de 4 a 16 unidades de repetição (coletivamente para o domínio WD40), todos os quais são pensados formar uma estrutura de beta-propelente circularizado. A função comum subjacente de todas as proteínas WD40 são montagens de complexo multi-proteína coordenante, quando as unidades de repetição como uma estrutura rígida para

as interações de proteína. A especificidade das proteínas é determinada pelas seqüências fora das repetições por si só.

O pedido de patente europeu publicado EP 1033405 fornece uma seqüência de DNA (SEQ ID N° : 39903) de *Arabidopsis thaliana* que codifica um polipeptídeo FBXW parcial (seqüência de aminoácido de terminal N, em torno de 350 nucleotídeos de comprimento).

Proteína de ligação de RAN (RANBP)

RAN é uma GTPase de sinalização pequena (proteína de ligação de pGTP), que está envolvido no transporte nucleocitoplásmico. As proteínas de ligação de Ran em *Arabidopsis thaliana*, o At-RanBP1a, At-RanBP1b, AtRanbp1c foram relatados interagir com as formas ligadas ao GTP das proteínas Ran1, Ran2 e Ran3 de *Arabidopsis thaliana* (Haizel T, Merkle T, Pay A, Fejes E, Nagy F. Characterization of proteins that interact with the GTP-bound form of the regulatory GTPase Ran in Arabidopsis. Plant J. 1997 Jan;11(1):93-103). Todas as proteínas RanBP1 contém um domínio de ligação de Ran de resíduo de aproximadamente 150 aminoácidos. O Ran BP1 liga-se diretamente ao RanGTP com afinidade alta. Este domínio estabiliza a forma de ligação de GTP de Ran (a GTPase pequena nuclear semelhante ao Ras).

O pedido de Patente Internacional WO 02/18538 descreve as plantas transgênicas tendo processos celulares mediados por proteína de ligação de RAN/Ran perturbado; isto é relatado originar plantas com rendimento e biomassa aumentados.

Fator de transcrição semelhante a GOLDEN2 (GLK)

Em plantas, dois tipos de ciclos fotossintéticos podem ocorrer: o mais comum é o ciclo de Calvin de plantas C3, em que o 3-fosfoglicerato é o primeiro produto estável e a bisfosfato ribulose é o receptor de CO_2. O segundo ciclo é o caminho de Hatch-Smack em plantas C4, em que o oxaloacetato é o primeiro produto estável e o fosfoenolpiruvato é o receptor

de CO_2. Em gramas C3, apenas as células mesófilas são fotossintéticas, visto que em plantas C4 tanto a folha de feixes quanto as células mesófilas são fotossintéticas. As plantas C4 apresentam fotossíntese compartimentalizada com as células mesófilas realizando a fixação de carbono por intermédio da fosfoenolpiruvato carboxilase, piruvato fosfato diquinase e malato desidrogenase e transportanto o malato para as células da superfície de feixes, em que o malato é decarboxilado ao piruvato, o carbono liberado é então processado no ciclo de Calvin. Como uma conseqüência, três tipos de cloroplastos estão presentes em plantas C4: os cloroplastos típicos da planta C3 existe em certos tecidos e em certos estágios de crescimento, enquanto na bainha do feixe e nas células mesófilas os cloroplastos morfologicamente distintos estão presentes. A gênese de cloroplastos em milho requer o envolvimento de dois reguladores transcripcionais: Golden2 (G2) e semelhante a Golden2 (GLK) (Rossini et al., Plant Cell 13, 1231-1244, 2001).

Os fatores de transcrição são usualmente definidos como as proteínas que apresentam ligação de DNA específica da seqüência e que são capazes de ativar e/ou reprimir a transcrição. O genoma de *Arabidopsis* codifica por pelo menos 1533 reguladores transcripcionais, que contam com ~5,9% de seu número total estimado de genes (Riechmann et al., Science 290, 2105-2109, 2000). O gene do milho GOLDEN2 (G2) é um representante do grupo de fatores de transcrição de GARP definidos por, além do GOLDEN2, o Regulador-B de Resposta de Receptor de *Arabidopsis* (ARR-B) e Psr1 de Chlamydomonas. Todas as proteínas GLK classificam-se como membros da família GARP. Os genes de *GLK* são duplicações monofiléticas e genéticas que ocorreram independentemente em plantas monocotiledôneas e dicotiledôneas. As proteínas GLK tipicamente compreendem o domínio de ligação de GARP DNA e um GOLDEN2 box de terminal C. As proteínas GLK de *Arabidopsis* atuam de maneira redundante na regulação do crescimento de cloroplasto (Fitter et al., Plant J. 31, 713-727, 2002). Além

disso, a função genética é conservada entre o musgo *Physcomitrella patens* e *Arabidopsis thaliana*, indicando que a regulação mediada por GLK de crescimento de cloroplasto é um mecanismo regulador antigo entre as plantas (Yasumura et al., Plant Cell 17, 1894-1907, 2005). No caso de G2, três das quatro características de definição da maioria dos fatores de transcrição foi verificada experimentalmente em sistemas heterólogos. O G2 é nuclear localizado (Hall et al., 1998), é capaz de transativar a expressão do gene repórter e pode tanto homo-dimerizar e heterodimerizar com ZmGLK1 (Rossini et al., 2001).

Domínio zíper de leucina de hemodomínio REV delta (Δ) (REV ΔHDZip/START)

A presente invenção diz respeito ao aumento do rendimento em plantas usando-se um tipo particular de fator de transcrição. Os polipeptídeos de fator de transcrição são usualmente definidos como proteínas que mostram afinidade de ligação de DNA específico de seqüência e que são capazes de ativar e/ou reprimir a transcrição. Os polipeptídeos *zíper* de leucina de homodomínio (HDZip) constitui uma família de fatores de transcrição caracterizados pela presença de um domínio de ligação de DNA (homeodomínio, HD) e um motivo de *zíper* de leucina adjacente (Zip). O homeodomínio usualmente consiste de aproximadamente 60 resíduos de aminoácido conservado que formam um helix1-loop-helix2-turn-helix3 que liga o DNA. Este local de ligação de DNA é usualmente pseudopalindrômico. O *zíper* de leucina, adjacente à extremidade de terminal C do homeodomínio (em alguns exemplos, sobrepondo-se por alguns aminoácidos), consiste de diversas repetições heptado de aminoácido (pelo menos quatro) em que, usualmente uma leucina (ocasionalmente uma valina ou um isoleucina) aparece a cada setenta aminoácidos. O *zíper* de leucina é importante para a dimerização de proteína. Esta dimerização é um pré-requisito para a ligação de DNA (Sessa *et al.* (1993) EMBO J 12(9): 3507-3517) e pode proceder

entre dois polipeptídeos HDZip idênticos (homodímero) ou entre dois polipeptídeos HDZip diferentes (heterodímero).

Os genes que codificam homeodomínio estão presentes em todos os eucariotas e constituem uma família de gene de pelo menos 89 membros em *Arabidopsis thaliana*. O *zíper* de leucina também é observado por si só em polipeptídeos de eucariotas que não outras plantas. Entretanto, a presença simultânea tanto de um homeodomínio quanto um *zíper* de leucina compreendido dentro do mesmo polipeptídeo é específico de planta (observado em pelo menos 47 fora dos 89 membros em *Arabidopsis*) e foi encontrado em musgo além das plantas vasculares (Sakakibara *et al.* (2001) Mol Biol Evol 18(4): 491-502).

Os polipeptídeos HDZip de *Arabidopsis* foram classificados em quatro classes diferentes, HDZip I a IV, com base em critérios de similaridade de seqüência (Sessa *et al.* (1994) *In*: Puigdomene P, Coruzzi G (ed), Springer, Berlin Heidelberg New York, pp 411–426). Em *Arabidopsis thaliana,* existem pelo menos cinco polipeptídeos de HDZip de classe III (*REVOLUTA* (REV/IFL), *PHABULOSA* (PHB), *PHAVOLUTA* (PHV), *CORONA* (CNA/AtHB15) e AtHB8), todos tipicamente muito grandes (mais do que 800 aminoácidos). Similarmente, em *Oryza sativa*, pelo menos 5 polipeptídeos HDZip classe III foram identificados.

Além do homeodomínio e o *zíper* de leucina, polipeptídeos HDZip de classe III também compreende o terminal C ao *zíper* de leucina um domínio de START (Transferência de lipídeo relacionado com (STAR) regulador agudo esteroidogênico) para a ligação de lipídeo/esterol e uma região de terminal C extensivo (CTR, mais do que metade do comprimento de polipeptídeo) de função desconhecida (Schrick *et al.* (2004) Genome Biology 5: R41). Além disso, um local complementar para microRNA (MIR165/166) é encontrado dentro da região de transcrição que codifica quanto ao domínio START de transcrições de mRNA que codificam quanto a polipeptídeos

HDZip de classe III, para a regulação por intermédio da clivagem de transcrição mediada por miRNA (Williams *et al.* (2005) Development 132: 3657-3668).

Em *Arabidopsis thaliana*, os polipeptídeos REV, PHB e PHV foram mostrados estar envolvidos na formação e função de broto ápico e meristamas axilares, padronização de estruturas tri-dimensionais (tais como embriões ou folhas) e crescimento vascular (diferenciação de cundutor lignificado e tecidos de suporte). Em *Arabidopsis thaliana*, a análise filogenética revela que estes três polipeptídeos HDZip de classe III intimamente relacionado forma o clado REV, os dois polipeptídeos HDZip de classe III remanescentes (CNA e AtHB8) que pertence ao clado CNA (Floyd *et al.* (2006) Genetics 173: 373-388). Quando combina-se arroz e genes de *Arabidopsis* em uma análise filogenética, dois grupos de polipeptídeos HDZip de classe III de arroz com polipeptídeo REV, os polipeptídeos OsHox10 e OsREV.

Os mutantes perda-de-função *phb, phv, cna* e *athb8* *Arabidopsis thaliana* são afenotípicos (Baima *et al.* (2001) Plant Physiol 126: 643 -655; Prigge *et al.* (2005) Plant Cell 17: 61 -76), mas mutantes *rev* formam meristemas laterais e florais defeituosos e desenvolve vasculatura de tronco aberrante bem como folhas onduladas (enrolada) (Otsuga *et al.* (2001) Plant J. 25,223 -236; Talbert *et al.* (1995) Development 121: 2723-2735).

Os alelos dominantes *rolled leaf1* (*rld1*) em milho (Juarez *et al.* (2004) Nature 428: 84-88), *phb-d* e *phv-d* em *Arabidopsis* (McConnell *et al.* (2001) Nature 411: 709-713), e *rev-d* em *Arabidopsis* (Emery *et al.* (2003) Curr Biol 13: 1768-1774) fenotipos mutantes similares presentes (folhas enroladas), devido a uma substituição de nucleotídeo na junção da seqüência do local de ligação MIR165/166 (no domínio START).

Na patente US US7056739 concedida (e pedido de patente internacional correspondente WO01/33944), são plantas e células de planta

descritas transformadas com um transgene que compreende uma seqüência de ácido nucleico REV de *Arabidopsis thaliana* que codifica um polipeptídeo REV representado por SEQ ID N° 2 (do padrão concedido). O transgene ainda é relatado compreender um promotor heterólogo operacionalmente ligado a uma seqüência de ácido nucleico que codifica o polipeptídeo de SEQ ID N°: 2. As plantas de *Arabidopsis* transgênicas que super-expressam a seqüência de ácido nucleico REV usando-se o promotor CaMV, apresentou tamanho de folha, tronco e semente aumentado. As seqüências genômicas de HDZip classe III parcial e terminal de cDNA 5' de tomate, arroz, milho e cevada são fornecidas. Os exemplos de vetores projetados para reduzir a expressão de *Arabidopsis* endógena, polipeptídeos HDZip de classe III de arroz e milho em arroz e milho respectivamente são descritos, para reproduzir o fenótipo de perda-de-função *rev*.

No pedido de patente internacional WO2004/063379, são fornecidas seqüências de ácido nucleico de dois polipeptídeos HDZip de classe III de milho ortólogo ao polipeptídeo REV de *Arabidopsis*. Um método para modular o nível de polipeptídeos HDZip de classe III pela inibição da expressão das transcrições de mRNA na célula de planta é descrito.

Em diversos pedidos internacionais e de patente US depositados por Mendel Biotechnology, Inc., o polipetídeo VER de *Arabidopsis* tem uma referência interna G438. A atividade de REVOLUTA reduzida após a inserção de T-DNA no local de REV locus resultou em plantas de *Arabidopsis* transgênicas com ramificação reduzida e teor de lignina reduzido, visto que a atividade de REVOLUTA aumentada (usando-se o promotor CaMV viral) resultou em plantas de *Arabidopsis* transgênicas das quais, em torno de metade, desenvolveram folhas mais planas levemente maiores do que as plantas do tipo selvagem em estágios posteriores.

Polipeptídeo semelhante ao CLE

Embora a comunicação de célula para célula em plantas ocorra

na maior parte através dos fito-hormônios, tais como auxina, citocinina, ácido abscísico, brassinostróides, ácido giberélico, etileno, hormônios de peptídeo são agora reconhecidos como mediadores importantes de eventos de sinalização. O grupo de hormônios de peptídeo incluem, por exemplo, sistemina, fitosulfocinas, peptídeos ENOD40, RALF, CLAVATA3, SCR e POLARIS.

Os polipetídeos semelhantes a CLE formam uma família de polipeptídeos que abrangem e dividem homologia com a Arabidopsis CLAVATA3 e polipeptídeos ESR do milho. Estes polipeptídeos são postulados serem desenvolvidos como ligandos em eventos de sinalização.

As raízes e as partes aéreas de uma planta são derivadas respectivamente da atividade do meristema ápico da raiz (RAM) e o meristema ápico da raiz (SAM). Estas estruturas contém células pluripotentes que permitem a produção de todos os tipos de células de planta e órgãos. Dentro do SAM, existe um equilíbrio entre a divisão das células tronco na zona central e nas células de diferenciação na zona periférica, embora a célula na zona central divida-se mais lentamente do que na zona periférica. CLAVATA3 (CLV3) é um membro da família do gene de ligando CLV3/ESR (CLE) e é secretado pelas células tronco do SAM desse modo ativando o dímero do receptor CLAVATA1-CLAVATA2 e levando à expressão restrita do gene WUSCHEL (WUS). O WUS é um fator de transcrição de homeodomínio e promove a formação das células tronco; isto é requerido para manter a população de célula tronco. O CLV3 por outro lado é requerido para evitar a proliferação não controlada de células tronco. O CLV3 e o WUS desta maneira formam um arco de regeneração que controla o número de células tronco e a organização do. Os mutantes *wus* falham em desenvolver um meristema ápico, visto que os mutantes *clv3* desenvolvem um meristema ápico de raiz grandemente aumentado.

Um outro membro da família do gene do ligando CLE é o

CLE2, que pode funcionar como o CLV3 quando a molécula sinalizadora secretada atua em diversos caminhos durante o crescimento e o crescimento. O CLE2 e outros membros da família CLE foram primeiro caracterizados por Cock and McCormick (Plant Physiol. 126, 939-942, 2001). Todos os membros da família CLE são polipeptídeos curtos (em torno de 7 a 9 kDa) com seqüência de terminal N hidrofóbico (peptídeos sinalizadores postulados ou âncoras de sinal). A maioria dos polipeptídeos maduros preditos é altamente básica (média pI 9,49 ± 1,57) e hidrofílico por todo o seu comprimento com uma região conservada em ou próximo da extremidade de terminal C. Esta região conservada pode estar envolvida nas interações de proteína-proteína. É postulado que os membros da família CLE, tais como CLV3 e CLE2, são processados por uma protease em um peptídeo curto que é secretado. Embora as proteínas dividam uma semelhança total no comprimento, carga e hidrofilicidade no nível da seqüência de aminoácido estes são altamente divergentes. O CLE2 foi relatado ser induzido pela adição de NO_3^- (Scheible et al., Plant Physiol. 136, 2483-2499, 2004). Além disso, a caracterização funcional (Strabala et al., Plant Physiol. 140, 1331-1344, 2006) revelou que a superexpressão de *CLE2* em *Arabidopsis* resultou em plantas impedidas de crescer com um atraso forte no período de florescimento (aproximadamente 40 dias vs 20 dias para plantas de controle), mas com raízes mais longas em comparação com os controles. A superexpressão de CLE2 também imitou um fenotipo *wus*. Um fenotipo similar como para o CLE2 foi observado para CLE3, CLE5, CLE6 e CLE7. Estas proteínas também são estruturalmente relacionadas. Entretanto, nenhuma informação fenotípica está disponível até hoje em mutantes *cle2* ou na sub-regulação de *CLE2*.

O WO 01/96582 divulga o uso de proteínas semelhantes ao ligando (LLPs) ou fragmentos funcionais destes para a modulação do fenotipo ou arquitetura da planta. Os LLPs preferidos ou fragmentos compreendem o

motivo de aminoácido XRXXXXGXXXXHX (em que X pode ser qualquer aminoácido), os LLPs mais preferidos de fragmentos compreendem o motivo de aminoácido KRXXXXGXXPXHX. O documento também descreve que a expressão ectópica de vários LLPs resulta em plantas transgênicas estéreis ou, na melhor das hipóteses, plantas com fertilidade reduzida. Também, a expressão anti-sentido de um LLP resultou em uma planta transgênica com fertilidade reduzida (número grandemente reduzido de semente por síliqua).

O WO 03/093450 divulga o uso de peptídeos semelhantes a CLAVATA3 para modular a divisão, diferenciação ou crescimento celular em plantas, em particular para a modulação do crescimento de meristema. Foi postulado que a atividade diminuída do peptídeo semelhante ao CLV3 deve resultar na geração de folhas adicionais antes do florescimento começar, desse modo, fornecendo plantas tendo produção de energia maior e, desta maneira, aumentando o rendimento ou em um número aumentado de carpelos que carregam semente ou na geração de caule mais espesso ou em uma alteração do fruto das plantas. Entretanto, apenas exemplos hipotéticos foram fornecidos com respeito à sub-regulação da expressão de CLV3, nenhum dado experimental real foi dado. Além disso, os peptídeos semelhantes a CLV3 divulgados estão dentro da classe de LLPs descrita no WO 01/96582, visto que estes compreendem o motivo XRXXXXGXXXXHX, para os quais foi mostrado que a expressão sub-regulada resultou em plantas que reduziram a fertilidade.

Similarmente, quando uma proteína semelhante ao CLE do nematódeo parasítico da planta *Heterodera glycines* foi super-expressada em *Arabidopsis*, as plantas transgênicas produziram flores que não abriram ou perderam o ginoécio central e o sistema de raízes foi impedido.

O problema, desta maneira permanece como os polipeptídeos semelhantes a CLE, tais como LLPs ou semelhantes a CLV3 podem ser usados para melhorar as características relacionadas com o rendimento.

Regulador do Rendimento Semente

O SYR é uma nova proteína que, até agora, não foi caracterizado. O SYR mostra alguma homologia (em torno de 48% de identidade de seqüência no nível do DNA, em torno de 45% no nível de proteína) a uma proteína de Arabidopsis denominada ARGOS (Hu et al., Plant Cell 15, 1951-1961, 2003; US 2005/0108793). Hu et al. Postularam que ARGOS é uma proteína de função única e é codificada por um gene simples. Os fenotipos principais de super-expressão de ARGOS em *Arabidopsis* são biomassa folhosa aumentada e florescimento atrasado. Em contraste, a super-expressão de SYR no arroz, primariamente aumenta o rendimento de semente, visto que a biomassa folhosa e o tempo de florescimento não são obviamente afetados.

Sumário da Invenção

Surpreendentemente, foi observado que a expressão da modulação em uma planta de um ácido nucleico codifica uma quinase semelhante ao receptor de extensina (ERLK) ou uma parte deste que compreende pelo menos o domínio de quinase e o domínio transmembrana, dá plantas tendo tratos relacionados com o crescimento intensificado com relação às plantas de controle.

Portanto, a invenção fornece um método para intensificar as características relacionadas com o rendimento em plantas com relação às plantas de controle, que compreendem modular a expressão em uma planta de um ácido nucleico que codifica uma proteína ERLK ou uma parte da mesma.

Surpreendentemente, foi observado agora que o aumento da expressão em uma planta de um ácido nucleico que codifica um polipeptídeo FBXW dá plantas tendo rendimento aumentado com relação às plantas de controle adequadas. Portanto, de acordo com uma outra forma de realização da invenção é fornecido um método para aumentar o rendimento em plantas com relação às plantas de controle adequadas, que compreende aumentar a

expressão em uma planta de um ácido nucleico que codifica um polipeptídeo FBXW.

Na investigação do uso de proteínas e ligação de RAN para intensificar as características relacionadas com o rendimento, os inventores observaram a escolha do promotor ser uma consideração importante. Estes observaram que a expressão das proteínas de ligação de RAN em uma planta (arroz) sob o controle de um promotor constituído não teve qualquer efeito nos fenotipos relacionados como o rendimento. Estes, surpreendentemente, observaram que a planta pode ser, de maneira bem sucedida, melhorada pela expressão de proteínas de ligação de RAN em uma planta sob o controle de um promotor específico de semente, particularmente um promotor específico de endosperma.

A presente invenção, portanto, também fornece um método para intensificar as características relacionadas com o rendimento em plantas com relação às plantas de controle, que compreende, preferivelmente modular a expressão em uma semente vegetal ou partes de sementes de um ácido nucleico que codifica um RANBP.

Além disso, ainda foi observado que a modulação da expressão em uma planta de um ácido nucleico que codifica uma proteína semelhante a Golden2 (GLK) dá plantas tendo características relacionadas com o rendimento intensificado com relação às plantas de controle.

Portanto, ainda em uma outra forma de realização da invenção é fornecido um método para intensificar as características relacionadas com o rendimento com relação às plantas de controle, que compreende modular a expressão em uma planta de um ácido nucleico que codifica uma proteína de GLK ou uma parte da mesma.

Também observamos surpreendentemente que a redução da expressão em um planta de um gene REV endógeno usando-se um domínio de *zíper* de leucina de homeodomínio REV delta (HDZip)/seqüência de ácido nucleico (START) domínio de transferência de lipídeo relacionado com o

regulador agudo esteroidogênico (STAR) dá plantas tendo rendimento aumentado com relação às plantas de controle. A presente invenção portanto fornece, em uma outra forma de realização, métodos para aumentar o rendimento em plantas com relação às plantas de controle, reduzindo a expressão em uma planta de um gene REV endógeno usando-se uma seqüência de ácido nucleico REV ΔHDZip/START.

Também foi observado que modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante ao CLE dá plantas tendo características relacionadas com o rendimento intensificadas com relação às plantas de controle.

Portanto, a invenção fornece em uma forma de realização adicional, um método para intensificar as características relacionadas com o rendimento em plantas com relação às plantas de controle, que compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo semelhante ao CLE ou uma parte da mesma.

Surpreendentemente, foi observado agora que modular a expressão em uma planta de um ácido nucleico que codifica uma proteína Reguladora de Rendimento de Semente (a seguir denominado SYR) dá plantas, quando desenvolve-se sob condições de tensão abióticas, tendo tolerância à tensão abiótica intensificada com relação às plantas de controle.

Portanto, a presente invenção fornece um método para intensificar características relacionadas com o rendimento em plantas sob condições de tensão abiótica, com relação às plantas de controle, que compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo SYR.

Definições

Polipeptídeo(s)/Proteína(s)

Os termos "polipeptídeo" e "proteína" são usados de maneira intercambiável e referem-se a aminoácidos em uma forma polimérica de

qualquer comprimento, ligados juntos por ligações de peptídeo.

Polinucleotídeo(s) / Ácido(s) nucleico(s) /Seqüência(s) de ácido nucleico/ Seqüência(s) de nucleotídeo

Os termos "polinucleotídeos", "seqüências de ácido nucleico ", "Seqüências de nucleotídeo", "ácido nucleicos" "molécula de ácido nucleico" são usados de maneira intercambiável neste e refere-se aos nucleotídeos, ribonucleotídeos ou desoxiribonucleotídeos ou uma combinação de ambos, em uma forma não ramificada polimérica de qualquer comprimento.

Planta(s) de controle

A escolha de plantas de controle adequadas é uma parte de rotina de um ajuste experimental e pode incluir plantas do tipo selvagem correspondentes ou plantas correspondentes sem o gene de interesse. A planta de controle é tipicamente da mesma espécie vegetal ou ainda da mesma variedade como a planta a ser estimada. A planta de controle também pode ser um nulizigoto da planta a ser estimada. Os nulizigotos são indivíduos que não compreendem o transgene por segregação. Uma "planta de controle" como usado neste refere-se não apenas às plantas totais, mas também a partes de plantas, incluindo semente e partes de sementes.

Homólogo(s)

Os "homólogos" de uma proteína abrange peptídeos, oligopeptídeos, polipeptídeos, proteínas e enzimas tendo substituições, anulações e/ou inserções de aminoácido com relação à proteína não modificada em questão e tendo atividade biológica e funcional similar como a proteína não modificada da qual estes são derivados.

A anulação refere-se à remoção de um ou mais aminoácidos de uma proteína.

Uma inserção refere-se a um ou mais resíduos de aminoácido sendo introduzidos em um local predeterminado em uma proteína. As inserções podem compreender fusões de terminal N e/ou de terminal C bem

como inserções intra-seqüência de aminoácidos simples ou múltiplos. No geral, inserções dentro da seqüência de aminoácido serão menores do que as fusões de terminal N ou C, da ordem de cerca de 1 a 10 resíduos. Os exemplos de proteínas de fusão de terminal N- ou C ou peptídeos incluem o domínio de ligação ou o domínio de ativação de um ativador de transcrição como usado no sistema de dois híbridos de levedura, proteínas de revestimento de fago, (histidina)-6-tag, glutationa S-transferase-tag, proteína A, proteína de ligação de maltose, diidrofolato redutase, epítopo Tag•100, epítopo c-myc, FLAG®-epítopo, lacZ, CMP (peptídeo de ligação de calmodulina), epítopo de HA, epítopo de proteína C e epítopo VSV.

Uma substituição refere-se à substituição de aminoácidos da proteína com outros aminoácidos tendo propriedades similares (tal como hidrofobicidade, hidrofilicidade, antigenicidade, propensão similar para a formação ou quebra de estruturas α-hélicas ou estruturas β-chapa). As substituições de aminoácido são tipicamente de resíduos simples, ms podem ser agrupados dependendo das restrições funcionais no polipeptídeo; as inserções são, usualmente da ordem de cerca de 1 a 10 resíduos de aminoácidos. A substituições de aminoácido são preferivelmente substituições de aminoácido conservativas. As tabelas de substituição conservativa são bem conhecidas na técnica (ver, por exemplo, Creighton (1984) Proteins. W.H. Freeman and Company (Eds) e Tabela 1 abaixo).

Tabela 1: Exemplos de substituições de aminoácido conservados

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

As substituições, anulações e/ou inserções de aminoácido podem ser feitas facilmente usando-se técnicas sintéticas de peptídeo bem conhecidas na técnica, tais como síntese de peptídeo de fase sólida e outros ou pela manipulação do DNA recombinante. Os métodos para a manipulação de seqüências de DNA para a produção de variações de substituição, inserção ou anulação de uma proteína são bem conhecidos na técnica. Por exemplo, as técnicas para realizar mutações de substituição em locais predeterminados no DNA são bem conhecidos por aqueles habilitados na técnica e incluem a mutagênese de M13, mutagênese in vitro de T7-Gen (USB, Cleveland, OH), Mutagênese direcionada ao local QuickChange (Stratagene, San Diego, CA), mutagênese direcionada ao local mediada por PCR ou outros protocolos de mutagênese direcionada ao local.

Derivados

"Derivados" incluem peptídeos, oligopeptídeos, polipeptídeos que podem, em comparação com a seqüência de aminoácido da forma de ocorrência natural da proteína, tal como a proteína de interesse, compreende substituições de aminoácidos com resíduos de aminoácido que não ocorrem naturalmente ou adições de resíduos de aminoácido de ocorrência não natural. Os "derivados" de uma proteína também abrangem peptídeos, oligopeptídeos, polipeptídeos que compreendem resíduos de aminoácido alterados de ocorrência natural (glicosilados, acilados, prenilados, fosforilados, miristoilados, sulfatados etc.) ou alterados de ocorrência não natural em comparação com a seqüência de aminoácido de uma forma de ocorrência natural do polipeptídeo. Um derivado também pode compreender um ou mais substituintes que não aminoácido ou adições em comparação com a seqüência de aminoácido da qual este é derivado, por exemplo, uma molécula repórter ou outro ligando, covalentemente ou não covalentemente ligado para a seqüência de aminoácido, tal como uma molécula repórter que é ligada para facilitar sua detecção e resíduos de aminoácido que não ocorrem naturalmente

com relação à seqüência de aminoácido de uma proteína de ocorrência natural. Além disso, "derivados" também incluem fusões da forma de ocorrência natural da proteína com peptídeos de rotulação, tais como FLAG, HIS6 ou tioredoxina (para uma revisão de peptídeos de rotulação, ver Terpe, Appl. Microbiol. Biotechnol. 60, 523-533, 2003).

Ortólogo(s)/Parálogo(s)

Ortólogos e parálogos abrangem conceitos evolucionários usados para descrever as relações ancestrais de genes. Os parálogos são genes dentro das mesmas espécies que se originaram através da duplicação de um gene ancestral; os ortólogos são genes de organismos diferentes que se originaram através da evolução e também são derivados de um gene ancestral comum.

Domínio

O termo "domínio" refere-se a um grupo de aminoácidos conservados em posições específicas junto com um alinhamento de seqüências de proteínas relacionadas com a evolução. Enquanto os aminoácidos em outras posições podem variar entre os homólogos, os aminoácidos que são altamente conservados em posições específicas indicam aminoácidos que também são essenciais na estrutura, estabilidade ou função de uma proteína. Identificado por seu alto grau de conservação em seqüências alinhadas de uma família de homólogos de proteína, estes podem ser usados como identificadores para determinar se qualquer polipeptídeo em questão pertence a uma família de polipeptídeo previamente identificada.

Motivo/Seqüência de Consenso/Assinatura

O termo "motivo" ou "seqüência de consenso" ou "assinatura" refere-se a uma região conservada curta na seqüência de proteínas evolucionariamente relacionadas. Os motivos são, freqüentemente, partes altamente conservadas de domínios, mas também incluem apenas parte do domínio ou estão localizados fora do domínio conservado (se todos os

aminoácidos dos motivos estiverem fora de um domínio definido).

Hibridização

O termo "hibridização" como definido neste é um processo em que seqüências de nucleotídeo complementares substancialmente homólogas fortalecem-se uma com a outra. O processo de hibridização pode ocorrer totalmente na solução, isto é, ambos os ácidos nucleicos complementares estão na solução. O processo de hibridização também pode ocorrer com um dos ácidos nucleicos complementares imobilizados em uma matriz, tais como contas magnéticas, contas de Sepharose ou qualquer outra resina. O processo de hibridização também pode ocorrer com um dos ácidos nucleicos complementares imobilizados em um suporte sólido, tal como uma membrana de nitro-celulose ou náilon ou imobilizado, por exemplo, por fotolitografia, por exemplo, em um suporte de vidro silicoso (o último conhecido como séries ou microsséries de ácido nucleico ou como chips de ácido nucleico). A fim de permitir que a hibridização ocorra, as moléculas de ácido nucleico são, no geral, térmica ou quimicamente desnaturadas para fundir um filamento duplo m dois filamentos simples e/ou para remover estruturas de grampo de cabelo ou outras secundárias dos ácidos nucleicos de filamento simples.

O termo "estringência" refere-se às condições sob as quais uma hibridização acontece. A estringência de hibridização é influenciada pelas condições, tais como temperatura, concentração de sal, força iônica e composição do tampão de hibridização. No geral, as condições de baixa estringência são selecionadas estarem em torno de 30° C menor do que o ponto de fusão térmico (T_m) para a seqüência específica em uma força iônica e pH definidos. As condições de estringência médias ocorre quando a temperatura está 20° C abaixo do T_m e as condições de alta estringência ocorrem quando a temperatura está 10° C abaixo do T_m. as condições de hibridização de estringência alta são tipicamente usadas para isolar seqüências de hibridização que têm similaridade de seqüência alta à seqüência de ácido

nucleico alvo. Entretanto, os ácidos nucleicos podem desviar na seqüência e ainda codificar um polipeptídeo substancialmente idêntico, devido à degeneração do código genético. Portanto, as condições de hibridização de estringência média podem ser, algumas vezes necessárias para a identificação de tais moléculas de ácido nucleico.

O T_m é a temperatura sob força iônica e pH definidos, em que 50% da seqüência alvo hibridiza-se a uma sonda perfeitamente comparada. O T_m é dependente das condições de solução e da composição de base e do comprimento da sonda. Por exemplo, as seqüências mais longas hibridizam-se especificamente em temperaturas mais altas. A taxa máxima de hibridização é obtida de cerca de 16° C até 32° C abaixo do T_m. A presença de cátions monovalentes na solução de hibridização reduz a repulsão eletrostática entre os dois filamentos de ácido nucleico, desse modo, promovendo a formação híbrida; este efeito é visível para as concentrações de sódio de até 0,4 M (para concentrações mais altas, este efeito pode ser ignorado). A formamida reduz a temperatura de fusão de duplexes de DNA-DNA e DNA-RNA com 0,6 a 0,7° C para cada porcentagem de formamida e adição de 50% de formamida permite que a hibridização seja realizada de 30 a 45° C, embora a taxa de hibridização seja diminuída. Os desacordos de pares de base reduzem a taxa de hibridização e a estabilidade térmica dos duplexes. Em sondas na média e grandes, o Tm diminui cerca d 1° C por% de desacordo de base. O Tm pode ser calculado usando-se as seguintes equações, dependendo dos tipos de híbridos:

1) híbridos DNA-DNA (Meinkoth and Wahl, Anal. Biochem., 138: 267-284, 1984):

$T_m = 81{,}5° C + 16{,}6 \times \log_{10}[Na^+]^a + 0{,}41 \times \%[G/C^b] - 500 \times [L^c]^{-1} - 0{,}61 \times \%$ formamida

2) híbridos de DNA-RNA ou RNA-RNA:

$Tm = 79{,}8 + 18{,}5 (\log_{10}[Na^+]^a) + 0{,}58 (\%G/C^b) + 11{,}8 (\%G/C^b)^2 - 820/L^c$

3) híbridos de ligo-DNA ou oligo-RNA[d]:

para <20 nucleotídeos: $T_m = 2 (l_n)$

para 20 a 35 nucleotídeos: $T_m = 22 + 1,46 (l_n)$

[a] ou para outro cátion monovalente, mas apenas exato na faixa 0,01 a 0,4 M.

[b] apenas preciso para %GC na faixa de 30% a 75%.

[c] L = comprimento de duplex em pares de base.

[d] oligo, oligonucleotídeo; l_n, = comprimento eficaz de iniciador = 2×(no. de G/C)+(no. de A/T).

A ligação não específica pode ser controlada usando-se qualquer uma de diversas técnicas conhecidas, tais como, por exemplo, bloqueio da membrana com soluções contendo proteína, adições de RNA, DNA e SDS heterólogos ao tampão de hibridização e tratamento com Rnase. Para sondas não homólogas, uma série de hibridizações pode ser realizadas pela variação de (i) diminuição progressiva da temperatura de recozimento (por exemplo, de 68° C a 42° C) ou (ii) diminuindo progressivamente a concentração de formamida (por exemplo, de 50% a 0%). O técnico habilitado está ciente de vários parâmetros que podem ser alterados durante a hibridização e que manterá ou mudará as condições de estringência.

Além das condições de hibridização, a especificidade de hibridização tipicamente também depende da função de lavagens de pós-hibridização. Para remover a base que resulta da hibridização não específica, as amostras são lavadas com soluções salinas dilutas. Os fatores críticos de tais lavagens incluem a força iônica e a temperatura da solução de lavagem final: quanto menor a concentração de sal maior a temperatura de lavagem, maior a estringência da lavagem. As condições de lavagem são tipicamente realizadas em ou abaixo da estringência de hibridização. Uma hibridização positiva dá um sinal que é pelo menos duas vezes aquela da base. No geral, as condições estringentes adequadas para os ensaios de hibridização de ácido nucleico ou procedimentos de detecção de amplificação de gene são como

apresentado acima. Condições mais ou menos estringentes também podem ser selecionadas. O técnico habilitado está ciente de vários parâmetros que podem ser alterados durante a lavagem e que manterá ou mudará as condições de estringência.

Por exemplo, as condições de hibridização de estringência alta típicas para os híbridos de DNA mais longos do que 50 nucleotídeos abrange a hibridização a 65° C e, 1x SSC ou em 42° C em 1x SSC e 50% de formamida, seguido pela lavagem a 65° C em 0.3x SSC. Os exemplos de condições de hibridização de estringência média para os híbridos de DNA mais longos do que 50 nucleotídeos abrange a hibridização a 50° C em 4x SSC ou em 40° C em 6x SSC e 50% de formamida, seguido pela lavagem a 50° C em 2x SSC. O comprimento do híbrido é o comprimento antecipado para o ácido nucleico hibridizador. Quando os ácidos nucleicos de seqüência conhecida são hibridizados, o comprimento do híbrido pode ser determinado pelo alinhamento das seqüências e identificação das regiões conservadas descritas neste. 1×SSC é 0,15M de NaCl e 15 mM de citrato de sódio; a solução de hibridização e as soluções de lavagem podem incluir, adicionalmente 5x de reagente de Denhardt, 0,5 a 1,0% de SDS, 100 μg/ml de DNA de esperma de salmão fragmentado desnaturado, 0,5% de pirofosfato de sódio.

Para os propósitos de definição do nível de estringência, referência pode ser feita a Sambrook *et al.* (2001) Molecular Cloning: a laboratory manual, 3rd Edition, Cold Spring Harbor Laboratory Press, CSH, New York or a Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989 e atualizações anuais).

Variante de união

O termo "variante de união" como usado neste abrange variantes de uma seqüência de ácido nucleico em que os introns e/ou exons selecionados foram excisados, substituídos, deslocados ou adicionados ou em

que os introns foram encurtados ou alongados. Tais variantes serão umas em que a atividade biológica da proteína é substancialmente retida; isto pode ser atingido retendo-se seletivamente os segmentos funcionais da proteína. Tais variantes de união podem ser encontradas em estado natural ou podem ser feitas pelo homem. Os métodos para predizer e isolar tais variantes de união são bem conhecidos na técnica (ver, por exemplo, Foissac and Schiex (2005) BMC Bioinformatics 6: 25).

Variante alélica

Os alelos e as variantes alélicas são formas alternativas de um dado gene, localizado na mesma posição cromossômica. As variantes alélicas abrangem os polimorfismos de nucleotídeo simples (SNPs), bem como polimorfosmos de inserção/anulação pequenos (INDELs). O tamanho de INDELs é usualmente menor do que 100 bp. O SNPs e o INDELs formam o grupo maior de variantes de seqüência em cepas polimórficas de ocorrência natural da maioria dos organismos.

Mistura de gene/Evolução direcionada

A mistura de gene ou a evolução direcionada consiste de interações de mistura de DNA seguido pela avaliação e/ou seleção apropriada para a geração de variantes de ácidos nucleicos e porções destes que codificam proteínas tendo uma atividade biológica modificada (Castle et al., (2004) Science 304(5674): 1151-4; Patentes US 5.811.238 e 6.395.547).

Elemento Regulador/Seqüência de controle/Promotor

Os termos "elemento regulador", "seqüência de controle" e "promotor" são todos usados de maneira intercambiável e devem ser usados em um contexto amplo para a referência às seqüências de ácido nucleico reguladoras capazes de realizar a expressão das seqüências às quais estas estão ligadas. O termo "promotor" tipicamente refere-se a uma seqüência de controle de ácido nucleico localizada a montante do início transcripcional de um gene e que está envolvido no reconhecimento e ligação de RNA

polimerase e outras proteínas, desse modo direcionando a transcrição de um ácido nucleico operacionalmente ligado. Abrangidas pelos termos já mencionados estão as seqüências reguladoras transcripcionais derivadas de um gene genômico eucariótico clássico (incluindo TATA box que é requerido para a iniciação de transcrição precisa, com ou sem uma seqüência de CCAAT box) e elementos reguladores adicionais (isto é, seqüências ativadoras a montante, intensificadores e silenciadores) que alteram a expressão genética em resposta aos estímulos de crescimento e/ou externos ou de uma maneira específica do tecido. Também incluído dentro do termo está uma seqüência reguladora de um gene procariótico clássico, caso no qual pode incluir a seqüência –35 box e/ou as seqüências reguladoras transcripcionais –10 box. O termo "elemento regulador" também abrange uma molécula de fusão sintética ou derivado que confere, ativa ou intensifica a expressão de uma molécula de ácido nucleico em uma célula, tecido ou órgão.

Um "promotor vegetal" compreende elementos reguladores, que mediam a expressão de um segmento de seqüência codificadora em células de planta. Conseqüentemente, um promotor vegetal não necessita ser de origem vegetal, mas pode originar-se de tecidos ou micro-organismos, por exemplo, de vírus que atacam células de planta. O "promotor vegetal" também pode originar-se de uma célula de planta, por exemplo, da planta que é transformada com a seqüência de ácido nucleico a ser expressada no processo inventivo e descrito neste. isto também aplica-se a outros sinais reguladores "vegetais", tais como terminadores "vegetais". Os promotores a montante das seqüências de nucleotídeo úteis nos métodos da presente invenção podem ser modificados por uma ou mais substituições, inserções e/ou anulação de nucleotídeo sem interferir com a funcionalidade ou atividade dos promotores, a estrutura de leitura aberta (ORF) ou a região 3'-reguladora tal como terminadores ou outras regiões 3' reguladoras que estão localizadas além do ORF. Também é possível que a atividade dos promotores seja

aumentada pela modificação de sua seqüência ou que estes são substituídos completamente pro promotores mais ativos, mesmo promotores de organismos heterólogos. Para a expressão em plantas, a molécula de ácido nucleico deve, como descrito acima, ser ligada operacionalmente a ou compreender um promotor adequado que expressa o gene no ponto certo no tempo e com o padrão de expressão espacial requerido.

Para a identificação de promotores funcionalmente equivalentes, a força do promotor e/ou o padrão de expressão de um promotor candidato podem ser analisados, por exemplo, ligando-se operavelmente o promotor a um gene repórter e estimando-se o nível e o padrão de expressão do gene repórter em vários tecidos da planta. Os genes repórteres bem conhecidos adequados incluem, por exemplo, beta-glucuronidase ou beta-galactosidase. A atividade do promotor é estimada medindo-se a atividade enzimática da beta-glucuronidase ou beta-galactosidase. A força e/ou o padrão de expressão do promotor pode ser então comparado com aquele do promotor de referência (tal como o usado nos métodos da presente invenção). Alternativamente, a força do promotor pode ser estimada pela quantificação dos níveis de mRNA ou pela comparação dos níveis de mRNA do ácido nucleico usado nos métodos da presente invenção, com os níveis de mRNA de genes de manutenção, tais como 18S rRNA, usando-se os métodos conhecidos na técnica, tais como Northern blotting com análise densitométrica de autorradiogramas, PCR ou RT-PCR de tempo real quantitativo (Heid et al., 1996 Genome Methods 6: 986-994). No geral, por "promotor fraco" é entendido um promotor fraco que conduz a expressão de uma seqüência codificadora em um nível baixo. Por "nível baixo" é entendido de níveis de 1/10.000 transcrições a cerca de 1/100.000 transcrições, a cerca de 1/500.0000 transcrições por célula. De maneira oposta, um "promotor forte" conduz a expressão de uma seqüência codificadora em nível alto ou de cerca de 1/10 transcrições a cerca de 1/100 transcrições a cerca de 1/1000

transcrições por célula.

Operacionalmente ligado

O termo "operacionalmente ligado" como usado neste refere-se a uma ligação funcional entre a seqüência de promotor e o gene de interesse, tal que a seqüência de promotor seja capaz de iniciar a transcrição do gene de interesse.

Promotor Constitutivo

Um "promotor constitutivo" refere-se a um promotor que é transcripcionalmente ativo durante a maioria, mas não necessariamente todas as fases de crescimento e crescimento e sob a maioria das condições ambientais, em pelo menos uma célula, tecido ou órgão. A tabela 2a abaixo dá exemplos de promotores constitutivos.

Tabela 2a: Exemplos de promotores constitutivos

Fonte do gene	Referência
Actin	McElroy et al, Plant Cell, 2: 163-171, 1990
HMGB	WO 2004/070039
CAMV 35S	Odell et al, Nature, 313: 810-812, 1985
CaMV 19S	Nilsson et al., Physiol. Plant. 100:456-462, 1997
GOS2	de Pater et al, Plant J Nov;2(6):837-44, 1992, WO 2004/065596
Ubiquitina	Christensen et al, Plant Mol. Biol. 18: 675-689, 1992
Ciclofilina do arroz	Buchholz et al, Plant Mol Biol. 25(5): 837-43, 1994
H3 histona do milho	Lepetit et al, Mol. Gen. Genet. 231:276-285, 1992
H3 histona da Alfafa	Wu et al. Plant Mol. Biol. 11:641-649, 1988
Actin 2	An et al, Plant J. 10(1); 107-121, 1996
34S FMV	Sanger et al., Plant. Mol. Biol., 14, 1990: 433-443
Subunidade pequena de Rubisco	US 4,962,028
OCS	Leisner (1988) Proc Natl Acad Sci USA 85(5): 2553
SAD1	Jain et al., Crop Science, 39 (6), 1999: 1696
SAD2	Jain et al., Crop Science, 39 (6), 1999: 1696
Nos	Shaw et al. (1984) Nucleic Acids Res. 12(20):7831-7846
V-ATPase	WO 01/14572
Super promotor	WO 95/14098
Proteínas G-box	WO 94/12015

Promotor Ubiquitoso promoter

Um promotor ubiquitoso é ativo substancialmente em todos os tecidos ou células de um organismo.

Promotor desenvolventemente regulado

Um promotor desenvolventemente regulado é ativo durante certos estágios de crescimento ou em partes da planta que sofre mudanças de crescimento.

Promotor indutível

Um promotor indutível induziu ou aumentou a iniciação da transcrição em resposta a um produto químico ambiental (para uma revisão ver Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48:89-108), ou estímulo físico ou pode ser "indutível por tensão", isto é, ativado quando uma planta é exposta a várias condições de tensão ou um "indutível por patogênio" isto é, ativado quando uma planta é exposta à exposição à vário patogênios.

Promotor específico de órgão/específico de tecido

Um promotor específico orgânico ou específico de tecido é um que é capaz de iniciar preferencialmente a transcrição em certos órgãos ou tecidos, tais como as folhas, raízes, tecido de semente etc. Por exemplo, um "promotor específico de raiz" é um promotor que é transcripcionalmente ativo, predominantemente em raízes de plantas, substancialmente para a exclusão de quaisquer outras partes de uma planta, enquanto ainda permite-se qualquer expressão ruim nestas outras partes de plantas. Os promotores capazes de iniciar a transcrição e certas células apenas são referidos aqui como "específico de célula".

Um promotor específico de semente é transcripcionalmente ativo, predominantemente no tecido da semente, mas não necessariamente exclusivo no tecido da semente (em casos de expressão ruim). O promotor específico de semente pode estar ativo durante o crescimento da semente e/ou durante a germinação. Alguns promotores específicos de semente podem ser específicos para o endosperma, aleurona e/ou embrião. Os exemplos de promotores específicos de semente são mostrados na Tabela 2b abaixo e de promotores específicos de endosperma na Tabela 2c. os exemplos adicionais

de promotores específicos de semente são dados em Qing Qu e Takaiwa (Plant Biotechnol. J. 2, 113-125, 2004), cuja divulgação é incorporada por referência neste como se totalmente apresentado.

Tabela 2b: Exemplos de promotores específicos de semente

Fonte de Gene	Referência
Genes específicos de semente	Simon et al., Plant Mol. Biol. 5: 191, 1985;
	Scofield et al., J. Biol. Chem. 262: 12202, 1987.;
	Baszczynski et al., Plant Mol. Biol. 14: 633, 1990.
Albumina de Brazil Nut	Pearson et al., Plant Mol. Biol. 18: 235-245, 1992.
Legumina	Ellis et al., Plant Mol. Biol. 10: 203-214, 1988.
glutelina (arroz)	Takaiwa et al., Mol. Gen. Genet. 208: 15-22, 1986;
	Takaiwa et al., FEBS Letts. 221: 43-47, 1987.
Zeina	Matzke et al Plant Mol Biol, 14(3):323-32 1990
NapA	Stalberg et al, Planta 199: 515-519, 1996.
Glutelina-1 de trigo LMW e HMW	Mol Gen Genet 216:81-90, 1989; NAR 17:461-2, 1989
Trigo SPA	Albani et al, Plant Cell, 9: 171-184, 1997
trigo α, β, γ-gliadinas	EMBO J. 3:1409-15, 1984
Promotor de cevada Itr1	Diaz et al. (1995) Mol Gen Genet 248(5):592-8
Cevada B1, C, D, hordeina	Theor Appl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750-60, 1996
Cevada DOF	Mena et al, The Plant Journal, 116(1): 53-62, 1998
blz2	EP99106056.7
Promotor sintético	Vicente-Carbajosa et al., Plant J. 13: 629-640, 1998.
prolamina NRP33 de arroz	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
α-globulina Glb-1 de arroz	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
Arroz OSH1	Sato et al, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996
α-globulina REB/OHP-1 de arroz	Nakase et al. Plant Mol. Biol. 33: 513-522, 1997
ADP-glicose pirofos-forilase de arroz	Trans Res 6:157-68, 1997
Família do gene ESR do milho	Plant J 12:235-46, 1997
α-kafirina de sorgo	DeRose et al., Plant Mol. Biol 32:1029-35, 1996
KNOX	Postma-Haarsma et al, Plant Mol. Biol. 39:257-71, 1999
Oleosina de arroz	Wu et al, J. Biochem. 123:386, 1998
Oleosina de girassol	Cummins et al., Plant Mol. Biol. 19: 873-876, 1992
PRO0117, proteína ribossômica 40S de arroz putativo	WO 2004/070039
PRO0136, alanina amino-transferase de	não publicado

arroz	
PRO0147, inibidor de tripsina ITR1 (cevada)	não publicado
PRO0151, arroz WSI18	WO 2004/070039
PRO0175, arroz RAB21	WO 2004/070039
PRO005	WO 2004/070039
PRO0095	WO 2004/070039
α-amilase (Amy32b)	Lanahan et al, Plant Cell 4:203-211, 1992; Skriver et al, Proc Natl Acad Sci USA 88:7266-7270, 1991
Gene semelhante à catepsina β	Cejudo et al, Plant Mol Biol 20:849-856, 1992.
Cevada Ltp2	Kalla et al., Plant J. 6:849-60, 1994
Chi26	Leah et al., Plant J. 4:579-89, 1994
milho B-Peru	Selinger et al., Genetics 149;1125-38,1998

Tabela 2c: Exemplos de promotores específicos de endosperma

Fonte de gene	**Referência**
Glutelina (arroz)	Takaiwa et al. (1986) Mol Gen Genet 208:15-22; Takaiwa et al. (1987) FEBS Letts. 221:43-47
Zeina	Matzke et al., (1990) Plant Mol Biol 14(3): 323-32
Glutelina de trigo LMW e HMW	Colot et al. (1989) Mol Gen Genet 216:81-90, Anderson et al. (1989) NAR 17:461-2
Trigo SPA	Albani et al. (1997) Plant Cell 9:171-184
gliadinas do trigo	Rafalski et al. (1984) EMBO 3:1409-15
Promotor Itr1 da cevada	Diaz et al. (1995) Mol Gen Genet 248(5):592-8
Hordeína B1, C, D, da cevada	Cho et al. (1999) Theor Appl Genet 98:1253-62; Muller et al. (1993) Plant J 4:343-55; Sorenson et al. (1996) Mol Gen Genet 250:750-60
Cevada DOF	Mena et al, (1998) Plant J 116(1): 53-62
blz2	Onate et al. (1999) J Biol Chem 274(14):9175-82
Promotor sintético	Vicente-Carbajosa et al. (1998) Plant J 13:629-640
Prolamina NRP33 do arroz	Wu et al, (1998) Plant Cell Physiol 39(8) 885-889
globulina Glb-1 do arroz	Wu et al. (1998) Plant Cell Physiol 39(8) 885-889
globulina REB/OHP-1 do arroz	Nakase et al. (1997) Plant Molec Biol 33: 513-522
ADP-glicose pirofosforilase do arroz	Russell et al. (1997) Trans Res 6:157-68
Família do gene ESR do milho	Opsahl-Ferstad et al. (1997) Plant J 12:235-46
Cafirina do sorgo	DeRose et al. (1996) Plant Mol Biol 32:1029-35

Um promotor específico de tecido verde como definido neste é um promotor que é transcripcionalmente ativo de maneira predominante no tecido verde, substancialmente para a exclusão de quaisquer outras partes de

uma planta, enquanto ainda permite-se qualquer expressão ruim nestas outras partes vegetais.

Um outro exemplo de um promotor específico de tecido é um promotor específico de meristema, que transcripcionalmente ativo, predominantemente no tecido meristemático, substancialmente para a exclusão de quaisquer outras partes de uma planta, enquanto ainda permite-se qualquer expressão ruim nestas outras partes vegetais.

Terminador

O termo "terminador" abrange uma seqüência de controle que é uma seqüência de DNA no final de uma unidade transcripcional que sinaliza o processamento 3' e a poliadenilação de uma transcrição primária e a terminação da transcrição. O terminador pode ser derivado do gene natural, a partir de uma variedade de outros genes vegetais ou de T-DNA. O terminador a ser adicionado pode ser derivado, por exemplo, dos genes de nopalina sintase ou octopina sintase ou, alternativamente, de um outro gene vegetal ou menos preferivelmente de qualquer outro gene eucariótico.

Modulação

O termo "modulação" significa, em relação à expressão ou expressão de gene, um processo em que o nível de expressão é mudado pela dita expressão genética em comparação com a planta de controle, preferivelmente, o nível de expressão é aumentado. A expressão não modulada original pode ser de qualquer tipo de expressão de um RNA (rRNA, tRNA) ou mRNA estrutural com a tradução subseqüente. O termo "modulação da atividade" deve significar qualquer mudança da expressão das seqüências de ácido nucleico inventivas ou proteínas codificadas, que leva ao rendimento aumentado e/ou crescimento aumentado das plantas.

Expressão

O termo "expressão" ou "expressão genética" significa a transcrição de um gene específico ou genes específicos ou construção de gene

específico. O termo "expressão" ou "expressão genética" em particular significa a transcrição de um gene ou de genes ou construção genética em RNA (rRNA, tRNA) ou mRNA estruturais com ou sem a tradução subseqüente do último em uma proteína. O processo inclui a transcrição de DNA e o processamento do produto de mRNA resultante.

Expressão aumentada/superexpressão

O termo "expressão aumentada" ou "superexpressão" como usado neste significa qualquer forma de expressão que é adicional ao nível de expressão do tipo selvagem original.

Os métodos para aumentar a expressão de genes ou produtos de genes são bem documentados na técnica e incluem, por exemplo, super-expressão conduzida por promotores apropriados, o uso de intensificadores de transcrição ou intensificadores de transcrição. Ácidos nucleicos isolados que servem como elementos promotores ou intensificadores podem ser introduzidos em uma posição apropriada (tipicamente a montante) de uma forma não heteróloga de um polinucleotídeo a fim de super-regular a expressão de um ácido nucleico que codifica o polipeptídeo de interesse. Por exemplo, os promotores endógenos podem ser alerados *in vivo* por mutação, anulação e/ou substituição (ver, Kmiec, US 5,565,350; Zarling et al., WO9322443) ou promotores isolados podem ser introduzidos em uma célula de planta na orientação e na distância apropriadas de um gene da presente invenção a fim de controlar a expressão do gene.

Se a expressão de polipeptídeo for desejada, é, no geral, desejável incluir uma região de poliadenilação na extremidade 3' de uma região codificadora de polinucleotídeo. A região de poliadenilação pode ser derivada do gene natural, de uma variedade de genes vegetais ou de T-DNA. A seqüência de extremidade 3' a ser adicionada pode ser derivada, por exemplo, dos genes de nopalina sintase ou octopina sintase ou alternativamente de um outro gene vegetal ou menos preferivelmente de

qualquer outro gene eucariótico.

Uma seqüência de íntron também pode ser adicionada à região não traduzida 5' (UTR) ou a seqüência codificadora da seqüência codificadora parcial para aumentar a quantidade da mensagem madura que acumula-se no citosol. A inclusão de um íntron que pode ser unido na unidade de transcrição tanto nas construções de expressão vegetal ou animal foi mostrada aumentar a expressão genética tanto no mRNA quanto nos níveis de proteína até 1000 vezes (Buchman and Berg (1988) Mol. Cell biol. 8: 4395-4405; Callis et al. (1987) Genes Dev 1:1183-1200). Tal intensificação de íntron de expressão genética é tipicamente maior quando colocado próximo da extremidade 5' da unidade de transcrição. O uso dos íntron de milho, íntron Adh1-S 1, 2 e 6, o íntron de Bronze-1 são conhecidos na técnica. Para informação geral, ver: The Maize Handbook, Chapter 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

Gene Endógeno

Aqui, referência a um gene "endógeno" não apenas refere-se ao gene em questão como observado em uma planta em sua forma natural (isto é, em haver qualquer intervenção humana), mas também refere-se àquele mesmo gene (ou um ácido nucleico/gene substancialmente homólogo) em uma forma isolada subseqüentemente (re)introduzida em uma planta (um transgene). Por exemplo, uma planta transgênica contendo um tal transgene pode encontrar uma redução substancial da expressão de transgene e/ou redução substancial de expressão do gene endógeno.

Gene Isolado

O gene isolado pode ser isolado de um organismo ou pode ser artificial, por exemplo, pela síntese química.

Expressão diminuída

Referência neste à "expressão diminuída" ou "redução ou eliminação substancial" de expressão é considerada significar uma diminuição

na expressão de gene endógeno e/ou níveis de polipeptídeo e/ou atividade de polipeptídeo com relação às plantas de controle. A redução ou a eliminação substancial está em ordem crescente, de preferência pelo menos 10%, 20%, 30%, 40% ou 50%, 60%, 70%, 80%, 85%, 90% ou 95%, 96%, 97%, 98%, 99% ou mais reduzido em comparação com aquela de plantas de controle.

Para a redução ou eliminação substancial de expressão, um gene endógeno em uma planta, um comprimento suficiente de nucleotídeos substancialmente contíguos de uma seqüência de ácido nucleico é requerido. A fim de realizar o silenciamento do gene, este pode ser tão pequeno quanto 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10 ou menos nucleotídeos, alternativamente, isto pode ser tanto quanto o gene total (incluindo o 5' e/ou 3' UTR, em parte ou em todo). O estiramento dos nucleotídeos substancialmente contíguos pode ser derivado do ácido nucleico que codifica a proteína de interesse (gene alvo) ou de qualquer ácido nucleico capaz de codificar um ortólogo, parálogo, ou homólogo da proteína de interesse. Preferivelmente, o estiramento de nucleotídeos substancialmente contíguos é capaz de formar ligações de hidrogênio com o gene alvo (filamento sentido ou anti-sentido), mais preferivelmente, o estiramento de nucleotídeos substancialmente contíguos tem, em ordem crescente, de preferência, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 100% de identidade de seqüência ao gene alvo (filamento sentido ou anti-sentido). Uma seqüência de ácido nucleico que codifica um polipeptídeo funcional (funcional) não é um requerimento para os vários métodos debatidos neste para a redução ou eliminação substancial de expressão de um gene endógeno.

Esta redução ou eliminação substancial de expressão pode ser atingida usando-se ferramentas e técnicas de rotina. Um método preferido para a redução ou eliminação substancial da expressão de gene endógeno ocorre pela introdução e expressão em uma planta, uma construção genética em que o ácido nucleico (neste caso, um estiramento de nucleotídeos

substancialmente contíguos derivados do gene de interesse ou de qualquer ácido nucleico capaz de codificar um ortólogo, parálogo ou homólogo de qualquer uma das proteínas de interesse) é clonado como uma repetição invertida (em parte ou completamente), separado por um espaçador (DNA não codificador).

Em um tal método preferido, a expressão do gene endógeno é reduzido ou substancialmente eliminado através do silenciamento mediado por RNA usando-se uma repetição invertida de um ácido nucleico ou uma parte da mesma (neste caso um estiramento de nucleotídeos substancialmente contíguos derivados do gene de interesse ou de qualquer ácido nucleico capaz de codificar um homólogo, parálogo ou homólogo da proteína de interesse), preferivelmente capaz de formar uma estrutura de grampo de cabelo. A repetição invertida é clonada em um vetor de expressão que compreende seqüências de controle. Uma seqüência de ácido nucleico de DNA não codificadora (um espaçador, por exemplo, um fragmento de região de ligação de matriz (MAR), um íntron, um poliligador, etc.) está localizada entre os dois ácidos nucleicos invertidos que formam a repetição invertida. Após a transcrição da repetição invertida, um RNA quimérico com uma estrutura auto-complementar é formada (parcial ou completa). Esta estrutura de RNA de filamento duplo é referida como o RNA grampo de cabelo (hpRNA). O hpRNA é processado pela planta em siRNAs que são incorporados e um complexo silenciador induzido por RNA (RISC). O RISC ainda cliva as transcrições de mRNA, desse modo substancialmente reduzindo o número de transcrições de mRNA a serem traduzidas em polipeptídeos. Para detalhes gerais adicionais ver, por exemplo, Grierson *et al.* (1998) WO 98/53083; Waterhouse *et al.* (1999) WO 99/53050).

O desempenho dos métodos da não contam com a introdução e a expressão em uma construção genética vegetal em que o ácido nucleico é clonado como uma repetição invertida, mas qualquer um ou mais de diversos

métodos "silenciadores genéticos" bem conhecidos podem ser usados para atingir os mesmos efeitos.

Um tal método para a redução da expressão de gene endógeno é o silenciamento mediado por RNA da expressão genética (sub-regulação). O silenciamento neste caso é disparado em uma planta por uma seqüência de RNA de filamento duplo (dsRNA) que é substancialmente similar ao gene endógeno alvo. Este dsRNA ainda é processado pela planta em cerca de 20 a cerca de 26 nucleotídeos denominados RNAs de interferência curta (siRNAs). Os siRNAs são incorporados em um complexo de silenciamento induzido por RNA (RISC) que cliva a transcrição de mRNA do gene alvo endógeno, desse modo substancialmente, reduzindo o número de transcrições de mRNA a serem traduzidas em um polipeptídeo. Preferivelmente, a seqüência de RNA de filamento duplo corresponde a um gene alvo.

Um outro exemplo de um método silenciador de RNA envolve a introdução de seqüências de ácido nucleico ou partes destas (neste caso um estiramento de nucleotídeos substancialmente contíguos derivados do gene de interesse ou de qualquer ácido nucleico capaz de codificar um homólogo, parálogo ou homólogo da proteína de interesse) em uma orientação sentido em uma planta. A "orientação sentido" refere-se a uma seqüência de DNA que é homóloga a uma transcrição de mRNA desta. Portanto, deve ser introduzido em uma planta, pelo menos uma cópia da seqüência de ácido nucleico. A seqüência de ácido nucleico adicional reduzirá a expressão do gene endógeno, dando origem a um fenômeno conhecido como co-supressão. A redução da expressão genética será mais pronunciada se diversas cópias adicionais de uma seqüência de ácido nucleico foram introduzidas na planta, como existe uma correlação positiva entre os níveis de transcrição altos e o disparo da co-supressão.

Um outro exemplo de um método de silenciamento envolve o uso de seqüências de ácido nucleico anti-sentido. Uma seqüência de ácido

nucleico "anti-sentido" compreende uma seqüência de nucleotídeo que é complementar a uma seqüência de ácido nucleico "sentido" que codifica uma proteína, isto é, complementar ao filamento codificador de uma molécula de cDNA de filamento duplo ou complementar a uma seqüência de transcrição de mRNA. A seqüência de ácido nucleico anti-sentido é preferivelmente complementar ao gene endógeno a ser silenciado. A complementaridade pode estar localizada na "região codificadora" e/ou na "região não codificadora" de um gene. O termo "região codificadora" refere-se a uma região da seqüência de nucleotídeo que compreende códons que são traduzidas em resíduos de aminoácido. O termo "região não codificadora" refere-se às seqüências 5' e 3' que flanqueiam a região codificadora que são descritas mas não traduzidas em aminoácidos (também referidos como regiões não traduzidas 5' e 3').

As seqüências de ácido nucleico anti-sentido podem ser projetadas de acordo com as regras de formação de pares de base de Watson and Crick. A seqüência de ácido nucleico anti-sentido pode ser complementar à seqüência de ácido nucleico total (neste caso um estiramento de nucleotídeos substancialmente contíguos derivados do gene de interesse ou de qualquer ácido nucleico capaz de codificar um homólogo, parálogo ou homólogo da proteína de interesse), mas também pode ser um oligonucleotídeo que é anti-sentido a apenas uma parte da seqüência de ácido nucleico (incluindo o mRNA 5' e 3' UTR). Por exemplo, a seqüência de oligonucleotídeo anti-sentido pode ser complementar à região que circunda o local de início de translação de uma transcrição de mRNA que codifica um polipeptídeo. O comprimento de uma seqüência de oligonucleotídeo anti-sentido adequado é conhecido na técnica e pode começar de cerca de 50, 45, 40, 35, 30, 25, 20, 15 ou 10 nucleotídeos de comprimento ou menos. Uma seqüência de ácido nucleico anti-sentido de acordo com a invenção pode ser construída usando-se síntese química e reações de ligação enzimática causando-se os métodos conhecidos na técnica. Por exemplo, uma seqüência

de ácido nucleico anti-sentido (por exemplo, uma seqüência de oligonucleotídeo anti-sentido) pode ser quimicamente sintetizado usando-se nucleotídeos de ocorrência natural ou nucleotídeos variadamente modificados projetados para aumentar a estabilidade biológica das moléculas ou para aumentar a estabilidade física do duplex formado entre as seqüências de ácido nucleico anti-sentido e sentido, por exemplo, derivados de fosforotioato e nucleotídeos substituídos por acridina podem ser usados. Os exemplos de nucleotídeos modificados que podem ser usados para gerar a seqüência de ácido nucleico anti-sentidos são bem conhecidos na técnica. As modificações de nucleotídeo conhecidas incluem metilação, ciclização e 'coberturas' e substituição de um ou mais dos nucleotídeos de ocorrência natural com um análogo, tal como inosina. Outras modificações de nucleotídeos são bem conhecidas na técnica.

A seqüência de ácido nucleico anti-sentido pode ser produzida biologicamente usando-se um vetor de expressão em que uma seqüência de ácido nucleico foi subclonada em uma orientação anti-sentido (isto é, RNA transcrito a partir do ácido nucleico inserido será de uma orientação anti-sentido a um ácido nucleico alvo de interesse). Preferivelmente, produção de seqüências de ácido nucleico anti-sentido em plantas ocorre por meio de construção de ácido nucleico estavelmente integrada que compreende um promotor, um oligonucleotídeo anti-sentido operacionalmente ligado e um terminador.

As moléculas de ácido nucleico usadas para o silenciamento nos métodos da invenção (introduzidas em uma planta ou geradas *in situ*) hibridizam-se ou ligam-se às transcrições de mRNA e/ou DNA genômico que codifica um polipeptídeo para, desse modo, inibir a expressão da proteína, por exemplo, pela inibição da transcrição e/ou tradução. A hibridização pode ocorrer pela complementaridade de nucleotídeo convencional para formar um duplex estável ou, por exemplo, no caso de uma seqüência de ácido nucleico

anti-sentido que liga-se aos duplexes de DNA, através das interações específicas no sulco principal da hélice dupla. As seqüências de ácido nucleico anti-sentido podem ser introduzidas em uma planta por transformação ou injeção direta em um local de tecido específico. Alternativamente, as seqüências de ácido nucleico anti-sentido podem ser modificadas para alvejar células selecionadas e depois administrado sistemicamente. Por exemplo, para a administração sistêmica, as seqüências de ácido nucleico anti-sentido podem ser modificadas, tal que estas liguem-se especificamente aos receptores ou antígenos expressados em uma superfície celular selecionada, por exemplo, pela ligação da seqüência de ácido nucleico anti-sentido aos peptídeos ou anticorpos que ligam-se aos receptores de superfície celular ou antígenos. A seqüência de ácido nucleico anti-sentidos também pode ser liberada às células usando-se os vetores escritos neste.

De acordo com um outro aspecto, a seqüência de ácido nucleico anti-sentido é uma seqüência de ácido nucleico a-anomérica. Uma seqüência de ácido nucleico a-anomérica forma híbridos de filamento duplo específicos com RNA complementar em que, ao contrários das unidades b usuais, os filamentos correm em paralelo um com relação ao outro (Gaultier et al. (1987) Nucl Ac Res 15: 6625-6641). A seqüência de ácido nucleico anti-sentido também pode compreender um 2'-o-metilribonucleotídeo (Inoue et al. (1987) Nucl Ac Res 15, 6131-6148) ou um análogo de RNA-DNA quimérico (Inoue et al. (1987) FEBS Lett. 215, 327-330).

A redução ou a eliminação substancial de expressão de gene endógeno também pode ser realizada usando-se ribozimas. As ribozimas são moléculas de RNA catalíticas com atividade de ribonuclease que são capazes de clivar uma seqüência de ácido nucleico de filamento simples, tal como um mRNA, aos quais estes tem uma região complementar. Desta maneira, as ribozimas (por exemplo, ribozimas cabeça de martelo (descrito em Haselhoff and Gerlach (1988) Nature 334, 585-591) podem ser usadas para clivar

cataliticamente as transcrições de mRNA que codifica um polipeptídeo, desse modo reduzindo substancialmente o número de transcrições de mRNA a serem traduzidas em um polipeptídeo. Uma ribozima tendo especificidade para uma seqüência de ácido nucleico pode ser projetada (ver, por exemplo: Cech et al. Patente U. S. N° 4.987.071 e Cech *et al.* Patente U. S. N° 5,116,742). Alternativamente, as transcrições de mRNA que correspondem a uma seqüência de ácido nucleico podem ser usadas para selecionar um RNA catalítico tendo uma atividade de ribonuclease específica de um grupo de moléculas de RNA (Bartel and Szostak (1993) Science 261, 1411-1418). O uso de ribozimas para o silenciamento de gene em plantas é conhecido na técnica (por exemplo, Atkins et al. (1994) WO 94/00012; Lenne et al. (1995) WO 95/03404; Lutziger et al. (2000) WO 00/00619; Prinsen et al. (1997) WO 97/13865 e Scott et al. (1997) WO 97/38116).

O silenciamento de gene também pode ser atingido pela mutagênese de inserção (por exemplo, inserção de T-DNA ou inserção de transposon) ou pelas estratégias como descrito por, entre outros, Angell and Baulcombe ((1999) Plant J 20(3): 357-62), (Amplicon VIGS WO 98/36083) ou Baulcombe (WO 99/15682).

O silenciamento de gene também pode ocorrer se houver uma mutação ou um gene endógeno e/ou uma mutação em um gene/ácido nucleico isolado subseqüentemente introduzido em uma planta. A redução ou a eliminação substancial pode ser causada por um polipeptídeo não funcional. Por exemplo, um polipeptídeo pode ligar-se a várias proteínas de interação; uma ou mais mutações e/ou truncações pode, portanto, fornecer um polipeptídeo que ainda é capaz de interagir proteínas (tais como proteínas receptoras) mas que não pode apresentar sua função normal (tal como ligando sinalizador).

Um método adicional para o silenciamento do gene ocorre pelo alvejamento das seqüências de ácido nucleico complementares à região

reguladora do gene (por exemplo, o promotor e/ou intensificadores) para a formação de estruturas hélicas triplas que evitam a transcrição do gene em células alvo. Ver Helene, C., Anticancer Drug Res. 6, 569-84, 1991; Helene et al., Ann. N.Y. Acad. Sci. 660, 27-36 1992; and Maher, L.J. Bioassays 14, 807-15, 1992.

Outros métodos, tais como o uso de anticorpos direcionados a um polipeptídeo endógeno para inibir sua função na planta ou interferência no caminho sinalizador em que um polipeptídeo está envolvido, será bem conhecido pela pessoa habilitada. Em particular, pode ser considerado que moléculas artificiais podem ser úteis para a inibição da função biológica de um polipeptídeo alvo ou para interferir com o caminho de sinalização em que o polipeptídeo alvo está envolvido.

Alternativamente, um programa de avaliação pode ser ajustado para identificação em variantes naturais de população vegetal de um gene, cujas variantes codificam polipeptídeos com atividade reduzida. Tais variantes naturais também podem ser usadas, por exemplo, para realizar a recombinação homóloga.

Os microRNAs artificiais e/ou naturais (miRNAs) podem ser usados para nocautear a expressão genética e/ou a tradução de mRNA. Os miRNAs endógenos são RNAs pequenos de filamento simples de, tipicamente 19 a 24 nucleotídeos de comprimento. Este funciona primariamente para regular a expressão genética e/ou tradução de mRNA. A maioria dos microRNAs vegetais (miRNAs) tem complementaridade perfeita ou quase perfeita com suas seqüências alvo. Entretanto, existem alvos naturais com até cinco má combinações. Estes são processados a partir de RNAs não codificadores mais longos com estruturas dobradas novamente características pelas RNases específicas de filamento duplo da família Dicer. No processamento, estes são incorporados no complexo silenciador induzido por RNA (RISC) pela ligação ao seu componente principal, uma proteína

Argonaute. MiRNAs serve como os componentes de especificidade de RISC, visto que estes formam pares de base para alvejar ácidos nucleicos, a maioria dos mRNAs, no citoplasma. Os eventos reguladores subseqüentes incluem a clivagem e a destruição de mRNA alvo e/ou inibição de tradução. Os efeitos da super-expressão de miRNA são, desta maneira, freqüentemente refletidos em níveis de mRNA diminuído de genes alvo.

Os microRNAs artificiais (amiRNAs), que são tipicamente 21 nucleotídeos de comprimento, podem ser geneticamente projetados especificamente para regular negativamente a expressão genética de interesse. Determinantes da seleção alvo de microRNA vegetal são bem conhecidos na técnica. Os parâmetros empíricos para o reconhecimento alvo foram definidos e podem ser usados para auxiliar no projeto de amiRNAs específicos, (Schwab et al., Dev. Cell 8, 517-527, 2005). Ferramentas convenientes para o projeto e geração de amiRNAs e seus precursores também estão disponíveis ao público (Schwab et al., Plant Cell 18, 1121-1133, 2006).

Para o desempenho ótimo, as técnicas silenciadoras de gene usadas para a redução da expressão em uma planta de um gene endógeno requer o uso de seqüências de ácido nucleico para plantas monocotiledôneas para a transformação de plantas monocotiledôneas e de plantas dicotiledôneas para a transformação de plantas dicotiledôneas. Preferivelmente, uma seqüência de ácido nucleico de qualquer espécie vegetal dada é introduzida naquela mesma espécie. Por exemplo, uma seqüência de ácido nucleico para arroz é transformada em uma planta de arroz. Entretanto, este não é requerimento absoluto que a seqüência de ácido nucleico a ser introduzida origine-se da mesma espécie vegetal como a planta em que esta será introduzida. É suficiente que exista homologia substancial entre o gene alvo endógeno e o ácido nucleico a ser introduzido.

São descritos acima, exemplos de vários métodos para a redução ou eliminação substancial de expressão em uma planta de um gene

endógeno. Uma pessoa habilitada na técnica seria facilmente capaz de adaptar os métodos já mencionados para o silenciamento a fim de atingir a redução da expressão de um gene endógeno em uma planta total ou em partes desta através do uso de um promotor apropriado, por exemplo.

Marcador selecionável (gene)/Gene Repórter

"Marcador selecionável", "gene marcador selecionável" ou "gene repórter" inclui qualquer gene que confere um fenotipo em uma célula em que esta é expressada para facilitar a identificação e/ou a seleção de células que são transfectadas ou transformadas com uma construção de ácido nucleico da invenção. Estes genes marcadores permitem a identificação de uma transferência bem sucedida das moléculas de ácido nucleico por intermédio de uma série de princípios diferentes. Os marcadores adequados podem ser selecionados a partir de marcadores que conferem resistência antibiótica ou herbicida, que introduz uma nova característica metabólica ou que permite a seleção visual. Os exemplos de marcadores de gene selecionáveis incluem genes que conferem resistência a antibióticos (tais como nptII que fosforila neomicina e canamicina ou hpt, fosforilam higromicina ou genes que conferem resistência a, por exemplo, bleomicina, estreptomicina, tetraciclina, cloranfenicol, ampicilina, gentamicina, geneticina (G418), espectinomicina ou blasticidina), a herbicidas (por exemplo, bar que fornece resistência a Basta®; aroA ou gox que fornece resistência contra glifosato ou os genes que conferem resistência a, por exemplo, imidazolinona, fosfinotricina ou sulfoniluréia), ou genes que fornecem uma característica metabólica (tal como manA que permitem às plantas usar manose como a fonte de carbono simples ou xilose isomerase para a utilização de xilose ou marcadores anti-nutritivos tal como a resistência a 2-desoxiglicose). A expressão dos genes marcadores visuais resulta na formação de cor (por exemplo, β-glucuronidase, GUS ou β-galactosidase com seus substratos coloridos, por exemplo, X-Gal), luminescência (tal como o sistema

luciferina/luceferase) ou fluorescência (Proteína Fluorescente Verde, GFP e derivados destes). Esta lista representa apenas um pequeno número marcadores possíveis. O trabalhador habilitado está familiarizado com tais marcadores. Marcadores diferentes são preferidos, dependendo do e do método selecionado.

É conhecido que integração estável ou transitória de ácidos nucleicos em células de planta, apenas uma minoria das células absorve o DNA estranho e, se desejado, integra-o em seu genoma, dependendo do vetor de expressão usado e da técnica de transfecção usada. Para identificar e selecionar estes integrantes, um gene codificador de um marcador selecionável (tal como os descritos acima) é usualmente introduzido nas células hospedeiras junto com o gene de interesse. Estes marcadores pode ser, por exemplo, usados em mutantes em que estes genes não são funcionais, por exemplo, pela anulação pelos métodos convencionais. Além disso, as moléculas de ácido nucleico que codificam um marcador selecionável podem ser introduzidas em uma célula hospedeira no mesmo vetor que compreende a seqüência que codifica os polipeptídeos da invenção ou usada nos métodos da invenção ou ainda em um vetor separado. As células que foram estavelmente transfectadas com o ácido nucleico introduzido podem ser identificadas, por exemplo, por seleção (por exemplo, células que integraram o marcador selecionável sobrevivem, visto que as outras células morrem).

Visto que os genes marcadores, particularmente genes para a resistência a antibióticos e herbicidas, não são mais requeridos ou são indesejados na célula hospedeira transgênica uma vez que os ácidos nucleicos foram introduzidos de maneira bem sucedida, o processo de acordo com a invenção para a introdução dos ácidos nucleicos vantajosamente utilizam técnicas que permitem a remoção ou excisão destes genes marcadores. Um tal método é aquele que é conhecido como co-transformação. O método de co-transformação utiliza dois vetores simultaneamente para a transformação, um

vetor carrega o ácido nucleico de acordo com a invenção e um segundo carrega os genes marcadores. Uma grande proporção de transformantes recebe ou, no caso de plantas, compreende (até 40% ou mais dos transformantes), ambos os vetores. No caso de transformação com Agrobacteria, os transformantes usualmente recebem apenas uma parte do vetor, isto é, a seqüência flanqueada pelo T-DNA, que usualmente representa o cassete de expressão. Os genes marcadores podem ser subseqüentemente removidos da planta transformada realizando-se cruzamentos. Em um outro método, os genes marcadores integrados em um transposon são usados para a transformação junto com o ácido nucleico desejado (conhecido como a tecnologia Ac/Ds). Os transformantes podem ser cruzados com uma fonte de transposase ou os transformantes são transformados com uma construção de ácido nucleico que confere a expressão de uma transposase, de maneira transitória ou estável. Em alguns casos (aproximadamente 10%), o transposon pula fora do genoma da célula hospedeira uma vez que a transformação acontece de maneira bem sucedida e é perdida. Em diversos casos adicionais, o transposon pula para um local diferente. Nestes casos, o gene marcador deve ser eliminado realizando-se cruzamentos. Na microbiologia, técnicas foram desenvolvidas tornando possível ou facilitando a detecção de tais eventos. Um outro método vantajoso conta com o que é conhecido como sistemas de recombinação; cuja vantagem é que a eliminação por cruzamento pode ser dispensada. O melhor sistema conhecido deste tipo é aquele que é conhecido como o sistema Cre/lox. Cre1 é uma recombinase que remove as seqüências localizadas entre as seqüências loxP. Se o gene marcador for integrado entre as seqüências loxP, este é removido uma vez que a transformação aconteceu de maneira bem sucedida, pela expressão da recombinase. Os sistemas de recombinação adequados são o sistema HIN/HIX, FLP/FRT e REP/STB (Tribble et al., J. Biol. Chem., 275, 2000: 22255-22267; Velmurugan et al., J. Cell Biol., 149, 2000: 553-566). Uma

integração específica de local no genoma vegetal das seqüências de ácido nucleico de acordo com a invenção é possível. Naturalmente, estes métodos também podem ser aplicados aos microorganismos, tais como levedura, fungos ou bactérias.

Transgênico/Transgene/Recombinante

Para os propósitos da invenção, "transgênico", "transgene" ou "recombinante" significa com respeito a, por exemplo, uma seqüência de ácido nucleico, um cassete de expressão, construção de gene ou um vetor que compreende a seqüência de ácido nucleico ou um organismo transformado com as seqüências de ácido nucleico, cassetes de expressão ou vetores de acordo com a invenção, todas aquelas construções realizadas pelos métodos recombinantes em que as seqüências de ácido nucleico que codificam proteínas úteis nos métodos da invenção ou seqüências de controle genético que são operacionalmente ligadas com a seqüência de ácido nucleico de acordo com a invenção, por exemplo, um promotor e a) e b) não estão localizados em seu ambiente genético natural ou foram modificados por métodos recombinantes, isto sendo possível para a modificação tomar a forma de, por exemplo, uma substituição, adição, anulação, inversão ou inserção de um ou mais resíduos de nucleotídeo. O ambiente genético natural é entendido como significando o local genômico ou cromossômico natural na planta original ou a presença em uma biblioteca genômica. No caso de uma biblioteca genômica, o ambiente genético natural da seqüência de ácido nucleico é preferivelmente retido, pelo menos em parte. O ambiente flanqueia a seqüência de ácido nucleico pelo menos em um lado e tem um comprimento de seqüência de pelo menos 50 bp, preferivelmente pelo menos 500 bp, especialmente preferível pelo menos 1000 bp, mais preferivelmente pelo menos 5000 bp. Um cassete de expressão de ocorrência natural– por exemplo a combinação de ocorrência natural do promotor natural das seqüências de ácido nucleico com a seqüência de ácido nucleico correspondente que

codifica um polipeptídeo úteis nos métodos da presente invenção, como definido acima – torna-se um cassete de expressão transgênico quando é modificado por métodos sintéticos não naturais ("artificiais") tais como, por exemplo, tratamento mutagênico. Os métodos adequados são descritos, por exemplo, em US 5.565.350 ou WO 00/15815.

Uma planta transgênica para os propósitos da invenção é, desta maneira entendido significar, como acima, que os ácidos nucleicos usados no método da invenção não estão em seu local natural no genoma da dita planta, sendo possível para os ácidos nucleicos serem expressados de maneira homóloga ou heteróloga. Entretanto, como mencionado, transgênico também significa que, enquanto os ácidos nucleicos de acordo com a invenção ou usados no método inventivo estão em sua posição natural no genoma de uma planta, a seqüência foi modificada com respeito à seqüência natural e/ou que as seqüências reguladoras das seqüências naturais foram modificadas. Transgênico é preferivelmente entendido como significando a expressão dos ácidos nucleicos de acordo com a invenção em um local não neural no genoma, isto é, a expressão homóloga ou, preferivelmente, heteróloga dos ácidos nucleicos acontece. As plantas transgênicas preferidas são mencionadas aqui.

Transformação

O termo "introdução" ou "transformação" como referido aqui abrange a transferência de um polinucleotídeo exógeno em uma célula hospedeira, sem consideração ao método usado para a transferência. O tecido vegetal capaz de propagação clonal subseqüente, por organogênese ou embriogênese, pode ser transformado com uma construção genética da presente invenção e uma planta total regenerada a partir deste. O tecido particular escolhido variará dependendo dos sistemas de propagação clonais disponíveis e melhor adaptados para as espécies particulares sendo transformadas. Os tecidos alvo exemplares incluem discos de filhas, pólen,

embriões, cotiledôneas, hipocotilas, megagametófitos, tecidos de calos, tecido meristemático existente (por exemplo, meristema ápico, botões axilares e meristemas de raiz) e tecido de meristema induzido (por exemplo, meristema de colilédone e meristema de hipocotila). O polinucleotídeo pode ser transitória ou estavelmente introduzido em uma célula hospedeira e pode ser mantido não integrado, por exemplo, como um plasmídeo. Alternativamente, este pode ser integrado no genoma hospedeiro. A célula de planta transformada resultante pode ser então usada para regenerar uma planta transformada de uma maneira conhecida pelas pessoas habilitadas na técnica.

A transferência de genes estranhos no genoma de uma planta é denominada transformação. A transformação de uma espécie vegetal é agora uma técnica de rotina considerável. Vantajosamente, qualquer um de diversos métodos de transformação podem ser usados para a introdução de gene de interesse em uma célula predecessora adequada. Os métodos descritos para a transformação e regeneração de plantas a partir dos tecidos vegetais ou células de planta podem ser utilizados para a transformação transitória ou estável. Os métodos de transformação incluem o uso de lipossomas, eletroporação, produtos químicos que aumentam a absorção de DNA livre, injeção do DNA diretamente na planta, bombardeamento com atirador de partícula, transformação usando-se vírus ou pólen e microprojeção. Os métodos podem ser selecionados os métodos podem ser selecionados a partir dos métodos de cálcio/polietileno glicol para os protoplastos (Krens, F.A. et al., (1982) Nature 296, 72-74; Negrutiu I *et al.* (1987) Plant Mol Biol 8: 363-373); eletroporação de protoplastos (Shillito R.D. *et al.* (1985) Bio/Technol 3, 1099-1102); microinjeção em material vegetal (Crossway A *et al.*, (1986) Mol. Gen Genet 202: 179-185); bombardeamento de partícula revestido com DNA ou RNA (Klein TM et al., (1987) Nature 327: 70) infecção com vírus (não-integrativo) e outros. As plantas transgênicas, incluindo plantas de safra transgênicas, são preferivelmente produzidas por intermédio de transformação

mediada por *Agrobacterium*. Um método de transformação vantajoso é a transformação *in planta*. Para este fim, é possível, por exemplo, deixar as agrobactérias atuarem em sementes de plantas ou inocular o meristema vegetal com agrobactérias. Mostrou-se particularmente conveniente de acordo com invenção permitir que uma suspensão de agrobactérias transformadas atuem na planta intacta ou pelo menos nos primórdios da flor. A planta está, subseqüentemente desenvolvendo-se até as sementes da planta tratada serem obtidas (Clough and Bent, Plant J. (1998) 16, 735–743). Os métodos para a transformação mediada por *Agrobacterium* de arroz incluem métodos bem conhecidos para a transformação do arroz, tais como aqueles descritos em qualquer um dos seguintes: pedido de patente europeu EP 1198985 A1, Aldemita and Hodges (Planta 199: 612-617, 1996); Chan et al. (Plant Mol Biol 22 (3): 491-506, 1993), Hiei et al. (Plant J 6 (2): 271-282, 1994), cujas divulgações são incorporadas neste por referência como se totalmente apresentado. No caso da transformação do milho, o método preferido é como descrito e, Ishida et al. (Nat. Biotechnol 14(6): 745-50, 1996) ou Frame et al. (Plant Physiol 129(1): 13-22, 2002), cujas divulgações são incorporadas por referência neste como se totalmente apresentado. Os ditos métodos ainda são descritos por meio de exemplo em B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S.D. Kung and R. Wu, Academic Press (1993) 128-143 e em Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205-225). Os ácidos nucleicos ou as construções a serem expressadas são preferivelmente clonadas em um vetor, que é adequado para a transformação de *Agrobacterium tumefaciens*, por exemplo pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984) 8711). As agrobactérias transformadas por tal vetor podem ser então usadas de uma maneira conhecida para a transformação das plantas, tais como as plantas usadas como um modelo, como *Arabidopsis* (*Arabidopsis thaliana* está dentro do escopo da presente invenção não considerada como uma planta de safra)

ou plantas de safra, tais como, por meio de exemplo, plantas de tabaco, por exemplo, pela imersão de folhas esmagadas ou folhas cortadas em uma solução agrobacteriana e então cultivando-as em meio adequado. A transformação de plantas por meio de *Agrobacterium tumefaciens* é descrita, por exemplo, por Höfgen and Willmitzer in Nucl. Acid Res. (1988) 16, 9877 ou é conhecida inter alia de F.F. White, Vetors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S.D. Kung and R. Wu, Academic Press, 1993, pp. 15-38.

Além da transformação de células somáticas, que foram então regeneradas em plantas intactas, também é possível transformar as células dos meristemas vegetais e, em particular, aquelas células que desenvolvem-se em gametas. Neste caso, os gametas transformados seguem o crescimento vegetal natural dando origem à plantas transgênicas. Desta maneira, por exemplo, as sementes de *Arabidopsis* são tratadas com agrobactérias e as sementes são obtidas para o crescimento de plantas cuja certa proporção é transformada e, desta maneira, transgênica [Feldman, KA and Marks MD (1987). Mol Gen Genet 208:274-289; Feldmann K (1992). In: C Koncz, N-H Chua and J Shell, eds, Methods in Arabidopsis Research. Word Scientific, Singapore, pp. 274-289]. Os métodos alternativos são fundamentados na remoção repetida de inflorescências e incubação do local de excisão no centro da roseta com agrobactéria transformada, desse modo, as sementes transformadas também podem ser obtidas em um ponto posterior no tempo (Chang (1994). Plant J. 5: 551-558; Katavic (1994). Mol Gen Genet, 245: 363-370). entretanto, um método especialmente eficaz é o método de filtração a vácuo com suas modificações tais como i método de "imersão floral". No caso da infiltração a vácuo de *Arabidopsis*, as plantas intactas sob pressão reduzida são tratadas com uma suspensão agrobacteriana [Bechthold, N (1993). C R Acad Sci Paris Life Sci, 316: 1194-1199], enquanto no caso do método de "imersão floral" o crescimento do tecido floral é incubado brevemente com uma suspensão de

agrobactéria tratada com tensoativo [Clough, SJ and Bent AF (1998) The Plant J. 16, 735-743]. Uma certa proporção de sementes transgênicas é colhida em ambos os casos e estas sementes podem ser distinguidas das sementes não transgênicas pelo crescimento sob condições seletivas descritas acima. Além disso, a transformação estável de plastídeos é vantajosa porque os plastídeos são herdados maternalmente na maioria das safras reduzindo ou eliminando o risco de que o transgene flua através do pólen. A transformação do genoma do cloroplasto é, no geral, atingido por um processo que foi esquematicamente apresentado em Klaus et al., 2004 [Nature Biotechnology 22 (2), 225-229]. Resumidamente, as seqüências a serem transformadas são clonadas junto com um gene marcador selecionável entre as seqüências de flanqueamento homólogas ao genoma do cloroplasto. Estas seqüências flanqueadoras homólogas direcionam a integração específica ao local no plastoma. A transformação plastidal foi descrita para muitas espécies vegetais diferentes e uma visão geral é dada em Bock (2001) Transgenic plastids in basic research and plant biotechnology. J Mol Biol. 2001 Sep 21; 312 (3):425-38 or Maliga, P (2003) Progress towards commercialization of plastid transformation technology. Trends Biotechnol. 21, 20-28. Além disso, o progresso biotecnológico foi recentemente relatado na forma de transformantes de plastídeo isentos de marcador, que pode ser produzido por um gene marcador co-integrado transitório (Klaus et al., 2004, Nature Biotechnology 22(2), 225-229).

Rotulação de ativação de T-DNA

A rotulação de ativação de T-DNA (Hayashi et al. Science (1992) 1350-1353), envolve a inserção de T-DNA, usualmente contendo um promotor (também pode ser um intensificador de tradução ou um íntron), na região genômica do gene de interesse ou 10 kb até ou a jusante da região codificadora de um gene em uma configuração tal que o promotor direcione a expressão do gene alvejado. Tipicamente, a regulação da expressão do gene

alvejado por seu promotor natural é interrompido e o gene está sob o controle do promotor recentemente introduzido. O promotor é tipicamente embutido em um T-DNA. Este T-DNA é aleatoriamente inserido nogenoma vegetal, por exemplo, através da infecção por *Agrobacterium* e leva à expressão modificada dos genes aproximando-se do T-DNA inserido. As plantas transgênicas resultantes apresentam fenotipos dominantes devido à expressão modificada dos genes próximos ao promotor introduzido.

TILLING

O termo "TILLING" é uma abreviação de "Lesões Locais Induzidas Alvejadas Em Genomas" e refere-se a uma tecnologia de mutagênese útil para gerar e/ou identificar ácidos nucleicos que codificam proteínas com expressão e/ou atividade modificadas. TILLING também permite a seleção de plantas que carregam tais variantes mutantes. Estas variantes mutantes podem apresentar expressão modificada, em concentração ou em local ou em regulação de tempo (se as mutações afetarem o promotor, por exemplo). Estas variantes mutantes podem apresentar atividade mais alta do que aquela apresentada pelo gene em sua forma natural. TILLING combina mutagênese de densidade alta com métodos de avaliação de resposta alta. As etapas tipicamente seguidas em TILLING são: (a) mutagênese de EMS (Redei GP and Koncz C (1992) In Methods in Arabidopsis Research, Koncz C, Chua NH, Schell J, eds. Singapore, World Scientific Publishing Co, pp. 16–82; Feldmann et al., (1994) In Meyerowitz EM, Somerville CR, eds, Arabidopsis. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 137-172; Lightner J and Caspar T (1998) In J Martinez-Zapater, J Salinas, eds, Methods on Molecular Biology, Vol. 82. Humana Press, Totowa, NJ, pp 91-104); (b) preparação de DNA e união de indivíduos; (c) amplificação de PCR de uma região de interesse; (d) desnaturação e recozimento para permitir a formação de heteroduplexes; (e) DHPLC, onde a presença de um heteroduplex em um grupo é detectado como um pico extra no cromatograma;

(f) identificação do indivíduo mutante e (g) sequenciamento do produto de PCR mutante. Os métodos para TILLING são bem conhecidos na técnica (McCallum et al., (2000) Nat Biotechnol 18: 455-457; revisto por Stemple (2004) Nat Rev Genet 5(2): 145-50).

Recombinação homóloga

A recombinação homóloga permite a introdução em um genoma de um ácido nucleico selecionado em uma posição selecionada definida. A recombinação homóloga é uma tecnologia padrão rotineiramente usada em ciências biológicas para organismos inferiores, tais como levedura ou o fungo *Physcomitrella*. Os métodos para realizar a recombinação homóloga em foram descritos não apenas para plantas modelo (Offringa et al. (1990) EMBO J 9(10): 3077-84) mas também para plantas de safra, por exemplo, arroz (Terada et al. (2002) Nat Biotech 20(10): 1030-4; Iida and Terada (2004) Curr Opin Biotech 15(2): 132-8).

Rendimento

O termo "rendimento", no geral, significa uma produção mensurável de valor econômico, tipicamente relacionado a uma safra, a uma área e a um período de tempo especificados. As partes vegetais individuais contribuem diretamente para o rendimento com base em seu número, tamanho e/ou peso ou o rendimento real é o rendimento por acre para uma safra e ano, que é determinado dividindo-se a produção total (inclui tanto a produção colhida quanto a estimada) por acres plantados. O termo "rendimento" de uma planta diz respeito à biomassa vegetativa (biomassa de raiz e/ou broto), aos órgãos reprodutivos e/ou à *propagules* (tais como sementes) daquela planta.

Vigor precoce

"Vigor precoce" refere-se ao crescimento bem equilibrado saudável ativo especialmente durante os estágios precoces de crescimento da planta e pode resultar da aptidão vegetal aumentada, por exemplo, devido às plantas estarem melhor adaptadas ao seu ambiente (isto é, otimizando-se o

uso de recursos de energia e divisão entre broto e raiz). As plantas tendo vigor precoce também apresentam sobrevivência de muda adequada e um estabelecimento melhor da safra, que freqüentemente resulta em campos altamente uniformes (com a safra desenvolvendo-se de maneira uniforme, isto é, com a maioria das plantas atingindo os vários estágios de crescimento substancialmente ao mesmo tempo) e, freqüentemente, rendimento melhor e mais alto. Portanto, o vigor precoce pode ser determinado medindo-se vários fatores, tais como peso de semente singular, porcentagem de germinação, porcentagem de emergência, crescimento de muda, altura da muda, comprimento da raiz, biomassa de raiz e de broto e muito mais.

Aumento/Melhora/Intensificação

Os termos "aumento", "melhora" ou "intensificação" são intercambiáveis e devem significar, no sentido da aplicação pelo menos 5%, 6%, 7%, 8%, 9% ou 10%, preferivelmente pelo menos 15% ou 20%, mais preferivelmente 25%, 30%, 35% ou 40% mais rendimento e/ou crescimento em comparação com as plantas de controle como definido neste.

Rendimento de semente

O rendimento de semente aumentado pode manifestar-se por si só como um ou mais dos seguintes: a) um aumento na biomassa da semente (peso de semente total) que pode estar em uma base de semente individual e/ou por planta e/ou por hectare ou acre; b) número aumentado de flores por planta; c) número aumentado de sementes (enchidas); d) taxa de enchimento de semente aumentada (que é expressada como a razão entre o número de sementes enchidas dividido pelo número total de sementes); e) índice de colheita aumentado, que é expressado como uma razão do rendimento de plantas colhíveis, tais como sementes, dividido pela biomassa total e f) peso de semente singular aumentado (TKW), que é extrapolado a partir do número de sementes enchidas contadas e seu peso total. Um TKW aumentado pode resultar de um tamanho de semente e/ou peso de semente aumentados e

também pode resultar de um aumento no tamanho de embrião e/ou endosperma.

Um aumento no rendimento de semente também pode ser manifestado como um aumento no tamanho de semente e/ou volume de semente. Além disso, um aumento no rendimento de semente também pode se manifestar por si só como um aumento na área de semente e/ou comprimento de semente e/ou largura de semente e/ou perímetro de semente. O rendimento aumentado também pode resultar em arquitetura modificada ou pode ocorrer por causa da arquitetura modificada.

Índice de verdor

O "índice de verdor" como usado neste é calculado a partir de imagens digitais de plantas. Para cada pixel que pertence à planta objeto na imagem, a razão do valor de verde versus o valor de vermelho (no modelo RGB para a codificação de cor) é calculada. O índice de verdor é expressado como a porcentagem de pixéis para a qual a razão verde-para-vermelho excede um dado limiar. Sob condições de crescimento normais, sob as condições de crescimento com tensão ao sal e sob condições de crescimento com disponibilidade de nutriente reduzida, o índice de verdor de plantas é medido na última formação de imagem antes do florescimento. Em contraste, sob condições de crescimento com tensão à seca, o índice de verdor das plantas é medido na primeira formação de imagem após a seca.

Planta

O termo "planta" como usado neste abrange plantas totais, predecessores e progênie das plantas e partes de plantas, incluindo sementes, brotos, caules, folhas, raízes (incluindo tubérculos), flores e tecidos e órgãos, em que cada um dos mencionados acima compreende o gene/ácido nucleico de interesse. O termo "planta" também abrange células de planta, culturas de suspensão, tecido de calo, embriões, regiões meristemáticas, gametófitos, esporófitos, pólen e microesporos, novamente em que cada um dos já

mencionados compreende o gene/ácido nucleico de interesse.

As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular plantas monocotiledôneas e dicotiledôneas incluindo legumes de forragem e alimentação, plantas ornamentais, safras alimentícias, árvores ou arbustos selecionados da lista que compreende *Acer* spp., *Actinidia* spp., *Abelmoschus* spp., *Agave sisalana, Agropyron* spp., *Agrostis stolonifera, Allium* spp., *Amaranthus* spp., *Ammophila arenaria, Ananas comosus, Annona* spp., *Apium graveolens, Arachis spp, Artocarpus* spp., *Asparagus officinalis, Avena* spp. *(por exemplo, Avena sativa, Avena fatua, Avena byzantina, Avena fatua var. sativa, Avena hybrida), Averrhoa carambola, Bambusa sp., Benincasa hispida, Bertholletia excelsea, Beta vulgaris, Brassica* spp. *(por exemplo, Brassica napus, Brassica rapa* ssp. [canola, óleo de semente de colza, nabo]), *Cadaba farinosa, Camellia sinensis, Canna indica, Cannabis sativa, Capsicum* spp., *Carex elata, Carica papaya, Carissa macrocarpa, Carya* spp., *Carthamus tinctorius, Castanea* spp., *Ceiba pentandra, Cichorium endivia, Cinnamomum* spp., *Citrullus lanatus, Citrus* spp., *Cocos* spp., *Coffea* spp., *Colocasia esculenta, Cola* spp., *Corchorus sp., Coriandrum sativum, Corylus* spp., *Crataegus* spp., *Crocus sativus, Cucurbita* spp., *Cucumis* spp., *Cynara* spp., *Daucus carota, Desmodium* spp., *Dimocarpus longan, Dioscorea* spp., *Diospyros* spp., *Echinochloa* spp., *Elaeis (por exemplo, Elaeis guineensis, Elaeis oleifera), Eleusine coracana, Erianthus sp., Eriobotrya japonica, Eucalyptus sp., Eugenia uniflora, Fagopyrum* spp., *Fagus* spp., *Festuca arundinacea, Ficus carica, Fortunella* spp., *Fragaria* spp., *Ginkgo biloba, Glycine* spp. *(por exemplo, Glycine max, Soja hispida ou Soja max), Gossypium hirsutum, Helianthus* spp. *(por exemplo, Helianthus annuus), Hemerocallis fulva, Hibiscus* spp., *Hordeum* spp. *(por exemplo, Hordeum vulgare), Ipomoea batatas, Juglans* spp., *Lactuca sativa, Lathyrus* spp., *Lens culinaris, Linum*

usitatissimum, Litchi chinensis, Lotus spp., *Luffa acutangula, Lupinus* spp., *Luzula sylvatica, Lycopersicon* spp. *(por exemplo, Lycopersicon esculentum, Lycopersicon lycopersicum, Lycopersicon pyriforme), Macrotyloma* spp., *Malus* spp., *Malpighia emarginata, Mammea americana, Mangifera indica, Manihot* spp., *Manilkara zapota, Medicago sativa, Melilotus* spp., *Mentha* spp., *Miscanthus sinensis, Momordica* spp., *Morus nigra, Musa* spp., *Nicotiana* spp., *Olea* spp., *Opuntia* spp., *Ornithopus* spp., *Oryza* spp. *(por exemplo, Oryza sativa, Oryza latifolia), Panicum miliaceum, Panicum virgatum, Passiflora edulis, Pastinaca sativa, Pennisetum sp., Persea* spp., *Petroselinum crispum, Phalaris arundinacea, Phaseolus* spp., *Phleum pratense, Phoenix* spp., *Phragmites australis, Physalis* spp., *Pinus* spp., *Pistacia vera, Pisum* spp., *Poa* spp., *Populus* spp., *Prosopis* spp., *Prunus* spp., *Psidium* spp., *Punica granatum, Pyrus communis, Quercus* spp., *Raphanus sativus, Rheum rhabarbarum, Ribes* spp., *Ricinus communis, Rubus* spp., *Saccharum* spp., *Salix sp., Sambucus* spp., *Secale cereale, Sesamum* spp., *Sinapis sp., Solanum* spp. *(por exemplo, Solanum tuberosum, Solanum integrifolium ou Solanum lycopersicum), Sorghum bicolor, Spinacia* spp., *Syzygium* spp., *Tagetes* spp., *Tamarindus indica, Theobroma cacao, Trifolium* spp., *Triticosecale rimpaui, Triticum* spp. *(por exemplo, Triticum aestivum, Triticum durum, Triticum turgidum, Triticum hybernum, Triticum macha, Triticum sativum ou Triticum vulgare), Tropaeolum minus, Tropaeolum majus, Vaccinium* spp., *Vicia* spp., *Vigna* spp., *Viola odorata, Vitis* spp., *Zea mays, Zizania palustris, Ziziphus* spp., entre outros.

Descrição detalhada da invenção

ERLK

Foi agora observado que modular a expressão em uma planta de um ácido nucleico que codifica um quinase semelhante ao receptor de extensina (ERLK) ou uma parte deste que compreende pelo menos o domínio de quinase e o domínio de transmembrana, dá plantas tendo características

intensificadas relacionadas com o rendimento com relação às plantas de controle.

Portanto, de acordo com uma primeira forma de realização, a invenção fornece um método para intensificar características relacionadas com o rendimento em plantas com relação às plantas de controle, que compreende modular a expressão em uma planta de um ácido nucleico que codifica uma proteína ERLK ou uma parte da mesma.

Um método preferido para a modulação da (preferivelmente, aumento) expressão de um ácido nucleico que codifica um quinase semelhante ao receptor de extensina (ERLK) ou uma parte deste que compreende pelo menos o domínio de quinase e o domínio transmembrana, ocorre pela introdução e expressão em uma planta, um ácido nucleico que codifica uma tal proteína de ERLK.

Qualquer referência a seguir a uma "proteína útil nos métodos da invenção" é entendido significar um polipeptídeo de ERLK como definido neste. Qualquer referência a seguir a um "ácido nucleico útil nos métodos da invenção" é entendido significar um ácido nucleico capaz de codificar um tal polipeptídeo de ERLK. O ácido nucleico a ser introduzido em uma planta (e portanto útil na realização dos métodos da invenção) é qualquer ácido nucleico que codifica o tipo de proteína que não será descrito, a seguir, também denominado "ácido nucleico ERLK" ou "gene ERLK".

A proteína de ERLK úteis nos métodos da presente invenção é uma proteína de ERLK como definido por Shiu and Bleeker (2001). O termo "proteína de ERLK" ou "quinase semelhante ao receptor de extensina" refere-se a uma proteína que compreende um domínio de quinase e no termina N deste um domínio de transmembrana (ver Figura 1 e Figura 2 para uma visão esquemática geral). As proteínas de ERLK preferivelmente, também compreendem uma seqüência de sinal de secreção de terminal N e, opcionalmente um domínio extracelular. Preferivelmente o domínio de

quinase (e/ou outros domínios) de uma proteína de ERLK útil na presente invenção classifica como uma quinase semelhante ao receptor de extensina como definido por Shiu and Bleeker (2001).

O termo "domínio" e "motivo" são definidos na seção "definições" neste. os bancos de dados especialistas existem para a identificação de domínios, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-244, InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318, Prosite (Bucher and Bairoch (1994), uma sintaxe de perfil generalizada para motivos de seqüências biomoleculares e sua função na interpretação da seqüência automática. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004), or Pfam (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002). Uma série de ferramentas para a análise *in silico* de seqüências de proteína está disponível no servidor ExPASy proteomics (Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31:3784-3788(2003)). Os domínios também podem ser identificados usando-se técnicas de rotina, tais como pelo alinhamento de seqüência.

Os métodos para o alinhamento de seqüências para comparação são bem conhecidos na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. GAP usa o algoritmo de Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar o alinhamento global (isto é, medindo-se as seqüências completas) de duas seqüências que maximizam o número de comparações e minimiza o número de fendas. O algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-10) calcula a porcentagem de identidade de seqüência e realiza uma análise estatística da

similaridade entre as duas seqüências. O software para realizar a análise está publicamente disponível através do National Centre for Biotechnology Information (NCBI). Os homólogos podem ser facilmente identificados usando-se, por exemplo, o algoritmo de alinhamento de seqüência múltiplo ClustalW (versão 1.83), com os parâmetros de alinhamento em pares padrão e um método de registro em porcentagem. As porcentagens globais de similaridade e identidade também podem ser determinados usando-se um dos métodos disponíveis no pacote de software MatGAT (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando-se proteína ou seqüências de DNA.). A edição manual menor pode ser realizada para otimizar o alinhamento entre os motivos conservados, como deve estar evidente a uma pessoa habilitada na técnica. Além disso, em vez de usar seqüências de comprimento total para a identificação de homólogos, domínios específicos (tais como o domínio de quinase) também podem ser usados. Os valores de identidade de seqüência podem ser determinados sobre a seqüência de ácido nucleico ou aminoácido total ou sobre os domínios selecionados ou motivos conservados, usando-se os programas mencionados acima usando-se os parâmetros padrão.

O domínio de quinase nas proteínas de ERLK úteis nos métodos da presente invenção é um domínio tipo proteína Tyr quinase (entrada Pfam PF07714, entrada InterPro IPR001245). O local ativo corresponde à assinatura PROSITE PS00109, com o seguinte padrão de consenso:

[LIVMFYC]-{A}-[HY]-x-D-[LIVMFY]-[RSTAC]-{D}-{PF}-N-[LIVMFYC] (3),

em que D é parte do local ativo. A sintaxe deste padrão é de acordo com as convenções usadas nos bancos de dados Prosite e é explicado no manual PROSITE.

Preferivelmente, o domínio de quinase ainda é caracterizado pela presença do motivo de seqüência 1 (SEQ ID N° : 6):

(M/L)L(S/G/R)R(L/M)(H/R/Q)(H/S/C)(R/P)(N/Y)L(L/V)XL(I/L/V)G

em que X pode ser qualquer aminoácido, preferivelmente um de K,N,A,S,G,E. Preferivelmente, o motivo 1 tem a seqüência LLSR(L/M)(H/R/Q)(C/S)PYL(L/V)(E/G/A)L(L/I); mais preferivelmente o motivo 1 tem a seqüência LLSRLQCPYLVELLG.

Preferivelmente, o domínio de quinase também compreende um ou mais de motivo de seqüência 2 (SEQ ID N° : 7):

L(Y/D/N)(W/F)X(A/V/T)R(L/M)(L/R/G)IA(L/V)

em que X pode ser qualquer aminoácido, preferivelmente um de D,N,E,K,P,Q ou G,

motivo de seqüência 3 (SEQ ID N° : 8):

A(R/K)(A/G)L(A/E)(Y/F)LHE,

motivo de seqüência 4 (SEQ ID N° : 9):

(V/I)IHR(D/N)(F/L)K(S/A/G/C)(S/T)N(I/V)LL(E/D)

em que o aminoácido na posição 6 não é preferivelmente I,V ou M,

motivo de seqüência 5 (SEQ ID N° : 10):

(K/R)V(S/A/T)DFG(L/M/S)

Preferivelmente, o motivo de seqüência 2 tem a seqüência LDW(G/Q/P/E)(A/T)R(L/M)(R/G)IA(L/V), mais preferivelmente, o motivo de seqüência 2 tem a seqüência LDW(G/Q)(T/A)RL(R/G)IAL, mais preferivelmente, motivo de seqüência 2 tem a seqüência LDWGARLRIAL. Motivo de seqüência 3 preferivelmente tem a seqüência ARALEFLHE. Motivo de seqüência 4 preferivelmente tem a seqüência VIHR(D/N)(F/L)K(S/C)(S/T)NILLD, mais preferivelmente, a seqüência é VIHRNFKCTNILLD. Motivo de seqüência 5 preferivelmente tem a seqüência (K/R)VSDFG(L/M), mais preferivelmente a seqüência é KVSDFGL.

Preferivelmente, o domínio de quinase de proteínas de ERLK úteis nos métodos da presente invenção tem, em ordem crescente, de

preferência, pelo menos 39%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de seqüência para o domínio de quinase de SEQ ID N° : 2 (como dado na SEQ ID N° : 57). Um domínio de quinase pode ser identificado usando-se os bancos de dados e as ferramentas para a identificação de proteína listada acima e/ou métodos para o alinhamento de seqüências para a comparação. Em alguns exemplos, os parâmetros padrão podem ser ajustados para modificar a estringência da pesquisa. Por exemplo, usando-se BLAST, o limiar de significância estatística (denominado valor "esperado") para relatar comparações contra as seqüências de bancos de dados podem ser aumentado para mostrar comparações menos estringentes. Desta maneira, as comparações quase exatas curtas podem ser identificadas.

Os domínios transmembrana são de cerca de 15 a 30 aminoácidos de comprimento e são usualmente compostos de resíduos hidrofóbicos que formam uma alfa hélice. Estes são usualmente preditos na base de hidrofobicidade (por exemplo Klein et al., Biochim. Biophys. Acta 815, 468, 1985; ou Sonnhammer et al., In J. Glasgow, T. Littlejohn, F. Major, R. Lathrop, D. Sankoff, and C. Sensen, editors, Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology, pages 175-182, Menlo Park, CA, 1998, AAAI Press).

O domínio extracelular de uma proteína de ERLK, se presente, pode (mas não necessita necessariamente) ter um ou mais motivos SP_x. A estrutura de seqüências de sinal de secreção e a predição de seus locais de clivagem são bem conhecidos na técnica.

Além disso, as proteínas de ERLK úteis nos métodos da presente invenção (pelo menos em sua forma natural) tipicamente, mas não necessariamente, tem atividade de quinase. Portanto, proteínas de ERLK com atividade de quinase reduzida ou sem atividade de quinase pode ser igualmente úteis nos métodos da presente invenção. Uma pessoa habilitada na

técnica pode determinar facilmente a presença da atividade de quinase usando-se as ferramentas as técnicas de rotina. Para determinar a atividade de quinase de receptores como a quinase, vários ensaios estão disponíveis (por exemplo Current Protocols in Molecular Biology, Volumes 1 e 2, Ausubel et al. (1994), Current Protocols). Em resumo, um ensaio de quinase, no geral, envolve (1) colocar a proteína quinase em contato com um polipeptídeo do substrato contendo o local alvo a ser fosforilado; (2) permitir a fosforilação do local alvo em um tampão de quinase apropriado sob as condições apropriadas; (3) separando-se os produtos fosforilados do substrato não fosforilado após um período de reação adequado. A presença ou a ausência de atividade de quinase é determinada pela presença ou ausência um alvo fosforilado. Além disso, as medições quantitativas podem ser realizadas. Os receptores purificados como a quinase ou extratos celulares contendo ou enriquecidos no receptor como a quinase podem ser usados como fonte da proteína quinase. Alternativamente, o método de Zhao *et al.* (Plant Mol. Biol. 26, 791-803, 1994) pode ser usado, quando o domínio citoplásmico de um receptor de arroz como a quinase foi expressado em *Escherichia coli* estimado para a atividade da quinase. Como um substrato, os peptídeos pequenos são particularmente bem adaptados. O peptídeo deve compreender uma ou mais resíduos de serina, treonina ou tirosina em um motivo de local de fosforilação. Uma compilação de locais de fosforilação pode ser encontrada em Biochimica et Biophysica Acta 1314, 191-225, (1996). Além disso, os substratos de peptídeo podem ter vantajosamente uma carga positiva de rede para facilitar a ligação aos filtros de fosfocelulose, (permitindo a separação dos peptídeos fosforilados dos não fosforilados e para detectar os peptídeos fosforilados). Se um motivo de local de fosforilação não for conhecido, um substrato de tirosina quinase geral pode ser usado. Por exemplo, "peptídeo relacionado com Src" (RRLIEDAEYAARG) é um substrato para muitas tirosina quinases receptoras e não receptoras). Para determinar os parâmetros cinéticos para a

fosforilação do peptídeo sintético, uma faixa de concentração de peptídeo é requerida. Para reações iniciais, uma concentração de peptídeo de 0,7 a 1,5 mM pode ser usada. Para cada enzima quinase, é importante determinar o tampão ótimo, força iônica e o pH para a atividade. Um tampão de quinase 5x padrão, no geral, contém 5 mg/ml de BSA (Albumina de soro Bovino que evita a absorção da quinase no tubo de ensaio), 150 mM de Tris-Cl (pH 7,5), 100 mM de $MgCl_2$. Os cátions bivalentes são requeridos para a maioria das tirosina quinases, embora algumas tirosina quinases (por exemplo, quinases receptoras de insulina, IGF-1 e PDGF) requerem $MnCl_2$ em vez de $MgCl_2$ (ou em adição ao $MgCl_2$). As concentrações ótimas de cátions bivalentes devem ser determinadas empiricamente para cada proteína quinase. Um doador comumente usado para o grupo fosforila é radio-rotulado [gama-^{32}P]ATP (normalmente em 0,2 mM de concentração final). A quantidade de ^{32}P incorporados nos peptídeos pode ser determinada medindo-se a atividade nas almofadas secas de nitrocelulose em um contador de cintilação.

A presente invenção é ilustrada pela transformação das plantas com a seqüência de ácido nucleico representada por SEQ ID N° : 1, que codifica a seqüência de polipeptídeo da SEQ ID N° : 2. Entretanto, o desempenho da invenção não está restrita àquelas seqüências; os métodos da invenção podem ser vantajosamente realizados usando-se qualquer ácido nucleico que codifica ERLK ou polipeptídeo de ERLK como definido neste.

Os exemplos de ácidos nucleicos que codificam polipeptídeos ERLK são dados na Tabela A do Exemplo 1 neste. Tais ácidos nucleicos são úteis na realização dos métodos da invenção. As seqüências de aminoácido dadas na Tabela A do Exemplo 1 são seqüências exemplos de ortólogos e parálogos do polipeptídeo de ERLK representados pela SEQ ID N° : 2, os termos "ortólogos" e "parálogos" sendo como definido neste. Os ortólogos e parálogos podem ser facilmente identificados pela realização de uma denominada pesquisa blast recíproca. Tipicamente, esta envolve um primeiro

BLAST que envolve BLASTing uma seqüência em questão (por exemplo usando-se qualquer uma das seqüências listadas na Tabela A do Exemplo 1) contra qualquer banco de dados de seqüência, tais como o banco de dados NCBI publicamente conhecido. BLASTN ou TBLASTX (usando-se valores de default padrão) são, no geral, usados quando começa-se a partir da seqüência de nucleotídeo e BLASTP ou TBLASTN (usando-se os valores de default padrão) quando começa-se a partir de uma seqüência de proteína. Os resultados BLAST podem ser opcionalmente filtrados. As seqüências de comprimento total dos resultados filtrados ou dos resultados não filtrados são então novamente submetidos ao BLAST (segundo BLAST) contra as seqüências do organismo a partir do qual a seqüência em questão é derivada (quando a seqüência em questão é a SEQ ID N° : 1, SEQ ID N° : 2, SEQ ID N° : 11, SEQ ID N° : 12, SEQ ID N° : 13 ou SEQ ID N° : 14, o segundo BLAST, portanto, deve ser contra as seqüências de *Arabidopsis*). Os resultados do primeiro e segundo BLASTs são então comparados. Um parálogo é identificado se um acerto de classificação alta do primeiro blast for da mesma espécie como a partir do qual a seqüência em questão é derivada, novamente um BLAST então resulta idealmente na seqüência em questão entre os acertos mais altos; um ortólogo é identificado se um acerto de classificação mais alto no primeiro BLAST não for da mesma espécie como aquela que a seqüência em questão é derivada e, preferivelmente, resulta novamente no BLAST na seqüência em questão estando entre os acertos mais altos.

Os acertos de classificação alta são aqueles tendo um valor E baixo. Quanto menor o valor E, mais significante o registro (ou em outras palavras (quanto menor a oportunidade de o acerto ser encontrado por oportunidade). A computação do valor E é bem conhecido na técnica. Além dos valores E, as comparações também são registradas pela porcentagem de identidade. A porcentagem de identidade refere-se ao número de seqüências

de nucleotídeos idênticos (ou aminoácido) entre os dois ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. No caso de famílias grandes, o ClustalW pode ser usado, seguido por uma árvore de união vizinha, para ajudar à visualização do agrupamento de genes relacionados e para identificar ortólogos e parálogos.

As variantes de ácido nucleico também podem ser úteis na prática dos métodos da invenção. os exemplos de tais variantes incluem ácidos nucleicos que codificam homólogos e derivados de qualquer uma das seqüências de aminoácido dadas na Tabela A do Exemplo 1, os termos "homólogo" e "derivado" sendo como definido neste. Também são úteis nos métodos da invenção, os ácidos nucleicos que codificam homólogos e derivados de ortólogos ou parálogos de qualquer uma das seqüências de aminoácido dadas na Tabela A do Exemplo 1. Os homólogos e derivados úteis nos métodos da presente invenção tem, substancialmente a mesma atividade biológica e funcional como a proteína não modificada da qual estes são derivados.

As variantes de ácido nucleico adicionais úteis na prática dos métodos da invenção incluem porções de ácidos nucleicos que codificam polipeptídeos de ERLK, ácidos nucleicos que codificam a ácidos nucleicos para ácidos nucleicos que codificam polipeptídeos ERLK, variantes de união de ácidos nucleicos que codificam polipeptídeos de ERLK, variantes alélicas de ácidos nucleicos que codificam polipeptídeos de ERLK e variantes de ácidos nucleicos que codificam polipeptídeos de ERLK obtidos por mistura de gene. Os termos que hibridizam seqüência, variante de união, variante alélica e mistura de gene são como descritos neste.

Os ácidos nucleicos que codificam polipeptídeos de ERLK não devem ser ácidos nucleicos de comprimento total, visto que o desempenho dos métodos da invenção não conta com o uso de seqüências de ácido nucleico de comprimento total. De acordo com a presente invenção, é

fornecido um método para intensificar características relacionadas com o rendimento em plantas, que compreende introduzir e expressar em uma planta, uma porção de qualquer uma das seqüências de ácido nucleico dadas na Tabela A do Exemplo 1 ou uma porção de um ácido nucleico que codifica um homólogo, parálogo ou homólogo de qualquer uma das seqüências de aminoácido dadas na Tabela A do Exemplo 1. Preferivelmente, a porção codifica um polipeptídeo que compreende pelo menos, de terminal N ao terminal C, (i) um domínio de transmembrana e (ii) uma quinase semelhante ao receptor de domínio de quinase tipo extensina (tipo ERLK).

Uma porção de um ácido nucleico pode ser preparada, por exemplo, realizando-se uma ou mais anulações ao ácido nucleico. As porções podem ser usadas na forma isolada ou estes podem ser fundidos a outras seqüências de codificação (ou não codificadoras) a fim de, por exemplo, produzir uma proteína que combina diversas atividades. Quando fundido a outras seqüências codificadoras, o polipeptídeo resultante produzido na tradução pode ser maior do que aquele predito para a porção de proteína.

As porções úteis nos métodos da invenção, codificam um polipeptídeo de ERLK como definido neste, tendo um domínio de quinase (como descrito acima) e tendo substancialmente a mesma atividade biológica como a proteína de ERLK representada por qualquer uma das seqüências de aminoácido dadas na Tabela A do Exemplo 1. Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos dados na Tabela A do Exemplo 1 ou é uma porção de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das seqüências de aminoácido dadas na Tabela A do Exemplo 1. Preferivelmente, a porção é de pelo menos 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300 nucleotídeos consecutivos de comprimento, os nucleotídeos consecutivos sendo de qualquer uma das seqüências de ácido nucleico dadas na Tabela A do Exemplo 1 ou de um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das

seqüências de aminoácido dadas na Tabela A do Exemplo 1. Mais preferivelmente a porção é uma porção do ácido nucleico da SEQ ID N° : 1.

Uma outra variante de ácido nucleico nos métodos da invenção é um ácido nucleico capaz de hibridizar, sob condições de estringência reduzida, preferivelmente sob condições estringentes, com um ácido nucleico que codifica um polipeptídeo de ERLK como definido neste ou com uma porção como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas, que compreende introdução e expressão em uma planta de um ácido nucleico capaz de hibridizar a qualquer um dos ácidos nucleicos dados na Tabela A do Exemplo 1 ou que compreende introdução e expressão em uma planta de um ácido nucleico capaz de hibridizar a um ácido nucleico que codifica um homólogo, parálogo ou homólogo de qualquer uma das seqüências de ácido nucleico dadas na Tabela A do Exemplo 1.

As seqüências de hibridização úteis nos métodos da invenção codificam um polipeptídeo de ERLK como definido neste, tendo um domínio de quinase do tipo ERLK e um domínio de transmembrana (como descrito acima) e tendo substancialmente a mesma atividade biológica como as seqüências de aminoácido dadas na Tabela A do Exemplo 1. Seqüência de hibridização é tipicamente de pelo menos 800 nucleotídeos de comprimento, preferivelmente pelo menos 1000 nucleotídeos de comprimento, mais preferivelmente pelo menos 1200 nucleotídeos de comprimento e mais preferivelmente pelo menos 1300 nucleotídeos de comprimento. Preferivelmente, a seqüência de hibridização é capaz de hibridização a qualquer um dos ácidos nucleicos dados na Tabela A do Exemplo 1 ou a uma porção de qualquer uma destas seqüências, uma porção como sendo definida acima ou a seqüência de hibridização é capaz de hibridizar a um ácido nucleico que codifica um ortólogo ou parálogo de qualquer uma das

seqüências de aminoácido dadas na Tabela A do Exemplo 1 ou a sondas ou a sondas derivadas destes. Mais preferivelmente, a seqüência de hibridização é capaz de hibridizar a um ácido nucleico como representado pela SEQ ID N° : 1 ou uma porção ou sonda destes.

Os métodos para projetar sondas são bem conhecidos na técnica. As sondas são, no geral menores do que 1000 bp de comprimento, preferivelmente menores do que 500 bp de comprimento. Comumente, os comprimento da onda para hibridizações de DNA-DNA tais como Southern blotting, variam entre 100 e 500 bp, visto que a região de hibridização na sonda para hibridizações de DNA-DNA, tais como amplificação de PCR no geral, é mais curta do que 50 mas mais longa do que 10 nucleotídeos.

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante de união que codifica um polipeptídeo de ERLK como definido neste acima, a variante de união sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas, que compreende introduzir e expressar, em uma planta, uma variante de união de qualquer uma das seqüências de ácido nucleico dadas na Tabela A do Exemplo 1 ou a variante de união de um ácido nucleico que codifica um homólogo, parálogo ou homólogo de qualquer uma das seqüências dadas na Tabela A do Exemplo 1.

As variantes de união preferidas são variantes de união de um ácido nucleico representado por SEQ ID N° : 1 ou a variante de união de um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N° : 2.

Uma outra variante de ácido nucleico útil na realização dos métodos da invenção é uma variante alélica de um ácido nucleico que codifica um polipeptídeo de ERLK como definido neste acima, uma variante alélica sendo como definida neste.

De acordo com a presente invenção, é fornecido um método

para intensificar características relacionadas com o rendimento em plantas, que compreende introduzir e expressar em uma planta uma variante alélica de qualquer um dos ácidos nucleicos dados na Tabela A do Exemplo 1 ou que compreende introduzir e expressar em uma planta uma variante alélica de um ácido nucleico que codifica um homólogo, parálogo ou homólogo de qualquer uma das seqüências dadas na Tabela A do Exemplo 1.

As variantes alélicas úteis nos métodos da presente invenção tem, substancialmente, a mesma atividade biológica como o polipeptídeo de ERLK da SEQ ID N° : 2 e qualquer um dos aminoácidos descritos na Tabela A de Exemplo 1. As variantes alélicas existem em estado natural, e abrangido dentro dos métodos da presente invenção está o uso destes alelos naturais. Preferivelmente, a variante alélica é uma variante alélica da SEQ ID N° : 1 ou uma variante alélica de um ácido nucleico que codifica um ortólogo ou parálogo da SEQ ID N° : 2.

Uma variante de ácidos nucleico adicional úteis nos métodos da invenção é uma variante de ácido nucleico obtida por mistura de gene. A mistura de gene ou evolução direcionada também podem ser usadas para gerar variantes de ácidos nucleicos que codificam polipeptídeos de ERLK como definido acima; o termo "mistura de gene" sendo como definido neste.

De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas, que compreende introduzir e expressar em uma planta a variante de qualquer uma das seqüências de ácido nucleico dadas na Tabela A do Exemplo 1 ou que compreende introduzir e expressar em uma planta a variante de um ácido nucleico que codifica um homólogo, parálogo ou homólogo de qualquer uma das seqüências dadas na Tabela A do Exemplo 1, cujo ácido nucleico variante é obtido por mistura de gene.

Além disso, as variantes de ácido nucleico também podem ser obtidas por mutagênese direcionada ao local. Diversos métodos estão

disponíveis para atingir a mutagênese direcionada ao local, o mais comum sendo métodos com base em PCR (Current Protocols in Molecular Biology. Wiley Eds.).

Os ácidos nucleicos que codificam polipeptídeos de ERLK podem ser derivados de qualquer fonte natural ou artificial. O ácido nucleico pode ser modificado a partir de sua forma natural em composição e/ou ambiente genômico através da manipulação humana deliberada. Preferivelmente, o ácido nucleico que codifica o polipeptídeo de ERLK é de uma planta, ainda preferivelmente de uma planta dicotiledônea, mais preferivelmente da família Brassicaceae, mais preferivelmente o ácido nucleico é de *Arabidopsis thaliana*.

O desempenho dos métodos da invenção dá plantas tendo características relacionadas com o rendimento intensificadas. Em particular desempenho dos métodos da invenção dá plantas tendo rendimento aumentado, especialmente, rendimento de semente aumentado com relação às plantas de controle. Os termos "rendimento" e "rendimento de semente" como descrito em mais detalhes na seção de "definições" neste.

Referência neste a características intensificadas relacionadas com o rendimento é entendido significar um aumento na biomassa (peso) de uma ou mais partes de uma planta, que podem incluir partes (colhíveis) acima do solo e/ou partes (colhíveis) abaixo do solo. Em particular, tais partes colhíveis são biomassa de sementes e folhas e o desempenho dos métodos da invenção resulta em plantas tendo biomassa de folhas aumentada e rendimento de semente aumentado, com relação às plantas de controle.

Tomando o milho como um exemplo, um aumento no rendimento pode ser manifestado como um ou mais dos seguintes: aumento no número de plantas estabelecidas por hectare ou acre, um aumento no número de espigas por planta, um aumento no número de séries, número de sementes por série, peso de semente, peso de mil núcleos,

diâmetro/comprimento de espiga, aumento na taxa de enchimento de semente (que é o número de sementes enchidas dividido pelo número total de sementes e multiplicado por 100), entre outros. Tomando-se o arroz como um exemplo, um aumento no rendimento pode manifestar-se por si só como um aumento em um ou mais dos seguintes: número de plantas por hectare ou acre, número de panículas por planta, número de espigas por panícula, número de flores (flósculos) por panícula (que é expressado como uma razão do número de sementes enchidas sobre o número de panículas primárias), aumento na taxa de enchimento de semente (que é o número de sementes enchidas dividido pelo número total de sementes e multiplicado por 100), aumento no peso de mil núcleos, entre outros.

A presente invenção fornece um método para aumentar o rendimento, especialmente biomassa folhosa aumentada e rendimento aumentado de semente de plantas, com relação às plantas de controle, cujo método compreende modular a expressão, preferivelmente aumentar a expressão, em uma planta, de um ácido nucleico, que codifica um polipeptídeo de ERLK como definido neste.

Visto que as plantas transgênicas de acordo com a presente invenção tem rendimento aumentado, é provável que estas plantas apresentem uma taxa de crescimento aumentada (durante pelo menos parte de seu ciclo de vida), com relação à taxa de crescimento de plantas de controle em um estágio correspondente em seu ciclo de vida.

A taxa de crescimento aumentada pode ser específica a uma ou mais partes de uma planta (incluindo sementes) ou pode ser, substancialmente a planta toda. As plantas tendo uma taxa de crescimento aumentada podem ter um ciclo de vida mais curto. O ciclo de vida de uma planta pode significar o tempo necessário para o crescimento de uma semente madura seca até o estágio onde a planta produziu sementes maduras secas, similar ao material de partida. O ciclo de vida pode ser influenciado por fatores tais como vigor

precoce, taxa de crescimento, índice de verdor, tempo de florescimento e velocidade de maturação da semente. O aumento na taxa de crescimento pode acontecer em um ou mais estágios no ciclo de vida de uma planta ou substancialmente durante todo o ciclo de vida da planta. A taxa de crescimento aumentada durante os estágios precoces no ciclo de vida de uma planta pode refletir vigor intensificado. A taxa de crescimento aumentada durante os estágios precoces no ciclo de vida de uma planta pode refletir vigor intensificado. O aumento na taxa de crescimento pode alterar o ciclo de colheita de uma planta permitindo que as plantas sejam semeadas depois e/ou colhidas mais tarde do que de outra maneira seria possível (um efeito similar pode ser obtido with tempo de florescimento precoce). Se a taxa de crescimento for suficientemente aumentada, isto pode permitir a semeadura adicional de sementes da mesma espécie vegetal (por exemplo semeadura e colheita de plantas de arroz seguido pela semeadura e colheita de plantas de arroz adicionais todas dentro de um período de crescimento convencional). Similarmente, se a taxa de crescimento for suficientemente aumentada, isto pode permitir a semeadura adicional de sementes de espécies vegetais diferentes (por exemplo a semeadura e a colheita de plantas de milho seguido por, por exemplo, a semeadura e colheita opcional de soja, batata ou qualquer outra planta adequada). Períodos adicionais de colheita a partir do mesmo estoque de raiz no caso das mesmas plantas de safra também pode ser possível. A alteração do ciclo de colheita de uma planta pode levar a um aumento no na produção de biomassa anual por acre (devido a um aumento no número de vezes (dito em um ano) que qualquer planta particular pode ser desenvolvida e colhida). Um aumento no taxa de crescimento também pode permitir o cultivo de plantas transgênicas em uma área geográfica mais ampla do que suas contrapartes do tipo selvagem, visto que as limitações territoriais para o crescimento de uma safra são freqüentemente determinadas por condições ambientais adversas no período de plantio (estação precoce) ou no

período de colheita (estação posterior). Tais condições adversas podem ser evitadas se o ciclo de colheita for encurtado. A taxa de crescimento pode ser determinada derivando-se vários parâmetros de curvas de crescimento, tais parâmetros podem ser: T-Mid (o tempo tomado para as plantas atingirem 50% de seu tamanho máximo) e T-90 (tempo tomado para as plantas atingirem 90% de seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção dá plantas tendo uma taxa de crescimento aumentada com relação às plantas de controle. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a taxa de crescimento de plantas, cujo método compreende modular a expressão, preferivelmente aumentar a expressão, em uma planta, de um ácido nucleico, que codifica um polipeptídeo de ERLK como definido neste.

Um aumento no rendimento e/ou taxa de crescimento ocorre quando a planta não está sob condições de tensão ou quando a planta está exposta a várias tensões em comparação com plantas de controle. As Plantas respondem tipicamente à exposição à tensão desenvolvendo-se mais lentamente. Nas condições de tensão severas, a planta pode parar de se desenvolver completamente. Tensão branda, por outro lado é definido neste como sendo qualquer tensão à qual uma planta é exposta que não resulta na interrupção do crescimento da planta completamente sem a capacidade de retomar o crescimento. Tensão branda no sentido da invenção leva à redução no crescimento de plantas submetidas à tensão menor do que 40%, 35% ou 30%, preferivelmente menor do que 25%, 20% ou 15%, mais preferivelmente menor do que 14%, 13%, 12%, 11% ou 10% ou menos em comparação com a planta de controle sob condições sem tensão. Devido a avanços nas práticas agrícolas, (irrigação, fertilização, tratamentos com pesticida) tensões severas não são freqüentemente encontradas em plantas de safra cultivadas. Como uma conseqüência, o crescimento comprometido induzido pela tensão branda

é freqüentemente uma característica indesejável para a agricultura. As tensões brandas são as tensões bióticas e/ou abióticas diárias (ambientais) Às quais a planta está exposta. As tensões abióticas podem ocorrer devido à seca ou água em excesso, tensões anaeróbicas, tensões ao sal, toxicidade química, tensões oxidativas e calor, frio ou temperaturas congelantes. A tensão abiótica pode ser uma tensão osmótica causada por uma tensão por água (particularmente devido à seca), tensões ao sal, tensões oxidativas ou uma tensão iônica. As Tensões bióticas são tipicamente aquelas tensões causadas por patogênios, tais como bactérias, vírus, fungos e insetos.

Em particular, os métodos da presente invenção podem ser realizados sob condições sem tensão ou sob condições de seca branda a dadas plantas tendo rendimento aumentado com relação às plantas de controle. Como relatado em Wang et al. (Planta (2003) 218: 1-14), a tensão abiótica leva a uma série de mudanças morfológicas, fisiológicas, bioquímicas e moleculares que afetam adversamente o crescimento e a produtividade da planta. Tensões de seca, salinidade, temperaturas extremas e oxidativas are conhecidas estarem interconectadas e pode induzir o crescimento e o dano celular através de mecanismos celulares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve um grau particularmente alto de "*cross talk*" entre tensão por seca tensão por salinidade alta. Por exemplo, seca e/ou salinização são manifestadas primariamente como tensão osmótica, resultando na interrupção da homeostase e distribuição iônica na célula. Tensão oxidativa, que freqüentemente acompanha temperatura alta ou baixa, tensão por salinidade ou seca, pode causar a desnaturação de proteínas funcionais ou estruturais. Como uma conseqüência, estas tensões ambientais diversas freqüentemente ativam caminhos de sinalização celular similares e respostas celulares, tal como a produção de proteínas de tensão, super-regulação de anti-oxidantes, acúmulo de solutos compatíveis ou interrupção de crescimento. O termo condições "sem tensão" como usado neste são aquelas

condições ambientais que permitem o crescimento ótimo de plantas. As pessoas habilitadas na técnica estão cientes das condições de solo e condições climáticas normais para um dado local.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições sem tensão ou sob condições de seca brandas com rendimento aumentado com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas desenvolvidas sob condições sem tensão ou sob condições de seca brandas, cujo método compreende aumentar a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo de ERLK.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições de deficiência de nutriente, particularmente sob condições de deficiência de nitrogênio, rendimento aumentado com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas desenvolvidas sob condições de deficiência de nutriente, cujo método compreende aumentar a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo de ERLK. A deficiência de nutriente pode resultar de uma falta ou de um excesso de nutrientes tais como nitrogênio, fosfatos e outros compostos contendo fósforo, potássio, cálcio, cádmio, magnésio, manganês, ferro e boro, entre outros.

A presente invenção abrange plantas ou partes destas (incluindo sementes) obteníveis pelos métodos de acordo com a presente invenção. As plantas ou partes destas compreendem um transgene de ácido nucleico que codifica um polipeptídeo de ERLK como definido acima.

A invenção também fornece construções genéticas e vetores para facilitar a introdução e/ou expressão em plantas de ácidos nucleicos que codificam polipeptídeos de ERLK. As construções genéticas podem ser

inseridas em vetores, que podem ser comercialmente disponíveis, adequados para a transformação em plantas e adequados para a expressão do gene de interesse nas células transformadas. A invenção também fornece o uso de uma construção genética como definido neste nos métodos da invenção.

Mais especificamente, a presente invenção fornece uma construção que compreende:

(a) um ácido nucleico que codifica um polipeptídeo de ERLK como definido acima;

(b) uma ou mais seqüências de controle capazes de conduzir a expressão da seqüência de ácido nucleico de (a); e opcionalmente

(c) uma seqüência de terminação de transcrição.

Preferivelmente, o ácido nucleico que codifica um polipeptídeo de ERLK é como definido acima. O termo "seqüência de controle" e "seqüência de terminação" são como definidos neste.

As plantas são transformadas com um vetor que compreende qualquer um dos ácidos nucleicos descritos acima. O técnico habilitado está bem ciente dos elementos genéticos que deve estar presente no vetor a fim de transformar, de maneira bem sucedida,, selecionar e propagar células hospedeiras contendo a seqüência de interesse. A seqüência de interesse está operacionalmente ligada a uma ou mais seqüências de controle (pelo menos a um promotor).

Vantajosamente, qualquer tipo de promotor, natural ou sintético, pode ser usado para conduzir a expressão da seqüência de ácido nucleico. Um promotor constitutivo é particularmente úteis nos métodos. Ver a seção de "Definições" aqui para definição dos vários tipos promotores. Um promotor constitutivo preferido é um promotor constitutivo que também é substancialmente expressado ubiquitosamente. Além disso, preferivelmente, o promotor é derivado de uma planta, mais preferivelmente uma planta monocotiledônea. Mais preferido é o uso de um promotor GOS2,

substancialmente similar ou idêntico para o promotor GOS2 do arroz (SEQ ID N° : 5 ou SEQ ID N° : 58).

Deve estar claro que aplicabilidade da presente invenção não está restrita ao ácido nucleico que codifica o polipeptídeo de ERLK representado por SEQ ID N° : 1, é a aplicabilidade da invenção restrita à expressão de um polipeptídeo de ácido nucleico codificador de ERLK quando conduzido por um promotor constituído. Ver a tabela 2a na seção de "definições" neste para exemplos adicionais de promotor constituídos.

Opcionalmente, uma ou mais seqüências terminadoras podem ser usadas na construção introduzida em uma planta. Os Elementos reguladores adicionais podem incluir intensificadores transcripcionais bem como translacionais. Aqueles habilitados na técnica estarão cientes de seqüências terminadoras e intensificadoras que podem ser adequados para o uso na realização da invenção. Uma seqüência de íntron também pode ser adicional à região não traduzida 5' (UTR) ou na seqüência codificadora para aumentar a quantidade da mensagem madura que acumula no citosol, como descrito na seção de definições. Outras seqüências de controle (além do promotor, intensificador, silenciador, seqüências de íntron, regiões 3'UTR e/ou 5'UTR) podem ser proteína e/ou elementos estabilizadores de RNA. Tais seqüências devem ser conhecidas ou podem ser facilmente obtidas por uma pessoa habilitada na técnica.

As construções genéticas da invenção ainda podem incluir uma origem de seqüência de replicação que é requerida par a manutenção e/ou replicação em um tipo celular específico. Um exemplo é quando uma construção genética é requerida para ser mantida em uma célula bacteriana como um elemento genético epissômico (por exemplo, molécula de plasmídeo ou cosmídeo). As origens preferidas de replicação inclui, mas não são limitados ao f1-ori e colE1.

Para a detecção da transferência bem sucedida das seqüências

de ácido nucleico como usado nos métodos da invenção e/ou seleção de plantas transgênicas que compreende estes ácidos nucleicos, é vantajoso usar genes marcadores (ou genes repórteres). Portanto, a construção genética pode compreender opcionalmente um gene marcador selecionável. Os marcadores selecionáveis como descrito em mais detalhes na seção de "definições" neste. Os genes marcadores podem ser removidos ou excisados a partir da célula transgênica uma vez que estes são serem mais necessários. As técnicas para a remoção de gene marcador são conhecidos na técnica, as técnicas úteis são descritas acima na seção de definições.

A invenção também fornece um método para a produção de plantas transgênicas tendo características intensificadas relacionadas com o rendimento com relação às plantas de controle, que compreende introdução e expressão em uma planta de qualquer ácido nucleico que codifica um polipeptídeo de ERLK como definido neste acima.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo características aumentadas relacionadas com o rendimento intensificadas, particularmente biomassa folhosa aumentada e rendimento de semente, cujo método compreende:

(i) introduzir e expressar em uma planta ou célula de planta, um polipeptídeo de ácido nucleico codificador de ERLK e

(ii) cultivar a célula de planta sob condições que promovem crescimento e crescimento vegetal.

O ácido nucleico de (i) pode ser qualquer um dos ácidos nucleicos capazes de codificar um polipeptídeo de ERLK como definido neste.

O ácido nucleico pode ser introduzindo diretamente em uma célula de planta ou na planta por si só (incluindo introdução em um tecido, órgão ou qualquer outra parte de uma planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é

preferivelmente introduzida em uma planta por transformação. O termo "transformação" é descrito em mais detalhes na seção de "definições" neste.

As células de planta geneticamente modificadas podem ser regeneradas por intermédio de todos os métodos com os quais o trabalhador habilitado está familiarizado. Os métodos adequados podem ser encontrados nas publicações mencionadas acima por S. D. Kung and R. Wu, Potrykus or Höfgen and Willmitzer.

No geral, após a transformação, as células de planta ou os agrupamentos celulares são selecionados para a presença de um ou mais marcadores que são codificados por genes expressíveis em plantas co-transferida com o gene de interesse, seguindo que o material transformado é regenerado em uma planta total. Para selecionar as plantas transformadas, o material vegetal obtido na transformação é, como uma regra, submetido às condições seletivas de modo que as plantas transformadas podem ser distinguidas das plantas transformadas. Por exemplo, as sementes obtidas da maneira descrita acima podem ser plantadas e após um período de crescimento inicial, submetido a uma seleção adequada por pulverização. Uma possibilidade adicional consiste do crescimento de sementes, se apropriado após a esterilização, em placas de ágar usando-se um agente de seleção adequado de modo que apenas as células transformadas podem se desenvolver em plantas.

Alternativamente, as plantas transformadas são avaliadas para a presença de um marcador selecionável tal como os descritos acima.

Seguindo a transferência e a regeneração de DNA, as plantas putativamente transformadas também podem ser avaliadas, por exemplo, usando-se a análise de Southern, para a presença do gene de interesse, número de cópia e/ou organização genômica. Alternativa ou adicionalmente, os níveis de expressão do DNA novamente introduzido pode ser monitorado usando-se a análise Northern e/ou Western, ambas as técnicas sendo bem conhecidas das

pessoas tendo habilidade comum na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de meios, tal como pelas técnicas de criação por propagação ou clássica. Por exemplo, uma planta transformada de primeira geração (ou T1) pode ser personalizada e transformantes de segunda geração homozigota (ou T2) selecionadas, e as plantas T2 podem ser então propagadas através da técnica de criação clássica. Os organismos transformados gerados podem tomar uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e células não transformadas; transformantes clonais (por exemplo, todas as células transformadas para conter o cassete de expressão); enxertos de tecidos transformados e não transformados (por exemplo, em plantas, um estoque de raiz transformado enxertado em um descendente não transformado).

A presente invenção claramente estende-se a qualquer célula de planta ou planta produzida por qualquer um dos métodos descritos neste e em todas as partes de plantas e *propagules* destes. A presente invenção ainda estende-se para abranger a progênie de um célula, tecido, órgão ou planta total transformada ou transfectada primária que foram produzidos por qualquer um dos métodos mencionados acima, o requerimento sendo apenas aquele progênie que apresenta as mesmas características genotípicas e/ou fenotípicas como aqueles produzidos pelo precursor nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucleico isolado que codifica um polipeptídeo de ERLK como definido neste acima. As células hospedeiras preferidas de acordo com a invenção são células de planta. As plantas hospedeiras para ácidos nucleicos ou o vetor usado no método de acordo com a invenção, o cassete ou construção ou vetor de expressão são, em princípio, vantajosamente todas as plantas, que são capazes de sintetizar os polipeptídeos usados no método inventivo.

Os métodos da invenção são vantajosamente aplicáveis a qualquer planta. A presente invenção também abrange plantas obteníveis pelos métodos de acordo com a presente invenção. A presente invenção portanto fornece plantas, partes de plantas ou células de planta destes obteníveis pelo método de acordo com a presente invenção, cujas plantas, partes ou células destes compreendem um transgene de ácido nucleico que codifica uma proteína ERLK como definido acima. As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular plantas monocotiledôneas e dicotiledôneas incluindo legumes de alimentação ou forragem, plantas ornamentais, safras alimentícias, árvores ou arbustos. De acordo com uma forma de realização preferida da presente invenção, a planta é uma planta de safra. Os exemplos de plantas de safra incluem soja, girassol, canola, alfafa, semente de colza, algodão, tomate, batata e tabaco. Além disso, preferivelmente, a planta é uma planta monocotiledônea. Os exemplos de plantas monocotiledôneas incluem cana de açúcar. Mais preferivelmente, a planta é um cereal. Exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

A invenção também estende-se às partes colhíveis de uma planta tal como, mas não limitado a sementes, folhas, frutos, flores, caules, raízes, rizomas, tubérculos e bulbos. A invenção além disso diz respeito a produtos derivados, preferivelmente diretamente derivados, de uma parte colhível de uma tal planta, tais como grânulos secos ou pós, óleo, gordura e ácidos graxos, amidos ou proteínas.

De acordo com uma característica preferida da invenção, a expressão modulada é a expressão aumentada. Métodos para aumentar a expressão de ácidos nucleicos ou genes ou produtos de gene, são bem documentados na técnica e exemplos são fornecidos na seção de definições.

Como mencionado acima, um método preferido para a

modulação da (preferivelmente, aumento) expressão de um ácido nucleico que codifica um polipeptídeo de ERLK ocorre pela introdução e expressão em uma planta a ácido nucleico que codifica um polipeptídeo de ERLK; entretanto, os efeitos de realizar o método, isto é, intensificar características relacionadas com o rendimento também podem ser atingidos usando-se outras técnicas bem conhecidas, que incluem, mas não limitam-se a rotulação de ativação de T-DNA, TILLING, recombinação homóloga. Uma descrição destas técnicas é fornecida na seção de definições.

A presente invenção também abrange o uso de ácidos nucleicos que codificam polipeptídeos de ERLK como descritos neste e uso destes polipeptídeos de ERLK na intensificação de qualquer uma das características mencionadas acima relacionadas com o rendimento em plantas.

Os ácidos nucleicos que codificam polipeptídeo de ERLK descritos neste ou os polipeptídeos de ERLK por si só, podem encontrar uso nos programas de criação em que um marcador de DNA é identificado sendo geneticamente ligado a um gene que codifica o polipeptídeo de ERLK. Os ácidos nucleicos/genes ou os polipeptídeos de ERLK por si só podem ser usados para definir um marcador molecular. Este DNA ou marcador de proteína podem ser usados nos programas de criação para selecionar plantas tendo características intensificadas relacionadas com o rendimento como definido neste acima nos métodos da invenção.

As variantes alélicas de um polipeptídeo de ácido nucleico codificador de ERLK/gene também pode encontrar uso nos programas de criação auxiliados por marcador. Tais programas de criação algumas vezes requerem a introdução de variação alélica pelo tratamento mutagênico das plantas, usando-se por exemplo mutagênese EMS; alternativamente, o programa pode começar com uma coleção de variantes alélicas de denominada origem "natural" causada não intencionalmente. A identificação de variantes alélicas então acontece, por exemplo, por PCR. Isto é seguido por

uma etapa para a seleção de variantes alélicas superiores das seqüências em questão e que dá rendimento aumentado. A seleção é tipicamente realizada pelo monitoramento do desempenho de crescimento de plantas contendo variantes alélicas diferentes das seqüências em questão. O Desempenho de crescimento pode ser monitorado em uma estufa ou no campo. As etapas opcionais adicionais incluem o cruzamento de plantas em que a variante alélica superior foi identificada com uma outra planta. Isto pode ser usado, por exemplo, para realizar uma combinação de características fenotípicas diferentes.

Os ácidos nucleicos que codificam polipeptídeos de ERLK também podem ser usados como sondas para mapear genética e fisicamente os genes que são uma parte de, e como marcadores para características ligadas àqueles genes. Tal informação pode ser útil na criação vegetal a fim de desenvolver linhas com fenotipos desejados. Tal uso de polipeptídeos de ácido nucleico codificador de ERLK requer apenas uma seqüência de ácido nucleico de pelo menos 15 nucleotídeos de comprimento. Os polipeptídeos de ácido nucleico codificador de ERLK podem ser usados como marcadores de polimorfismo de comprimento de fragmento de restrição (RFLP). Southern blots (Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico vegetal digerido por restrição podem ser sondados com os ácidos nucleicos codificadores de ERLK. Os padrões formadores de faixa resultantes podem ser então submetidos à análise genética usando-se programas de computador, tais como MapMaker (Lander et al. (1987) Genomics 1: 174-181) a fim de construir um mapa genético. Além disso, os ácidos nucleicos podem ser usados para a sonda Southern blots contendo DNAs genômicos tratados com endonuclease de restrição de uma série de indivíduos que representa precursores e progênie de cruzamento genético definido. A segregação dos polimorfismos de DNA é observado e usado para calcular a posição do ácido nucleico que codifica o polipeptídeo de

ERLK no mapa genético previamente obtido usando-se esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

A produção e o uso de sondas derivadas do gene vegetal para o uso no mapeamento genético são descritos em Bernatzky and Tanksley (1986) Plant Mol. Biol. Reporter 4: 37-41. Numerosas publicações descrevem o mapeamento de clones de cDNA específicos usando-se a metodologia resumida acima ou variações desta. Por exemplo, Populações intercruzadas F2, populações retrocruzadas, populações unidas aleatoriamente, linhas isogênicas próximas e outras séries de indivíduos podem ser usadas para o mapeamento. Tais metodologias são bem conhecidas por aqueles habilitados na técnica.

As sondas de ácido nucleico também podem ser usadas para o mapeamento físico (isto é, substituição de seqüências em mapas físicos; ver Hoheisel et al. In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346 e referências citadas neste).

Em uma outra forma de realização, as sondas de ácido nucleico podem ser usadas no mapeamento de hibridização in situ por fluorescência direta (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora métodos correntes de mapeamento de FISH favorece o uso de clones grandes (diversos kb a diversas centenas de kb; see Laan et al. (1995) Genome Res. 5:13-20), melhoras na sensibilidade podem permitir desempenho de mapeamento de FISH usando-se sondas mais curtas.

Uma variedade de métodos com base na amplificação e ácido nucleico para mapeamento genético e físico pode ser realizada usando-se os ácidos nucleicos. Os exemplos incluem amplificação específica de alelo (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismos de fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332), ligação específica de alelo (Landegren et al. (1988) Science 241:1077-1080), reação de extensão de nucleotídeo (Sokolov (1990) Nucleic Acid Res.

18:3671), Mapeamento de Híbrido por Radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e Mapeamento de Happy (Dear and Cook (1989) Nucleic Acid Res. 17:6795-6807). Para estes métodos, a seqüência de um ácido nucleico é usada para projetar e produzir pares de iniciador para o uso na reação de amplificação ou nas reações de extensão de iniciador. O projeto de tais iniciadores é bem conhecido por aqueles habilitados na técnica. Nos métodos que utilizam mapeamento genético com base em PCR, pode ser necessário identificar diferenças de seqüência de DNA entre os parentes do mapeamento de cruzamento na região correspondente à presente seqüência de ácido nucleico. Entretanto, no geral, não são necessários métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo características relacionadas com o rendimento intensificado, como descrito anteriormente. estas características também podem ser combinadas com outras características economicamente vantajosas, tal como características de intensificação de rendimento adicionais, tolerância a outras tensões abióticas e ióticas, características que modificam várias características arquitetônicas e/ou características bioquímicas e/ou fisiológicas.

FBW40

Surpreendentemente, foi observado agora que aumentar a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo FBXW dá plantas tendo características intensificadas relacionadas com o rendimento com relação às plantas de controle. Portanto, a invenção fornece um método para intensificar rendimento relacionado em plantas com relação às plantas de controle, que compreende aumentar a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo FBXW.

Um método preferido para aumentar a expressão de um ácido nucleico que codifica um polipeptídeo FBXW ocorre pela introdução e expressão em uma planta, de um ácido nucleico que codifica um polipeptídeo

FBXW.

Qualquer referência a seguir a uma "proteína útil nos métodos da invenção" é entendida significar um polipeptídeo FBXW como definido neste. Qualquer referência a seguir a um "ácido nucleico útil nos métodos da invenção" é entendido significar um ácido nucleico capaz de codificar um tal polipeptídeo FBXW. O ácido nucleico a ser introduzido em uma planta (e portanto útil na realização dos métodos da invenção) é qualquer ácido nucleico que codifica o tipo de proteína que não será descrito, a seguir, também denominado "ácido nucleico *FBXW*" ou "gene *FBXW*".

O termo "polipeptídeo FBXW" como definido neste refere-se a um polipeptídeo que compreende: (i) um F-box; (ii) um domínio WD40 que compreende pelo menos uma repetição de WD40; (iii) Motivo 1 como representado por SEQ ID N° : 97 e (iv) Motivo 2 como representado por SEQ ID N° : 98.

Preferivelmente, a seqüência de Motivo 1 é:

WK(E/K)(F/V/L)Y(C/R/G)ERWGXP, X representa qualquer aminoácido.

Os aminoácidos mais conservados dentro do Motivo 1 são XLXFGXXXYFXWKXXYXERWGXP e dentro do Motivo 2 SLXFEXPWLVSXSXDG (onde X é uma sub-série especificada de aminoácidos diferentes para cada posição, como apresentado na SEQ ID N° : 97 e SEQ ID N° : 98). Dentro do Motivo 1 e Motivo 2, são permitidas uma ou mais mudanças conservativas em qualquer posição, e/ou uma, duas ou três mudanças não conservativas em qualquer posição.

Opcionalmente, o polipeptídeo FBXW pode compreender qualquer um ou mais dos seguintes: (a) Motivo 3 como representado por SEQ ID N° : 99; (b) Motivo 4 como representado por SEQ ID N° : 100 e (c) Motivo 5 como representado por SEQ ID N° : 101. Dentro dos Motivos 3 a 5, são permitidas uma ou mais mudanças conservativas em qualquer posição,

e/ou uma ou duas mudanças não conservativas em qualquer posição.

Um exemplo de um polipeptídeo FBXW como definido neste acima que compreende (i) um F-box; (ii) um domínio WD40 que compreende pelo menos uma repetição de WD40; (iii) Motivo 1 como representado por SEQ ID N° : 97 e (iv) Motivo 2 como representado por SEQ ID N° : 98; e opcionalmente que compreende qualquer um ou mais dos seguintes: (a) Motivo 3 como representado por SEQ ID N° : 99; (b) Motivo 4 como representado por SEQ ID N° : 100 e (c) Motivo 5 como representado por SEQ ID N° : 101, é representado como em SEQ ID N° : 60 (Figura 5 é um desenho que representa os domínios diferentes e sua posição relativa em SEQ ID N° : 60). Tais exemplos adicionais são representados por qualquer um de SEQ ID N° : 62, SEQ ID N° : 64, SEQ ID N° : 66, SEQ ID N° : 68 ou SEQ ID N° : 70 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. A invenção é ilustrada pela transformação das plantas com a seqüência de *Arabidopsis thaliana* representada por SEQ ID N° : 59, que codifica o polipeptídeo da SEQ ID N° : 60. SEQ ID N° : 62 (codificado por SEQ ID N° : 61, de *Oryza sativa*), SEQ ID N° : 64 (codificado por SEQ ID N° : 63, de *Medicago trunculata*), SEQ ID N° : 66 (codificado por SEQ ID N° : 65, de *Triticum aestivum*), SEQ ID N° : 68 (codificado por SEQ ID N° : 67, de *Populus tremuloides*) e SEQ ID N° : 70 (codificado por SEQ ID N° : 69, de *Zea mays*) são ortólogos do polipeptídeo da SEQ ID N° : 60.

Os ortólogos e parálogos (os termos como sendo definidos acima) podem ser facilmente encontrados pela realização de uma denominada pesquisa blast recíproca. Isto pode ser realizado por um BLAST que envolve realizar BLAST em uma seqüência em questão (por exemplo, SEQ ID N° : 59 ou SEQ ID N° : 60) contra quaisquer bancos de dados de seqüência, tais como os bancos de dados NCBI publicamente disponíveis. BLASTN ou TBLASTX (usando-se valores default padrão) podem ser usados quando começa-se a partir de uma seqüência de nucleotídeo e BLASTP ou TBLASTN (usando-se

valores default padrão) podem ser usados quando começa-se a partir de uma seqüência de polipeptídeo. os resultados de BLAST podem ser opcionalmente filtrados. as seqüências de comprimento total dos resultados filtrados ou resultados não filtrados são então submetidos ao BLAST novamente (segundo BLAST) contra seqüências do organismo a partir do qual a seqüência em questão é derivada (onde a seqüência em questão é SEQ ID N° : 59 ou SEQ ID N° : 60, o segundo BLAST portanto, deve ser contra seqüências de *Arabidopsis*). Os resultados o primeiro ou segundo BLASTs são então comparados. Um parálogo é identificado se um acerto de classificação alta do primeiro BLAST é da mesma espécie como a partir do qual a seqüência em questão é derivada, um BLAST novamente, então resulta idealmente na seqüência em questão como o acerto mais alto (além disso por si só); um ortólogo é identificado se um acerto de classificação alta no primeiro BLAST não é da mesma espécie como a partir do qual a seqüência em questão é derivada e preferivelmente resulta em BLAST novamente na seqüência em questão entre os acertos mais altos. Os acertos de classificação alta são aqueles tendo um valor E. Quanto menor o valor E, mais significante o registro (ou em outras palavras (quanto menor a oportunidade de o acerto ser encontrado por oportunidade). A computação do valor E é bem conhecido na técnica. Além dos valores E, as comparações também são registradas pela porcentagem de identidade. A porcentagem de identidade refere-se ao número de seqüências de nucleotídeos idênticos (ou aminoácido) entre os dois ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. No caso de famílias grandes, o ClustalW pode ser usado, seguido por uma árvore de união vizinha, para ajudar à visualização do agrupamento de genes relacionados e para identificar ortólogos e parálogos. Preferivelmente, os polipeptídeos FBXW úteis nos métodos da invenção compreendem, em ordem crescente de preferência, pelo menos 45%, 50%, 55%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ou 98% de identidade de seqüência para a SEQ

ID N° : 60 (cálculos mostrados no Exemplo 9). Os polipeptídeos FBXW apresentam identidade de aminoácido relativamente baixa de conservação de seqüência entre estes, embora sua estrutura de polipeptídeo (incluindo o domínio F-box e o WD40) também seja conservada. A Seqüência de conservação entre as duas regiões mais conservadas de polipeptídeos FBXW como representado por SEQ ID N° : : 102 e SEQ ID N° : : 103 (ambas compreendidas dentro da SEQ ID N° : : 60) está em ordem crescente de preferência, de pelo menos 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ou 98% de identidade (cálculos mostrados no Exemplo 9).

Os polipeptídeos representados por qualquer um de SEQ ID N° : 60, SEQ ID N° : 62, SEQ ID N° : 64, SEQ ID N° : 66, SEQ ID N° : 68, SEQ ID N° : 70 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima, todos compreendem (i) um F-box; (ii) um domínio WD40 que compreende pelo menos uma repetição de WD40; (iii) Motivo 1 como representado por SEQ ID N° : 97 e (iv) Motivo 2 como representado por SEQ ID N° : 98.

Os termos "domínio" e "motivo" são definidos acima. Bancos de dados especiais existem para a identificação de domínios. As repetições F-box e WD40 em um polipeptídeo FBXW podem ser identificados usando-se, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-244; hospedado pelo EMBL em Heidelberg, Alemanha), InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318; hospedado pelo European Bioinformatics Institute (EBI) no Reino Unido), Prosite (Bucher and Bairoch (1994), Uma sintaxe de perfil generalizado para motivos de seqüências biomoleculares e sua função em interpretação de seqüência automática. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAIPress, Menlo Park; Hulo et al., Nucl. Acids. Res. 32: D134-D137,

(2004), O servidor ExPASy proteomics é fornecido como um serviço pela comunidade cientifica (hospedado pelo Swiss Institute of Bioinformatics (SIB) na Suíça) ou Pfam (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002), hospedado pelo Sanger Institute in the United Kingdom). O F-box compreendem de 40 a 50 resíduos, em que existem muito poucas posições não variantes. Esta perda de consenso estrita realiza a identificação usando-se algoritmos de busca essenciais. No banco de dados InterPro, o F-box é projetado por IPR001810, PF00646 no banco de dados Pfam e PS50181 no banco de dados PROSITE. As repetições de WD40 compreendidas dentro do domínio WD40 estão tipicamente em torno de 40 aminoácidos. Logo para o F-box, existem algumas posições não variantes e as extremidades freqüentemente repetem (mas não necessariamente) com o dipeptídeo Trp-Asp (W-D). A identificação usando-se algoritmos de pesquisa é igualmente essencial. No banco de dados InterPro, a repetição WD40 é designada por IPR001680, PF00400 no banco de dados Pfam e PS50082 no banco de dados PROSITE. O domínio WD40 tipicamente compreende de 4 a 16 repetições, preferivelmente de 5 a 10, mais preferivelmente de 6 a 8, mais preferivelmente 7 repetições de acordo com o algoritmo PFAM (repetições PF00400). O domínio WD40 é designado por IPR0011046 no banco de dados InterPro.

Os métodos para o alinhamento de seqüências para comparação incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. O GAP usa o algoritmo de Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar o alinhamento completo de duas seqüências que maximizam o número de comparações e minimiza o número fendas. O algoritmo BLAST (Altschul *et al.* (1990) J Mol Biol 215: 403-10) calcula o percentual de identidade de seqüência e realiza uma análise estatística da similaridade entre as duas seqüências. O software para a análise BLAST está publicamente disponível através do National Centre for Biotechnology Information. Os

homólogos podem ser facilmente identificados usando-se, por exemplo, o algoritmo de alinhamento de seqüência múltiplo ClustalW (versão 1.83) disponível no serviço GenomeNet no Kyoto University Bioinformatics Center, com os parâmetros de alinhamento em pares default e um método de registro em porcentagem. A edição manual menor pode ser realizada para otimizar o alinhamento entre os motivos conservados, como deveria estar evidente para a pessoa habilitada na técnica. Em alguns exemplos, os parâmetros de default podem ser ajustados para modificar a estringência da pesquisa. Por exemplo usando-se BLAST, a resposta de significância estatística (denominado valor "esperado") para relatar comparações contra seqüências de bancos de dados pode ser aumentada para mostrar comparações menos estringentes. Desta maneira, comparações quase exatas curtas podem ser identificadas. Motivo 1 como representado por SEQ ID N° : 97 e Motivo 2 como representado por SEQ ID N° : 98 ambos compreendidos nos polipeptídeos FBXW úteis nos métodos da invenção podem ser identificados desta maneira (Figura 6). Dentro do Motivo 1 e Motivo 2, são permitidas uma ou mais mudanças conservativas em qualquer posição, e/ou uma, duas ou três mudanças não conservativas em qualquer posição. Os motivos 3 a 5 (representados respectivamente por SEQ ID N° : 99, SEQ ID N° : 100 e SEQ ID N° : 101) podem da mesma forma ser identificado (Figure 6). Dentro dos Motivos 3 to 5, são permitidas uma ou mais mudanças conservativas em qualquer posição, e/ou uma ou duas mudanças não conservativas em qualquer posição.

O ácido nucleico que codifica os polipeptídeos representados por qualquer um de SEQ ID N° : 60, SEQ ID N° : 62, SEQ ID N° : 64, SEQ ID N° : 66, SEQ ID N° : 68, SEQ ID N° : 70 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima, não necessitam ser ácidos nucleicos de comprimento total, visto que o desempenho dos métodos da invenção não conta com o uso de seqüência de ácido nucleico de

comprimento total. Além disso, exemplos de ácidos nucleicos adequados para a realização de métodos da invenção incluem mas não são limitados àqueles representados por qualquer uma de: SEQ ID N° : 59, SEQ ID N° : 61, SEQ ID N° : 63, SEQ ID N° : 65, SEQ ID N° : 67 ou SEQ ID N° : 69. As variantes de ácido nucleico também podem ser úteis na prática dos métodos da invenção. Os exemplos de tais variantes incluem porções de ácidos nucleicos, seqüências hibridizadoras, variantes de união, variantes alélicas de ocorrência natural ou obtidos por manipulação de DNA.

Uma porção pode ser preparada, por exemplo, realizando-se uma ou mais anulações a um ácido nucleico que codifica um polipeptídeo FBXW como definido neste acima. as porções podem ser usadas na forma isolada ou estas podem ser fundidas a outras seqüências codificadoras (ou não codificadoras) a fim de, por exemplo, produzir uma proteína que combine diversas atividade. Quando fundido a outras seqüências codificadoras, o polipeptídeo resultante produzido na tradução pode ser maior do que aquele predito a partir da porção de FBXW. As porções úteis nos métodos da invenção, codificam um polipeptídeo FBXW (como descritos acima) e tendo substancialmente a mesma atividade biológica como o polipeptídeo FBXW representado por qualquer um de SEQ ID N° : 60, SEQ ID N° : 62, SEQ ID N° : 64, SEQ ID N° : 66, SEQ ID N° : 68, SEQ ID N° : 70 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. Os exemplos de porções podem incluir nucleotídeos que codificam um polipeptídeo que compreende: (i) um F-box; (ii) um domínio WD40 que compreende pelo menos uma repetição de WD40; (iii) Motivo 1 como representado por SEQ ID N° : 97 e (iv) Motivo 2 como representado por SEQ ID N° : 98. As porções podem opcionalmente compreender qualquer um ou mais dos seguintes: (a) Motivo 3 como representado por SEQ ID N° : 99; (b) Motivo 4 como representado por SEQ ID N° : 100 e (c) Motivo 5 como representado por SEQ ID N° : 101. A porção é tipicamente de pelo menos 500 nucleotídeos

de comprimento, preferivelmente de pelo menos 750 nucleotídeos de comprimento, mais preferivelmente de pelo menos 1000 nucleotídeos de comprimento e mais preferivelmente de pelo menos 1500 nucleotídeos de comprimento. Preferivelmente, a porção é uma porção de um ácido nucleico como representado por qualquer um de SEQ ID N° : 59, SEQ ID N° : 61, SEQ ID N° : 63, SEQ ID N° : 65, SEQ ID N° : 67 ou SEQ ID N° : 69. Mais preferivelmente a porção é uma porção de um ácido nucleico como representado por SEQ ID N° : 59.

Uma outra variante de ácido nucleico útil nos métodos da invenção, é a ácido nucleico capaz de hibridizar sob condições de estringência reduzidas, preferivelmente sob condições estringentes, com um ácido nucleico que codifica um polipeptídeo FBXW como definido neste acima ou a com uma porção como definido neste acima.

O termo "hibridização" é definido na seção de definições acima. As seqüências de hibridização úteis nos métodos da invenção, codificam um polipeptídeo que compreende: (i) um F-box; (ii) um domínio WD40 que compreende pelo menos uma repetição de WD40; (iii) Motivo 1 como representado por SEQ ID N° : 97 e (iv) Motivo 2 como representado por SEQ ID N° : 98 e tendo substancialmente a mesma atividade biológica como os polipeptídeos FBXW representados por SEQ ID N° : 60, SEQ ID N° : 62, SEQ ID N° : 64, SEQ ID N° : 66, SEQ ID N° : 68, SEQ ID N° : 70 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. A seqüência de hibridização é tipicamente de pelo menos 250 nucleotídeos de comprimento, preferivelmente de pelo menos 500 nucleotídeos de comprimento, mais preferivelmente de pelo menos 750 nucleotídeos de comprimento, ainda preferivelmente de pelo menos 1000 nucleotídeos de comprimento, mais preferivelmente a seqüência de hibridização é 1500 nucleotídeos de comprimento. Preferivelmente, a seqüência de hibridização é uma que é capaz de hibridização a qualquer dos ácidos nucleicos representado

por SEQ ID N° : 59, SEQ ID N° : 61, SEQ ID N° : 63, SEQ ID N° : 65, SEQ ID N° : 67 ou SEQ ID N° : 69 ou a uma porção de qualquer das seqüências mencionadas acima, uma porção como sendo definida acima. Mais preferivelmente a seqüência de hibridização é capaz de hibridização à SEQ ID N° : 59 ou a porções desta.

As porções que codificam um polipeptídeo FBXW precisa de um ou mais ou de parte de: (i) um F-box; (ii) um domínio WD40 que compreende pelo menos uma repetição de WD40; (iii) Motivo 1 como representado por SEQ ID N° : 97 e (iv) Motivo 2 como representado por SEQ ID N° : 98, pode ser usado, por exemplo, como uma sonda no processo de hibridização como descrito abaixo, para obter porções úteis na realização dos métodos da invenção, que compreende todos de: (i) um F-box; (ii) um domínio WD40 que compreende pelo menos uma repetição de WD40; (iii) Motivo 1 como representado por SEQ ID N° : 97 e (iv) Motivo 2 como representado por SEQ ID N° : 98. Os exemplos úteis para o processo de hibridização são representados por SEQ ID N° : 71 de *Vitis vinifera* (contig de NCBI ESTs CF210354, CF413646 e CF213082), SEQ ID N° : 73 de *Senecio cambrensis* (NCBI EST DY662683.1), SEQ ID N° : 75 de *Helianthus annuus* (NCBI EST DY916708), SEQ ID N° : 77 de *Euphorbia esula* (NCBI EST DV129599), SEQ ID N° : 79 de *Lycopersicon esculentum* (NCBI EST BI931509), SEQ ID N° : 81 de *Aquilegia formosa x Aquilegia pubescens* (NCBI EST DT753991.1), SEQ ID N° : 83 de *Gossypium hirsutum* (NCBI EST DT466472), SEQ ID N° : 85 de *Sorghum bicolor* (NCBI EST CF770159), SEQ ID N° : 87 de *Ipomea nil* (NCBI EST BJ574759.1), SEQ ID N° : 89 de *Solanum tuberosum* (NCBI EST CX161187), SEQ ID N° : 91 de *Zamia fischeri* (NCBI EST DY032229), SEQ ID N° : 93 de *Persea americana* (NCBI EST CK756534) e SEQ ID N° : 95 de *Glycine max* (NCBI EST CD418593.1).

Uma outra variante de ácido nucleico útil nos métodos da

invenção é uma variante de união que codifica um polipeptídeo FBXW como definido neste acima. As variantes de união preferidas são variantes de união de um ácido nucleico que codificam o polipeptídeo FBXW representado por qualquer um de SEQ ID N° : 60, SEQ ID N° : 62, SEQ ID N° : 64, SEQ ID N° : 66, SEQ ID N° : 68, SEQ ID N° : 70 ou variantes de união que codifica ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. ainda são preferidas variantes de união de ácidos nucleicos representado por qualquer um de SEQ ID N° : 59, SEQ ID N° : 61, SEQ ID N° : 63, SEQ ID N° : 65, SEQ ID N° : 67 ou SEQ ID N° : 69. É mais preferida uma variante de união de um ácido nucleico como representado por SEQ ID N° : 59.

Uma outra variante de ácido nucleico útil na realização dos métodos da invenção é uma variante alélica de um ácido nucleico que codifica um polipeptídeo FBXW como definido neste acima. Variantes alélicas existem em estado natural e abrangido dentro dos métodos da presente invenção está o uso destes alelos naturais. As variantes alélicas úteis nos métodos da presente invenção tem, substancialmente, a mesma atividade biológica como o polipeptídeo FBXW da SEQ ID N° : 60 e qualquer um dos aminoácidos descritos na Tabela G de Exemplo 8. A variante alélica pode ser uma variante alélica de um ácido nucleico que codifica um polipeptídeo FBXW representado por qualquer da SEQ ID N° : 60, SEQ ID N° : 62, SEQ ID N° : 64, SEQ ID N° : 66, SEQ ID N° : 68, SEQ ID N° : 70 ou uma variante alélica de um ácido nucleico codificando ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. Ainda preferidas são variantes alélicas de ácidos nucleicos representados por qualquer um de SEQ ID N° : 59, SEQ ID N° : 61, SEQ ID N° : 63, SEQ ID N° : 65, SEQ ID N° : 67 ou SEQ ID N° : 69. Mais preferida é uma variante alélica de um ácido nucleico como representado por SEQ ID N° : 59.

Uma variante de ácidos nucleico adicionais úteis nos métodos da invenção é uma variante de ácido nucleico obtidos por mistura de gene. A

mistura de gene ou evolução direcionada também podem ser usadas para gerar variantes de ácidos nucleicos que codificam polipeptídeos FBXW como definido acima.

Além disso, variantes de ácido nucleico também podem ser obtidas, por exemplo, por mutagênese direcionada ao local. Diversos métodos estão disponíveis para atingir a mutagênese direcionada ao local, o mais comum sendo métodos com base em PCR (Current Protocols in Molecular Biology. Wiley (Eds)).

Também úteis nos métodos da invenção são ácidos nucleicos que codificam homólogos de qualquer um dos aminoácidos representados por SEQ ID N° : 60, SEQ ID N° : 62, SEQ ID N° : 64, SEQ ID N° : 66, SEQ ID N° : 68, SEQ ID N° : 70 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. Os termos "homólogos", "ortólogos" e "parálogos" são como definidos acima.

Também úteis nos métodos da invenção são os ácidos nucleicos que codificam derivados de qualquer um dos aminoácidos representados por SEQ ID N° : 60, SEQ ID N° : 62, SEQ ID N° : 64, SEQ ID N° : 66, SEQ ID N° : 68, SEQ ID N° : 70 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima.

Os ácidos nucleicos que codificam polipeptídeos FBXW podem ser derivados de qualquer fonte natural ou artificial. O ácido nucleico pode ser modificado a partir de sua forma natural em composição e/ou ambiente genômico através da manipulação humana deliberada. Preferivelmente o ácido nucleico que codifica o polipeptídeo FBXW é de uma planta, ainda preferivelmente de uma planta dicotiledônea, mais preferivelmente da família *Brassicaceae*, mais preferivelmente o ácido nucleico é de *Arabidopsis thaliana*.

A invenção também fornece construções genéticas e vetores para facilitar a introdução e/ou expressão das seqüências de ácido nucleico

útil nos métodos de acordo com a invenção, em uma planta.

Portanto, é fornecida uma construção genética que compreende:

(i) Um ácido nucleico que codifica um polipeptídeo FBXW como definido neste acima;

(ii) Uma ou mais seqüências de controle operacionalmente ligadas ao ácido nucleico de (i).

As construções úteis nos métodos de acordo com a presente invenção podem ser construídas usando-se tecnologia de DNA recombinante DNA bem conhecida pela pessoa habilitada na técnica. As construções genéticas podem ser inseridas em vetores, que podem ser comercialmente disponíveis, adequados para a transformação em plantas e adequados para a expressão do gene de interesse nas células transformadas. A invenção portanto fornece o uso de uma construção genética como definida neste acima nos métodos da invenção.

As plantas são transformadas com um vetor que compreende a seqüência de interesse (isto é, um ácido nucleico que codifica um polipeptídeo FBXW). O técnico habilitado está bem ciente dos elementos genéticos que deve estar presente no vetor a fim de transformar, de maneira bem sucedida, selecionar e propagar células hospedeiras contendo a seqüência de interesse. A seqüência de interesse está operacionalmente ligada a uma ou mais seqüências de controle (pelo menos a um promotor).

Vantajosamente, qualquer tipo de promotor, natural ou sintético, pode ser usado para conduzir a expressão da seqüência de ácido nucleico.

De acordo com um aspecto preferido da invenção, o ácido nucleico que codifica um polipeptídeo FBXW está operacionalmente ligada a um promotor constitutivo (a seqüência de controle). O promotor constituído é preferivelmente um promotor GOS2 (também denominado SUI1 ou eIF1

(fator de iniciação eucariótica 1), mais preferivelmente o promotor constituído é um promotor GOS2 de arroz, ainda preferivelmente, o promotor constituído é representado por uma seqüência de ácido nucleico substancialmente similar à SEQ ID N° : 104 ou SEQ ID N° : 58, mais preferivelmente o promotor constituído é como representado por SEQ ID N° : 104 ou SEQ ID N° : 58.

Deve estar claro que aplicabilidade da presente invenção não está restrita ao ácido nucleico que codifica um polipeptídeo FBXW como representado por SEQ ID N° : 59, é a aplicabilidade da invenção restringida à expressão de um tal ácido nucleico que codifica um polipeptídeo FBXW quando conduzido por um promotor GOS2.

Elementos reguladores adicionais para aumentar a expressão de ácidos nucleicos ou genes ou produtos de gene, podem incluir intensificadores transcripcionais bem como translacionais. Aqueles habilitados na técnica estarão cientes de seqüências terminadoras e intensificadoras que podem ser adequados para o uso na realização da invenção. um exemplo de tal elemento regulador é um íntron introduzido na região não traduzida 5'. Opcionalmente, uma ou mais seqüências terminadoras (também um seqüência de controle) podem ser usadas na construção introduzida em uma planta.

Outras seqüências de controle (além do promotor, intensificador, silenciador, seqüências de íntron, regiões 3'UTR e/ou 5'UTR) podem ser elementos estabilizadores de proteína e/ou RNA. Tais seqüências devem ser conhecidas ou podem ser facilmente obtidas por uma pessoa habilitada na técnica.

As construções genéticas da invenção ainda podem incluir uma origem de seqüência de replicação que é requerida par a manutenção e/ou replicação em um tipo celular específico. Um exemplo é quando uma construção genética é requerida para ser mantida em uma célula bacteriana como um elemento genético epissômico (por exemplo, molécula de plasmídeo

ou cosmídeo). Origens preferidas de replicação inclui, mas não são limitados a, o f1-ori e o colE1.

Para a detecção da transferência bem sucedida das seqüências de ácido nucleico como usado nos métodos da invenção e/ou seleção de plantas transgênicas que compreende estes ácidos nucleicos, é vantajoso usar genes marcadores (ou genes repórteres). Portanto, a construção genética pode compreender opcionalmente um gene marcador selecionável. Os genes marcadores podem ser removidos ou excisados a partir da célula transgênica uma vez que estes são serem mais necessários. As técnicas para a remoção de marcador são conhecidas na técnica, técnicas úteis são descritas acima na seção de definições.

A invenção também fornece um método para a produção de plantas transgênicas tendo rendimento aumentado com relação às plantas de controle adequadas, que compreende introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo FBXW como definido neste acima.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo rendimento aumentado com relação às plantas de controle adequadas, cujo método compreende:

(i) introduzir e expressar a ácido nucleico que codifica um polipeptídeo FBXW em uma célula de planta e

(ii) cultivar a célula de planta sob condições que promovem crescimento e crescimento vegetal.

O ácido nucleico pode ser introduzindo diretamente em uma célula de planta ou na planta por si só (incluindo introdução em um tecido, órgão ou qualquer outra parte de uma planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é preferivelmente introduzido em uma planta por transformação. O termo "transformação" como referido neste é definido acima.

No geral após a transformação, as células de planta ou os agrupamentos celulares são selecionados para a presença de um ou mais marcadores que são codificados por genes expressíveis em plantas co-transferida com o gene de interesse, seguindo que o material transformado é regenerado em uma planta total. As células de planta geneticamente modificadas podem ser regeneradas por intermédio de todos os métodos com os quais o trabalhador habilitado está familiarizado. Os métodos adequados podem ser encontrados nas publicações mencionadas acima por S.D. Kung and R. Wu, Potrykus or Höfgen and Willmitzer. Para selecionar as plantas transformadas, o material vegetal obtido na transformação é, como uma regra, submetido às condições seletivas de modo que as plantas transformadas podem ser distinguidas das plantas transformadas. Por exemplo, as sementes obtidas da maneira descrita acima podem ser plantadas e após um período de crescimento inicial, submetido a uma seleção adequada por pulverização. Uma possibilidade adicional consiste do crescimento de sementes, se apropriado após a esterilização, em placas de ágar usando-se um agente de seleção adequado de modo que apenas as células transformadas podem se desenvolver em plantas. Alternativamente, as plantas transformadas são avaliadas para a presença de um marcador selecionável tal como os descritos acima.

Seguindo a transferência e a regeneração de DNA, as plantas putativamente transformadas podem ser avaliadas, por exemplo, usando-se a análise de Southern, para a presença do gene de interesse, número de cópia e/ou organização genômica. Alternativa ou adicionalmente, níveis de expressão do DNA novamente introduzido pode ser monitorado usando-se análise Northern e/ou Western ou PCR quantitativo PCR, todas as técnicas sendo bem conhecidas das pessoas tendo habilidade na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de meios, tal como pelas técnicas de criação por propagação

ou clássica. Por exemplo, uma planta transformada de primeira geração (ou T1) pode ser auto-polinizada para dar transformantes de segunda geração homozigotos (ou T2), e as plantas T2 ainda propagadas através de técnicas de criação clássicas.

Os organismos transformados gerados podem tomar uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e células não transformadas; transformantes clonais (por exemplo, todas as células transformadas para conter o cassete de expressão); enxertos de tecidos transformados e não transformados (por exemplo, em plantas, um estoque de raiz transformado enxertado em um descendente não transformado).

A presente invenção claramente estende-se a qualquer célula de planta ou planta produzida por qualquer um dos métodos descritos neste e em todas as partes de plantas e *propagules* destes. A presente invenção ainda estende-se para abranger a progênie de um célula, tecido, órgão ou planta total transformada ou transfectada primária que foram produzidos por qualquer um dos métodos mencionados acima, o requerimento sendo apenas aquele progênie que apresenta as mesmas características genotípicas e/ou fenotípicas como aqueles produzidos pelo precursor nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucleico isolado que codifica um polipeptídeo FBXW como definido neste acima. As células hospedeiras preferidas de acordo com a invenção são células de planta. Plantas hospedeiras para ácidos nucleicos ou o vetor usado no método de acordo com a invenção, o cassete ou construção ou vetor de expressão são, em princípio, vantajosamente todas as plantas, que são capazes de sintetizar os polipeptídeos usados no método inventivo.

A invenção também estende-se às partes colhíveis de uma planta tal como, mas não limitado a sementes, folhas, frutos, flores, caules,

raízes, rizomas, tubérculos e bulbos. A invenção além disso diz respeito a produtos derivados, preferivelmente diretamente derivados, de uma parte colhível de uma tal planta, tais como grânulos secos ou pós, óleo, gordura e ácidos graxos, amidos ou proteínas.

De acordo com uma característica preferida da invenção, a expressão modulada é a expressão aumentada. Os métodos para aumentar a expressão de ácidos nucleicos ou genes ou produtos de gene, são bem documentados na técnica e exemplos são fornecidos na seção de definições.

Como mencionado acima, um método preferido para a modulação da (preferivelmente, aumento) expressão de um ácido nucleico que codifica um polipeptídeo FBXW ocorre pela introdução e expressão em uma planta a ácido nucleico que codifica um polipeptídeo FBXW; entretanto, os efeitos de realizar o método, isto é, intensificar características relacionadas com o rendimento também podem ser atingidos usando-se outras técnicas bem conhecidas, que incluem, mas não limitam-se a rotulação de ativação de T-DNA, TILLING, recombinação homóloga. Uma descrição destas técnicas é fornecida na seção de definições.

O desempenho dos métodos da invenção dá plantas tendo características relacionadas com o rendimento intensificadas. Em particular desempenho dos métodos da invenção dá plantas tendo rendimento aumentado, especialmente, rendimento de semente aumentado com relação às plantas de controle. Os termos "rendimento" e "rendimento de semente" como descrito em mais detalhes na seção de "definições" neste.

Referência neste a características intensificadas relacionadas com o rendimento é entendida significar um aumento na biomassa (peso) de uma ou mais partes de uma planta, que podem incluir partes (colhíveis) acima do solo e/ou partes (colhíveis) abaixo do solo. Em particular, tais partes colhíveis são sementes e o desempenho dos métodos da invenção resulta em plantas tendo rendimento de semente aumentado com relação ao rendimento

de semente de plantas de controle.

Tomando o milho como um exemplo, um aumento no rendimento pode ser manifestado como um ou mais dos seguintes: aumento no número de plantas per hectare ou acre, um aumento no número de espigas por planta, um aumento no número de séries, número de sementes por série, peso de semente, peso de mil núcleos, diâmetro/comprimento de espiga, aumento na taxa de enchimento de semente (que é o número de sementes enchidas dividido pelo número total de sementes e multiplicado por 100), entre outros. Tomando-se o arroz como um exemplo, um aumento no rendimento pode manifestar-se por si só como um aumento em um ou mais dos seguintes: número de plantas por hectare ou acre, número de panículas por planta, número de espigas por panícula, número de flores (flósculos) per panícula (que é expressado como uma razão do número de sementes enchidas sobre o número de panículas primárias), aumento na taxa de enchimento de semente (que é o número de sementes enchidas dividido pelo número total de sementes e multiplicado por 100), aumento no peso de mil núcleos, entre outros.

Visto que as plantas transgênicas de acordo com a presente invenção tem rendimento aumentado, é provável que estas plantas apresentem uma taxa de crescimento aumentada (durante pelo menos parte de seu ciclo de vida), com relação à taxa de crescimento de plantas de controle em um estágio correspondente em seu ciclo de vida.

A taxa de crescimento aumentada pode ser específica a uma ou mais partes de uma planta (incluindo sementes) ou pode ser, substancialmente a planta toda. As plantas tendo uma taxa de crescimento aumentada podem ter um ciclo de vida mais curto. O ciclo de vida de uma planta pode significar o tempo necessário para o crescimento de uma semente madura seca até o estágio onde a planta produziu sementes maduras secas, similar ao material de partida. O ciclo de vida pode ser influenciado por fatores tais como vigor

precoce, taxa de crescimento, índice de verdor, tempo de florescimento e velocidade de maturação da semente. O aumento na taxa de crescimento pode acontecer em um ou mais estágios no ciclo de vida de uma planta ou substancialmente durante todo o ciclo de vida da planta. A taxa de crescimento aumentada durante os estágios precoces no ciclo de vida de uma planta pode refletir vigor intensificado. O aumento na taxa de crescimento pode alterar o ciclo de colheita de uma planta permitindo que as plantas sejam semeadas depois e/ou colhidas mais tarde do que de outra maneira seria possível (um efeito similar pode ser obtido com o tempo de florescimento precoce). Se a taxa de crescimento for suficientemente aumentada, isto pode permitir a semeadura adicional de sementes da mesma espécie vegetal (por exemplo semeadura e colheita de plantas de arroz seguido pela semeadura e colheita de plantas de arroz adicionais todas dentro de um período de crescimento convencional). Similarmente, se a taxa de crescimento for suficientemente aumentada, isto pode permitir a semeadura adicional de sementes de espécies vegetais diferentes (por exemplo a semeadura e colheita de plantas de milho seguido, por exemplo, pela semeadura e colheita opcional de soja, batata ou qualquer outra planta adequada). Períodos adicionais de colheita a partir do mesmo estoque de raiz no caso das mesmas plantas de safra também pode ser possível. A alteração do ciclo de colheita de uma planta pode levar a um aumento na produção de biomassa anual por acre (devido a um aumento no número de vezes (dito em um ano) que qualquer planta particular pode ser desenvolvida e colhida). Um aumento no taxa de crescimento também pode permitir o cultivo de plantas transgênicas em uma área geográfica mais ampla do que suas contrapartes do tipo selvagem, visto que as limitações territoriais para o crescimento de uma safra são freqüentemente determinadas por condições ambientais adversas no período de plantio (estação precoce) ou no período de colheita (estação posterior). Tais condições adversas podem ser evitadas se o ciclo de colheita for

encurtado. A taxa de crescimento pode ser determinada derivando-se vários parâmetros de curvas de crescimento, tais parâmetros podem ser: T-Mid (o tempo tomado para as plantas atingirem 50% de seu tamanho máximo) e T-90 (tempo tomado para as plantas atingirem 90% de seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, desempenho dos métodos da invenção dá plantas tendo uma taxa de crescimento aumentada com relação às plantas de controle. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a taxa de crescimento de plantas, cujo método compreende modular a expressão, preferivelmente aumentar a expressão, em uma planta, de um ácido nucleico, que codifica um polipeptídeo FBXW como definido neste.

Um aumento no rendimento e/ou taxa de crescimento ocorre quando a planta não está sob condições de tensão ou quando a planta está exposta a várias tensões em comparação com plantas de controle. Plantas respondem tipicamente à exposição à tensão desenvolvendo-se mais lentamente. Nas condições de tensão severas, a planta pode parar de se desenvolver completamente. Tensão branda, por outro lado é definido neste como sendo qualquer tensão à qual uma planta é exposta que não resulta na interrupção do crescimento da planta completamente sem a capacidade de retomar o crescimento. Tensão branda no sentido da invenção leva à redução no crescimento de plantas submetidas à tensão menor do que 40%, 35% ou 30%, preferivelmente menor do que 25%, 20% ou 15%, mais preferivelmente menor do que 14%, 13%, 12%, 11% ou 10% ou menos em comparação com a planta de controle sob condições sem tensão. Devido a avanços nas práticas agrícolas (irrigação, fertilização, tratamentos com pesticida) tensões severas não são freqüentemente encontradas em plantas de safra cultivadas. Como uma conseqüência, o crescimento comprometido induzido pela tensão branda é freqüentemente uma característica indesejável para a agricultura. Tensões

brandas são as tensões bióticas e/ou abióticas diárias (ambientais) às quais a planta está exposta. As tensões abióticas podem ocorrer devido à seca ou água em excesso, tensões anaeróbicas, tensões ao sal, toxicidade química, tensões oxidativas e calor, frio ou temperaturas congelantes. A tensão abiótica pode ser uma tensão osmótica causada por uma tensão por água (particularmente devido à seca), tensões ao sal, tensões oxidativas ou uma tensão iônica. Tensões bióticas são tipicamente aquelas tensões causadas por patogênios, tais como bactérias, vírus, fungos e insetos.

Em particular, os métodos da presente invenção podem ser realizados sob condições sem tensão ou sob condições de seca branda a dadas plantas tendo rendimento aumentado com relação às plantas de controle. Como relatado em Wang et al. (Planta (2003) 218: 1-14), tensão abiótica leva a uma série de mudanças morfológicas, fisiológicas, bioquímicas e moleculares que afetam adversamente o crescimento e a produtividade da planta. Tensões de seca, salinidade, temperaturas extremas e oxidativas são conhecidas estarem interconectadas e podem induzir o crescimento e o dano celular através de mecanismos celulares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve um grau particularmente alto de "*cross talk*" entre tensão por seca tensão por salinidade alta. Por exemplo, seca e/ou salinização são manifestadas primariamente como tensão osmótica, resultando na interrupção da homeostase e distribuição iônica na célula. Tensão oxidativa, que freqüentemente acompanha temperatura alta ou baixa, tensão por salinidade ou seca, pode causar a desnaturação de proteínas funcionais ou estruturais. Como uma conseqüência, estas tensões ambientais diversas freqüentemente ativam caminhos de sinalização celular similares e respostas celulares, tal como a produção de proteínas de tensão, super-regulação de anti-oxidantes, acúmulo de solutos compatíveis ou interrupção de crescimento. O termo condições "sem tensão" como usado neste são aquelas condições ambientais que permitem o crescimento ótimo de plantas. Pessoas

habilitadas na técnica estão cientes das condições de solo e condições climáticas normais para um dado local.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições sem tensão ou sob condições de seca brandas rendimento aumentado com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas desenvolvidas sob condições sem tensão ou sob condições de seca brandas, cujo método compreende aumentar a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo FBXW.

O desempenho dos métodos da invenção dá plantas desenvolvidas sob condições de deficiência de nutriente, particularmente sob condições de deficiência de nitrogênio, rendimento aumentado com relação às plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas desenvolvidas sob condições de deficiência de nutriente, cujo método compreende aumentar a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo FBXW. A deficiência de nutriente pode resultar da perda de nutrientes tais como nitrogênio, fosfatos e outros compostos contendo fósforo, potássio, cálcio, cádmio, magnésio, manganês, ferro e boro, entre outros.

Os métodos da invenção são vantajosamente aplicáveis a qualquer planta. Plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular plantas monocotiledôneas e dicotiledôneas incluindo legumes de alimentação ou forragem, plantas ornamentais, safras alimentícias, árvores ou arbustos. De acordo com uma forma de realização preferida da presente invenção, a planta é uma planta de safra. Os exemplos de plantas de safra incluem soja, girassol, canola, alfafa, semente de colza,

algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta monocotiledônea. Os exemplos de plantas monocotiledôneas incluem cana de açúcar. Mais preferivelmente a planta é um cereal. Os exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

A presente invenção também abrange plantas obteníveis pelos métodos de acordo com a presente invenção. A presente invenção portanto fornece plantas, partes e células de tais plantas obteníveis pelos métodos de acordo com a presente invenção, cujas plantas ou partes ou células compreendem um transgene de ácido nucleico que codifica um polipeptídeo FBXW como definido acima.

A presente invenção também abrange o uso de ácidos nucleicos que codificam polipeptídeos FBXW em rendimento crescente em uma planta comparado com o rendimento em uma planta de controle adequada.

Um tal uso diz respeito ao rendimento crescente de plantas, o rendimento sendo definido como definido neste acima. O rendimento pode, em particular incluir um ou mais dos seguintes: rendimento de semente aumentado, número aumentado de sementes (enchidas), peso de semente singular aumentado (TKW), índice de colheita aumentado e taxa de enchimento de semente aumentada.

Os ácidos nucleicos que codificam polipeptídeos FBXW podem encontrar uso nos programas de criação em que um marcador de DNA é identificado sendo geneticamente ligado a um gene que codifica o polipeptídeo FBXW. Os ácidos nucleicos que codificam polipeptídeos FBXW podem ser usados para definir um marcador molecular. Este marcador pode ser usado nos programas de criação para selecionar plantas tendo rendimento de semente aumentado. Os ácidos nucleicos que codificam polipeptídeos FBXW podem ser, por exemplo, um ácido nucleico como representado por qualquer um de SEQ ID N° : 59, SEQ ID N° : 61, SEQ ID N° : 63, SEQ ID

N° : 65, SEQ ID N° : 67 ou SEQ ID N° : 69.

As variantes alélicas de um ácido nucleico que codifica um polipeptídeo FBXW também pode encontrar uso nos programas de criação auxiliados por marcador. Tais programas de criação algumas vezes requerem a introdução de variação alélica pelo tratamento mutagênico das plantas, usando-se, por exemplo, mutagênese EMS; alternativamente, o programa pode começar com uma coleção de variantes alélicas de denominada origem "natural" causada não intencionalmente. Identificação de variantes alélicas então acontece, por exemplo, por PCR. Isto é seguido por uma etapa para a seleção de variantes alélicas superiores das seqüências em questão e que dá rendimento aumentado de semente. A seleção é tipicamente realizada pelo monitoramento do desempenho de crescimento de plantas contendo variantes alélicas diferentes das seqüências em questão, por exemplo, variantes alélicas diferentes de qualquer um de SEQ ID N° : 59, SEQ ID N° : 61, SEQ ID N° : 63, SEQ ID N° : 65, SEQ ID N° : 67 ou SEQ ID N° : 69. O desempenho de crescimento pode ser monitorado em uma estufa ou no campo. As etapas opcionais adicionais incluem o cruzamento de plantas, em que a variante alélica superior foi identificado, com uma outra planta. Isto pode ser usado, por exemplo, para realizar uma combinação de características fenotípicas diferentes.

Os ácidos nucleicos que codificam polipeptídeos FBXW também podem ser usados como sondas para mapear genética e fisicamente os genes que são uma parte de, e como marcadores para características ligadas àqueles genes. Tal informação pode ser útil na criação vegetal a fim de desenvolver linhas com fenotipos desejados. Tal uso de ácidos nucleicos que codificam polipeptídeos FBXW requer apenas uma seqüência de ácido nucleico de pelo menos 15 nucleotídeos de comprimento. Os ácidos nucleicos que codificam polipeptídeos FBXW podem ser usados como marcadores de polimorfismo de comprimento de fragmento de restrição (RFLP). Southern

blots (Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico vegetal digerido por restrição pode ser sondado com um ácido nucleico que codifica o polipeptídeo FBXW. Os padrões formadores de faixa resultantes podem ser então submetidos à análise genética usando-se programas de computador, tais como MapMaker (Lander *et al.* (1987) Genomics 1: 174-181) a fim de construir um mapa genético. Além disso, o ácido nucleico pode ser usado para a sonda Southern blots contendo DNAs genômicos tratados com endonuclease de restrição de uma série de indivíduos que representa precursores e progênie de cruzamento genético definido. Segregação dos polimorfismos de DNA é observado e usado para calcular a posição do ácido nucleico codificando polipeptídeo FBXW no mapa genético previamente obtido usando-se esta população (Botstein *et al.* (1980) Am. J. Hum. Genet. 32: 314-331).

A produção e o uso de sondas derivadas do gene vegetal para o uso no mapeamento genético é descrita em Bernatzky and Tanksley (GENETICS 112 (4): 887-898, 1986). Numerosas publicações descrevem o mapeamento genético de clones de cDNA específicos usando-se a metodologia resumida acima ou variações desta. Por exemplo, populações intercruzadas F2, populações retrocruzadas, populações unidas aleatoriamente, linhas isogênicas próximas (NIL) e outras séries de indivíduos podem ser usadas para o mapeamento. Tais metodologias são bem conhecidas por aqueles habilitados na técnica.

As sondas de ácido nucleico também podem ser usadas para o mapeamento físico (isto é, substituição de seqüências em mapas físicos; ver Hoheisel *et al.* In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346 e referências citadas neste).

Em uma outra forma de realização, as sondas de ácido nucleico podem ser usadas na fluorescência direta no mapeamento da hibridização in situ (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora

métodos correntes de mapeamento de FISH favoreçam o uso de clones grandes (diversos kb a diversas centenas de kb; see Laan et al. (1995) Genome Res. 5:13-20), melhoras na sensibilidade podem permitir desempenho de mapeamento de FISH usando-se sondas mais curtas.

Uma variedade de métodos com base na amplificação e ácido nucleico para mapeamento genético e físico pode ser realizada usando-se os ácidos nucleicos. Os exemplos incluem amplificação específica de alelo (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismos de fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332), ligação específica de alelo (Landegren *et al.* (1988) Science 241:1077-1080), reação de extensão de nucleotídeo (Sokolov (1990) Nucleic Acid Res. 18:3671), Mapeamento de Híbrido por Radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e Mapeamento de Happy (Dear and Cook (1989) Nucleic Acid Res. 17:6795-6807). Para estes métodos, a seqüência de um ácido nucleico é usada para projetar e produzir pares de iniciador para o uso na reação de amplificação ou nas reações de extensão de iniciador. O projeto de tais iniciadores é bem conhecido por aqueles habilitados na técnica. Nos métodos que utilizam mapeamento genético com base em PCR, pode ser necessário identificar diferenças de seqüência de DNA entre os parentes do mapeamento de cruzamento na região correspondente à presente seqüência de ácido nucleico. entretanto, no geral, não são necessários métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo rendimento aumentado, como descrito anteriormente. Estas características também podem ser combinadas com outras características economicamente vantajosas, tais como características que aumentam o rendimento, tolerância a outras tensões abióticas e bióticas, características que modificam várias características arquitetônicas e/ou características bioquímicas e/ou fisiológicas.

RANBP

Na investigação do uso de proteínas de ligação de RAN para intensificar características relacionadas com o rendimento, os inventores mencionaram neste pedido observaram a escolha de promotor ser uma consideração importante. Estes observaram que a expressão de proteínas de ligação de RAN em uma planta (arroz) sob o controle de um promotor constituinte não teve qualquer efeito nos fenotipos relacionados com o rendimento. Estes surpreendentemente observaram que o rendimento da planta pode ser bem sucedidamente aumentado pela expressão de proteínas de ligação de RAN em uma planta sob o controle de um promotor específico de semente, particularmente um promotor específico de endosperma.

A presente invenção portanto fornece um método para intensificar características relacionadas com o rendimento em plantas com relação às plantas de controle, que compreende preferivelmente modular a expressão em sementes ou partes de sementes de plantas de um ácido nucleico que codifica um RANBP.

Um método preferido para a modulação da (preferivelmente, aumento) expressão em semente ou partes de semente de planta de um ácido nucleico que codifica um RANBP ocorre pela introdução e expressão em uma planta a ácido nucleico que codifica um RANBP sob o controle de um promotor específico de semente.

Qualquer referência a seguir a uma "proteína útil nos métodos da invenção" é entendido significar um polipeptídeo RANBP como definido neste. Qualquer referência a seguir a um "ácido nucleico útil nos métodos da invenção" é entendido significar um ácido nucleico capaz de codificar um tal polipeptídeo RANBP. O ácido nucleico a ser introduzido em uma planta (e portanto útil na realização dos métodos da invenção) é qualquer ácido nucleico que codifica o tipo de proteína que não será descrito, a seguir, também denominado "ácido nucleico de *RANBP*" ou "gene *RANBP*".

Os ácidos nucleicos adequados para a introdução em uma planta (e portanto útil na realização dos métodos da invenção) incluem qualquer ácido nucleico que codifica um RANBP tendo o motivo I: KSC V/L WHAXDF A/S DGELK D/E EXF, onde 'X' é qualquer aminoácido, deixado zero ou uma mudança conservativa em qualquer posição e/ou zero uma, duas ou três mudanças não conservativas em qualquer posição.

No caso de RANBPs de plantas monocotiledôneas, o terminal C do Motivo I freqüentemente termina em 'AIRFG' e no caso de RANBPs de plantas dicotiledôneas, o terminal C do Motivo I freqüentemente termina em 'CIRFA'.

Ácidos nucleicos que codificam RANBP úteis nos métodos da invenção também podem compreender (além do Motivo I) qualquer um ou mais dos seguintes motivos.

1. O Motivo II como representado por SEQ ID N° : 139 ou 145 ou um motivo tendo em ordem crescente de preferência pelo menos 60%, 70%, 80%, 90% ou mais porcentagem de identidade de seqüência para o Motivo II representado por SEQ ID N° : 139 or 145;

2. O Motivo III como representado por representado por SEQ ID N° : 140 ou 146 ou um motivo tendo em ordem crescente de preferência pelo menos 70%, 80%, 90% ou mais porcentagem de identidade de seqüência para o Motivo III como representado por SEQ ID N° : 140 or 146;

3. O Motivo IV como representado por SEQ ID N° : 141 ou 147 seguindo de zero a uma mudanças conservativas em qualquer posição e/ou zero ou uma mudanças não conservativas em qualquer posição;

4. O Motivo V como representado por SEQ ID N° : 142 ou 148 ou um motivo tendo em ordem crescente de preferência pelo menos 70%, 80%, 90% ou mais porcentagem de identidade de seqüência para o Motivo V como representado por SEQ ID N° : 142 ou 148;

5. O Motivo VI como representado por SEQ ID N° : 143 ou

149 seguindo de zero a uma mudanças conservativas em qualquer posição e/ou zero ou uma mudanças não conservativas em qualquer posição;

6. O Motivo VII como representado por SEQ ID N° : 144 ou 150 ou um motivo tendo em ordem crescente de preferência pelo menos 60%, 70%, 80%, 90% ou mais porcentagem de identidade de seqüência para o Motivo VII representado por SEQ ID N° : 144 ou 150.

Os motivos já mencionados representam aminoácidos conservados em posições específicas junto com o alinhamento de seqüências de proteínas evolucionariamente relacionadas. Enquanto os aminoácidos em outras posições podem ser variar entre os homólogos, os aminoácidos que são altamente conservados em posições específicas indica aminoácidos que são, provavelmente essenciais para a estrutura, estabilidade ou atividade da proteína. Identificado por seu alto grau de conservação em seqüências alinhadas de uma família de homólogos de proteína, estes podem ser usados como identificadores para determinar se qualquer polipeptídeo em questão pertence a uma família de polipeptídeo previamente identificada (neste caso, a família de RNABPs).

Os vários motivos mencionados acima podem ser facilmente identificados usando-se métodos para o alinhamento de seqüências para comparação. Em alguns exemplos, os parâmetros de default podem ser ajustados para modificar a estringência da pesquisa. Por exemplo usando-se BLAST, a resposta de significância estatística (denominado valor E) para relatar comparações contra seqüências de bancos de dados pode ser aumentada para mostrar menos comparações estringentes. Desta maneira, comparações quase exatas curtas podem ser identificadas.

Os métodos para o alinhamento de seqüências para a comparação são bem conhecidos na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. O GAP usa o algoritmo de Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar o

alinhamento global (sobre o total da seqüência) de duas seqüências que maximizam o número de comparações e minimiza o número de fendas. O algoritmo BLAST (Altschul *et al.* (1990) J Mol Biol 215: 403-10) calcula o percentual de identidade de seqüência e realiza uma análise estatística de similaridade entre as duas seqüências. O software para realizar a análise BLAST está publicamente disponível através do National Centre for Biotechnology Information (NCBI). Os homólogos podem ser facilmente identificados usando-se, por exemplo, o algoritmo de alinhamento de seqüência múltiplo ClustalW (versão 1.83), com os parâmetros de alinhamento em pares default e um método de registro em porcentagem. Os homólogos podem ser facilmente identificados usando-se, por exemplo, o algoritmo de alinhamento de seqüência múltiplo ClustalW (versão 1.83), com os parâmetros de alinhamento em pares default e um método de registro em porcentagem. As porcentagens globais de similaridade e identidade também podem ser determinadas suando-se um dos métodos disponíveis no pacote de software MatGAT (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: uma aplicação que gera matrizes de similaridade / identidade ou seqüências de DNA.). A edição manual menor pode ser realizada para otimizar o alinhamento entre os motivos conservados, como pode estar evidentes para uma pessoa habilitada na técnica. Além disso, em vez de usar seqüências de comprimento total para a identificação de homólogos, os domínios específicos também podem ser usados. A identidade dos valores de seqüência pode ser determinada sobre o ácido nucleico inteiro ou seqüência de aminoácido ou sobre domínios selecionados ou motivos conservados, usando-se os programas mencionados acima usando-se os parâmetros default.

Todas as proteínas RanBP1 contém um domínio de ligação de Ran com aproximadamente 150 aminoácidos. Os bancos de dados especialistas existem para a identificação de domínios. Os domínios em RANBPs podem ser identificados usando-se, por exemplo, SMART (Schultz

et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic *et al.* (2002) Nucleic Acids Res 30, 242-244), InterPro (Mulder *et al.*, (2003) Nucl. Acids. Res. 31, 315-318), Prosite (Bucher and Bairoch (1994), Uma sintaxe de perfil generalizado para motivos de seqüências biomoleculares e sua função em interpretação de seqüência automática. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAIPress, Menlo Park; Hulo *et al.*, Nucl. Acids. Res. 32:D134-D137, (2004),) ou Pfam (Bateman *et al.*, Nucleic Acids Research 30(1): 276-280 (2002). Uma série de ferramentas para a análise *in silico* de seqüências de proteína está disponível no servidor ExPASy proteomics (Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31:3784-3788(2003)). Domínios ou motivos também podem ser identificados usando-se técnicas de rotina, al como pelo alinhamento de seqüência.

A invenção é ilustrada (ver a seção Exemplos) transformando-se as plantas com um RANBP de *Zea mays* como representado por SEQ ID N° : 113, que codifica o polipeptídeo da SEQ ID N° : 114 ou SEQ ID N° : 115. A invenção também é ilustrada pela transformação de plantas com um RANBP de *Arabidopsis thaliana* como representado por SEQ ID N° : 116, que codifica o polipeptídeo da SEQ ID N° : 117 ou SEQ ID N° : 118.

É claro, o desempenho dos métodos da invenção não é restrito ao uso das seqüências mencionadas acima, mas pode ser realizado usando-se qualquer ácido nucleico que codifica um RANBP que compreende o Motivo I, como definido neste acima. Os exemplos de tais ácidos nucleicos que codificam RANBPs que compreende o Motivo I inclui ácidos nucleicos que codificam homólogos, ortólogos e parálogos da SEQ ID N° : 114, SEQ ID N° : 115, SEQ ID N° : 117 e SEQ ID N° : 118; os termos homólogos, ortólogos e parálogos como sendo definida acima. Os exemplos de tais homólogos,

ortólogos e parálogos incluem as seqüências listadas na Tabela P do Exemplo 14.

Os ortólogos e parálogos podem ser facilmente encontrados pela realização de uma denominada pesquisa blast recíproca. Tipicamente, esta envolve um primeiro BLAST que envolve submeter uma seqüência em questão ao BLAST (por exemplo, SEQ ID N° : 113, SEQ ID N° : 114, SEQ ID N° : 115, SEQ ID N° : 116, SEQ ID N° : 117 ou SEQ ID N° : 118) contra quaisquer bancos de dados de seqüência, tais como os bancos de dados NCBI publicamente disponíveis. BLASTN ou TBLASTX (usando-se valores default padrão) são no geral usados quando começa-se a partir de uma seqüência de nucleotídeo e BLASTP ou TBLASTN (usando-se valores default padrão) quando começa-se de uma seqüência de proteína. Os resultados de BLAST podem ser opcionalmente filtrados. As seqüências de comprimento total dos resultados filtrados ou resultados não filtrados são então submetidas ao BLAST novamente (segundo BLAST) contra seqüências do organismo a partir do qual a seqüência em questão é derivada (onde a seqüência em questão é SEQ ID N° : 113, SEQ ID N° : 114 ou SEQ ID N° : 115, o segundo BLAST deve ser contra as seqüências de *Zea mays*; onde a seqüência em questão é SEQ ID N° : 116, SEQ ID N° : 117 ou SEQ ID N° : 118, o segundo BLAST deve ser contra *Arabidopsis thaliana* seqüências). Os resultados o primeiro ou segundo BLASTs são então comparados. Um parálogo é identificado se um acerto de classificação alta do primeiro blast é da mesma espécie como a partir do qual a seqüência em questão é derivada, a BLAST novamente, então resulta idealmente na seqüência em questão como o acerto mais alto; um ortólogo é identificado se um acerto de classificação alta no primeiro BLAST não é da mesma espécie como a partir do qual a seqüência em questão é derivada e preferivelmente resulta no BLAST novamente na seqüência em questão estando entre os acertos mais altos.

Acertos de classificação alta são aqueles tendo um valor E.

quanto menor o valor E, mais significante o registro (ou em outras palavras menor a oportunidade de que o acerto seja encontrado por oportunidade). A computação do valor E é bem conhecido na técnica. Além dos valores E, as comparações também são registradas pela identidade de porcentagem. A identidade de porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas seqüências de ácidos nucleicos comparadas (ou polipeptídeo) em um comprimento particular. No caso de famílias grandes, o ClustalW pode ser usado, seguido por uma árvore de união vizinha, para ajudar a visualizar o agrupamento de genes relacionados e para identificar ortólogos e parálogos.

As variantes de ácido nucleico também podem ser úteis na prática dos métodos da invenção. os exemplos de tais variantes incluem ácidos nucleicos que codificam homólogos e derivados de qualquer uma das seqüências de aminoácido dadas na Tabela P de Exemplo 14, os termos "homólogo" e "derivado" sendo como definido neste. Também úteis nos métodos da invenção são os ácidos nucleicos que codificam homólogos e derivados de ortólogos ou parálogos de qualquer uma das seqüências de aminoácido dadas na Tabela P de Exemplo 14. Os homólogos e derivados úteis nos métodos da presente invenção tem, substancialmente, uma mesma atividade biológica e funcional como a proteína não modificada a partir da qual estes são derivados.

Tipicamente, os ácidos nucleicos que codificam RANBPs que compreendem pelo menos o Motivo I tem, em ordem crescente de preferência, pelo menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ou mais identidade de seqüência para a seqüência de ácido nucleico representada pela SEQ ID N° : 113 ou SEQ ID N° : 116.

Também úteis nos métodos da invenção são os ácidos nucleicos que codificam derivados dos aminoácidos representados por SEQ ID N° : 114, SEQ ID N° : 115, SEQ ID N° : 117 ou SEQ ID N° : 118 ou

ácidos nucleicos que codificam derivados dos ortólogos ou parálogos da SEQ ID N° : 114, SEQ ID N° : 115, SEQ ID N° : 117 ou SEQ ID N° : 118.

Os ácidos nucleicos que codificam os polipeptídeos representados pela seqüência na Tabela P ou ácidos nucleicos que codificam ortólogos e parálogos de qualquer um desta SEQ ID N°, não será necessário os ácidos nucleicos de comprimento total, visto que o desempenho dos métodos da invenção não conta com o uso das seqüências de ácido nucleico de comprimento total. Exemplos de ácidos nucleicos adequados para o uso no desempenho dos métodos da invenção incluem, mas não são limitados aqueles representados na Tabela P do Exemplo 14. As variantes de ácido nucleico também podem ser úteis na prática dos métodos da invenção. Exemplos de tais variantes de ácido nucleico incluem porções de ácidos nucleicos que codificam um RANBP, variantes de união de ácidos nucleicos que codificam um RANBP, hibridização de seqüência por ácidos nucleicos que codificam um RANBP, variantes alélicas de ácidos nucleicos que codificam um RANBP e variantes de ácidos nucleicos que codificam um RANBP projetado pela mistura do gene. Os termos variante de união e variante alélica são descritos acima.

Uma porção de um ácido nucleico que codifica um RANBP pode ser preparado, por exemplo, pela fabricação de uma ou mais anulações ao ácido nucleico. As porções podem ser usadas na forma isolada ou estas podem ser fundidas a outras seqüências codificadoras (ou não codificadoras) a fim de, por exemplo, produzir uma proteína que combina com diversas atividades. Quando fundiu as outras seqüências codificadoras, o polipeptídeo resultante produz na tradução podendo ser bem maior do que o predito para a porção RANBP.

As porções úteis nos métodos da invenção, codificam um polipeptídeo que compreende motivo I como descrito acima e tendo substancialmente a mesma atividade biológica como o RANBP representado

pelas seqüências listadas na Tabela P ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. A porção é tipicamente pelo menos 200 nucleotídeos consecutivos de comprimento, preferivelmente pelo menos 300 nucleotídeos consecutivos de comprimento, mais preferivelmente pelo menos 400 nucleotídeos consecutivos de comprimento. Preferivelmente, a porção é uma porção é um ácido nucleico como representado por qualquer uma das seqüências listadas na Tabela P. Mais preferivelmente a porção é uma porção é um ácido nucleico como representado pela SEQ ID N° : 113 ou SEQ ID N° : 116.

De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas, que compreende a introdução e preferivelmente que expressa na semente da planta ou partes da semente uma porção que é um ácido nucleico representado por qualquer uma das seqüências listadas na Tabela P.

Uma outra variante de ácido nucleico útil nos métodos da invenção, é um ácido nucleico capaz de hibridizar sob condições de estringência reduzida, preferivelmente sob condições estringentes reduzidas, como um ácido nucleico que codifica um RANBP como definido neste ou como uma porção como definido neste.

As seqüências de hibridização úteis nos métodos da invenção, codificam um polipeptídeo que compreende Motivo I e tendo substancialmente a mesma atividade biológica como o RANBP representado por qualquer uma das seqüências listadas na Tabela P ou tendo substancialmente a mesma atividade biológica como ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. A seqüência de hibridização é tipicamente pelo menos 200 nucleotídeos consecutivos de comprimento, preferivelmente pelo menos 300 nucleotídeos consecutivos de comprimento, mais preferivelmente pelo menos 400 nucleotídeos consecutivos de comprimento. Preferivelmente, a seqüência de hibridização é

uma que é capaz de hibridização a qualquer um dos ácidos nucleicos representados pelas seqüências listadas na Tabela P ou a uma porção de qualquer uma das seqüências mencionadas acima, uma porção como sendo definida acima. Mais preferivelmente a seqüência de hibridização é capaz de hibridizar um ácido nucleico como representado pela SEQ ID N° : 113 ou 116 ou porções destas.

De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas, que compreende a introdução e preferivelmente que expressa na semente da planta ou partes da semente um ácido nucleico capaz de hibridizar por um ácido nucleico que codifica um RANBP representado por qualquer uma das seqüências listadas na Tabela P ou que compreende a introdução e preferivelmente que expressa na semente da planta ou partes da semente um ácido nucleico capaz de hibridizar a um ácido nucleico que codifica um homólogo, parálogo ou ortólogo de qualquer uma das SEQ ID N° mencionadas acima.

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante de união que codifica um RANBP como definido neste acima. De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas, que compreende a introdução e preferivelmente que expressa na semente da planta ou partes da semente a variante de união de um ácido nucleico que codifica um RANBP representado por qualquer uma das seqüências listadas na Tabela P ou a variante de união de um ácido nucleico que codifica um homólogo, parálogo ou ortólogo de qualquer uma das SEQ ID N° mencionadas acima.

As variantes de união preferidas são as variantes de união de um ácido nucleico que codifica um RANBP representado por qualquer uma das SEQ ID N° : 114, SEQ ID N° 115, SEQ ID N° : 117 ou SEQ ID N° : 118. Ainda preferidos são as variantes de união de ácidos nucleicos representados

por qualquer uma das seqüências listadas na Tabela P. Mais preferida é uma variante de união e é um ácido nucleico como representado pela SEQ ID N° : 113 ou SEQ ID N° : 116.

Uma outra variante de ácido nucleico útil na realização dos métodos da invenção é uma variante alélica de um ácido nucleico que codifica um RANBP como definido neste acima. Variantes alélicas existe na natureza, e abrange os métodos da presente invenção e é usada por estes alelos naturais.

De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas, que compreende a introdução e preferivelmente que expressa na semente da planta ou partes da semente uma variante alélica de um ácido nucleico que codifica um RANBP representado por qualquer uma das seqüências listadas na Tabela P ou que compreende a introdução e expressão na planta de uma alélica de um ácido nucleico que codifica um homólogo, parálogo ou ortólogo de qualquer uma das SEQ ID N° mencionadas acima.

A variante alélica pode ser uma variante alélica de um ácido nucleico que codifica um RANBP representado por qualquer uma das SEQ ID N° : 114, SEQ ID N° 115, SEQ ID N° : 117 ou SEQ ID N° : 118 ou uma variante alélica que é um ácido nucleico que codifica ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. Ainda preferidos são as variantes alélicas de ácidos nucleicos representados por qualquer uma das seqüências listadas na Tabela P. Mais preferida é uma variante alélica que é um ácido nucleico como representado pela SEQ ID N° : 113 ou 116.

Uma variante de ácido nucleico adicional útil nos métodos da invenção é uma variante de ácido nucleico projetada e/ou obtida por mistura do gene. Mistura do gene ou evolução direta podem ser usadas para gerar variantes de ácidos nucleicos que codificam RANBP como definido acima.

De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas,

que compreende a introdução e preferivelmente que expressa na semente da planta ou partes da semente uma variante e é um ácido nucleico representado por qualquer uma das seqüências listadas na Tabela P, pelo qual o ácido nucleico da variante é projetado e/ou obtido por mistura do gene.

Além disso, as variantes de ácido nucleico também podem ser obtidas por mutagênese direcionada ao local. Diversos métodos são disponíveis para atingir a mutagêne direcionada ao local, o mais comum sendo os métodos com base em PCR (current protocols in molecular biology. Wiley Eds.).

Os ácidos nucleicos que codificam RANBPs podem ser derivados a partir de qualquer fonte natural ou artificial. O ácido nucleico pode ser modificado a partir da forma natural em ambiente de composição e/ou genômico através da manipulação humana deliberada. De acordo com uma forma de realização preferida do RANBP que codifica o ácido nucleico é a partir de uma planta, ainda preferivelmente de uma monocotiledônia, mais preferivelmente da família Poaceae, mais preferivelmente do gênero *Zea*, mais preferivelmente de *Zea mays*.

De acordo com uma forma de realização adicional preferida, o RANBP que codifica o ácido nucleico é a partir de uma planta, ainda preferivelmente de uma planta dicotiledônea, ainda preferivelmente da família *Brassicaceae*, mais preferivelmente o ácido nucleico é de *Arabidopsis thaliana*.

A presente invenção também abrange plantas ou partes destas obtidas pelos métodos de acordo com a presente invenção. As plantas ou partes destas compreendem um transgene de ácido nucleico que codifica um RANBP operavelmente ligado a um promotor específico de semente.

A invenção também fornece construtos genéticos e vetores para facilitar a introdução e/ou expressão das seqüências de ácido nucleico úteis nos métodos de acordo com a invenção, na planta.

Portanto, aqui é fornecido um construto de gene que compreende:

(i) Um ácido nucleico que codifica um RANBP que compreende Motivo I como definido neste acima;

(ii) Um promotor específico de semente operavelmente ligado a um ácido nucleico de (i).

As Construções úteis nos métodos de acordo com a presente invenção podem ser construídos usando tecnologia de DNA recombinante bem conhecido as pessoas habilitadas na técnica. Os construtos dos genes podem ser inseridos nos vetores, que podem ser comercialmente disponíveis, adequados para a transformação nas plantas e adequados para a expressão do gene de interesse nas células transformadas. A invenção portanto fornece o uso de um construto do gene como definido neste acima nos métodos da invenção.

As plantas são transformadas com um vetor que compreende a seqüência de interesse (isto é, um ácido nucleico que codifica um RANBP). A pessoa habilitada é bem consciente dos elementos genéticos que podem estar muito presente no vetor a fim de sucessivamente transformar, selecionar e propragar as células hospedeiras contendo uma seqüência de interesse. A seqüência de interesse é operavelmente ligada por uma ou mais seqüências de controle (pelo menos por um promotor). Os termos "elemento regulador", "seqüência de controle" e "promotor" são todas usadas permutavelmente neste e são definidas acima.

O ácido nucleico que codifica um RANBP é operavelmente ligado a um promotor específico da semente, isto é, um promotor que é expressado predominantemente no tecido da semente, mas que pode ter a expressão residual entre outra parte na planta devido à expressão do promotor de vazamento. Ainda preferivelmente, o promotor específico para semente é isolado de um gene que codifica uma proteína que armazena a semente,

especialmente um promotor específico de esdosperma. Um promotor específico de esdosperma refere-se a qualquer promotor capaz de preferencialmente conduzir a expressão do gene de interesse no tecido do endosperma. A referência neste a expressão conduzida "preferencialmente" no tecido do endosperma é tornada significar a expressão conduzida de uma qualquer seqüência operavelmente ligado também no tecido do endosperma substancialmente à exclusão da expressão conduzida entre outra parte na planta, a parte de qualquer expressão residual devido à expressão do promotor de vazamento. Por exemplo, o promotor prolamina mostra a expressão forte no endosperma, com vazamento no meristema, mais especificamente no meristema de broto e/ou centro de descriminação no meristema. Mais preferivelmente o promotor específico do endosperma é isolado de um gene prolamina, tal como um RP6 de prolamina de arroz (Wen *et al.*, (1993) Plant Physiol 101(3): 1115-6) promotor como representado pela SEQ ID N° : 155 ou um promotor de comprimento similar e/ou um promotor com uma origem de expressão similar como o promotor de prolamina de arroz. Exemplos de outros promotores específicos de endosperma que também podem ser usados para realizar os métodos da invenção são mostrados na Tabela 2c acima.

Será claro que a aplicabilidade da presente invenção não é restrita ao RANBP que codifica o ácido nucleico representado pela SEQ ID N° : 113 ou SEQ ID N° : 116, ou é a aplicabilidade da invenção restrita a expressão de um tal RANBP que codifica o ácido nucleico quando conduzido por um promotor prolamina. Exemplos de outros promotores específicos de semente que também podem ser usados para realizar os métodos da invenção que são mostrados na Tabela 2b acima.

Opcionalmente, uma ou mais seqüências terminadoras (também uma seqüência de controle) podem ser usadas no construto introduzido na planta. Reguladores de elemento adicionais podem incluir intensificadores de transcrição bem como de tradução. Aqueles habilitados na

técnica estarão atentos as seqüências terminadora e intensificadora que podem ser adequadas para o uso no desempenho da invenção. Outras seqüências de controle, além disso promotor, intensificador, silenciador, seqüências de íntron, regiões de extremidade 3'UTR e/ou 5'UTR, podem ser proteína e/ou elementos estabilizadores de RNA. Tais seqüências serão conhecidas ou podem realmente ser obtidas por uma pessoa habilitada na técnica.

Os construtos genéticos da invenção ainda podem incluir uma origem de seqüência de replicação que é requerida pela manutenção e/ou replicação em um tipo celular específico. Um exemplo é quando um construto genético é requerido a ser mantido em uma célula bacteriana como um elemento genético episssomal (por exemplo molécula de plasmídio ou cosmídio). As origens preferidas da replicação incluem, mas não são limitados a, o f1-ori e colE1.

Para a detecção da transferência bem sucedida das seqüências de ácido nucleico como usado nos métodos da invenção e/ou seleção de plantas transgênicas que compreendem estes ácidos nucleicos, é vantajoso para o uso dos genes marcadores (ou genes repórter). Portanto, o construto genético pode opcionalmente compreender um gene marcador selecionável. Os genes marcadores podem ser removidos ou taxados a partir da célula transgênica uma vez que estes não são mais longos do que o necessário. As técnicas para a remoção dos marcadores são conhecidas na técnica, as técnicas úteis são descritas acima na seção de definição.

A invenção também fornece um método para a produção de plantas transgênicas tendo características intensificadas relacionadas com o rendimento com relação às plantas de controle, que compreende a introdução e expressão preferencial na semente da planta ou partes da semente de um ácido nucleico que codifica um RANBP que compreende Motivo I como definido neste acima.

Mais especificamente, a presente invenção fornece um método

para a produção de plantas transgênicas tendo traços relacionados de rendimento intensificados, pelo quais métodos compreende:

(i) introdução e expressão na célula da planta ao ácido nucleico que codifica um RANBP que compreende Motivo I (como definido neste) operavelmente ligado ao promotor específico de semente; e

(ii) cultivo da célula da planta sob condições que promovem o crescimento e crescimento da planta.

O ácido nucleico pode ser introduzido diretamente na célula da planta ou na própria planta (incluindo a introdução no tecido, órgão ou qualquer outra parte da planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é preferivelmente introduzido na planta pela transformação.

O termo "transformação" como referido neste é descrito acima.

As células das plantas modificadas geneticamente podem ser regeneradas por intermédio de todos os métodos com pelo qual o habilitado na técnica seja familiar. Métodos adequados podem ser observados nas publicações mencionadas acima por S.D. Kung e R. Wu, Potrykus ou Höfgen e Willmitzer.

Geralmente após a transformação, células das plantas ou agrupamentos celulares são selecionados pela presença de um ou mais marcadores que são codificados pelos genes expressíveis da planta co-transferidos com o gene de interesse, seguindo que o material transformado é regenerado na planta total. Para selecionar as plantas transformadas, o material da planta obtido na transformação é, como uma regra, submetido às condições seletivas de modo que as plantas transformadas podem ser distinguidas das plantas não transformadas. Por exemplo, as sementes obtidas na maneira descrita acima podem ser plantadas e, após um período de crescimento inicial, submetido a uma seleção adequada por pulverização.

Uma possibilidade adicional consiste no crescimento das sementes, se apropriado após a esterilização, em placas de ágar usando um agente de seleção adequado de modo que apenas as sementes transformadas podem desenvolver-se nas plantas. Alternativamente, as plantas transformadas são avaliadas pela presença de um marcador selecionável tal como um dos descritos acima.

Seguindo a transferência e regeneração do DNA, as plantas transformadas putativamente podem ser avaliadas, por exemplo usando análise Southern, pela presença do gene de interesse, número de cópias e/ou organização genômica. Alternativamente ou adicionalmente, os níveis de expressão dos DNA introduzidos novamente podem ser monitorados usando análise Northern e/ou Western ou PCR quantitativo, todas as técnicas sendo bem conhecidas a pessoa tendo habilidade comum na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de significados, tal como pela propagação clonal ou técnicas de geração clássica. Por exemplo, uma primeira geração de planta transformada (ou T1) pode ser por si mesma para dar a segunda geração de transformantes homozigoto (ou T2), e as plantas T2 ainda propagadas através das técnicas de geração clássica.

Os organismos transformados gerados pode tomar uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e células não transformadas; transformantes clonais (por exemplo, todas as células transformadas por conter o cassete de expressão); enxertos de tecidos transformados e não transformados (por exemplo, em plantas, uma origem transformada enxertada por uma muda não transformada).

A presente invenção claramente estende-se a qualquer célula de planta ou planta produzida por qualquer um dos métodos descritos neste, e por todas as partes das plantas e *propagules* destes. A presente invenção

estende-se ainda por abranger a origem de uma célula primária transfectada ou transformada, tecido, órgão ou planta total que tem sido produzida por qualquer um dos métodos mencionados acima, o requerimento apenas sendo a origem exibe a mesma característica genotípica e/ou fenotípica como aquela produzida pela origem nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucleico isolado que codifica um RANBP que compreende Motivo I como definido neste acima. Células hospedeiras preferidas de acordo com a invenção são células das plantas. Plantas hospedeiras para os ácidos nucleicos ou os vetores usados no método de acordo com a invenção, o cassete de expressão ou construto ou vetor são, em princípio, vantajosamente todas as plantas, que são capazes de sintetizar os polipeptídeos usados no método inventivo.

A invenção também estende-se por partes colhíveis da planta tal como, mas não limitadas a sementes, folhas, frutos, flores, hastes, raízes, rizomas, tubérculos e bulbos. A invenção além disso diz respeito aos produtos derivados, preferivelmente diretamente derivados, de uma parte colhível de uma tal planta, tal como grânulos ou pós secos, óleo, gordura e ácidos graxos, amido ou proteínas.

De acordo com uma característica preferida da invenção, a expressão modulada é expressão aumentada. Os métodos para aumentar a expressão de ácidos nucleicos ou genes ou produtos do gene, são bem documentados na técnica e Exemplos são fornecidos na seção de definição.

Como mencionado acima, um método preferido para a expressão de modulação preferencialmente (preferivelmente, aumento) na semente da planta ou partes da semente de um ácido nucleico que codifica um RANBP ocorre pela introdução e expressão na planta a ácido nucleico que codifica um RANBP que compreende Motivo I; entretanto os efeitos do desempenho dos métodos, isto é, características intensificadoras relacionadas

com o rendimento também podem ser atingidas usando outras técnicas bem conhecidas, incluindo mas não limitadas a alvejamento da ativação de T-DNA, TILLING, recombinação homóloga. Uma descrição de algumas destas técnicas é fornecida na seção de definição.

Os efeitos da invenção também podem ser reproduzidas usando recombinação homóloga. O ácido nucleico a ser alvejado é preferivelmente a região de controle da expressão natural de um ácido nucleico que codifica um RANBP na planta. Um promotor específico de semente é introduzido nesta região, substituindo substancialmente parte ou toda esta.

O desempenho dos métodos da invenção dão plantas tendo traços relacionados de rendimento intensificados. Em particular o desempenho dos métodos da invenção dão plantas tendo rendimento aumentado, especialmente biomassa aumentada e rendimento de semente com relação às plantas de controle. Os termos "rendimento" e "rendimento de semente" são descritos em mais detalhes na seção de "definição" deste.

Uma referência neste de uma característica intensificada relacionada com o rendimento é entendido significar um aumento na biomassa (peso) de uma ou mais partes da planta, que podem incluir partes acima do solo (colhíveis) e/ou partes baixo desenvolvidas (colhíveis). Em particular, tais partes colhíveis são acima desenvolvidas de biomassa e sementes, e desempenho dos métodos da invenção resultam em plantas tendo biomassa aumentada e rendimento aumentado da semente com relação aos rendimentos da semente de plantas de controle.

Tomando milho como um exemplo, um aumento do rendimento pode ser manifestado como um ou mais dos seguintes: aumento do número de plantas estabilizadas por hectare ou acre, um aumento no número de espigas por planta, um aumento no número de séries, número de sementes por série, peso da semente, cem no peso da semente,

comprimento/diâmetro de espiga, aumento na taxa enchedora da semente (que é o número de sementes enchedoras divididas pelo número total das sementes e multiplicadas por 100), entre outros. Tomando o arroz como um exemplo, um aumento do rendimento pode manifestar em si mesmo como um aumento em um ou mais dos seguintes: número de plantas por hectare ou acre, número de panículas por planta, número de espigas por panícula, número de flores (flóculos) por panícula (que é expressado como uma razão do número de sementes enchedoras durante o número de panículas primárias), aumento na taxa enchedora da semente (que é o número de sementes enchedoras divididas pelo número total das sementes e multiplicadas por 100), aumento em cem no peso da semente, entre outros.

A presente invenção fornece um método para aumentar o rendimento, especialmente o rendimento da semente das plantas, com relação às plantas de controle, pelo quais os métodos compreendem a expressão da modulação, preferivelmente aumento da expressão, em uma planta de um ácido nucleico que codifica um polipeptídeo RANBP como definido neste.

Visto que as plantas transgênicas de acordo com a presente invenção tem o rendimento aumentado, é semelhante que estas plantas exibam uma taxa de crescimento aumentado (durante pelo menos parte de seu ciclo de vida), com relação a taxa de crescimento das plantas de controle em um estágio correspondente em seu ciclo de vida.

A taxa de crescimento aumentada pode ser específica por uma ou mais partes da planta (incluindo sementes) ou podem ser completamente substancialmente na planta total. As plantas tendo uma taxa de crescimento aumentada pode ter um ciclo de vida curto. O ciclo de vida da planta pode ser tornado por significar a necessidade de tempo para desenvolver de uma semente madura seca até o estágio onde a planta tem produzido sementes maduras secas, similar ao material de partida. Este ciclo de vida pode ser influenciado pelo fatores tal como vigor precoce, taxa de crescimento, índice

de vetor, período de florescência e velocidade da maduração da semente. O aumento em taxa de crescimento pode acontecer em um ou mais estágios no ciclo de vida da planta ou durante substancialmente o ciclo de vida total da planta. Aumento na taxa de crescimento durante os estágios facilmente no ciclo de vida da planta pode refletir vigor intensificado. O aumento em taxa de crescimento pode alterar no ciclo colhível da planta permitindo as plantas a serem semeadas depois e/ou quanto antes colhida do que seria de outra maneira possível (um efeito similar pode ser obtido com o período de florescência antecipado). Se a taxa de crescimento é suficientemente aumentada, pode permitir por uma semeadura adicional das sementes da mesma espécie de plantas (por exemplo semeadura e colheita de plantas de arroz seguido pela semeadura e colheita de plantas de arroz adicionais todas dentro de um período de crescimento convencional). Similarmente, se a taxa de crescimento é suficientemente aumentada, pode permitir por uma semeadura adicional das sementes de diferentes espécies de plantas (por exemplo a semeadura e colheita de plantas de milho seguido pela, por exemplo, a semeadura e colheita adicional de soja, batata ou qualquer outra planta adequada). Os períodos adicionais colhidos da mesma origem no caso de algumas plantas de safra também podem ser possíveis. Alterando o ciclo de colheita da planta pode levar a um aumento em produção de biomassa anual por acre (devido ao um aumento no número de vezes (dito em um ano) que qualquer planta particular pode ser desenvolvida e colhida). Um aumento na taxa de crescimento também pode permitir o cultivo de plantas transgênicas em uma área geográfica ampla do que na sua contra parte de tipo selvagem, visto que as limitações territoriais para o crescimento de uma safra são freqüentemente determinados pelas condições ambientais adversas no período de plantio (época precoce) ou no período de colheita (época tardia). Tais condições adversas podem ser evitadas se o ciclo colhido é mais curto. A taxa de crescimento pode ser determinado pela variação de vários parâmetros

para curvas de crescimento, tais parâmetros podem ser: T-Mid (o período para retirar as plantas para atingir 50% de seu tamanho máximo) e T-90 (tempo de retirado das plantas para atingir 90% de seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção dão plantas tendo uma taxa de crescimento aumentada com relação às plantas de controle. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a taxa de crescimento de plantas, pelo quais os métodos compreendem uma expressão de modulação, preferivelmente aumento da expressão, em uma planta de um ácido nucleico que codifica um polipeptídeo RANBP como definido neste.

Um aumento em rendimento e/ou taxa de crescimento ocorre se a planta está sob condições de não tensão ou se a planta está exposta a várias tensões comparadas as plantas de controle. As plantas tipicamente respondem a exposição a tensão pelo crescimento mais lentamente. Em condições de diversas tensões, a planta ainda pode interromper o crescimento todo. A tensão branda em outra maneira é definida neste como sendo qualquer tensão ao que uma planta está exposta que não resulta na planta cessando para desenvolver-se todo dentro da capacidade para o crescimento resumido. Uma tensão branda no sentido da invenção leva a uma redução no crescimento das plantas tensas de menos do que 40%, 35% ou 30%, preferivelmente menos do que 25%, 20% ou 15%, mais preferivelmente menos do que 14%, 13%, 12%, 11% ou 10% ou menos em comparação à planta de controle sob condições de não tensão. Devido ao avanço na prática da agricultura (irrigação, fertilização, tratamentos de pesticidas) diversas tensões não são freqüentemente encontradas no cultivo das plantas de safra. Como uma conseqüência, o crescimento compromissado induzido pela tensão branda é freqüentemente uma característica não desejado para a agricultura. As tensões brandas são as tensões biótica e/ou abiótica todos os dias (ambiental) ao que uma planta está

exposta. As tensões abióticas podem ser devido à secura ou excesso de água, tensão anaeróbico, tensão da solução salina, toxicidade química, tensão oxidativa e temperaturas de calor, frio ou congelada. As tensões abiótica podem ser uma tensão osmótica causada por uma tensão de água (particularmente devido à secura), tensão da solução salina, tensão oxidativa ou uma tensão iônica. Tensões bióticas são tipicamente aquelas tensões causadas pelos patogênios, tal como bactéria, vírus, fungo e insetos.

Em particular, os métodos da presente invenção podem ser realizados sob condições de não tensão ou sob condições de secura branda para dar as plantas tendo rendimento aumentado com relação às plantas de controle. Como relatado em Wang et al. (Planta (2003) 218: 1-14), tensão abiótica leva a uma série de mudanças morfológicas, fisiológicas, bioquímicas e moleculares que adversamente afetam o crescimento da planta e produtividade. Secura, salinidade, temperaturas extremas e tensão oxidativa são conhecidas a ser interconectadas e podem induzir o crescimento e dano celular através de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve um alto grau particularmente de "*cross talk*" entre tensão de secura e tensão de alta salinidade. Por exemplo, a secura e/ou salinização são manifestados primariamente como tensão osmótica, resultando no rompimento de homeostase e distribuição de íon na célula. Tensão oxidativa, pelo qual freqüentemente acompanha temperatura alta ou baixa, salinidade ou tensão de secura, pode causar desnaturação de proteínas funcional ou estrutural. Como uma conseqüência, estas tensões ambientais diversas freqüentemente célula similar ativa dos caminhos de sinalização e respostas celulares, tal como a produção de proteínas de tensão, sub-regulação de anti-oxidantes, acumulação dos solutos compatíveis e detenção do crescimento. O termo condições de "não tensão" como usado neste são aquelas condições ambientais que permite o crescimento ótimo das plantas. As pessoas habilitadas na técnica são cautelosas das condições de solo normal

e condições climáticas para uma posição dada.

O desempenho dos métodos da invenção dão plantas desenvolvidas sob condições de não tensão ou sob condições de secura branda de rendimento aumentado com relação às plantas de controle desenvolvida sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas desenvolvidas sob condições de não tensão ou sob condições de secura branda, pelo quais métodos compreende aumento da expressão em uma planta de um ácido nucleico que codifica um polipeptídeo RANBP.

O desempenho dos métodos da invenção dão plantas desenvolvidas sob condições de deficiência de nutriente, particularmente sob condições de deficiência de nitrogênio, rendimento aumentado com relação às plantas de controle desenvolvida sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas desenvolvidas sob condições de deficiência de nutriente, pelo quais métodos compreende aumento da expressão em uma planta de um ácido nucleico que codifica um polipeptídeo RANBP. A deficiência de nutriente pode resultar de uma falta de nutrientes tal como nitrogênio, fosfatos e outros compostos contendo fosforosos, potássio, cálcio, cádmio, magnésio, manganês, ferro e boro, entre outros.

Os métodos da invenção são vantajosamente aplicável a qualquer planta. As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular plantas monocotiledôneas e dicotiledôneas incluindo legumes de alimentação ou forragem, plantas ornamentais, safras de alimentação, árvores ou arbustos. De acordo com uma forma de realização preferida da presente invenção, a planta é uma planta de safra. Exemplos de plantas de safra inclui soja, girassol, canola, alfafa, colza, algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta

monocotiledônia. Exemplos de plantas monocotiledôneas inclui cana de açúcar. Mais preferivelmente a planta é uma cereal. Exemplos de cereais inclui arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

A presente invenção também abrange o uso de ácidos nucleicos que codifica RANBPs e uso o RANBPs por si mesmo em características intensificadoras relacionadas com o rendimento em plantas.

Os ácidos nucleicos que codificam RANBPs ou RANBPs por si mesmo, pode observar o uso nos programas de geração em que um marcador de DNA é identificado pelo qual podem ser geneticamente ligados a um RANBP que codifica o gene. Os ácidos nucleicos/genes ou os RANBPs por si mesmo podem ser usados por definir um marcador molecular. Este marcador de DNA ou proteína pode então ser usado nos programas de geração para selecionar plantas tendo o rendimento aumentado como definido neste acima nos métodos da invenção.

As variantes alélicas de um RANBP que codifica um ácido/gene também pode observar o uso em programas de geração assistidos pelo marcador. Tais programas de geração às vezes a introdução requer a variação alélica pelo tratamento mutagênico das plantas, usando por exemplo mutagênese EMS; alternativamente, o programa pode começar com uma coleção de variantes alélicas deste modo denominado origem "natural" não causados intencionalmente. A identificação de variantes alélicas então apresenta, por exemplo, por PCR. Este é seguido por uma etapa para a seleção de variantes alélicas superiores da seqüência em questão e que dão o rendimento aumentado. A seleção é tipicamente realizado pelo monitoramento desenvolvido do desempenho de plantas contente diferente variantes alélicas da seqüência em questão. Desempenho desenvolvido pode ser monitorado em uma estufa para plantas ou no campo. Etapas opcionais adicionais inclui plantas cruzadas em que a variante alélica superior foi identificada com uma outra planta. Este poderá ser usado, por exemplo, por

fazer uma combinação de características de fenotípico de interesse.

Um ácido nucleico que codifica um RANBP também pode ser usado como sondas por mapeamento geneticamente e físico dos genes que estes são uma parte de, e como marcadores por traços ligados aqueles genes. Tal informação pode ser útil na geração da planta a fim de linhas desenvolvidas com fenotipos desejados. Um tal uso de RANBP que codifica os ácidos nucleicos requer apenas uma seqüência de ácido nucleico de pelo menos 15 nucleotídeos de comprimento. O RANBP que codifica os ácidos nucleicos podem ser usados como marcadores de poliformismo de comprimento de fragmento de restrição (RFLP). *Southern blots* (Sambrook J, Fritsch EF e Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico da planta digerida por restrição pode ser sondado com o RANBP que codifica os ácidos nucleicos. As origens de faixa resultante podem então ser submetidas à análise genética usando programas de computador tal como MapMaker (Lander *et al.* (1987) Genomics 1: 174-181) a fim de um mapa genético de um construto. Além disso, os ácidos nucleicos podem ser usados por sonda de Southern blots contendo restrição de DNA genômico tratados com endonuclease de uma série de origem e descendência representando indivíduos de um cruzado genético definido. A segregação do poliformismo de DNA é notado e usado para calcular a posição de RANBP que codifica o ácido nucleico no mapa genético previamente obtido usando esta população (Botstein *et al.* (1980) Am. J. Hum. Genet. 32:314-331).

A produção e o uso de plantas derivadas de gene para o uso no mapeamento genético é descrito em Bematzky e Tanksley (1986) Plant Mol. Biol. Reporter 4: 37-41. Publicações numerosas descrevem o mapeamento genético de clone de cDNA específico usando a metodologia resumida acima ou variações destes. Por exemplo, populações intercruzada F2, populações posterior cruzada, populações unidas aleatoriamente, linhas isogênicas próximas, e outras séries de indivíduos podem ser usados para o mapeamento.

Tais metodologias são bem conhecidas aqueles habilitados na técnica.

As sondas de ácido nucleico também pode ser usado para o mapeamento físico (isto é, disposições das seqüências em mapas físicas; ver Hoheisel *et al. In*: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346, e referências citadas neste).

Em uma outra forma de realização, as sondas de ácido nucleico podem ser usadas no mapeamento de hibridização de fluorescência direta in situ (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora os métodos correntes do mapeamento de FISH a favor do uso de clones amplos (diversos kb por diversos cem kb; ver Laan *et al.* (1995) Genome Res. 5:13-20), melhora na sensibilidade pode permitir o desempenho do mapeamento de FISH usando sondas mais curtas.

Uma variedade de ácido nucleico de métodos com base em amplificação pelo mapeamento genético ou físico podem ser realizado usando os ácidos nucleicos. Exemplos inclui amplificação específica alélica (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polymorphism of PCR-amplified fragments (CAPS; Sheffield *et al.* (1993) Genomics 16:325-332), allele-specific ligation (Landegren *et al.* (1988) Science 241:1077-1080), nucleotide extension reactions (Sokolov (1990) Nucl. Acids Res. 18:3671), Radiation Hybrid Mapping (Walter *et al.* (1997) Nat. Genet. 7:22-28) e Happy Mapping (Dear and Cook (1989) Nucl. Acids Res. 17:6795-6807). Para estes métodos, a seqüência é um ácido nucleico que é usado para projetar e produzir pares iniciadores para o uso na reação de amplificação ou nas reações de extensão iniciadora. O projeto de tais iniciadores é bem conhecido aqueles habilitados na técnica. Em métodos utilizados No mapeamento genético com base em PCR, pode ser necessário para identificar as Diferenças da seqüência de DNA entre as origens do mapeamento cruzado na região correspondente ao instante da seqüência de ácido nucleico. Este, entretanto, geralmente não é necessário para os métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam nas plantas tendo traços relacionados de rendimento intensificados, como descrito neste acima. Estes traços também podem ser combinados com outros traços vantajosos economicamente, tal como traços de intensificação de rendimento adicional, tolerância a outras tensões abióticas e bióticas, características arquitetural de várias modificações de traços e/ou características bioquímicas e fisiológicas.

GLK

Será agora observado que a expressão modular em uma planta de um ácido nucleico que codifica uma proteína semelhante a Golden2 (GLK) dão plantas tendo características intensificadas relacionadas com o rendimento com relação às plantas de controle.

Portanto, a invenção fornece um método para intensificar características relacionadas com o rendimento em plantas com relação às plantas de controle, que compreende a expressão modular em uma planta de um ácido nucleico que codifica uma proteína de GLK ou uma parte da mesma.

Um método preferido para a modulação da (preferivelmente, aumento) expressão de um ácido nucleico que codifica uma proteína semelhante a Golden2 (GLK) ocorre pela introdução e expressão em uma planta, um ácido nucleico que codifica uma tal proteína GLK.

Qualquer referência a seguir a uma "proteína útil nos métodos da invenção" é entendido significar um polipeptídeo GLK como definido neste. Qualquer referência a seguir a um "ácido nucleico útil nos métodos da invenção" é entendido significar um ácido nucleico capaz de que codificar um tal polipeptídeo GLK. O ácido nucleico a ser introduzido em uma planta (e portanto útil na realização dos métodos da invenção) é qualquer ácido nucleico que codifica o tipo de proteína que não será descrito, a seguir, também denominado "ácido nucleico GLK" ou "gene GLK".

O ácido nucleico a ser introduzido em uma planta (e portanto útil na realização dos métodos da invenção) é qualquer ácido nucleico que codifica uma proteína de GLK (Figura 11). O termo "proteína GLK" ou "proteína semelhante a Golden2" refere-se a proteínas reguladoras transcripcionais que compreendem um domínio de ligação DNA GARP (Tamai et al., Plants of cell Physiol. 43, 99-107, 2002). É pressuposto que o domínio GARP é um domínio multifuncional responsável tanto pelo localização nuclear quanto pela ligação de DNA (Hosoda et al., Plants of cell 14, 2015-2021, 2002). As proteínas GLK preferivelmente também compreendem uma região de terminal N que é rica em aminoácidos ácidos, uma parte central de cerca de 100 aminoácidos enriquecido nos aminoácidos básicos e um domínio de terminal C enriquecido no pró- resíduo. A região de terminal C preferivelmente também compreende um domínio de terminal C GARP (GCT) (Rossini et al. 2001).

Os termos "domínio" e "motivo" são definidos na seção de definição neste. As bases de dados especializadas existem para uma identificação de domínios. O domínio GARP em um Regulador de transcrição semelhante a Gloden 2 pode ser identificado usando, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucl. Acids Res 30, 242-244), InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318), Prosite (Bucher e Bairoch (1994). Um índice de perfil generalizado pelos motivos das seqüências biomoleculares e esta função em interpretação da seqüência automática. (In) ISMB-94; Proceedings 2° International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp.53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004)) ou Pfam (Bateman et al., Nucl. Acids Research 30(1): 276-280 (2002)). Uma série de ferramentas para análise in silico das seqüências de proteína é disponível nos servidores proteômicos ExPASY (hospedado pelo Swiss

Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucl. Acids Res. 31:3784-3788(2003)). Os domínios ou motivos também podem ser identificados usando técnicas de rotina, tal como pelo alinhamento de seqüência.

O domínio de DNA de ligação GARP (Tamai et al. 2002) preferivelmente tem três ou mais das seqüências de consenso seguintes:

Seqüência de consenso GARP 1 (SEQ ID N° : 161):
(K/R)(P/M/V/A)(R/K/M)(V/L)(V/D)W(S/T/I/N)(V/AP/S/T/C/H/Q/D)(E/Q/T/D/S)L(H/D)(R/K/Q/A/H/D/E/L/I)(K/R/Q/S/C/V/A/H)F(V/L/I)(A/K/Q/E/H/D/N/R/S)(A/V/C)(V/G/L/I)(N/E/A/Q/D/G/T/I/K/H)(Q/E/H/L/I/M/K/R/S)L

Seqüência de consenso GARP 2 (SEQ ID N° : 162):
G(I/V/L/P/S/H/Q/A/G)(D/E/K/H/Q/N/A)

Seqüência de consenso GARP 3 (SEQ ID N° : 163):
(A/T)(V/I/Y/F/T)P(K/S)(K/R/T/Q/L/S/G/A)(I/V/L)(L/M/R/K)(D/E/Q/K/R/S)(L/I/F/H/V/M/T/R/A)(M/I/L)(N/K/G/S/D/Q/E)

Seqüência de consenso GARP 4 (SEQ ID N° : 164):
(V/I/M/T/E/L/S/N)(E/D/G/N/Y/K/H/Q/P)(N/G/K/T/S/C/R/D)(I/L)(T/D/A/S)(R/N/I/L/V)(E/H/D/S/A/Y/F)(N/E/H)

Seqüência de consenso GARP 5 (SEQ ID N° : 165):
(V/I/L)(A/K)SHLQ(K/M/I)(Y/F)(R/V)

Mais preferivelmente, as seqüências de consenso GARP tem respectivamente as seqüências seguintes:

1:
(K/R)(P/M/V/A)(R/K/M)(V/L)(V/D)W(S/T/I)(V/A/P)(E/Q)LH(R/K/Q)(K/R/Q)FV(A/K/Q/E/H/D)A(V/G)(N/E/A)(Q/E/H)L

2: G(I/V/L)(D/E/K)

3: A(V/I/Y/F)P(K/S)(K/R/T)I(L/M)(D/E/Q)(L/I)M(N/K/G/S)

4: (V/I/M/T/E)(E/D/G/N/Y/K/H/Q/P)(N/G/K/T/S/C/R)(I/L)(T/D)R (E/H)N

5: (V/I)ASHLQK(Y/F)R

Ainda mais preferivelmente, as seqüências de consenso GARP tem respectivamente as seqüências seguintes:

1: K(P/V/A)KVDWTPELHR(K/R)FV(Q/E/H)A(V/G)E(Q/E)L

2: G(I/V/L)(D/E)

3: A(V/Y/F)PSRILE(L/I)M(N/G)

4: (V/I/M/T/E)(E/D/N/Y/K/H/Q)(S/C/R)LTRHN

5: (V/I)ASHLQKYR

Ainda mais preferivelmente adicional, as seqüências de consenso GARP tem respectivamente as seqüências seguintes:

1: K(V/A)KVDWTPELHRRFVQA(V/G)E(Q/E)L

2: G(I/V/L)D

3: AVPSRILE(L/I)MG

4: (I/M/T/E)(E/D/N/Y)(S/C/R)LTRHN

5: IASHLQKYR

Mais preferivelmente, as seqüências de consenso GARP tem respectivamente as seqüências seguintes:

1: KAKVDWTPELHRRFVQAVEQL

2: GID

3: AVPSRILEIMG

4: IDSLTRHN

5: IASHLQKYR

Opcionalmente, a seqüência de consenso GARP 5 é seguida por um outro motivo conservado (seqüência de consenso 6, SEQ ID N° : 166):

SHR(K/R)H(L/M)(L/A/M/I)ARE(A/G/V)EA(A/G)(S/N/T)W

Preferivelmente esta seqüência de consenso 6 tem a seqüência: SHRKH(L/M)(L/M/I)ARE(A/G/V)EA(A/G)(S/N)W

Mais preferivelmente a seqüência de consenso 6 tem a seqüência:

SHRKHMIAREAEAASW

Um motivo de domínio MYB pode, mas não necessita estar presente no domínio GARP (Figura 12). Este domínio MYB pode corresponder à entrada Pfam PF00249 e entrada InterPro IPR001005, e pode compreender a origem Prosite PS00037 (W-[ST]-{W}-{PTLN}-E-[DE]-{GIYS}-{GYPH}-[LIV].) ou origem Prosite PS00334 (W-x(2)-[LI]-[SAG]-x(4,5)-R-{RE}-x(3)-{AG}-x(3)-[YW]-x(3)-[LIVM].) ou origem Prosite PS50090.

As proteínas GLK úteis na presente invenção preferivelmente (mas não necessariamente) também compreendem um domínio GCT (Rossini et al., 2001). A seqüência de consenso para este domínio GCT é dado na SEQ ID N° : 167:

(H/Q)(P/L)S(N/K/S)E(S/V)(I/V/L)DAAIG(D/E)(V/A)(I/L)(S/T/A/V)(N/K/R)PW(L/T)P(L/P)PLGL(K/N)PP(S/A)(V/M/L)(D/E/G)(G/S)V(M/I)(T/A/S/G)EL(Q/H/E)(R/K)(Q/H)G(V/I)(S/N/P/A)(N/E/T/K)(V/I)P(P/Q)

Preferivelmente, este domínio de consenso GCT tem a seqüência:

(H/Q)PS(N/K/S)ESIDAAIGD(V/A)L(S/T/V)KPW(L/T)PLPLGLKPPS(V/L)(D/G)SV(M/I)(S/G)EL(Q/H/E)RQG(V/I)(P/A)(N/K)(V/I)P(P/Q)

Mais preferivelmente, este domínio de consenso GCT tem a seqüência:

QPSSESIDAAIGDVLSKPWLPLPLGLKPPSVDSVMGELQRQGVANVPP

As proteínas GLK também conhecida por ter um conteúdo médio mais alto do que o dos aminoácidos ácidos (D e E) na região do terminal N (a partir do terminal N ao começo do domínio GARP, Figura 11, Figura 12), preferivelmente o conteúdo é uma ordem aumentada de preferência, mais alta do que 12%, 15%, 20%, mais inferior do que 30%. Tipicamente o conteúdo da região D e E do terminal N é cerca de 23%, considerando o conteúdo médio de D e E em proteínas é cerca de 11,9%

(Tabela 3). Similarmente, a partida da região de terminal C no final do domínio GARP e incluindo o domínio GCT é enriquecido de resíduo Pro. Considerando uma proteína média tem um conteúdo Pro de 4,8%, o conteúdo Pro nesta região do terminal C é de 25,4% para a SEQ ID N° : 157. O conteúdo P pode variar nesta região entre 10 e 30%. (faixa: PpGLK1:11,23, PpGLK2: 11,17, ZmG2: 20,73, ZmGLK1: 23,30, AtGLK2, 17,6%, AtGLK1: 20,13.

Tabela 3: Significa composição de aminoácido de proteínas em SWISS PROT (Julho 2004):

Resíduo	Mol%	Resíduo	Mol%
A = Ala	7,80	M = Met	2,37
C = Cys	1,57	N = Asn	4,22
D = Asp	5,30	P = Pro	4,85
E = Glu	6,59	Q = Gln	3,93
F = Phe	4,02	R = Arg	5,29
G = Gly	6,93	S = Ser	6,89
H = His	2,27	T = Thr	5,46
I = Ile	5,91	V = Val	6,69
K = Lys	5,93	W = Trp	1,16
L = Leu	9,62	Y = Tyr	3,09

Exemplos de proteínas GLK como definido neste inclui a proteína representada pela SEQ ID N° 157, mas o termo "proteínas GLK" também abrange ortólogos ou parálogos da SEQ ID N° mencionada acima: 157. A invenção é ilustrada pelas plantas em transformação com a seqüência Oryza sativa representada pela SEQ ID N° : 156, que codifica o polipeptídeo da SEQ ID N° : 157. SEQ ID N° : 169 (de *Oryza sativa*, codificado pela SEQ ID N° : 168) é um parálogo do polipeptídeo da SEQ ID N° : 157 considerando SEQ ID N° : 171 e 173 de um *Arabidopsis thaliana* (codificado pela SEQ ID N° : 170 e 172), SEQ ID N° : 175 e 177 de *Physcomitrella patens* (codificado pela SEQ ID N° : 174 e 176), SEQ ID N° : 179 e 181 de *Zea mays* (codificado pela SEQ ID N° : 178 e 180), SEQ ID N° : 183, uma seqüência parcial de *Triticum aestivum*, e SEQ ID N° : 189, uma seqüência parcial de *Sorgo bicolor*, são Exemplos de ortólogos da proteína da SEQ ID N° : 157. SEQ ID N° : 193 que representa uma variante da proteína da SEQ ID N° : 157.

Ortólogos e parálogos podem ser facilmente observados pela realização denominada uma busca de blast recíproca. Estes podem ser feito por um primeiro BLAST envolvendo explosão de uma seqüência de questão (por exemplo, SEQ ID N° : 156 ou SEQ ID N° : 157) contra qualquer base de dados de seqüência, tal como a base de dados NCBI disponível publicamente. BLASTN ou TBLASTX (usando valores defeituosos padrões) podem ser usados quando a partida de uma seqüência de nucleotídeo e BLASTP ou TBLASTN (usando valores defeituosos padrões) podem ser usados quando a partida de uma seqüência de proteína. Os resultados de BLAST podem opcionalmente ser filtrados. As seqüências de comprimento total dos resultados filtrados ou resultados não filtrados são então posteriormente estragados (segundo BLAST) contra as seqüências dos organismos de que a seqüência em questão é derivada (onde a seqüência em questão é a SEQ ID N° : 156 ou SEQ ID N° : 157, o segundo BLAST seria portanto contra as seqüências de arroz). Os resultados do primeiro e segundo BLASTs são então comparados. Um parálogo é identificado se um acerto alto de classificação a partir do segundo BLAST é da mesma espécie como da seqüência em questão é derivada; um ortólogo é identificado se um acerto alto de classificação não é da mesma espécie como da seqüência em questão é derivada. Ortólogos preferidos são ortólogos de SEQ ID N° : 156 ou SEQ ID N° : 157. Acertos de alta classificação são aqueles tendo um valor E inferior. O valor E inferior, o mais significante da contagem (ou em outras palavras a chance inferior que o acerto foi observado pela escolha). A computação do valor E é bem conhecido na técnica. Além disso os valores E, as comparações também são contadas pela identidade da porcentagem. A identidade da porcentagem refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas seqüências de ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. Preferivelmente a contagem é maior do que 50, mais preferivelmente maior do que 100; e preferivelmente o valor E é menos do

que e-5, mais preferivelmente menos do que e-6. No caso de famílias maiores, ClustalW pode ser usado, seguido pela geração de uma árvore de ligação próxima, para ajudar a visualizar o agrupamento dos genes relatados e para identificar ortólogos e parálogos.

Os homólogos (ou proteínas de homólogos, abrangem ortólogos e parálogos) podendo ser facilmente identificado usando técnicas de rotina bem conhecida na técnica, tal como pelo alinhamento de seqüência. Os métodos para o alinhamento das seqüências para comparação são bem conhecidas na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. GAP usa o algoritmo de Needleman e Wunsch ((1970) J Mol Biol 48: 443-453) para observar o alinhamento de duas seqüências completas que maximizam diversos pontos e minimizam diversas fendas. O algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-410) calcula o percentual da identidade de seqüência e realiza uma análise estatística da similaridade entre as duas seqüências. O software para a realização da análise BLAST é publicamente disponível do National Centre for Biotechnology Information. Homólogos podendo ser facilmente identificado usando, por exemplo, o algoritmo de alinhamento de seqüência múltipla ClustalW (versão 1.83), com os parâmetros de alinhamento em formas de pares defeituosos, e o método de contagem em porcentagem. As porcentagens globais de similaridade e identidade também podem ser determinadas usando um dos métodos disponível na embalagem do software MatGAT (Campanella et al., BMC Bioinformatics. 4, 29, 2003). A edição menor do manual pode ser realizada para otimizar o alinhamento entre motivos conservados, bem como apresentar para uma pessoa habilitada na técnica. Além disso, em vez do uso das seqüências de comprimento total para a identificação de homólogos, domínios específicos (tal como o domínio GARP ou o domínio GCT) também podem ser usados.

Preferivelmente, as proteínas GLK úteis nos métodos da

presente invenção tem, um aumento na ordem de preferência, pelo menos 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% da identidade de seqüência da proteína da SEQ ID N° : 157. Alternativamente, a identidade de seqüência entre homólogos podem ser determinados usando um domínio específico (tal como o domínio GARP ou o domínio GCT). Um GARP ou domínio GCT pode ser identificado e delineado usando a base de dados e ferramentas para identificação listada acima, e/ou os métodos para o alinhamento das seqüências para a comparação. Em alguns exemplos, os parâmetros defeituosos podem ser ajustados para modificar a estringência da busca. Por exemplo usando BLAST, a significância estatística começa (denominada "exceto" valor) por pontos de relatório contra as seqüências de base de dados que podem ser aumentadas para mostrar menos pontos estringentes. Nesta maneira, pontos exatos aproximadamente curtos podem ser identificados.

Um exemplo detalhado da identificação de homólogos é dado no Exemplo 21. As matrizes mostram no Exemplo 22 as similaridades e identidades (em ambos) durante o GARP ou domínio GCT, onde o curso dos valores são mais altos do que quando considerado a proteína de comprimento total.

O ácido nucleico que codifica o polipeptídeo representado por qualquer uma das SEQ ID N° 157, SEQ ID N° : 193 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima, não será necessário ácido nucleico de comprimento total, visto que o desempenho dos métodos da invenção não conta com o uso das seqüências de ácido nucleico de comprimento total. Além disso, Exemplos de ácidos nucleicos adequados para o uso no desempenho dos métodos da invenção incluem mas não são limitados aqueles listados na Tabela Q do Exemplo 21. As variantes de ácido nucleico também podem ser úteis na prática dos métodos da invenção. Exemplos de tais variantes incluem porções de ácidos nucleicos, seqüências

de hibridização, variantes de união, variantes alélicas ocorrência natural ou pela manipulação do DNA.

O termo "porção" como usado neste refere-se a uma parte do DNA que codifica um polipetídeo que compreende pelo menos um domínio GARP como descrito acima, e preferivelmente também, o terminal C ao terminal N, (i) uma região enriquecido em ácidos nucleicos ácidos (D ou E), precedido do domínio GARP e (ii) uma região de terminal C do domínio GARP que é enriquecido no resíduo Pro e preferivelmente compreende um domínio GCT.

Uma porção pode ser preparada, por exemplo, pela fabricação uma ou mais anulações ao ácido nucleico que codifica uma proteína de GLK como definido neste acima. As porções podem ser usadas na forma isolada ou estas podem ser fundidas a outras seqüências que codificam (ou não codificam) a fim de, por exemplo, produzir uma proteína que combina diversas atividades. Quando fundiu as outras seqüências codificadoras, o polipeptídeo resultante produz na tradução podendo ser bem maior do que o predito para a porção GLK. As porções úteis nos métodos da invenção, codificam um polipeptídeo tendo um domínio GARP (como descrito acima) e tendo substancialmente a mesma atividade biológica como a proteína GLK representado por qualquer uma das SEQ ID N° : 157, SEQ ID N° : 193 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. A porção é tipicamente pelo menos 800 nucleotídeos de comprimento, preferivelmente pelo menos 900 nucleotídeos de comprimento, mais preferivelmente pelo menos 1000 nucleotídeos de comprimento e mais preferivelmente pelo menos 1100 nucleotídeos de comprimento. Preferivelmente, a porção é uma porção é um ácido nucleico como representado por qualquer uma das seqüências listadas na Tabela Q do Exemplo 21. Mais preferivelmente a porção é uma porção que é um ácido nucleico como representado pela SEQ ID N° : 156.

De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas, que compreende introdução e expressão na planta uma porção de qualquer uma da seqüência de ácido nucleico dada na Tabela Q do Exemplo 21 ou uma porção de um ácido nucleico que codifica um homólogo, parálogo ou ortólogo de qualquer uma das seqüências de aminoácidos dadas na Tabela Q do Exemplo 21.

Uma outra variante de ácido nucleico útil nos métodos da invenção é um ácido nucleico capaz de hibridizar sob condições de estringência reduzida, preferivelmente sob condições estringentes reduzidas, como um ácido nucleico que codifica uma proteína de GLK como definido neste acima ou como uma porção como definido neste acima.

As seqüências de hibridização úteis nos métodos da invenção, codificam um polipeptídeo tendo uma região de terminal N enriquecida em ácidos nucleicos ácidos (D ou E), um domínio GARP e uma região de terminal C do domínio GARP que é enriquecido no resíduo Pro e que preferivelmente compreende um domínio GCT (como descrito acima) e tendo substancialmente a mesma atividade biológica como a proteína GLK representado por qualquer uma das SEQ ID N° : 157, SEQ ID N° 193 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. A seqüência de hibridização é tipicamente pelo menos 800 nucleotídeos de comprimento, preferivelmente pelo menos 900 nucleotídeos de comprimento, mais preferivelmente pelo menos 1000 nucleotídeos de comprimento e mais preferivelmente pelo menos 1100 nucleotídeos de comprimento. Preferivelmente, a seqüência de hibridização é uma que é capaz de hibridização a qualquer dos ácidos nucleicos representados pelas seqüências (ou por sondas derivadas de) listadas na Tabela Q do Exemplo 21 ou a uma porção de qualquer uma das seqüências mencionadas acima, uma porção como sendo definida acima. Mais preferivelmente a seqüência de hibridização

é capaz de hibridização a SEQ ID N° : 156 ou as porções (ou sondas) destas. Os métodos para as sondas projetadas são bem conhecida na técnica. As sondas são geralmente menos do que 1000 bp de comprimento, preferivelmente menos do que 500 bp de comprimento. Comumente, os comprimentos da sonda para a hibridização de DNA-DNA tal como *Southern blotting*, variam entre 100 e 500 bp, considerando a região de hibridização em sondas pela hibridização de DNA-DNA tal como na amplificação de PCR geralmente são mais curtos do que 50 mais longos do que 10 nucleotídeos. De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas, que compreende introdução e expressão em uma planta de um ácido nucleico capaz de hibridizar a qualquer um dos ácidos nucleicos dados na Tabela Q do Exemplo 21 ou que compreende introdução e expressão em uma planta de um ácido nucleico capaz de hibridizar a um ácido nucleico que codifica um homólogo, parálogo ou ortólogo de qualquer uma das seqüências de ácido nucleico dadas na Tabela Q do Exemplo 21.

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante de união que codifica uma proteína GLK como definido neste acima.

As variantes de união preferidas são as variantes de união que é um ácido nucleico que codifica as proteínas GLK representadas por qualquer uma das SEQ ID N° : 157, SEQ ID N° 193 ou variantes de união que codifica ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. Ainda preferidos são as variantes de união de ácidos nucleicos representados por qualquer uma das seqüências listadas na Tabela Q do Exemplo 21. Mais preferida é uma variante de união que é um ácido nucleico como representado pela SEQ ID N° : 156.

De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas,

que compreende introdução e expressão na planta a variante de união de qualquer uma da seqüência de ácido nucleicos dada na Tabela Q do Exemplo 21 ou a variante de união de um ácido nucleico que codifica um homólogo, parálogo ou ortólogo de qualquer uma das seqüências de aminoácido dadas na Tabela Q do Exemplo 21.

Uma outra variante de ácido nucleico útil na realização dos métodos da invenção é uma variante alélica de um ácido nucleico que codifica uma proteína de GLK como definido neste acima. Variantes alélicas existe na natureza, e abrange os métodos da presente invenção e é usada por estes alelos naturais. A variante alélica pode ser uma variante alélica de um ácido nucleico que codifica uma proteína de GLK representado por qualquer uma das SEQ ID N° : 157, SEQ ID N° 193 ou uma variante alélica que é um ácido nucleico que codifica ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima. Ainda preferidos são as variantes alélicas de ácidos nucleicos representados por qualquer uma das seqüências listadas na Tabela Q do Exemplo 21. Mais preferida é uma variante alélica que é um ácido nucleico como representado pela SEQ ID N° : 156, tal como SEQ ID N° : 192.

De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas, que compreende introdução e expressão na planta de uma alélica de qualquer um dos ácidos nucleicos dados na Tabela Q do Exemplo 21 ou que compreende a introdução e expressão na planta de uma alélica de um ácido nucleico que codifica um homólogo, parálogo ou ortólogo de qualquer uma das seqüências de aminoácido dadas na Tabela Q do Exemplo 21.

Uma variante de ácido nucleico adicional útil nos métodos da invenção é uma variante de ácido nucleico obtido pela mistura do gene. A mistura do gene ou evolução direta também pode ser usada para gerar variantes de ácidos nucleicos que codificam as proteínas GLK como definido

acima. De acordo com a presente invenção, é fornecido um método para intensificar características relacionadas com o rendimento em plantas, que compreende introdução e expressão na planta, uma variante de qualquer uma da seqüência de ácido nucleico dada na Tabela Q do Exemplo 21 ou que compreende introdução e expressão na planta de uma variante de um ácido nucleico que codifica um homólogo, parálogo ou ortólogo de qualquer uma das seqüências de aminoácido dadas na Tabela Q do Exemplo 21, pelo qual o ácido nucleico da variante é obtido pela mistura do gene.

Além disso, as variantes de ácido nucleico também podem ser obtidas por mutagêne direcionada ao local. Diversos métodos são disponíveis para atingir mutagêne direcionada ao local, a mais comum sendo métodos com base em PCR (Current Protocols in Molecular Biology. Wiley Eds.). Os mutantes preferidos são aqueles que resultam na mudança identidade do tecido da estrutura do tecido C3 à anatomia Kranz das plantas C4.

Também útil nos métodos da invenção são os ácidos nucleicos que codificam os homólogos de qualquer um dos aminoácidos representados pela SEQ ID N° 157, SEQ ID N° 193 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima.

Também útil nos métodos da invenção são os ácidos nucleicos que codificam os derivados de qualquer uma das seqüências de aminoácido representados pela SEQ ID N° 157, SEQ ID N° 193 ou de ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima.

Além disso, as proteínas GLK úteis nos métodos da presente invenção (pelo menos na sua forma natural) tipicamente, mas não necessariamente, tem atividade reguladora transcripcional. Portanto, as proteínas GLK com atividade reguladora transcripcional reduzida ou sem atividade reguladora transcripcional pode ser igualmente útil nos métodos da presente invenção. Uma pessoa habilitada na técnica pode determinar facilmente a presença da ligação da atividade de DNA ou ativação

transcripcional usando ferramentas e técnicas de rotina. Para determinar a ligação da atividade de DNA das proteínas GLK, diversos ensaios são disponíveis (por exemplo Current Protocols in Molecular Biology, Volumes 1 e 2, Ausubel et al. (1994), Current Protocols). Em particular, a ligação do ensaio de DNA para os fatores transcripcional que compreende um domínio GARP é descrito por Hosoda et al. (2002), incluindo uma seleção do local de ligação de DNA assistido por PCR e um ensaio de mudança em gel pela ligação do DNA. Alternativamente, o método de Tamai *et al.* (2002) será usado, onde um *Arabidopsis* GPRI1 como usado em um ensaio para conduzir a transcrição de um gene repórter *lacZ*; Rossini et al 2001 além disso descreve um ensaio de trans-ativação de levedura GAL4.

Os ácidos nucleicos que codificam as proteínas GLK podem ser derivados a partir de qualquer fonte natural ou artificial. O ácido nucleico pode ser modificado a partir da forma natural em ambiente de composição e/ou genômico através da manipulação humana deliberada. Preferivelmente a proteína GLK que codifica o ácido nucleico e é a partir de uma planta, ainda preferivelmente de uma planta monocotiledônea, mais preferivelmente da família de Poaceae, mais preferivelmente o ácido nucleico é de *Oryza sativa*.

A invenção também fornece construtos genéticos e vetores para facilitar a introdução e/ou expressão das seqüências de ácido nucleico útil nos métodos de acordo com a invenção, na planta.

Portanto, aqui é fornecido um construto de gene que compreende:

(i) um ácido nucleico que codifica uma proteína de GLK como definido neste acima;

(ii) uma ou mais seqüências de controle operavelmente ligado a um ácido nucleico de (i).

As construções úteis nos métodos de acordo com a presente invenção podem ser construídos usando tecnologia de DNA recombinante

bem conhecido as pessoas habilitadas na técnica. Os construtos dos genes podem ser inseridos nos vetores, que podem ser comercialmente disponíveis, adequados para a transformação nas plantas e adequados para a expressão do gene de interesse nas células transformadas. A invenção portanto fornece o uso de um construto do gene como definido neste acima nos métodos da invenção. Preferivelmente, o construto do gene é para a expressão de condução de GLK em plantas.

As plantas são transformadas com um vetor que compreende a seqüência de interesse (isto é, um ácido nucleico que codifica uma proteína de GLK). A pessoa habilitada é bem consciente dos elementos genéticos que podem estar muito presente no vetor a fim de sucessivamente transformar, selecionar e propragar células hospedeiras contendo uma seqüência de interesse. A seqüência de interesse é operavelmente ligado por uma ou mais seqüências de controle (pelo menos por um promotor).

Vantajosamente, qualquer tipo de promotor pode ser usado para conduzir a expressão da seqüência de ácido nucleico. Preferivelmente, o GLK que codifica o ácido nucleico ou variante deste é operavelmente ligado a um promotor constitutivo. Um promotor constitutivo é transcripcionalmente mais ativo em, mas não necessariamente todas as, fases deste crescimento e crescimento e sob condições mais ambientais em pelo menos uma célula, tecido ou órgão. Um promotor constitutivo preferido é um promotor constitutivo que é também substancialmente ubiquitosamente expressada. Ainda preferivelmente o promotor é derivado de uma planta, mais preferivelmente uma planta monocotiledônia. Mais preferida é o uso de um promotor GOS2 (de arroz) (SEQ ID N° : 160 ou SEQ ID N° : 58). Será claro que a aplicabilidade da presente invenção não é restrita ao GLK que codifica o ácido nucleico representado pela SEQ ID N° : 156 ou SEQ ID N° : 192, não é a aplicabilidade da invenção restrita da expressão de um ácido nucleico que codifica uma proteína de GLK quando conduzido por um promotor GOS2.

Exemplos de outros promotores constitutivos que também podem ser usados para conduzir a expressão de um ácido nucleico que codifica uma proteína de GLK são mostrados na seção de definição acima.

Opcionalmente, uma ou mais seqüências terminadoras podem ser usadas no construto introduzido na planta. Reguladores de elemento adicionais podem incluir intensificadores de transcrição bem como de tradução. Aqueles habilitados na técnica estarão atentos as seqüências terminadora e intensificadoras que podem ser adequadas para o uso no desempenho da invenção. Uma seqüência de íntron pode ser adicionado a região 5' não traduzida (UTR) ou na seqüência codificadora para aumentar a quantidade da mensagem madura que acumula no citosol, como escrito na seção de definição. Outras seqüências de controle (além disso promotor, intensificador, silenciador, seqüências de íntron, regiões de extremidade 3'UTR e/ou 5'UTR) podem ser proteína e/ou elementos estabilizadores de RNA. Tais seqüências serão conhecidas ou podem realmente ser obtidas por uma pessoa habilitada na técnica.

Os construtos genéticos da invenção ainda podem incluir uma origem de seqüência de replicação que é requerida pela manutenção e/ou replicação em um tipo celular específico. Um exemplo é quando um construto genético é requerido a ser mantido em uma célula bacteriana como um elemento genético episssomal (por exemplo molécula de plasmídio ou cosmídio). As origens preferidas da replicação incluem, mas não são limitados a, o f1-ori e colE1.

Para a detecção da transferência bem sucedida das seqüências de ácido nucleico como usado nos métodos da invenção e/ou seleção de plantas transgênicas que compreende este ácidos nucleicos, é vantajoso para o uso dos genes marcadores (ou genes repórter). Portanto, o construto genético pode opcionalmente compreender um gene marcador selecionável. Marcadores selecionáveis são descritos em mais detalhes na seção de

"definição" deste. Os genes marcadores podem ser removidos ou taxados a partir da célula transgênica uma vez que estes não são mais longos do que o necessário. As técnicas para a remoção dos marcadores são conhecidas na técnica, as técnicas úteis são descritos acima na seção de definição.

A invenção também fornece um método para a produção de plantas transgênicas tendo alterado o traço relacionado ao rendimento com relação às plantas de controle, que compreende a introdução e a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo GLK como definido neste acima.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo alterado os traços relacionados ao rendimento, pelo quais os métodos compreendem:

(i) introdução e expressão de um ácido nucleico que codifica uma proteína de GLK na célula da planta; e

(ii) cultivo da célula da planta sob condições que promovem o crescimento e crescimento da planta.

O ácido nucleico pode ser introduzido diretamente na célula da planta ou na própria planta (incluindo a introdução no tecido, órgão ou qualquer outra parte da planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é preferivelmente introduzido na planta pela transformação. O termo "transformação" é descrito em mais detalhes na seção de definição neste.

Geralmente após a transformação, as células das plantas ou agrupamentos celulares são selecionados pela presença de um ou mais marcadores que são codificados pelos genes expressíveis da planta co-transferidos com o gene de interesse, seguindo que o material transformado é regenerado na planta total. Para selecionar as plantas transformadas, o material da planta obtido na transformação é, como uma regra, submetido às condições seletivas de modo que as plantas transformadas podem ser

distinguidas das plantas não transformadas. Por exemplo, as sementes obtidas na maneira descrita acima podem ser plantadas e, após um período de crescimento inicial, submetido a uma seleção adequada por pulverização. Uma possibilidade adicional consiste no crescimento das sementes, se apropriado após a esterilização, em placas de ágar usando um agente de seleção adequado de modo que apenas as sementes transformadas podem desenvolver-se nas plantas. Alternativamente, as plantas transformadas são avaliadas pela presença de um marcador selecionável tal como um dos descritos acima.

Seguindo a transferência e regeneração do DNA, as plantas transformadas putativamente podem ser avaliadas, por exemplo usando análise Southern, pela presença do gene de interesse, número de cópias e/ou organização genômica. Alternativamente ou adicionalmente, os níveis de expressão dos DNA introduzidos novamente podem ser monitorados usando análise Northern e/ou Western ou PCR quantitativo, todas as técnicas sendo bem conhecidas a pessoa tendo habilidade comum na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de significados, tal como pela propagação clonal ou técnicas de geração clássica. Por exemplo, uma primeira geração planta transformada (ou T1) pode ser por si mesma para dar a segunda geração de transformantes homozigoto (ou T2), e as plantas T2 ainda propagadas através das técnicas de geração clássica.

Os organismos transformados gerados pode tomar uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e células não transformadas; transformantes clonais (por exemplo, todas as células transformadas por conter o cassete de expressão); enxertos de tecidos transformados e não transformados (por exemplo, em plantas, uma origem transformada enxertada por uma muda não transformada).

A presente invenção claramente estende-se a qualquer célula de planta ou planta produzida por qualquer um dos métodos descritos neste, e por todas as partes das plantas e *propagules* destes. A presente invenção estende-se ainda por abranger a origem de uma célula primária transfectada ou transformada, tecido, órgão ou planta total que tem sido produzida por qualquer um dos métodos mencionados acima, o requerimento apenas sendo a origem exibe a mesma característica genotípica e/ou fenotípica como aquela produzida pela origem nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucleico isolado que codifica uma proteína de GLK como definido neste acima. As células hospedeiras preferidas de acordo com a invenção são células das plantas. Plantas hospedeiras para os ácidos nucleicos ou os vetores usados no método de acordo com a invenção, o cassete de expressão ou construto ou vetor são, em princípio, vantajosamente todas as plantas, que são capazes de sintetizar os polipeptídeos usados no método inventivo.

A invenção também estende-se por partes colhíveis da planta tal como, mas não limitadas a sementes, folhas, frutos, flores, hastes, raízes, rizomas, raízes, tubérculos e bulbos. A invenção além disso diz respeito aos produtos derivados, preferivelmente diretamente derivados, de um parte colhível de uma tal planta, tal como grânulos ou pós secos, óleo, gordura e ácidos graxos, amido ou proteínas.

De acordo com uma característica preferida da invenção, a expressão modulada é expressão aumentada. Os métodos para aumentar a expressão de ácidos nucleicos ou genes ou produtos do gene, são bem documentados na técnica e Exemplos são fornecidos na seção de definição.

Como mencionado acima, um método preferido para a modulação da expressão (preferivelmente, aumento) de um ácido nucleico que codifica uma proteína de GLK que ocorre pela introdução e expressão na planta de um ácido nucleico que codifica uma proteína de GLK; entretanto os

efeitos do desempenho dos métodos, isto é, alterando as características relacionadas com o rendimento também podem ser atingidas usando outras técnicas bem conhecidas, incluindo mas não limitadas a alvejamento da ativação de T-DNA, TILLING, recombinação homóloga. Uma descrição de algumas destas técnicas é fornecida na seção de definição.

O desempenho dos métodos da invenção dão plantas tendo traços relacionados de rendimento intensificados. Em particular desempenho dos métodos da invenção dão plantas tendo rendimento aumentado, especialmente rendimento aumentado de semente com relação às plantas de controle. Os termos "rendimento" e "rendimento de semente" são descritos em mais detalhes na seção de "definição" deste.

A referência neste é uma característica intensificada relacionada com o rendimento que é entendido significar um aumento na biomassa (peso) de uma ou mais partes da planta, que podem incluir partes acima do solo (colhíveis) e/ou partes baixo desenvolvidas (colhíveis). Em particular, tais partes colhíveis são acima desenvolvidas biomassa e/ou sementes, e desempenho dos métodos da invenção resultam em plantas tendo rendimento aumentado de semente e/ou aumentado acima da trituração de biomassa, com relação aos rendimentos das sementes e/ou biomassa das plantas de controle.

Tomando o milho como um exemplo, um aumento do rendimento pode ser manifestado como um ou mais dos seguintes: aumento do número de plantas estabilizadas por hectare ou acre, um aumento no número de espigas por planta, um aumento no número de séries, número de sementes por série, peso da semente, cem no peso da semente, comprimento/diâmetro de espiga, aumento na taxa enchedora da semente (que é o número de sementes enchedoras divididas pelo número total das sementes e multiplicadas por 100), entre outros. Tomando o arroz como um exemplo, um aumento do rendimento pode manifestar em si mesmo como um aumento

em um ou mais dos seguintes: número de plantas por hectare ou acre, número de panículas por planta, número de espigas por panícula, número de flores (flóculos) por panícula (que é expressado como uma razão do número de sementes enchedoras durante o número de panículas primárias), aumento na taxa enchedora da semente (que é o número de sementes enchedoras divididas pelo número total das sementes e multiplicadas por 100), aumento em cem no peso da semente, entre outros.

A presente invenção fornece um método para aumentar o rendimento, especialmente acima da trituração da biomassa e/ou rendimento das semente das plantas, com relação às plantas de controle, pelo quais os métodos compreendem uma expressão da modulação, preferivelmente aumento da expressão, em uma planta de um ácido nucleico que codifica um polipeptídeo GLK como definido neste.

Visto que as plantas transgênicas de acordo com a presente invenção tem o rendimento aumentado, é semelhante que estas plantas exibam uma taxa de crescimento aumentado (durante pelo menos parte de seu ciclo de vida), com relação a taxa de crescimento das plantas de controle em um estágio correspondente em seu ciclo de vida.

A taxa de crescimento aumentada pode ser específica por uma ou mais partes da planta (incluindo sementes) ou podem ser completamente substancial a planta total. As plantas tendo uma taxa de crescimento aumentada pode ter um ciclo de vida curto. O ciclo de vida da planta pode ser tornado por significar a necessidade de tempo para desenvolver de um semente madura seca até o estágio onde a planta tem produzido sementes maduras secas, similar ao material de partida. Este ciclo de vida pode ser influenciado pelo fatores tal como vigor precoce, taxa de crescimento, índice de vetor, período de florescência e velocidade da maduração da semente. O aumento em taxa de crescimento pode acontecer em um ou mais estágios no ciclo de vida da planta ou durante substancialmente o ciclo de vida total da

planta. Aumento na taxa de crescimento durante os estágios facilmente no ciclo de vida da planta pode refletir vigor intensificado. O aumento em taxa de crescimento pode alterar no ciclo colhível da planta permitindo as plantas a serem semeadas depois e/ou quanto antes colhida do que seria de outra maneira possível (um efeito similar pode ser obtido com o período de florescência antecipado). Se a taxa de crescimento é suficientemente aumentada, pode permitir por uma semeadura adicional das sementes da mesma espécie de plantas (por exemplo semeadura e colheita de plantas de arroz seguido pela semeadura e colheita de plantas de arroz adicionais todas dentro de um período de crescimento convencional). Similarmente, se a taxa de crescimento é suficientemente aumentada, pode permitir por uma semeadura adicional das sementes de diferentes espécies de plantas (por exemplo a semeadura e colheita de plantas de milho seguido pela, por exemplo, a semeadura e colheita adicional de soja, batata ou qualquer outra planta adequada). Os períodos adicionais colhidos da mesma origem no caso de algumas plantas de safra também podem ser possíveis. Alterando o ciclo de colheita da planta pode levar a um aumento na produção de biomassa anual por acre (devido ao um aumento no número de vezes (dito em um ano) que qualquer planta particular pode ser desenvolvido e colhida). Um aumento em taxa de crescimento também pode permitir o cultivo de plantas transgênicas em uma área geográfica ampla do que na sua contra parte de tipo selvagem, visto que as limitações territoriais para o crescimento de uma safra são freqüentemente determinadas pela condições ambientais adversas no período de plantio (época precoce) ou no período de colheita (época tardia). Tais condições adversas podem ser evitadas se o ciclo colhido é mais curto. A taxa de crescimento pode ser determinado pela variação de vários parâmetros para curvas de crescimento, tais parâmetros podem ser: T-Mid (o período para retirar as plantas para atingir 50% de seu tamanho máximo) e T-90 (tempo de retirado das plantas para atingir 90% de seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção dão plantas tendo uma taxa de crescimento aumentada com relação às plantas de controle. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a taxa de crescimento de plantas, pelo quais métodos compreendem a expressão da modulação, preferivelmente aumento da expressão, em uma planta de um ácido nucleico que codifica um polipeptídeo GLK como definido neste.

Um aumento em rendimento e/ou taxa de crescimento ocorre se a planta está sob condições de não tensão ou se a planta está exposta a várias tensões comparadas as plantas de controle. As plantas tipicamente respondem a exposição a tensão pelo crescimento mais lentamente. Em condições de diversas tensões, a planta ainda pode interromper o crescimento todo. A tensão branda em outra maneira é definida neste como sendo qualquer tensão ao que uma planta está exposta que não resulta na planta cessando para desenvolver todo dentro da capacidade para o crescimento resumido. A tensão branda no sentido da invenção leva a uma redução no crescimento das plantas tensas de menos do que 40%, 35% ou 30%, preferivelmente menos do que 25%, 20% ou 15%, mais preferivelmente menos do que 14%, 13%, 12%, 11% ou 10% ou menos em comparação à planta de controle sob condições de não tensão. Devido ao avanço na prática da agricultura (irrigação, fertilização, tratamentos de pesticidas) diversas tensões não são freqüentemente encontradas no cultivo das plantas de safra. Como uma conseqüência, o crescimento compromissado induzido pela tensão branda é freqüentemente uma característica não desejada para a agricultura. As tensões brandas são as tensões biótica e/ou abiótica todos os dias (ambiental) ao que uma planta está exposta. As tensões abióticas podem ser devido à secura ou excesso de água, tensão anaeróbico, tensão da solução salina, toxicidade química, tensão oxidativa e temperaturas de calor, frio ou congelada. As tensões abiótica podem ser uma tensão osmótica causada por uma tensão de água

(particularmente devido à secura), tensão da solução salina, tensão oxidativa ou uma tensão iônica. Tensões bióticas são tipicamente aquelas tensões causadas pelos patogênios, tal como bactéria, vírus, fungo e insetos.

Em particular, os métodos da presente invenção podem ser realizados sob condições de não tensão ou sob condições de secura branda para dar as plantas tendo rendimento aumentado com relação às plantas de controle. Como relatado em Wang et al. (Planta (2003) 218: 1-14), tensão abiótica leva a uma série de mudanças morfológicas, fisiológicas, bioquímicas e moleculares que adversamente afetam o crescimento da planta e produtividade. Secura, salinidade, temperaturas extremas e tensão oxidativa são conhecidas a ser interconectadas e podem induzir o crescimento e dano celular através de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve um alto grau particularmente de "*cross talk*" entre tensão de secura e tensão de alta salinidade. Por exemplo, secura e/ou salinização são manifestados primariamente como tensão osmótica, resultando no rompimento de homeostase e distribuição de íon na célula. Tensão oxidativa, pelo qual freqüentemente acompanha temperatura alta ou baixa, salinidade ou tensão de secura, pode causar desnaturação de proteínas funcionais ou estruturais. Como uma conseqüência, estas tensões ambientais diversas freqüentemente célula similar ativa caminhos de sinalização e respostas celulares, tal como a produção de proteínas de tensão, sub-regulação de anti-oxidantes, acumulação dos solutos compatíveis e detenção do crescimento. O termo condições de "não tensão" como usado neste são aquelas condições ambientais que permite o crescimento ótimo de plantas. As pessoas habilitadas na técnica são cautelosas das condições de solo normal e condições climáticas para uma posição dada.

O desempenho dos métodos da invenção dão plantas desenvolvidas sob condições de não tensão ou sob condições de secura branda rendimento aumentado com relação às plantas de controle desenvolvida sob

condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas desenvolvidas sob condições de não tensão ou sob condições de secura branda, pelo quais métodos compreende aumento da expressão em uma planta de um ácido nucleico que codifica um polipeptídeo GLK.

O desempenho dos métodos da invenção dão plantas desenvolvidas sob condições de deficiência de nutriente, particularmente sob condições de deficiência de nitrogênio, rendimento aumentado com relação às plantas de controle desenvolvida sob condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas desenvolvidas sob condições de deficiência de nutriente, pelo quais métodos compreende aumento da expressão em uma planta de um ácido nucleico que codifica um polipeptídeo GLK. A deficiência de nutriente pode resultar de uma falta de nutrientes tal como nitrogênio, fosfatos e outros compostos contendo fosforosos, potássio, cálcio, cádmio, magnésio, manganês, ferro e boro, entre outros.

Os métodos da invenção são vantajosamente aplicável a qualquer planta. As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular plantas monocotiledôneas e dicotiledôneas incluindo legumes de alimentação ou forragem, plantas ornamentais, safras de alimentação, árvores ou arbustos. De acordo com uma forma de realização preferida da presente invenção, a planta é uma planta de safra. Exemplos de plantas de safra inclui soja, girassol, canola, alfafa, colza, algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta monocotiledônia. Exemplos de planta monocotiledôneas inclui cana de açúcar. Mais preferivelmente a planta é uma cereal. Exemplos de cereais inclui arroz, milho, trigo, cevada, painço, centeio, *triticale*, sorgo e aveias.

A presente invenção também abrange plantas obtidas pelos

métodos de acordo com a presente invenção. A presente invenção portanto fornece plantas, partes de plantas ou células das plantas destes obtidos pelos métodos de acordo com a presente invenção, pelo qual as plantas ou partes ou células destas compreendem um transgene de ácido nucleico que codifica uma proteína GLK como definido acima.

A presente invenção também abrange o uso de ácidos nucleicos que codificam proteínas GLK e o uso de polipeptídeos GLK que alteram os traços relatados de rendimento.

Os ácidos nucleicos que codificam os polipeptídeos GLK ou proteínas GLK por si mesmo, podem observar o uso nos programas de geração em que um marcador de DNA é identificado pelo qual podem ser geneticamente ligados a uma proteína GLK que codifica o gene. Os ácidos nucleicos/genes ou as proteínas GLK por si mesmas podem ser usadas por definir um marcador molecular. Este marcador de DNA ou proteína pode então ser usado nos programas de geração para selecionar plantas tendo rendimento aumentado como definido neste acima nos métodos da invenção.

As variantes alélicas de uma proteína GLK que codifica ácido/gene também pode observar o uso em programas de geração assistidos pelo marcador. Tais programas de geração as vezes a introdução requer a variação alélica pelo tratamento mutagênico das plantas, usando por exemplo mutagênese EMS; alternativamente, o programa pode começar com uma coleção de variantes alélicas deste modo denominado origem "natural" não causados intencionalmente. A identificação de variantes alélicas então apresenta, por exemplo, por PCR. Este é seguido por uma etapa para a seleção de variantes alélicas superiores da seqüência em questão e que dão rendimento aumentado. Seleção é tipicamente realizada pelo monitoramento desenvolvido do desempenho de conteúdo de plantas diferentes, variantes alélicas da seqüência em questão. Desempenho desenvolvido podem ser monitorado em uma estufa para plantas ou no campo. Etapas opcionais

adicionais inclui plantas cruzadas em que a variante alélica superior foi identificada com uma outra planta. Este poderá ser usado, por exemplo, por fazer uma combinação de características de fenotípico de interesse.

Um ácido nucleico que codifica uma proteína de GLK também pode ser usado como sondas por mapeamento geneticamente e físico os genes que estes são uma parte de, e como marcadores por traços ligados aqueles genes. Tal informação pode ser útil na geração da planta a fim de linhas desenvolvidas com fenotipos desejados. Um tal uso de GLK que codifica os ácidos nucleicos requer apenas uma seqüência de ácido nucleico de pelo menos 15 nucleotídeos de comprimento. O GLK que codifica os ácidos nucleicos podem ser usados como marcadores de poliformismo de comprimento de fragmento de restrição (RFLP). *Southern blots* (Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico da planta digerida por restrição podem ser sondado com o GLK que codificam os ácidos nucleicos. As origens de faixa resultante podem então ser submetidas à análise genética usando programas de computador tal como MapMaker (Lander et al. (1987) Genomics 1: 174-181) a fim de um mapa genético de um construto. Além disso, os ácidos nucleicos podem ser usados por sonda *Southern blots* contendo restrição DNA genômico tratados com endonuclease de uma série de origem e descendência representando indivíduos de um cruzado genético definido. A segregação do poliformismo de DNA é notado e usado para calcular a posição de GLK que codifica o ácido nucleico no mapa genético previamente obtido usando esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

A produção e o uso de plantas derivadas do gene para o uso no mapeamento genético é descrito em Bernatzky e Tanksley (Plant Mol. Biol. Reporter 4: 37-41, 1986). Publicações numerosas descrevem mapeamento genético de clone de cDNA específico usando a metodologia resumida acima ou variações destes. Por exemplo, populações intercruzada F2, populações

posterior cruzada, populações unidas aleatoriamente, linhas isogênicas próximas, e outras séries indivíduos podem ser usadas para o mapeamento. Tais metodologias são bem conhecidas aqueles habilitados na técnica.

As sondas de ácido nucleico também pode ser usadas para o mapeamento físico (isto é, disposições das seqüências em mapas físicas; ver Hoheisel et al. In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346, e referências citadas neste).

Em uma outra forma de realização, as sondas de ácido nucleico podem ser usada no mapeamento de hibridização de fluorescência direta in situ (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora os métodos correntes de Mapeamento de FISH favoreçam o uso de clones amplos (diversos kb por diversos cem kb; ver Laan et al. (1995) Genome Res. 5:13-20), melhora na sensibilidade pode permitir desempenho do mapeamento de FISH usando sondas mais curtas.

Uma variedade de métodos de ácido nucleico com base em amplificação pelo mapeamento genético ou físico podem ser realizado usando os ácidos nucleicos. Exemplos inclui amplificação específica alélica (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polymorphism of PCR-amplified fragments (CAPS; Sheffield et al. (1993) Genomics 16:325-332), allele-specific ligation (Landegren et al. (1988) Science 241:1077-1080), nucleotide extension reactions (Sokolov (1990) Nucl. Acids Res. 18:3671), Radiation Hybrid Mapping (Walter et al. (1997) Nat. Genet. 7:22-28) e Happy Mapping (Dear and Cook (1989) Nucl. Acids Res. 17:6795-6807). Para estes métodos, a seqüência é um ácido nucleico é usado para projetar e produzir pares iniciadores para o uso na reação de amplificação ou nas reações de extensão iniciadora. O projeto de tais iniciadores é bem conhecido aqueles habilitados na técnica. Em métodos utilizados no mapeamento genético com base em PCR, podem ser necessários para identificar as diferenças da seqüência de DNA entre as origens do mapeamento cruzado na região

correspondente ao instante seqüência de ácido nucleico. Este, entretanto, geralmente não é necessário para o métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam nas plantas tendo alterado os traços relatados de rendimento, como descrito neste acima. Estes traços também podem ser combinados com outros traços vantajosos economicamente, tal como traços de intensificação de rendimento adicional, tolerância a outras tensões abióticas e bióticas, características arquitetural de várias modificações de traços e/ou características bioquímicas e fisiológicas.

REV DHDZip/START

Tem sido agora surpreendentemente observado que a redução da expressão na planta de um gene REV endógeno usando um domínio zíper de leucina homeodomínio delta REV (HDZip) /Regulador agudo esteroidogênico (STAR) que diz respeito a uma seqüência de ácido nucleico de domínio de transferência de lipídio (START) dando plantas que tem aumento relativo para plantas de controle. A presente invenção portanto fornece métodos para aumentar a produção em plantas relativas para plantas de controle, pela redução da expressão na planta de um gene REV endógeno usando uma seqüência de ácido nucleico *REV ΔHDZip*/START.

Vantajosamente, o desempenho dos métodos de acordo com a presente invenção resulta em plantas que tem produção intensificada que diz respeito a traços, particularmente aumento da produção, mais particularmente semente aumentada na produção e/ou biomassa aumentada, que diz respeito as plantas de controle. Os termos "produção" e "produção da semente" são descritos em mais detalhes na seção de "definições" neste. A biomassa aumentada pode manifestar em si mesma como biomassa de raiz aumentada. A biomassa de raiz aumentada pode ser devido aos números aumentados de raízes, espessura de raiz aumentada e/ou comprimento de raiz aumentada.

O termo "produção aumentada" também refere-se a um

aumento na biomassa (peso) de uma ou mais partes de uma planta, que pode incluir partes acima do solo (colhíveis) e/ou partes (colhíveis) abaixo trituradas. Tais partes colhíveis incluem biomassa vegetativa e/ou sementes, e o desempenho dos métodos da invenção resultam em plantas tendo produção aumentada (em biomassa vegetativa e/ou semente) relativa à produção de plantas de controle.

Portanto, de acordo com a presente invenção, aqui é fornecido um método para aumentar a produção da semente e/ou biomassa da planta, que compreende métodos de redução da expressão na planta de um gene REV endógeno usando uma seqüência de ácido nucleico *REV ΔHDZip*/START.

Em particular, a semente aumentada na produção é selecionada a partir de um ou mais dos seguintes: (i) peso da semente aumentada; (ii) número aumentado de sementes enchedoras; (iii) taxa de abastecimento da semente aumentada; (iv) índice de revestimento aumentado; e (v) comprimento da semente individual aumentada.

Tomando o milho como um exemplo, um aumento da produção pode ser manifestada como um ou mais dos seguintes: aumento de diversas plantas estabelecidas por hectare ou acre, um aumento de diversas espigas por plantas, um aumento de diversas séries, diversas sementes por série, peso da semente, peso singular da semente, comprimento da espiga/diâmetro, aumento da taxa de enchimento da semente (que são diversas sementes enchedoras divididas pelo número total de sementes e multiplicadas por 100), entre outros. Tomando o arroz como um exemplo, um aumento da produção pode manifestar em si mesma como um aumento de um ou mais dos seguintes: diversas plantas por hectare ou acre, diversas panículas por plantas, diversas espigas por panícula, número de flores (flósculos) por panícula (que é expressado como uma razão de diversas sementes enchedoras em diversas panículas primárias), aumento da taxa de enchimento da semente (que são diversas sementes enchedoras divididas pelo número total de

sementes e multiplicadas por 100), aumento no peso singular da semente, entre outros.

A presente invenção fornece um método para aumentar a produção, especialmente produção da semente de plantas, que diz respeito as plantas de controle, que compreende métodos de modulação de expressão, preferivelmente expressão aumentada, na planta de um ácido nucleico codificando um polipeptídeo REV ΔHDZip/START como definido neste.

Visto que as plantas transgênicas de acordo com a presente invenção tem produção aumentada, isto é provavelmente que estas plantas exibem uma taxa de crescimento aumentada (durante pelo menos parte de seu ciclo de vida), relativa a taxa de crescimento de plantas de controle em estágios correspondentes em seu ciclo de vida.

A taxa de crescimento aumentado pode ser específica a uma ou mais partes de uma planta (incluindo sementes), ou pode ser completamente substancial na planta total. As plantas que tem uma taxa de crescimento aumentada pode ter um ciclo de vida curto. O ciclo de vida de uma planta pode ser retirado por significar o período necessário ao crescimento a partir de uma semente madura seca até o estágio onde as plantas tem produzido sementes maduras secas, similar ao material de partida. Este ciclo de vida pode ser influenciado pelos fatores tais como vigor precoce, taxa de crescimento, índice da verdura, período de florescência e rapidez da maturação da semente. O aumento na taxa de crescimento pode acontecer em um ou mais estágios no ciclo de vida de uma planta ou durante ou durante o ciclo de vida total na planta substancial. Aumento na taxa de crescimento durante os estágios precoces no ciclo de vida de uma planta pode refletir vigor intensificado. O aumento na taxa de crescimento pode alterar o ciclo de colheita de uma planta permitindo as plantas serem semeadas depois e/ou colhidas mais cedo do que seria possível de outra maneira (um efeito similar pode ser obtido com período de florescência precoce). Se a taxa de

crescimento é suficientemente aumentada, isto permite uma semeadura adicional das sementes das mesmas espécies de plantas (por exemplo semeadura e colheita de plantas de arroz seguidas pela semeadura e colheita de plantas de arroz adicionais todas dentro de um período de crescimento convencional). Similarmente, se a taxa de crescimento é suficientemente aumentada, isto permite uma semeadura adicional de sementes de diferentes espécies de plantas (por exemplo a semeadura e colheita de plantas de milho seguidas pela, por exemplo, a semeadura e colheita opcional de soja, batata ou qualquer outra planta adequada). Os períodos de colheita adicionais a partir da mesma fonte no caso de algumas plantas de colheita também pode ser possível. Alterando o ciclo de colheita de uma planta pode levar a um aumento na produção de biomassa anual por acre (devido a um aumento de diversos períodos (anuncia em um ano) que qualquer planta particular pode ser desenvolvida e colhida). Um aumento na taxa de crescimento também pode permitir para o cultivo de plantas transgênicas em uma área geográfica extensa do que suas contra-partes de tipo selvagem, visto que as limitações territoriais para crescimento de uma safra são freqüentemente determinadas pelas condições ambientais adversas no período do plantio (época precoce) ou no período da colheita (época atrasada). Tais condições adversas podem ser evitadas se o ciclo de colheita é encurtado. A taxa de crescimento pode ser determinada pelos vários parâmetros derivados a partir de curvas de crescimento, tais parâmetros podem ser: *T-Mid* (o período de retirada das plantas para atingir 50% de seu tamanho máximo) e T-90 (período de retirada da planta para atingir 90% de seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção dão as plantas que tem uma taxa de crescimento aumentada que diz respeito as plantas de controle. Portanto, de acordo com a presente invenção, aqui é fornecido um método para aumentar a taxa de crescimento de plantas, que compreende métodos de

modulação de expressão, preferivelmente expressão aumentada, na planta de um ácido nucleico codificando um polipeptídeo REV ΔHDZip/START como definido neste.

Um aumento na produção e/ou taxa de crescimento ocorre se a planta está sob condições sem-tensão ou se a planta está exposta a várias tensões comparadas as plantas de controle. As plantas tipicamente respondem a exposição por tensão pelo crescimento mais lentamente. Em condições de grave tensão, a planta ainda pode interromper o crescimento completamente. Tensão branda, por outro lado é definida neste como sendo qualquer tensão que uma planta está exposta pelo qual não resulta na interrupção do crescimento completamente da planta sem a capacidade de recuperar o crescimento. Tensão branda no sentido da invenção conduz a uma redução no crescimento das plantas submetidas à tensão de menos do que 40%, 35% ou 30%, preferivelmente menos do que 25%, 20% ou 15%, mais preferivelmente menos do que 14%, 13%, 12%, 11% ou 10% ou menos em comparação ao controle da planta sob condições sem-tensão. Devido a melhora nas práticas agrícolas (irrigação, fertilização, tratamentos com pesticidas) graves tensões não são freqüentemente encontradas em safra de plantas cultivadas. Como uma conseqüência, o crescimento compromissado induzido pela tensão branda é freqüentemente uma característica indesejável para a agricultura. Tensões brandas são as tensões (ambientais) abióticas ou bióticas diárias que uma planta está exposta. Tensões abióticas podem ser devido à aridez ou excesso de água, tensão anaeróbica, tensão ao sal, toxicidade química, tensão oxidativa e temperaturas de calor, frio ou congeladas. A tensão abiótica pode ser uma tensão osmótica causada por uma tensão à água (particularmente devido à aridez), tensão ao sal, tensão oxidativa ou uma tensão iônica. Tensões bióticas são tipicamente aquelas tensões causadas por patogênios, tal como bactéria, vírus, fungo e insetos.

Em particular, os métodos da presente invenção podem ser

realizados sob condições sem-tensão ou sob condições de aridez branda para dar as plantas que tem produção aumentada que diz respeito as plantas de controle. Como relatadas em Wang et al. (Plant (2003) 218: 1-14), tensão abiótica conduz a uma série de mudanças morfológica, fisiológica, bioquímica e molecular que adversamente afetam o crescimento da planta e produtividade. Aridez, salinidade, temperaturas extremas e tensão oxidativa são conhecidas a serem interconectadas e podem induzir o crescimento e dano celular através de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve um alto grau particularmente de "*cross talk*" entre tensão da aridez e tensão de alta salinidade. Por exemplo, aridez e/ou salinização são manifestadas primariamente como tensão osmótica, resultando no rompimento de homeostase e distribuição de íon na célula. Tensão oxidativa, que freqüentemente acompanha temperatura alta ou baixa, salinidade ou tensão da aridez, pode causar desnaturização das proteínas funcionais e estruturais. Como uma conseqüência, estas diversas tensões ambientais freqüentemente ativam os caminhos de sinalização celular similares e respostas celulares, tal como a produção de proteínas de tensão, super-regulação de anti-oxidantes, acumulação de solutos compatíveis e deter o crescimento. O termo condições "sem-tensão" como usado neste são aquelas condições ambientais que permitem ótimo crescimento de plantas. As pessoas habilitadas na técnica estão atentas as condições normais do solo e condições climáticas para uma dado local.

O desempenho dos métodos da invenção dão crescimento de plantas sob condições sem-tensão ou produção aumentada de sob condições de aridez branda que diz respeito as plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, aqui é fornecido um método para aumentar a produção em plantas desenvolvidas sob condições sem-tensão ou sob condições de aridez branda, que compreende métodos de expressão aumentada na planta de um ácido nucleico codificando

um polipeptídeo REV ΔHDZip/START.

O desempenho dos métodos da invenção dão o crescimento de plantas sob condições de deficiência de nutriente, particularmente sob condições de deficiência de nitrogênio, produção aumentada que diz respeito as plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, aqui é fornecido um método para aumentar a produção em plantas desenvolvidas sob condições de deficiência de nutriente, que compreende métodos expressão aumentada na planta de um ácido nucleico codificando um polipeptídeo REV ΔHDZip/START. A deficiência de nutriente pode resultar a partir de uma falta de nutrientes tal como nitrogênio, fosfatos e outros compostos contendo fosforosos, potássio, cálcio, cádmio, magnésio, manganês, ferro e boro, entre outros.

A referência neste a um gene REV "endógeno" não apenas refere-se a um gene REV como observado na planta nesta forma natural (isto é, sem ter qualquer intervenção humana), mas também refere-se à seqüência do ácido nucleico REV isolado subseqüentemente introduzido na planta. Por exemplo, uma planta transgênica que contém um transgene REV por encontrar uma redução substancial da expressão transgene e/ou redução substancial da expressão de um gene REV endógeno, de acordo com os métodos da invenção.

O termo "expressão" ou "expressão do gene" significa a transcrição de um gene específico ou genes específicos ou construto genético específico. O termo "expressão" ou "expressão do gene" em particular significa a transcrição de um gene ou genes ou construto genético em RNA estrutural (rRNA, tRNA) ou mRNA com ou sem tradução subseqüente do último na proteína. Os processos incluem transcrição de DNA e processamento do produto de mRNA resultante.

A "redução" ou "diminuição" da expressão são usados permutavelmente neste, e refere-se, para os métodos da presente invenção, a

uma diminuição, mas não eliminação, da expressão REV do gene endógeno e/ou níveis de polipeptídeo REV e/ou atividade de polipeptídeo REV, usando uma seqüência de ácido nucleico REV ΔHDZip/START, relativo respectivamente à expressão REV do gene e/ou nível de polipeptídeo REV e/ou atividade de polipeptídeo REV observado em plantas de controle. A redução da expressão REV do gene e/ou nível de polipeptídeo REV e/ou atividade de polipeptídeo REV é retirado por significar no sentido da aplicação pelo menos 10%, 20%, 30%, 40% ou 50%, preferivelmente pelo menos 60%, 70 ou 80%, mais preferivelmente 85%, 90%, ou 95% menos expressão REV do gene e/ou nível de polipeptídeo REV e/ou atividade de polipeptídeo REV em comparação a uma planta de controle como definido neste. Preferivelmente, a redução da expressão do REV do gene endógeno usando uma seqüência de ácido nucleico REV ΔHDZip/START leva ao aparecimento de um ou mais traços fenotípicos.

Esta redução da expressão REV do gene endógeno pode ser atingida pelo uso qualquer de um ou mais de métodos "silenciamento do gene" diversos bem conhecidos (ver seção de definições por mais detalhes). O termo "silenciamento" de um gene como usado neste refere-se à redução, mas não eliminação, da expressão REV do gene endógeno.

Um método preferido para redução da expressão na planta de um gene REV endógeno por intermédio do silenciamento mediado por RNA é pelo uso uma repetição invertida de uma seqüência de ácido nucleico REV ΔHDZip/START, preferivelmente capaz de formação de uma estrutura grampo de cabelo. A repetição invertida é clonada em um vetou de expressão que compreende seqüências de controle. Uma seqüência de ácido nucleico de DNA que codifica (um espaçador, por exemplo um fragmento de região de ligação de matriz (MAR), um íntron, um poliligador, etc.) é localizado entre as duas seqüências de ácido nucleico REV ΔHDZip/START invertidas formando uma repetição invertida. Após a transcrição de uma repetição

invertida, um RNA quimérico como uma estrutura complementar própria é formado (parcial ou completo). Esta estrutura de RNA filamentada dupla é referida como o RNA grampo de cabelo (hpRNA). O hpRNA é processado pela planta em siRNAs que são incorporados em um RISC. O RISC ainda cliva os transcritos de mRNA codificando um polipeptídeo REV, deste modo reduzindo o número de transcritos de mRNA a serem traduzidos em um polipeptídeo REV. Ver por exemplo, Grierson et al. (1998) WO 98/53083; Waterhouse et al. (1999) WO 99/53050).

A expressão de um gene REV endógeno também pode ser reduzido pela introdução de uma modificação genética, dentro do local de um gene REV ou em outra parte no genoma. O local de um gene como definido neste é retirado por significar uma região genômica, que inclui o gene de interesse e 10 kb a montante ou a jusante da região codificadora.

A modificação genética pode ser introduzida, por exemplo, por qualquer um (ou mais) dos seguintes métodos: identificação de T-DNA, TILLING, mutagênese direcionado ao local, evolução direcionada, recombinação homóloga. Seguindo a introdução da modificação genética, existe uma seguinte etapa de seleção pela redução da expressão de um gene REV endógeno, pelo qual a redução na expressão dá as plantas que tem produção aumentada comparada as plantas de controle.

A identificação de T-DNA envolve inserção de uma T-DNA, na região genômica do gene de interesse ou 10 kb a montante ou a jusante da região codificadora de um gene em uma configuração tal que o T-DNA reduz (mas não elimina) a expressão do gene alvejado.

A recombinação homóloga permite a introdução no genoma de um ácido nucleico selecionado na posição selecionada definida. Alternativamente, um programa de avaliação pode ser apresentado para identificar nas variantes naturais da população de plantas de um gene REV pelo qual REV de polipeptídeos codificam variantes com atividade reduzida.

Tais variantes naturais também podem ser usadas por exemplo, para realizar a recombinação homóloga.

A identificação de T-DNA, TILLING, mutagênese direcionado ao local e evolução direcionada são exemplos de tecnologias que intensificam a geração de novos alelos e variantes da seqüência de ácido nucleicos REV pelo qual o REV de polipeptídeos codificam variantes com atividade reduzida.

Outros métodos, tal como o uso de anticorpos direcionados a um polipeptídeo REV endógeno para inibir esta função na planta, ou interferência no caminho de sinalização em que um polipeptídeo REV está envolvido, será bem conhecido pela pessoa habilitada na técnica.

Para ótimo desempenho, o silenciamento mediado por técnicas de RNA usadas para redução da expressão na planta de um gene REV endógeno usando uma seqüência de ácido nucleico REV ΔHDZip/START, requere-se o uso da seqüência de ácido nucleicos a partir de plantas monocotiledôneas para transformação de plantas monocotiledôneas, e a partir de plantas dicotiletiledônias para transformação de plantas dicotiletiledônias. Preferivelmente, a seqüência de ácido nucleico REV ΔHDZip/START de qualquer dão espécies de plantas que são introduzida naquela mesma espécie. Por exemplo, a seqüência de ácido nucleico REV ΔHDZip/START a partir do arroz é transformada em uma planta de arroz. A seqüência de ácido nucleico REV ΔHDZip/START não necessita ser introduzida na mesma variedade de planta.

A referência neste a uma "seqüência de ácido nucleico" é retirado por significar uma forma polimérica de um polímero desoxiribonucleotídeo ou um ribonucleotídeo de qualquer comprimento, filamentado duplo ou simples, ou análogos destes, que tem a característica essencial de um ribonucleotídeo natural em que pode hibridizar a seqüência de ácido nucleicos em uma maneira similar aos polinucleotídeos de

ocorrência natural.

A referência neste a uma seqüência de ácido nucleico "REV ΔHDZip/START" é retirada por significar um comprimento suficiente de nucleotídeos contíguos substancialmente a partir de uma seqüência de ácido nucleico REV substancialmente excluindo a parte que codifica os domínios HDZip e START. A fim de realizar o silenciamento do gene, este pode ser tão pequeno quanto 20 ou nucleotídeos menores, alternativamente este pode ser tão maior quanto uma seqüência de ácido nucleico REV ΔHDZip/START (incluindo a extremidade 5' e/ou 3' UTR, em parte ou no total. Uma pessoa habilitada na técnica seria cauteloso que um comprimento suficiente de nucleotídeos contíguos substancialmente a partir de uma seqüência de ácido nucleico REV que codifica os domínios HDZip e START é ser excluído no desempenho dos métodos da invenção. Este pode ser tão pequeno quanto 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 ou mais nucleotídeos. Este pode ser tão maior quanto uma seqüência de ácido nucleico completa que codifica os domínios HDZip e START. Pela seqüência de ácido nucleico "que codifica os domínios HDZip e START" é entendido neste que a região da seqüência de ácido nucleico compreende códons que são traduzidos em resíduos de aminoácido dos domínios HDZip e START (pelo qual os domínios não necessitavam ser completa ou funcional). Uma pessoa habilitada na técnica seria cautelosa que os nucleotídeos contíguos substancialmente a partir de uma seqüência de ácido nucleico REV pode sobrepor limites de domínio HDZip e START por poucos nucleotídeos, tipicamente por não mais do que 20 nucleotídeos. Também nenhum desempenho excluído dos métodos da invenção são as seqüências de ácidos nucleicos que reduziriam simultaneamente a expressão de pelo menos um outro gene endógeno, com respeito se os outros genes endógenos codificados por um polipeptídeo que compreende um domínio HDZip e START ou não. A seqüência de ácido nucleico codificando um polipeptídeo (funcional) não é uma necessidade para

os vários métodos debatidos acima para a redução substancial da expressão de um gene REV endógeno.

Os genes REV são bem conhecidos na técnica (descrito recentemente em Floyd et al. ((2006) Genetics 173: 373-388) e úteis nos métodos da invenção são as seqüência de ácido nucleico REV ΔHDZip/STARTs.

Outra seqüência de ácido nucleico REV ΔHDZip/STARTs também pode ser usado os métodos da invenção, e pode ser facilmente identificado por uma pessoa habilitada na técnica. Os polipeptídeo REV podem ser identificados pela presença de um ou mais de diversas características bem conhecidas (ver abaixo). Na identificação de um polipeptídeo REV, uma pessoa habilitada na técnica pode facilmente deduzir, usando técnicas de rotina, o correspondente que codifica a seqüência de ácido nucleico REV ΔHDZip/START, e o uso de um comprimento suficiente de nucleotídeos contíguos do mesmo para realizar qualquer um ou mais dos métodos de silenciamento do gene descrito acima.

O termo "polipeptídeo REV" como definido neste refere-se a um polipeptídeo que tem queda na classe III dos polipeptídeos HDZip como delineado por essa et al. ((1994) Em : Puigdomene P, Coruzzi G (ed), Springer, Berlin Heidelberg New York, pp 411–426). O polipeptídeos REV que compreende a partir de um terminal N ao terminal C: (i) um domínio (HD) homeodomínio, para ligação de DNA; (ii) um zíper de leucina, para interação de proteína-proteína; (iii) um domínio START para ligação de lipídeo/esterol, e (iv) uma região terminal C (CTR), de função não definida.

Os polipeptídeos REV pode ser facilmente identificados usando técnicas de rotina bem conhecidos na técnica, tal como pelo alinhamento de seqüência. Os métodos para o alinhamento das seqüências para comparação são bem conhecidos na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. GAP usa o algoritmo de Needleman

e Wunsch ((1970) J Mol Biol 48: 443-453) para observar o alinhamento de duas seqüências completas que maximizam o número de pontos e minimiza o número de fendas. O algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-10) calcula o percentual da identidade da seqüência e realiza uma análise estatística de similaridade entre as duas seqüências. O software para a realização da análise BLAST é publicamente disponível através do National Centre for Biotechnology Information. Os polipeptídeos REV compreendem um homeodomínio, um zíper de leucina, um domínio START e um CTR pode ser facilmente identificado usando, por exemplo, o algoritmo de alinhamento da seqüência múltipla ClustalW (versão 1,83), com os parâmetros de alinhamento em pares padrão, e um método de contagem em porcentagem. A edição do manual menor pode ser realizado para otimizar o alinhamento, como seria apresentado a uma pessoa habilitada na técnica. Uma árvore filogenética, pelo qual é avaliada da filogenia (ou linhagem comum), pode ser construindo usando o algoritmo de construção de árvore Neighbour-Joining (at EBI), para ajudar a visualizar o agrupamento que diz respeito aos genes e para identificar ortólogos e parálogos. Os polipeptídeos REV como definido neste refere-se a qualquer polipeptídeo que, quando usado na construção de uma árvore filogenética de polipeptídeo HDZip de classe III, tal como uma descrita nas figuras 15 e 16, falha no clado REV (compreende REV, PHB e PHV) e não o clado CORONA (compreende ATHb8 e CNA), e mais especificamente, que baixa no ramificado REV (e não o ramificado PHB/PHV). Na identificação de um polipeptídeo REV (queda no ramificado REV), uma pessoa habilitada na técnica pode facilmente deduzir, usando técnicas de rotina, o correspondente que codifica a seqüência de ácido nucleico REV ΔHDZip/START e usa um comprimento suficiente de nucleotídeos contíguos do mesmo para realizar qualquer um ou mais dos métodos de silenciamento do gene descrito acima.

Os ortólogos e parálogos podem ser facilmente observado pela

realização de uma busca de blast recíproca denominada. Esta pode ser realizada por um primeiro BLAST que envolve a explosão da seqüência em questão (por exemplo, SEQ ID N°: 198 ou SEQ ID N°: 199) contra qualquer base de dados de seqüência, tal como a base de dados NCBI disponível publicamente. Os BLASTN ou TBLASTX (usando valores de descuido padrão) podem ser usados quando a partida a partir de uma seqüência de nucleotídeo e BLASTP ou TBLASTN (usando valores de descuido padrão) podem ser usados quando a partida a partir de uma seqüência de polipeptídeo. O resultado de BLAST pode ser opcionalmente filtrado. As seqüências de comprimento total dos resultados filtrados ou resultados não filtrados são então suporte estragado (segundo BLAST) contra as seqüências a partir de um organismo de que a seqüência em questão é derivada (onde a seqüência em questão é a SEQ ID N°: 198 ou SEQ ID N°: 199, o segundo BLAST seria portanto contra as seqüências de arroz). Os acertos de alta classificação são aqueles tendo um valor E baixo. Além disso à comparação aos valores E também são contados pela identidade de porcentagem. A identidade de porcentagem refere-se à diversos nucleotídeos idênticos (ou aminoácidos) entre as seqüências de dois ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. Um exemplo detalhando a identificação de ortólogos e parálogos é dado no Exemplo 27. Todos os polipeptídeos REV que compreende um CTR tendo, um aumento a fim da preferência, pelo menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ou 98% da seqüência da identidade ao CTR de um polipeptídeo REV como representado pela SEQ ID N°: 197. Preferivelmente, o CTR de um polipeptídeo REV é como representado pela SEQ ID N°: 197. Mais preferivelmente, SEQ ID N°: 194 codificando uma parte do CTR de um polipeptídeo REV é usado no desempenho dos métodos de acordo com a invenção. Na figura 16, os parálogos de polipeptídeo REV e ortólogos agrupados juntos.

Os termos "domínio" e “motivo” são definidos acima. O termo

"região" como definido neste refere-se à partida da seqüência de aminoácido na extremidade do domínio START e a conclusão no códon de interrupção do polipeptídeo REV.

As bases de dados especiais existe para a identificação dos domínios. Os domínios HD e o START em um polipeptídeo REV pode ser identificado usando, por exemplo, SMART, InterPro, Prosite, ou Pfam. O HD que compreende cerca de 60 resíduos, o domínio START cerca de 210 resíduos. Na base de dados InterPro, o HD é indicado por IPR001356, PF00046 na base de dados Pfam e PS50071 na base de dados PROSITE. Na base de dados InterPro, o domínio START é indicado por IPR002913, PF01852 na base de dados Pfam e PS50848 na base de dados PROSITE. Por exemplo na SEQ ID N°: 199, a entrada HD Pfam estende-se dos aminoácidos 27 a 87, e a entrada do domínio START Pfam de 163 a 376. O CTR portanto começa no aminoácido 377 e termina no códon de interrupção (no 840). A predição de zíper de leucina e a identificação heptado pode ser realizada usando software especializado tal como 2ZIP, que combina com um algoritmo de predição de espiral espiralado padrão com uma busca aproximada para a repetição de leucina característica (Bornberg-Bauer et al. (1998) Computational Approaches para identificar Leucine Zippers, Nucl. Acids Res., 26(11): 2740-2746), hospedado em um servidor no Max Planck Institute for Molecular Genetics in Berlin. Por exemplo, o zíper de leucina da SEQ ID N°: 199 que compreende cinco repetições de leucina (*heptads*), e aminoácidos expandidos 91 a 127.

Além disso, um polipeptídeo REV também pode ser identificável por esta capacidade de ligar-se ao DNA e para reagir com outras proteínas. A atividade de ligação de DNA e interações proteína-proteína pode ser facilmente determinado in vitro ou in vivo usando técnicas bem conhecidas na técnica. Exemplos de ensaios in vitro para ligação da atividade de DNA que inclui: análise de retardo de gel usando o domínio de ligação

DNA HD (West et al. (1998) Nucl Acid Res 26(23): 5277-87), ou ensaios híbridos de uma levedura. Um exemplo de um ensaio in vivo para interação de proteína-proteína é a análise híbrida de duas leveduras (Fields and Song (1989) Nature 340:245-6).

Portanto na identificação de um polipeptídeo REV usando uma ou diversas características como descritas acima, uma pessoa habilitada na técnica pode derivar facilmente o correspondente seqüência de ácido nucleico REV ΔHDZip/START e uso de um comprimento suficiente de nucleotídeos contíguos substancialmente do mesmo para realizar qualquer um ou mais dos métodos de silenciamento do gene descrito acima (para a redução substancial de uma expressão de gene REV endógeno).

Preferido para uso nos métodos da invenção é uma seqüência de ácido nucleico como representado pela SEQ ID: 194, que codifica parte do CTR de um polipeptídeo *VER Oryza sativa*, ou SEQ ID N°: 196, que codifica o CTR do mesmo polipeptídeo *VER Oryza sativa*. A seqüência de ácido nucleico REV ΔHDZip/STARTs são compreendidos na seqüência de ácidos nucleicos que codificam ortólogos ou parálogos de polipeptídeos REV. Um exemplo de um parálogo de polipeptídeo REV para a SEQ ID N°: 199 é representado pela SEQ ID N°: 201, ouyza sativa Orysa_HOX10 (codificado pela SEQ ID N°: 200, número de acessão NCBI AY425991.1). Exemplos de ortólogos de polipeptídeo REV são representados pela SEQ ID N°: 203 Arabidopsis thaliana Arath_REV (codificado pela SEQ ID N°: 202, número de acessão NCBI AF188994), SEQ ID N°: 205 Zea mays Zeama_HDIII RLD1 (folha enrolada; codificado pela SEQ ID N°: 204, número de acessão NCBI AY501430.1), SEQ ID N°: 207 Populus trichocarpa Poptr_HDIII (codificado pela SEQ ID N°: 206, número de acessão NCBI AY919617), SEQ ID N°: 209 Medicago trunculata Medtr_HDIII (codificado pela SEQ ID N°: 208, número de acessão NCBI AC138171.17), SEQ ID N°: 211 Saccharum officinarum Sacof_HDIII partial (codificado pela SEQ ID N°:

210, próximo do número de acessão NCBIs CA125167.1 CA217027.1 CA241276.1 CA124509.1), SEQ ID N°: 213 Triticum aestivum Triae_HDIII (parcial; codificado pela SEQ ID N°: 212, próximo do número de acessão NCBIs CD905903 BM135681.1 BQ578798.1 CJ565259.1), SEQ ID N°: 215 Hordeum vulgare Horvu_HDIII (parcial, codificado pela SEQ ID N°: 214, próximo do número de acessão NCBIs BU996988.1 BJ452342.1 BJ459891.1), e SEQ ID N°: 217 Phyllostachys praecox Phypr_HDIII (parcial; codificado pela SEQ ID N°: 216, número de acessão NCBI DQ013803). No exemplo 27 (e tabela Z) da presente aplicação é descrito um método para identificar a seqüência de ácidos nucleicos úteis no desempenho dos métodos da invenção.

A fonte da seqüência de ácido nucleico REV ΔHDZip/START útil no desempenho dos métodos da invenção pode ser qualquer fonte de planta ou fonte artificial. Para ótimo desempenho, as técnicas de silenciamento do gene usadas para a redução de uma expressão de gene REV endógeno requere-se o uso das seqüências de ácido nucleico REV ΔHDZip/START a partir de plantas monocotiledôneas para transformação de plantas monocotiledôneas, e uso das seqüências de ácido nucleico REV ΔHDZip/START de plantas dicotiletiledônias para transformação de plantas dicotiletiledônias. Preferivelmente, as seqüências de ácido nucleico REV ΔHDZip/START de plantas da família Poaceae são transformadas em plantas da família Poaceae. Ainda preferivelmente, a seqüência de ácido nucleico REV ΔHDZip/START a partir do arroz é transformado em uma planta de arroz. A seqüência de ácido nucleico REV ΔHDZip/START não necessita ser introduzida na mesma variedade de planta. Mais preferivelmente, o REV ΔHDZip/START a partir do arroz é um comprimento suficiente de nucleotídeos contíguos substancialmente da SEQ ID N°: 194 ou SEQ ID N°: 196, ou um comprimento suficiente de nucleotídeos contíguos substancialmente de uma seqüência de ácido nucleico REV ΔHDZip/START

a partir da seqüência de ácido nucleico que codifica ortólogos ou parálogos de polipeptídeo REV. Como mencionado acima, uma pessoa habilitada na técnica seria bem cauteloso do qual constituiria um comprimento suficiente de nucleotídeos contíguos substancialmente para realizar qualquer dos métodos de silenciamento do gene definido neste acima, este pode ser tão pequeno quanto 20 ou nucleotídeos contíguos menores substancialmente em alguns casos.

A invenção também fornece construto genético e vetores para facilitar a introdução e/ou expressão das seqüências de nucleotídeos úteis nos métodos de acordo com a invenção.

Portanto, aqui é fornecido um construto genético pela redução da expressão na planta de um gene REV endógeno compreende um ou mais seqüências de controle, a seqüência de ácido nucleico REV ΔHDZip/START, e opcionalmente uma seqüência de terminação de transcrição. Preferivelmente, a seqüência de controle é um promotor constitutivo.

Um construto preferido para redução da expressão na planta de um gene REV endógeno é um que compreende uma repetição invertida de uma seqüência de ácido nucleico REV ΔHDZip/START, preferivelmente capaz da formação e a estrutura grampo de cabelo, pelo qual a repetição invertida está sob o controle de um promotor constitutivo.

Os construtos úteis nos métodos de acordo com a presente invenção pode ser criados usando tecnologia de DNA recombinante bem conhecidas pelas pessoas habilitadas na técnica. Os construtos genéticos pode ser inseridos em vetores, que podem ser comercialmente disponíveis, adequado para a transformação em plantas e adequados para a expressão do gene de interesse nas células transformadas. A invenção portanto fornece o uso de um construto genético como definido neste acima nos métodos da invenção.

A seqüência de interesse é operavelmente ligada a um ou mais

seqüências de controle (pelo menos a um promotor) capaz da expressão aumentada na planta. Vantajosamente, qualquer tipo de promotor pode ser usado para conduzir a expressão da seqüência de ácido nucleico.

Em uma forma de realização, a seqüência de ácido nucleico REV ΔHDZip/START é operavelmente ligada a um promotor constitutivo. Preferivelmente o promotor é um promotor ubíquo e é expressado predominantemente completa da planta. Preferivelmente, o promotor constitutivo é substancialmente como representado pela SEQ ID N°: 218 ou SEQ ID N°: 58, ainda preferivelmente o promotor capaz da expressão preferencialmente da seqüência de ácido nucleico completamente a planta é um promotor GOS2, mais preferivelmente o promotor GOS2 é a partir do arroz (SEQ ID N°: 218 ou SEQ ID N°: 58). Será claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico REV ΔHDZip/START como representado pela SEQ ID N°: 194, ou é a aplicabilidade da invenção restrita para expressão de uma seqüência de ácido nucleico REV ΔHDZip/START quando conduzida por um promotor GOS2. Um promotor constitutivo alternativo que é útil nos métodos da presente invenção é o promotor de proteína do grupo de alta mobilidade (SEQ ID N°: 293, PRO0170). Exemplos de outros promotores constitutivos que também podem ser usados para conduzir a expressão de uma seqüência de ácido nucleico REV ΔHDZip/START são mostrados na seção de definições.

Opcionalmente, um ou mais seqüências terminadoras também podem ser usadas no construto introduzido na planta.

Os construtos genéticos da invenção ainda podem incluir uma origem da seqüência de replicação que é requerida para manutenção e/ou replicação em um tipo celular específico. Um exemplo é então um construto genético que é requerido a ser mantido em uma célula bacteriana como um elemento genético episomal (por exemplo moléculas de plasmídio ou cosmídio). As origens preferidas da replicação incluem, mas não limitam a, o

f1-ori e colE1.

Para a detecção de transferir bem sucedida a seqüência de ácido nucleico como usado os métodos da invenção e/ou seleção de plantas transgênicas compreende estes ácidos nucleicos, este é vantajoso para o uso em genes marcadores (ou genes repórter). Portanto, o construto genético pode opcionalmente compreender um gene marcador selecionável. Marcadores selecionáveis são descritos em mais detalhes nas seção de "definições" neste. Os genes marcadores podem ser removidos ou taxados a partir da célula transgênica uma vez que estes não são necessariamente longos. Técnicas para a remoção do marcador são conhecidas na técnica, técnicas úteis são descritas acima na seção de definições.

A presente invenção também abrange plantas incluindo partes das plantas e células das plantas obtidas pelos métodos de acordo com a presente invenção tendo reduzido a expressão de um gene REV endógeno usando uma seqüência de ácido nucleico REV ΔHDZip/START e que tem produção aumentada que diz respeito as plantas de controle.

A invenção também fornece um método para a produção de plantas transgênicas tendo produção aumentada que diz respeito as plantas de controle, pelo qual as plantas transgênicas tem reduzido a expressão de um gene REV endógeno usando uma seqüência de ácido nucleico REV ΔHDZip/START.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo produção aumentada que diz respeito as plantas de controle, que compreende métodos:

(i) introdução de uma expressão na planta, parte da planta ou célula da planta um construto genético que compreende um ou mais seqüências de controle para redução da expressão na planta de um gene REV endógeno usando uma seqüência de ácido nucleico REV ΔHDZip/START; e

(ii) cultivar a planta, parte da planta ou célula da planta sob

condições que promovem o crescimento da planta e crescimento.

Preferivelmente, o construto introduzido na planta é um que compreende uma repetição invertida (em parte ou completa) de uma seqüência de ácido nucleico REV ΔHDZip/START, preferivelmente capaz de formação da estrutura grampo de cabelo, pelo qual a repetição invertida está sob o controle de um promotor constitutivo.

O ácido nucleico pode ser introduzido diretamente na célula da planta ou na própria planta (incluindo introdução no tecido, órgão ou qualquer outra parte de uma planta). De acordo com uma característica preferida da presente invenção, o construto é introduzido na planta pela transformação.

As células das plantas geneticamente modificadas podem ser regeneradas por intermédio de todos os métodos com que o técnico habilitado seja familiar. Métodos adequados podem ser observados nas publicações mencionadas acima por S.D. Kung e R. Wu, Potrykus ou Höfgen e Willmitzer.

Geralmente após a transformação, células das plantas ou disposições celulares são selecionadas pela presença de um ou mais marcadores que são codificados por genes expressíveis em plantas co-transferido com o gene de interesse, seguindo pelo qual o material transformado é regenerado na planta total. Para plantas transformadas seletas, o material da planta obtido na transformação é, como uma regra, submetida às condições seletivas de modo que as plantas transformadas podem ser distinguidas das plantas não transformadas. Por exemplo, as sementes obtidas na maneira descrita acima pode ser plantada e, após um período de crescimento inicial, submetida a uma seleção adequada por pulverização. Uma possibilidade adicional consiste no crescimento das sementes, se apropriado após a esterilização, em placas ágar usando um agente de seleção adequado de modo que apenas as sementes transformadas podem desenvolver nas plantas. Alternativamente, as plantas transformadas são avaliadas na

presença de um marcador selecionável tal como o descrito acima.

Seguindo a regeneração e transferência do DNA, putativamente as plantas transformadas podem ser avaliadas, por exemplo usando análise de Southern, na presença do gene de interesse, organização de número de cópias e/ou genômicas. Alternativamente ou adicionalmente, os níveis de expressão do DNA introduzido novamente pode ser monitorado usando análise Northern e/ou Western, ou PCR quantitativo, todas as técnicas sendo bem conhecidas às pessoas comuns habilitadas na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de meios, tal como pela propagação clonal ou técnicas de geração clássica. Por exemplo, uma primeira geração (ou T1) a planta transformada pode ser tornada independente para dar transformantes de segunda geração de homozigotos (ou T2), e as plantas T2 propagadas adicionais através das técnicas de geração clássica.

Os organismos transformados gerados podem tornar-se uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e células não transformadas; transformadores de clone (por exemplo, todas as células transformadas que contém a expressão cassete); enxertados de tecidos transformados e não transformados (por exemplo, em plantas, um origem transformada enxertada por um muda não transformada).

As características desenvolvidas mencionadas acima podem ser vantajosamente modificado em qualquer planta. Os métodos da invenção são vantajosamente aplicáveis a qualquer planta. As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular monocotiledôneas e plantas dicotiletiledônias incluindo legumes de alimentação e de forragem, plantas ornamentais, safra de alimentação, árvores ou arbustos. De acordo com uma forma de realização preferida da presente invenção, a planta é uma planta de safra. Exemplos de plantas de safra incluem soja, girassol, canola,

alfafa, colza, algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta monocotiledônia. Exemplos de plantas monocotiledôneas incluem cana de açúcar. Mais preferivelmente a planta é um cereal. Exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, *triticale*, sorgo e aveias.

Outras plantas vantajosas são selecionadas a partir do grupo que consiste de Asteraceae tal como os gêneros Heliandeste modo, Tagetes por exemplo a espécie Helianthus annus [girassol], Tagetes lucida, Tagetes erecta ou Tagetes tenuifolia [cravo de defunto], Brassicaceae tal como o gênero Brassica, Arabadopsis por exemplo a espécie Brassica napus, Brassica rapa ssp. [canola, semente de óleo de colza, nabo] ou Arabidopsis thaliana; Fabaceae tal como o gênero Glycine por exemplo a espécie Glycine max, Soja hispida ou Soja max [soja]; Linaceae tal como o gênero Linum por exemplo a espécie Linum usitatissimum, [linho, linhaça]; Poaceae tal como o gênero Hordeum, Secale, Avena, Sorgo, ouyza, Zea, Triticum por exemplo a espécie Hordeum vulgare [cevada]; Secale cereale [centeio], Avena sativa, Avena fatua, Avena byzantina, Avena fatua var. sativa, Avena hybrida [aveia], Sorgo bicolor [Sorgo, painço], ouyza sativa, ouyza latifolia [arroz], Zea mays [milho] Triticum aestivum, Triticum durum, Triticum turgidum, Triticum hybernum, Triticum macha, Triticum sativum ou Triticum vulgare [trigo, trigo do pão, trigo comum]; Solanaceae tal como o gênero Solanum, Lycopersicon por exemplo a espécie Solanum tuberosum [batata], Lycopersicon esculentum, Lycopersicon lycopersicum, Lycopersicon pyriforme, Solanum integrifolium ou Solanum lycopersicum [tomate].

A presente invenção claramente estende-se a qualquer célula da planta ou plantas produzidas por quaisquer dos métodos descrito neste, e por todas as partes das plantas e *propagules* deste. A presente invenção estende-se ainda para abranger o resultado de uma célula transfectada ou transformada primária, tecido, órgão ou planta total que tem sido produzido

por qualquer dos métodos mencionados acima, a necessidade apenas sendo o resultado exibe a mesma característica genotípica e/ou fenotípica como aquela produzida pela origem nos métodos de acordo com a invenção.

A invenção também estende-se a partes colhíveis de uma planta tal como, mas não limitado a sementes, folhas, frutos, flores, hastes, raízes, rizomas, tubérculos e bulbos. A invenção além disso diz respeito aos produtos derivados, preferivelmente derivado diretamente, da parte colhível de uma tal planta, tal como grânulos ou pós secos, óleo, gordura e ácidos graxos, amido ou proteínas.

A presente invenção também abrange o uso de uma seqüência de ácido nucleico REV ΔHDZip/START, para redução da expressão REV do gene endógeno na planta, parte da planta, ou célula da planta para aumentar a produção da planta como definido neste acima.

CLE

Tem sido agora observado a modulação da expressão na planta de um ácido nucleico codificando um polipeptídeo semelhante a CLE dando plantas que tem produção intensificada que diz respeito a traços que diz respeito as plantas de controle.

Portanto, a invenção fornece um método para intensificar a produção que diz respeito a traços em plantas relativas para plantas de controle, compreende a modulação da expressão na planta de um ácido nucleico codificando um polipeptídeo semelhante a CLE, ou uma parte deste.

A "referência", "planta de referência", "controle", "planta de controle", "tipo selvagem" ou "planta do tipo selvagem" é em particular uma célula, um tecido, um órgão, uma planta, ou uma parte deste, que não foi produzido de acordo com o método da invenção. Conseqüentemente, os termos "tipo selvagem", "controle" ou "referência" são permutáveis e podem ser uma célula ou uma parte da planta tal como uma organela ou tecido, ou uma planta, que não foi modificada ou tratada de acordo com o método

descrito neste de acordo com a invenção. Conseqüentemente, a célula ou uma parte da planta tal como uma organela ou uma planta usada como tipo selvagem, controle ou referência corresponde a uma célula, planta ou parte destes tão maior quanto possível e é qualquer outra propriedade mas no resultado do processo da invenção como idêntico à origem do objeto da invenção como possível. Deste modo, o tipo selvagem, controle ou referência é tratada identicamente ou tão idêntica quanto possível, declarando que apenas as condições ou propriedades terão força diferentes que não influencia nba qualidade da propriedades testada. Que os meios em outras palavras que o tipo selvagem indica (1) uma planta, que carrega a forma não alterada ou não modelada de um gene ou alelo ou (2) o material de partida/planta de que uma planta produzida pelo processo ou método da invenção são derivados.

Preferivelmente, qualquer comparação entre as plantas de tipo selvagem plants e as plantas produzidas pelo método da invenção é realizada sob condições análogas. O termo "condições análogas" significa que todas as condições tal como, por exemplo, cultura ou condições de crescimento, condições de ensaio (tal como composição de tampão, temperatura, substratos, cepa de patogênio, concentrações e outros) são mantidos idênticos entre os experimentos a serem comparados.

A "referência", "controle", ou "tipo selvagem" é preferivelmente um objeto, por exemplo uma organela, uma célula, um tecido, uma planta, que não foi modulada, modificada ou tratada de acordo com o processo neste descrito da invenção e é qualquer outra propriedade como similar à origem do objeto da invenção como possível. A referência, controle ou tipo selvagem é neste genoma, transcriptoma, proteoma ou metaboloma tão similar quanto possível ao objeto da presente invenção. Preferivelmente, o termo "referência" "controle" ou "tipo selvagem" organela, célula, tecido ou planta, que diz respeito a uma organela, célula, tecido ou planta, que é aproximadamente idêntica geneticamente á organela, célula, tecido ou planta,

da presente invenção ou uma parte deste preferivelmente 95%, mais preferido são 98%, ainda mais preferido são 99,00%, em particular 99,10%, 99,30%, 99,50%, 99,70%, 99,90%, 99,99%, 99, 999% ou mais. Mais preferível a "referência", "controle", ou "tipo selvagem" são preferivelmente um objeto, por exemplo uma organela, uma célula, um tecido, uma planta, que é geneticamente idêntica á planta, organela celular usada de acordo com o método da invenção exceto que as moléculas de ácido nucleico ou o produto do gene codificado por estas sejam mudados, modulados ou modificados de acordo com o método inventivo.

O termo "modulação" significa em relação a expressão ou expressão do gene, um processo em que o nível de expressão é mudado pela dita expressão do gene em comparação ao controle da planta, preferivelmente um nível da expressão é diminuído. A expressão não modulada original pode ser de qualquer tipo da expressão de um RNA estrutural (rRNA, tRNA) ou mRNA que tem tradução subseqüente. O termo "modulação da atividade" deve significar qualquer mudança de uma expressão da seqüência de ácido nucleico inventiva ou proteínas codificadas, que levam a produção aumentada e/ou crescimento aumentado da plantas.

Um método preferido para expressão da modulação (preferivelmente, diminuição) de um ácido nucleico codificando de um polipeptídeo semelhante a CLE é por introdução de uma expressão na planta um construto genético em que o ácido nucleico que codifica um tal polipeptídeo semelhante a CLE é clonada como uma repetição invertida (em parte ou completamente), separada por um espaçador (DNA não codificado).

Qualquer referência depois disso às "proteínas úteis nos métodos da invenção" é retirado por significar um polipeptídeo POI como definido neste. Qualquer referência depois disso ao "ácido nucleico útil nos métodos da invenção" é retirado por significar um ácido nucleico capaz de codificar um tal polipeptídeo POI. O ácido nucleico a ser introduzido na

planta (e portanto úteis no desempenho os métodos da invenção) é qualquer ácido nucleico que codifica o tipo de proteína que será agora descrita, depois também chamada "ácido nucleico POI" ou "gene POI".

O polipeptídeo semelhante a CLE que codifica genes são conhecidos na técnica (ver por exemplo Cock e McCormick (Plant Physiol. 126, 939-942, 2001) e úteis nos métodos da invenção são os ácidos nucleicos codificando um polipeptídeo semelhante a CLE, ou uma parte deste.

O termo "polipeptídeo semelhante a CLE" como definido neste refere-se a um polipeptídeo homólogo a SEQ ID N°: 233. Os polipeptídeos semelhantes a CLE que compreendem uma seqüência de sinal do terminal N e um motivo conservado (Cock e McCormick, 2001: Figura 21), também conhecido como domínio CLE (Strabala et al., 2006), que é localizado perto do terminal C do polipeptídeo. os polipeptídeos não processados são geralmente entre 60 a 140 aminoácidos longos e tem um alto ponto isoelétrico, preferivelmente acima de pI 7,0, mais preferivelmente acima de pI 8,0, mais preferivelmente acima de pI 9, (por exemplo, a proteína representada pela SEQ ID N°: 233 tem um pI de 10,46). Após a clivagem da seqüência de sinal, o polipeptídeo semelhante a CLE pode ser ainda processado pela clivagem proteolítica na parte do terminal C, preferivelmente no resíduo Arg conservado na parte do terminal N do domínio CLE, deste modo gerando um peptídeo curto ativo biologicamente abrangendo mais do domínio CLE (Ni e Clark, Plant Physiol. 140, 726-733, 2006).

Os termos "domínio" e "motivo" são definidos na seção de definições neste. A base de dados especialista existe para a identificação dos domínios. O domínio CLE em um polipeptídeo semelhante a CLE pode ser identificado usando, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucl. Acids Res 30, 242-244), InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318), Prosite (Bucher e Bairoch (1994). Uma regra de perfil generalizado pelos

motivos das seqüências biomoleculares e esta função na interpretação da seqüência automática. (In) ISMB-94; Proceedings 2° International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp. 53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004)) or Pfam (Bateman et al., B] Nucl. Acids Research 30(1): 276-280 (2002)). Uma série de ferramentas para análise *in silico* das seqüências de proteína é disponível no servidor proteômico de ExPASY (hospedado por Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: os proteomics server for in-depth protein knowledge and analysis, Nucl. Acids Res. 31:3784-3788(2003)). Entretanto, o domínio CLE também pode ser facilmente identificado no alinhamento da seqüência de um polipeptídeo semelhante a CLE putativo com polipeptídeos semelhantes a CLE conhecidos na técnica (tal como aquele divulgado em Cock e McCormick (2001) ou Strabala et al (2006).

O domínio CLE preferivelmente tem as seqüências de consenso seguintes (SEQ ID N°: 237):
(S/E/R/M/P/L)(K/E/D/R/S)R(K/I/L/R/F/Q/V/M)(V/I/S/L)(P/L/R)(R/T/Q/C/N/G/K/S)(N/G)(S/P)(D/N/Y)P(I/L/Q/R/Y/H)(H/L/I)(H/N).

Mais preferivelmente, o domínio CLE tem as seqüências seguintes (SEQ ID N°: 238):
(S/P)(R/E/K)R(M/L/I)(V/S/I)P(Q/G/C/T/S)GP(N/D)P(L/Q/H)H(H/N).

Mais preferivelmente, o domínio CLE tem as seqüências seguintes:
SRRMVPQGPNPLHN.

O domínio CLE compreende um número de aminoácidos altamente conservados, incluindo o resíduo Arg necessário para o procedimento proteolítico (Arg73 na SEQ ID N°: 233, ver Figuras 21 e 22), e dois ou três resíduos Pro.

O preferido para uso nos métodos da invenção é um ácido

nucleico codificando pelo menos parte do polipeptídeo semelhante a CLE, como representado pela SEQ ID: 233, ou um ácido nucleico codificando pelo menos parte de um homólogo da SEQ ID N°: 233. Exemplos de polipeptídeos semelhantes a CLE incluem SEQ ID N°: 233 e também abrange homólogos (incluindo ortólogos e parálogos) da SEQ ID N°: 233. A invenção é ilustrada pela transformação de plantas de arroz com a seqüência *Saccharum officinarum* representada pela SEQ ID N°: 232, que codifica o polipeptídeo da SEQ ID N°: 233. SEQ ID N°: 240 de *Populus*, SEQ ID N°: 242 a partir do arroz, SEQ ID N°: 246 de *Arabidopsis* e SEQ ID N°: 248 de *Brassica napus* representa ortólogos da SEQ ID N°: 233. SEQ ID N°: 246 e SEQ ID N°: 250 são parálogos de cada um outro.

Os ortólogos e parálogos podem ser facilmente observado pela realização de uma busca de blast recíproca denominada. Este pode ser realizado por um primeiro BLAST que envolve explosão da seqüência em questão (por exemplo, SEQ ID N°: 241 ou SEQ ID N°: 242) contra qualquer base de dados de seqüência, tal como a base de dados NCBI disponível publicamente. BLASTN ou TBLASTX (usando valores de descuido padrão) pode ser usados quando a partida a partir de uma seqüência de nucleotídeo e BLASTP ou TBLASTN (usando valores de descuido padrão) pode ser usados quando a partida da seqüência da proteína. O resultado de BLAST pode ser opcionalmente filtrado. As seqüências de comprimento total dos resultados filtrados ou resultados não filtrados são então suportes estragados (segundo BLAST) contra as seqüências a partir de um organismo de que a seqüência em questão é derivada (onde a seqüência em questão é a SEQ ID N°: 241 ou SEQ ID N°: 242, o segundo BLAST seria portanto contra as seqüências de arroz). Os resultados do primeiro e segundo BLASTs são então comparados. Um parálogo é identificado se um acerto de alta classificação a partir do segundo BLAST é da mesma espécie como da seqüência em questão é derivada; um ortólogo é identificado se um acerto de alta classificação não é

da mesma espécie como da seqüência em questão é derivado. Os ortólogos preferidos são ortólogos da SEQ ID N°: 232 ou SEQ ID N°: 233. Os acertos de alta classificação são aqueles tendo um valor E baixo. O valor E inferior, o mais significante da contagem (ou em outras palavras o inferior a chance que o acerto como observado pela escolha). A computação do valor E é bem conhecidos na técnica. Além disso a comparação aos valores E também são contados pela identidade da porcentagem. Identidade de porcentagem refere-se à diversos nucleotídeos idênticos (ou aminoácidos) entre as seqüências de dois ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. Preferivelmente a contagem é maior do que 50, mais preferivelmente maior do que 100; e preferivelmente o valor E é menos do que e-5, mais preferivelmente menos do que e-6. No caso de grandes famílias, ClustalW pode ser usado, seguido pela geração de uma árvore de ligação próxima, para ajudar a visualizar o agrupamento que diz respeito um genes e para identificar ortólogos e parálogos.

Os homólogos (ou proteínas de homólogos, abrangendo ortólogos e parálogos) pode ser facilmente identificados usando técnicas de rotina bem conhecidas na técnica, tal como pelo alinhamento da seqüência. Métodos para o alinhamento das seqüências para comparação são bem conhecidas na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. GAP usa o algoritmo de Needleman e Wunsch ((1970) J Mol Biol 48: 443-453) para observar o alinhamento de duas seqüências completas que maximizam o número de pontos e minimiza o número de fendas. O algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-410) calcula o percentual da identidade da seqüência e realiza uma análise estatística de similaridade entre as duas seqüências. O software para a realização da análise BLAST é publicamente disponível através do National Centre for Biotechnology Information. Homólogos pode ser facilmente identificados usando, por exemplo, o algoritmo de alinhamento da seqüência

múltipla ClustalW (versão 1.83), com os parâmetros de alinhamento em pares padrão, e um método de contagem em porcentagem. As percentagens globais de similaridade e identidade também pode ser determinado usando um dos métodos disponíveis na embalagem de software MatGAT (Campanella et al., BMC Bioinformatics. 4, 29, 2003). A edição do manual menor pode ser realizado para otimizar o alinhamento entre motivos conservados, como seria apresentado a uma pessoa habilitada na técnica. Além disso, em vez do uso das seqüências de comprimento total para uma identificação de homólogos, domínios específicos (tal como o domínio CLE) pode ser usado bem como. Os valores da identidade da seqüência, que são indicados abaixo como uma porcentagem foram determinados na seqüência de ácido nucleico o aminoácido inteiro usando Os programas mencionados acima usando os parâmetros padrões.

Preferivelmente, os polipeptídeos semelhantes a CLE úteis nos métodos da presente invenção tem, um aumento a fim da preferência, pelo menos 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% da seqüência da identidade aos polipeptídeo da SEQ ID N°: 233. Alternativamente, a identidade de seqüência entre os homólogos podem ser determinados usando um domínio específico (tal como o domínio CLE). Um domínio CLE pode ser identificado e delineado usando a base de dados e ferramentas para identificação de proteína listada acima, e/ou métodos para o alinhamento das seqüências para comparação. Em alguns exemplos, os parâmetros padrões podem ser ajustados para modificar a severidade da busca. Por exemplo usando BLAST, o princípio de significância estatística (denominado valor "exceto") para relatar os pontos contra as seqüências de base de dados que podem ser aumentadas por parecer menos pontos estringentes. Nesta maneira, pontos exatos aproximadamente curtos podem ser identidificados.

O termo ácido nucleico semelhante a CLE como usado neste,

refere-se a qualquer ácido nucleico codificando um polipeptídeo semelhante a CLE como definido acima, ou o complemento deste. O ácido nucleico semelhante a CLE não necessita ser ácido nucleico de comprimento total, visto que o desempenho dos métodos da invenção não confiam no uso da seqüência de ácido nucleico de comprimento total. Além disso, os exemplos de ácido nucleicos adequados para o uso no desempenho dos métodos da invenção incluem mas não são limitados a aquele representado por qualquer um da: SEQ ID N°: 232, SEQ ID N°: 239, SEQ ID N°: 241, SEQ ID N°: 243, SEQ ID N°: 245 ou SEQ ID N°: 247. Variantes de ácidos nucleicos também pode ser úteis na prática dos métodos da invenção. Exemplos de tais variantes incluem porções de ácido nucleico, seqüências de hibridização, variantes unidas, variantes alélicas de ocorrência natural ou pela manipulação de DNA.

A referência neste a uma seqüência de ácido nucleico "semelhante a CLE" é retirado por significar um comprimento suficiente de nucleotídeos contíguos substancialmente a partir da seqüência de ácido nucleico semelhante a CLE. A fim de realizar o silenciamento do gene, este pode ser tão pequeno quanto 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10 ou nucleotídeos menores, alternativamente este pode ser tão maior quanto uma seqüência de ácido nucleico semelhante a CLE (incluindo a extremidade 5' e/ou 3' UTR, em parte ou no total). A seqüência de ácido nucleico codificando um polipeptídeo (funcional) não é uma necessidade para os vários métodos debatidos acima para a redução substancial da expressão de um gene semelhante a CLE endógeno.

O termo "porção" ou "parte" de um ácido nucleico semelhante a CLE como usado neste refere-se a uma porção de DNA codificando pelo menos parte de um polipeptídeo semelhante a CLE, ou o complemento deste, mas também pode ser uma parte da extremidade de 5' ou 3' de região não traduzida (UTR) de um polipeptídeo CLE que codifica cDNA, ou o complemento deste, ou pode ser a extremidade 5' ou 3' UTR inteira, ou este

complemento. O termo cDNA como usado neste é entendido para abranger não apenas as seqüências codificadoras, mas também as seqüências não codificadoras que correspondem à extremidade 5' e 3' UTRs do mRNA.

Os termos "fragmento", "fragmento de uma seqüência" ou "parte de uma seqüência" "porção" ou "porção deste" significa uma seqüência truncada da seqüência original referido. A seqüência truncada (ácido nucleico ou seqüência da proteína) pode variar amplamente no comprimento; o tamanho mínimo sendo uma seqüência do tamanho suficiente para fornecer uma seqüência com pelo menos função e/ou atividade comparável da seqüência original referida ou hibridização com a molécula de ácido nucleico da invenção ou usada nos processos da invenção sob condições estringentes, enquanto o tamanho máximo não é crítico. Em algumas aplicações, o tamanho máximo usualmente não é substancialmente maior do que o requerido para fornecer a atividade e/ou função desejada da seqüência original. Um meio da função comparável pelo menos 40%, 45% ou 50%, preferivelmente pelo menos 60%, 70%, 80% or 90% ou mais da seqüência original.

Uma porção pode ser preparada, por exemplo, pela fabricação de um ou mais anulações ao ácido nucleico codificando um polipeptídeo semelhante a CLE como definido neste acima. As porções pode ser usadas na forma isolada ou estas podem ser fundidas a outras seqüências codificadoras (ou não codificadoras). A porção é tipicamente pelo menos 100 nucleotídeos no comprimento, preferivelmente pelo menos 200, 250 nucleotídeos no comprimento, mais preferivelmente pelo menos 300, 350 nucleotídeos no comprimento e mais preferivelmente pelo menos 400 ou 450 nucleotídeos no comprimento. Preferivelmente, a porção é uma porção de um ácido nucleico como representado por qualquer um da SEQ ID N°: 232, SEQ ID N°: 239, SEQ ID N°: 241, SEQ ID N°: 243, SEQ ID N°: 245 ou SEQ ID N°: 247. Mais preferivelmente a porção é uma porção de um ácido nucleico como

representado pela SEQ ID N°: 232. Uma porção preferida de um ácido nucleico semelhante a CLE para o uso nos métodos da presente invenção é uma porção tendo alta homologia à seqüência transcrita do gene semelhante a CLE alvo endógeno ou ao complemento deste, enquanto tendo a homologia baixa ou nenhuma homologia as seqüências transcritas (ou o complemento seqüências deste) dos genes semelhantes a CLE não alvo endógeno.

Uma outra variante de ácido nucleico útil nos métodos da invenção, é um ácido nucleico capaz da hibridização sob condições de estringência reduzidas, preferivelmente sob condições estringentes, com um ácido nucleico codificando um polipeptídeo semelhante a CLE como definido neste acima, ou com uma porção como definido neste acima.

As seqüências de hibridização úteis nos métodos da invenção, codificadas pelo menos parte de um polipeptídeo semelhante a CLE como definido acima, ou são capazes da hibridização à extremidade 5' ou 3' UTR de um polipeptídeo semelhante a CLE que codifica mRNA ou cDNA. As seqüências de hibridização são tipicamente pelo menos 100 nucleotídeos no comprimento, preferivelmente pelo menos 200 nucleotídeos no comprimento, mais preferivelmente pelo menos 300 nucleotídeos no comprimento e mais preferivelmente pelo menos 400 nucleotídeos no comprimento. Preferivelmente, as seqüências de hibridização é que um é capaz da hibridização a qualquer do ácido nucleico representado pela (ou por sondas derivadas da) SEQ ID N°: 232, SEQ ID N°: 239, SEQ ID N°: 241, SEQ ID N°: 243, SEQ ID N°: 245 ou SEQ ID N°: 247, ou por uma porção de qualquer das seqüências mencionadas acima, uma porção sendo como definido acima. Mais preferivelmente as seqüências de hibridização é capaz da hibridização por SEQ ID N°: 232, ou por porções (ou sondas) deste. Métodos para sondas do desenho são bem conhecidos na técnica. As sondas são geralmente menos do que 500, 400, 300, 200 bp no comprimento, preferivelmente menos do que 100 bp no comprimento. Comumente, o comprimento da fonte pela

hibridização de DNA-DNA tal como Southern blotting, varia entre 100 e 500 bp, considerando a região de hibridização em sondas por hibridizações por dNA-DNA tal como na amplificação PCR geralmente são mais curtos do que 50 mais longo do que 10 nucleotídeos, preferivelmente estes são 15, 20, 25, 30, 35, 40, 45 ou 50 bp no comprimento.

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante unida codificando um polipeptídeo semelhante a CLE como definido neste acima. Variantes unidas preferidas são variantes unidas de um ácido nucleico que codifica os polipeptídeo semelhante a CLE representada pela SEQ ID N°: 233, ou variantes unidas que codificam ortólogos ou parálogos da SEQ ID N°: 233. Ainda preferidos são as variantes unidas do ácido nucleico representado por qualquer uma das SEQ ID N°: 232, SEQ ID N°: 239, SEQ ID N°: 241, SEQ ID N°: 243, SEQ ID N°: 245 ou SEQ ID N°: 247. Mais preferido é uma variante unida de um ácido nucleico como representado pela SEQ ID N°: 232.

Uma outra variante de ácido nucleico útil no desempenho dos métodos da invenção é uma variante alélica de um ácido nucleico codificando uma proteína semelhante a CLE como definido neste acima. Variantes alélicas existe na natureza, e abrange os métodos da presente invenção e o uso destes alelos naturais. A variante alélica pode ser uma variante alélica de um ácido nucleico codificando um polipeptídeo semelhante a CLE representada pela SEQ ID N°: 233, ou uma variante alélica de um ácido nucleico que codifica ortólogos ou parólogos de qualquer das SEQ ID N° mencionadas acima. Ainda preferidos são variantes alélicas do ácido nucleico representado por qualquer uma das SEQ ID N°: 232, SEQ ID N°: 239, SEQ ID N°: 241, SEQ ID N°: 243, SEQ ID N°: 245 ou SEQ ID N°: 247. Mais preferido é uma variante alélica de um ácido nucleico como representado pela SEQ ID N°: 232.

De acordo com a presente invenção, aqui é fornecido um

método para intensificar a produção que diz respeito a traços em plantas, que compreende a introdução de uma expressão uma porção da planta, seqüências de hibridização, variantes unidas, ou variantes alélicas de qualquer um dos ácidos nucleicos dados na Tabela HH do Exemplo 33, ou que compreendem a introdução de uma expressão de uma porção da planta, seqüências de hibridização, variantes unidas, ou variantes alélicas de um ácido nucleico codificando um ortólogo, parálogo ou homólogo de qualquer um das seqüências de aminoácido dados na Tabela HH do Exemplo 33.

Uma variante de ácido nucleico adicional útil nos métodos da invenção é uma variante de ácido nucleico obtida pela mistura do gene. A mistura do gene ou evolução direcionada também pode ser usadas às variantes geradas do ácido nucleico que codifica polipeptídeos semelhantes a CLE como definido acima. Além disso, as variantes de ácidos nucleico também podem ser obtidas pela mutagênese direcionada ao local. Diversos métodos são disponíveis para atingir a mutagênese direcionado ao local, a mais comum sendo métodos com base em PCR (Current Protocols in Molecular Biology. Wiley Eds.).

Também úteis nos métodos da invenção são os ácidos nucleicos que codificam homólogos de qualquer um dos aminoácidos representados pela SEQ ID N° 233, ou ortólogos ou parálogos destes; e ácidos nucleicos que codificam derivados dos polipeptídeos representados pela SEQ ID N° 233 ou ortólogos ou parálogos deste.

Além disso, ácidos nucleicos semelhante a CLE úteis nos métodos da presente invenção (pelo menos em sua forma natural) tipicamente, mas não necessariamente, codificam polipeptídeos tendo atividade de sinalização.

Preferivelmente, os ácidos nucleicos semelhante a CLE codificam os polipeptídeos que, quando super-expressados em plantas, causam um fenotipo semelhante a wus. Ainda preferivelmente, um

polipeptídeo semelhante a CLE, quando super-expressado em Arabidopsis, resulta em um fenotipo Aii como definido por Strabala et al. (2006). Mais preferivelmente, um ácido nucleico semelhante a CLE, quando expressado como uma repetição invertida sob controle do promotor representado pela SEQ ID N°: 236 em arroz resulta em semente aumentada na produção, tal como semente total da semente aumentada.

Os ácidos nucleicos que codificam os polipeptídeos semelhantes a CLE podem ser derivados de qualquer fonte natural ou artificial. O ácido nucleico pode ser modificado a partir da forma natural em ambiente de composição e/ou genômico através da manipulação humana deliberada. Preferivelmente os polipeptídeos semelhante a CLE que codificam os ácidos nucleicos são a partir da planta, ainda preferivelmente a partir da planta monocotiledônia, mais preferivelmente a partir da família de Poaceae, mais preferivelmente o ácido nucleico é de *Saccharum officinarum*.

A invenção também fornece construto genéticos e vetores para facilitar a introdução e/ou expressão da seqüência de ácido nucleico útil nos métodos de acordo com a invenção, na planta.

Portanto, aqui é fornecido um construto do gene que compreende:

(i) um ácido nucleico semelhante a CLE como definido neste acima, ou uma porção deste;

(ii) uma ou mais seqüências de controle operavelmente ligados ao ácido nucleico de (i).

Um construto preferido é um compreende uma repetição invertida de um ácido nucleico semelhante a CLE, preferivelmente capaz da formação da estrutura grampo de cabelo, que a repetição invertida está sob o controle de um promotor específico de semente.

Construtos úteis nos métodos de acordo com a presente invenção podem ser construídos usando tecnologia de DNA recombinante

bem conhecidas pelas pessoas habilitadas na técnica. Os construtos do gene podem ser inseridos em vetores, que podem ser comercialmente disponíveis, adequado para a transformação em plantas e adequados para transcrição do gene de interesse nas células transformadas. A invenção portanto fornece o uso de um construto do gene como definido neste acima nos métodos da invenção.

As plantas são transformadas com um vetor que compreende a seqüência de interesse. A pessoa habilitada é bem cauteloso dos elementos genéticos que podem estar presentes no vetor a fim de transformar sucessivamente, selecionado e propagado das células hospedeiras que contém uma seqüência de interesse. A seqüência de interesse é operavelmente ligada a uma ou mais seqüências de controle (pelo menos a um promotor).

Vantajosamente, qualquer tipo de promotor pode ser usado os métodos da presente invenção. Os promotores preferidos são em particular aqueles que trazem a expressão do gene em tecidos e órgãos, em células da semente, tal como células de endosperma e células do embrião em crescimento. Os promotores adequados são as sementes de óleo de colza do promotor de gene napin (US 5.608.152), o promotor Vicia faba USP (Baeumlein et al., Mol Gen Genet, 1991, 225 (3): 459-67), o promotor Arabidopsis oleosin (WO 98/45461), o promotor phaseolus vulgaris phaseolin (US 5.504.200), o promotor Brassica Bce4 (WO 91/13980), e o promotor de feijão arc5, o promotor de cenoura DcG3, ou o promotor de legume B4 (LeB4; Baeumlein et al., 1992, Plant Journal, 2 (2): 233-9), e promotores que trazem cerca da expressão específica da semente em plantas monocotiledôneas tal como milho, cevada, trigo, centeio, arroz e outros. Promotores específicos de semente vantajosa são a sacarose do promotor de proteína de ligação (WO 00/26388), o promotor phaseolin e o promotor napin. Os promotores adequados que são mais considerados são a cevada lpt2 ou promotor do gene lpt1 (WO 95/15389 e WO 95/23230), e os promotores

descritos no WO 99/16890 (promotores a partir do gene hordeína da cevada, o gene de glutelina de arroz, o gene de rizina de arroz, o gene de prolamina de arroz, o gene de gliadina de trigo, o gene de glutelina de trigo, o gene de zeína de milho, o gene de glutelina de aveia, o gene de kasirina de sorgo e o gene de secalina de centeio). Promotores adicionais adequados são Amy32b, Amy 6-6 e Aleuraína [US 5.677.474], Bce4 (semente de óleo de colza) [US 5.530.149], glicinina (soja) [EP 571 741], carboxilase de fosfoenolpiruvato (soja) [JP 06/62870], ADR12-2 (soja) [WO 98/08962], isocitrato liase (semente de óleo de colza) [US 5.689.040] ou α amilase (cevada) [EP 781 849]. Outros promotores que são disponíveis para a expressão de genes em plantas são promotores específico de folhas tal como aquele descrito em DE-A 19644478 ou promotores de luz regulada tal como, por exemplo, o promotor de ervilha petE.

Preferivelmente, o ácido nucleico semelhante a CLE ou variante deste é operavelmente ligado a um promotor específico de semente. Um promotor específico de semente é transcrito ativo predominantemente no tecido da semente, mas não necessariamente exclusivo no tecido da semente (em caso de expressão fraca). O promotor específico de semente pode ser ativo durante a germinação e/ou crescimento da semente. Os promotores específicos da semente são bem conhecidos na técnica. Preferivelmente, o promotor específico da semente é um promotor específico de endosperma. Mais preferivelmente, o promotor é uma prolamina de arroz RP6 ou um promotor equivalente funcionalmente. Mais preferivelmente, a seqüência do promotor é como representada pela SEQ ID N°: 236. Será claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico semelhante a CLE representada pela SEQ ID N°: 232, ou é a aplicabilidade da invenção restrita a transcrição do ácido nucleico semelhante a CLE quando conduzido por um promotor específico de semente. Exemplos de outros promotores específicos de semente (incluindo promotores específicos de

endosperma) são listados acima.

Opcionalmente, uma ou mais seqüências terminadoras podem ser usadas no construto introduzido na planta. Uma seqüência íntron também pode ser adicionada à extremidade da região 5' não traduzida (UTR) ou na seqüência codificadora para aumentar a quantidade da mensagem madura que acumula no citosol. Outras seqüências de controle (além disso promotor, intensificador, silenciador, seqüências de íntron, regiões 3'UTR e/ou 5'UTR) podem ser proteínas e/ou elementos estabilizadores de RNA. Tais seqüências serão conhecidas ou podem ser facilmente obtidas por uma pessoa habilitada na técnica.

Os construtos genéticos da invenção ainda podem incluir uma origem da seqüência de replicação que é requerida para manutenção e/ou replicação em um tipo celular específico. Um exemplo é então um construto genético é requerido a ser mantido em uma célula bacteriana como um elemento genético episssomal (por exemplo moléculas de plasmídio ou cosmídio). Origens preferidas da replicação incluem, mas não limitam a, o f1-ori e colE1.

Para a detecção e/ou seleção para transferir de seqüência de ácido nucleico bem sucedida como descrito no protocolo da seqüência e usado nos processos da invenção, este é vantajoso para o uso de genes marcadores (= genes repórter). Portanto construto genético pode opcionalmente compreender um gene marcador selecionável. Estes genes marcadores capazes da identificação de uma transferência sucessivamente das moléculas de ácido nucleico por intermédio de uma série de princípios diferentes, por exemplo por intermédio da identificação visual com a ajuda de fluorescência ou luminescência ou na faixa de comprimento de onda de luz que é perceptível para o olho humano, por uma resistência aos herbicidas ou antibióticos, por intermédio que estes são conhecidos como marcadores nutritivos (marcadores de auxotrofismo) ou marcadores antinutritivos, por

ensaios de enzimas intermediários ou por intermédio de fito-hormônios. Exemplos de tais marcadores que podem ser mencionados são GFP (= proteína fluorescente verde); o sistema luciferina/luciferase, a galactosidade β com os substratos coloridos, por exemplo X-Gal, a resistência herbicida a, por exemplo, imidazolinona, glifosato, fosfinotricina ou sulfoniluréia, a resistência antibiótica a, por exemplo, bleomicina, higromicina, estreptomicina, kanamicina, tetraciclina, cloranfenicol, ampicilina, gentamicina, geneticina (G418), espectinomicina ou blasticidina, a menção apenas a poucos, marcadores nutritivos tal como a utilização de manose ou xilose, ou marcadores antinutritivos tal como de resistência a 2-desoxiglicose. Marcadores selecionáveis preferidos nas plantas compreendem aqueles, que conferem a resistência a um herbicida tal como glifosato ou glufosinato. Outros marcadores adequados são, por exemplo, marcadores, que genes codificadores envolve nos caminhos biosintéticos de, por exemplo, açúcares ou aminoácidos, tal como galactosidase β, ura3 ou ilv2. Marcadores, que genes codificadores tal como luciferase, gfp ou outros genes de fluorescência, são de outra maneira adequados. Estes marcadores e os marcadores mencionados acima podem ser usados nos mutantes em quem estes genes não são funcionais visto que, por exemplo, estes tem sido anulados pelos métodos convencionais. Esta lista é um número menor de marcadores possíveis. Para a pessoa habilitada é muito familiar estes tais marcadores. Os marcadores diferentes são preferidos, dependendo do organismo e o método de seleção.

Em uma forma de realização preferida da presente invenção, a expressão modulada de uma proteína semelhante a CLE é diminuído a expressão de uma proteína semelhante a CLE, preferivelmente a expressão diminuída de uma proteína semelhante a CLE endógena.

A referência neste a um gene semelhante a CLE "endógeno" não apenas refere-se a um gene semelhante a CLE como observado na planta nesta forma natural (isto é, sem ter qualquer intervenção humana), mas

também refere-se as seqüências de ácido nucleico semelhante a CLE isoladas subseqüentemente introduzido na planta. Por exemplo, uma planta transgênica que contém um transgene semelhante a CLE por encontrar uma redução substancial da expressão transgene e/ou redução substancial da expressão de um gene semelhante a CLE endógeno, de acordo com os métodos da invenção.

Um método preferido para a diminuição da expressão de uma proteína semelhante a CLE é pelo uso de um vetou de expressão em que a seqüência de ácido nucleico semelhante a CLE que codifica o polipeptídeo semelhante a CLE foi clonado como uma repetição invertida (em parte ou completamente), separada por um espaçador (DNA não codificado).

A "redução" ou "diminuição" ou "baixa-regulação" da expressão ou "silenciamento do gene" são usados permutavelmente neste, e são definidos acima. Preferivelmente, a expressão diminuída não é eliminação completa da expressão. Preferivelmente, a redução da expressão do gene semelhante a CLE endógeno usando uma seqüência de ácido nucleico semelhante a CLE leva ao aparecimento de um ou mais traços fenotípicos.

Esta redução da expressão do gene semelhante a CLE endógena pode ser atingida pelo uso qualquer um ou mais de diversos "silenciamento do gene" bem conhecidos métodos, como descrito acima.

Um método preferido para redução da expressão na planta de um gene semelhante a CLE endógeno por intermédio silenciamento mediado por RNA é pelo uso uma repetição invertida de um ácido nucleico semelhante a CLE ou uma parte deste, preferivelmente capaz da formação da estrutura grampo de cabelo. A repetição invertida é clonada em um vetou de expressão que compreende as seqüências de controle. Uma seqüência de ácido nucleico de DNA que codifica (um espaçador, por exemplo um fragmento de região de ligação de matriz (MAR), um íntron, um poliligador, etc) é localizado entre os dois ácidos nucleicos semelhante a CLE invertidos formando uma repetição

invertida. Após a transcrição da repetição invertida, um RNA quimérico como uma estrutura complementar própria é formada (parcial ou completa). Esta estrutura de RNA filamentada dupla é referida como o RNA grampo de cabelo (hpRNA). O hpRNA é processado pela planta em siRNAs que são incorporados em um RISC. O RISC ainda cliva os transcritos de mRNA codificando um polipeptídeo semelhante a CLE, deste modo substancialmente reduzindo o número de transcritos de mRNA a serem traduzidos em um polipeptídeo semelhante a CLE. Ver por exemplo, Grierson et al. (1998) WO 98/53083; Waterhouse et al. (1999) WO 99/53050).

Preferivelmente, a seqüência de ácido nucleico semelhante a CLE de qualquer dada espécie de plantas é introduzida naquela mesma espécie. Por exemplo, a seqüência de ácido nucleico semelhante a CLE a partir do arroz é transformada em uma planta de arroz. Entretanto, não é um requerimento absoluto que a seqüência de ácido nucleico semelhante a CLE será introduzido originando-se da mesma espécie de plantas como as plantas que serão introduzidas, como mostrado na seção de exemplos onde a seqüência de cana de açúcar é introduzida no arroz para obter os efeitos desejados. Aqui é suficiente que a homologia substancial entre o gene alvo semelhante a CLE endógeno e o ácido nucleico semelhante a CLE ser introduzido.

A expressão de um gene semelhante a CLE endógeno também pode ser reduzido pela introdução de uma modificação genética, dentro do local do gene semelhante a CLE ou em outra parte no genoma. O local de um gene como definido neste é retirado por significar uma região genômica, que inclui o gene de interesse e 10 kb acima ou corrente abaixo da região codificadora.

A modificação genética pode ser introduzida, por exemplo, por qualquer um (ou mais) dos seguintes métodos: Identificação de T-DNA, TILLING, mutagênese direcionado ao local, evolução direcionada,

recombinação homóloga. Seguindo a introdução da modificação genética, existe uma seguinte etapa de seleção pela redução da expressão de um gene semelhante a CLE endógeno, pelo qual a redução na expressão dão as plantas que tem produção aumentada comparadas as plantas de controle.

A mutagênese direcionado ao local e mutagênese aleatória pode ser usada às variantes geradas das seqüências de ácidos nucleicos semelhante a CLE pelo qual as variantes codificam os polipeptídeos semelhantes a CLE com atividade reduzida. Diversos métodos são disponíveis para atingir mutagênese direcionado ao local, a mais comum sendo métodos com base em PCR (ver por exemplo Current Protocols in Molecular Biology. Wiley Eds).

A evolução direcionada também pode ser usada às variantes geradas das seqüências de ácido nucleico semelhante a CLE pelo qual as variantes codificam os polipeptídeos semelhantes a CLE com atividade reduzida.

A identificação de T-DNA, TILLING, mutagênese direcionado ao local e evolução direcionada são exemplos de tecnologias que intensificam a geração de novos alelos e variantes das seqüências de ácidos nucleicos semelhante a CLE pelo qual as variantes codificam polipeptídeos semelhantes a CLE com atividade reduzida.

A recombinação homóloga permite a introdução no genoma de um ácido nucleico selecionado na posição selecionada definida. O ácido nucleico a ser alvejado pode ser um alelo que codifica polipeptídeo semelhante a CLE com atividade reduzida, usado por substituir os genes endógenos, e necessários a serem alvejados ao local do gene semelhante a CLE.

Alternativamente, um programa de avaliação pode ser apresentado para identificar as variantes naturais da população de plantas de um gene semelhante a CLE pelo qual as variantes codificam polipeptídeos

semelhantes a CLE com atividade reduzida. Tais variantes naturais também pode ser usadas por exemplo, para realizar recombinação homóloga.

Os métodos da invenção são vantajosamente aplicáveis a qualquer planta. As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular monocotiledôneas e plantas dicotiletiledônias incluindo legumes de alimentação e de forragem, plantas ornamentais, safra de alimentação, árvores ou arbustos. De acordo com uma forma de realização preferida da presente invenção, a planta é uma planta de safra. Exemplos de plantas de safra incluem soja, girassol, canola, alfafa, colza, algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta monocotiledônia. Exemplos de plantas monocotiledôneas incluem cana de açúcar. Mais preferivelmente a planta é um cereal. Exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

Outras plantas vantajosas são selecionadas a partir do grupo que consiste de Asteraceae tal como o gênero Helianthus, Tagetes por exemplo a espécie Helianthus annus [girassol], Tagetes lucida, Tagetes erecta ou Tagetes tenuifolia [Cravo de defunto], Brassicaceae tal como o gênero Brassica, Arabadopsis por exemplo a espécie Brassica napus, Brassica rapa ssp. [canola, semente de óleo de colza, nabo] ou Arabidopsis thaliana. Fabaceae tal como o gênero Glycine por exemplo a espécie Glycine max, Soja hispida ou Soja max [soja]. Linaceae tal como o gênero Linum por exemplo a espécie Linum usitatissimum, [linho, linhaça]; Poaceae tal como o gênero Hordeum, Secale, Avena, Sorgo, ouyza, Zea, Triticum por exemplo a espécie Hordeum vulgare [cevada]; Secale cereale [centeio], Avena sativa, Avena fatua, Avena byzantina, Avena fatua var. sativa, Avena hybrida [aveia], Sorgo bicolor [Sorgo, painço], ouyza sativa, ouyza latifolia [arroz], Zea mays [milho] Triticum aestivum, Triticum durum, Triticum turgidum, Triticum hybernum, Triticum macha, Triticum sativum ou Triticum vulgare

[trigo, trigo do pão, trigo comum]; Solanaceae tal como o gênero Solanum, Lycopersicon por exemplo a espécie Solanum tuberosum [batata], Lycopersicon esculentum, Lycopersicon lycopersicum., Lycopersicon pyriforme, Solanum integrifolium ou Solanum lycopersicum [tomate].

A presente invenção também abrange plantas obtidas pelos métodos de acordo com a presente invenção. A presente invenção portanto fornece plantas, partes das plantas ou células das plantas deste obtido pelo método de acordo com a presente invenção, pelos qual as plantas ou partes ou células deste que compreendem um transgene de ácido nucleico codificando uma proteína semelhante a CLE como definido acima.

A invenção além disso fornece um método para a produção de plantas transgênicas tendo produção alterada que diz respeito ao traço que diz respeito as plantas de controle, compreende a introdução e a expressão na planta de um ácido nucleico semelhante a CLE como definido neste acima e úteis em um método para sub-regulação da expressão como debatido acima.

As plantas hospedeiras para os ácidos nucleicos ou o vetor usado no método de acordo com a invenção, a expressão cassete ou construto ou vetor são, em princípio, todas as plantas vantajosamente, que são capazes da sintetização dos polipeptídeos usados no método inventivo.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo produção alterada que diz respeito a traços, que compreende métodos:

(i) introdução de uma expressão de ácido nucleico semelhante a CLE em um construto para a sub-regulação da expressão do gene semelhante a CLE na célula da planta; e

(ii) cultivo da célula da planta sob condições que promovem o crescimento da planta e crescimento.

O ácido nucleico pode ser introduzido diretamente na célula da planta ou na própria planta (incluindo introdução no tecido, órgão ou qualquer

outra parte de uma planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é preferivelmente introduzido na planta pela transformação. Ainda preferivelmente o construto para a sub-regulação da expressão do gene semelhante a CLE e introduzido na célula da planta ou planta que compreende uma repetição invertida do ácido nucleico semelhante a CLE ou uma parte deste.

A transferência dos genes estranhos no genoma de uma planta é denominado transformação. Em feitos destes métodos descritos para a transformação e regeneração de plantas a partir do tecido das plantas ou células das plantas são utilizados para transformação transitória ou estável. Um método de transformação vantajoso é a transformação em planta. Nesta extremidade, é possível, por exemplo, permitir a agrobactéria atuar em sementes de plantas ou inocular os meristemas das plantas com agrobactéria. Tem fornecido particularmente expediente de acordo com a invenção permitindo uma suspensão de agrobactéria transformada atuar na planta intacta ou pelo menos a flor primordial. A planta é subseqüentemente desenvolvida até as sementes das plantas tratadas serem obtidas (Clough and Bent, Plant J. (1998) 16, 735–743). Para plantas transformadas seletas, o material da planta obtido na transformação é, como uma regra, submetida às condições seletivas de modo que plantas transformadas podem ser distinguidas das plantas não transformadas. Por exemplo, as sementes obtidas na maneira descrita acima pode ser plantada e, após um período de crescimento inicial, submetida a uma seleção adequada por pulverização. Uma possibilidade adicional consiste no crescimento das sementes, se apropriado após a esterilização, em placas de ágar usando um agente de seleção adequado de modo que apenas as sementes transformadas podem desenvolver-se nas plantas. Métodos de transformação vantajosos adicionais, em particular por plantas, são conhecidos para a pessoa habilitada e são descritos neste abaixo

Geralmente após a transformação, células das plantas ou disposições celulares são selecionadas na presença de um ou mais marcadores que são codificados pelos genes expressíveis em plantas co-transferidas com o gene de interesse, seguido pelo material transformado é regenerado na planta total.

Como as agrobactérias transformadas mencionadas com um vetou de expressão de acordo com a invenção também pode ser usado em uma maneira conhecida por, se para a transformação de plantas tal como plantas experimentais semelhantes a Arabidopsis ou plantas de safra, tal como, por exemplo, cereais, milho, aveias, centeio, cevada, trigo, soja, arroz, algodão, beterraba de açúcar, canola, girassol, linho, cânhamo, batata, tabaco, tomate, cenoura, pimentão, semente de óleo de colza, tapioca, cassava, araruta, tagetes, alfafa, alface e as várias árvores, noz, e espécies de videira, em particular plantas de safra que contém óleo tal como soja, amendoim, planta de óleo de mamona, girassol, milho, algodão, linho, semente de óleo de colza, coco, óleo de palma, cártamo (Carthamus tinctorius) ou sementes de cacau, por exemplo pelo banho de folhas escarificadas ou segmentos de folha em uma solução de agrobactéria e crescimento subseqüentemente destes em meio adequado.

As células das plantas geneticamente modificadas podem ser regeneradas por intermédio de todos os métodos com que o técnico habilitado seja familiar. Métodos adequados podem ser observados nas publicações mencionadas acima por S.D. Kung e R. Wu, Potrykus ou Höfgen e Willmitzer. Para plantas transformadas seletas, o material da planta obtida na transformação é, como uma regra, submetida às condições seletivas de modo que plantas transformadas podem ser distinguidas das plantas não transformadas. Por exemplo, as sementes obtidas na maneira descrita acima podem ser plantada e, após um período de crescimento inicial, submetida a uma seleção adequada por pulverização. Uma possibilidade adicional consiste

no crescimento das sementes, se apropriado após a esterilização, em placas de ágar usando um agente de seleção adequado de modo que apenas as sementes transformadas podem desenvolver-se nas plantas. Alternativamente, as plantas transformadas são avaliadas na presença de um marcador selecionável tal como em um descrito acima.

Seguindo a regeneração e transferência do DNA, putativamente as plantas transformadas podem ser avaliadas, por exemplo usando análise de Southern, na presença do gene de interesse, organização de número de cópias e/ou genômicas. Alternativamente ou adicionalmente, baixa-regulação dos níveis de expressão do gene semelhante a CLE alvejado pode ser monitorado usando a análise Northern e/ou Western, ou PCR quantitativo, todas as técnicas sendo bem conhecidas às pessoas comuns habilitadas na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de meios, tal como pela propagação clonal ou técnicas de geração clássica. Por exemplo, uma primeira geração (ou T1) planta transformada pode ser tornada independente para dar transformantes de segunda geração de homozigotos (ou T2), e as plantas T2 ainda propagadas através das técnicas de geração clássica.

Os organismos transformados gerados podem tornar-se uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e células não transformadas; transformadores de clone (por exemplo, todas as células transformadas que contém a expressão cassete); enxertados de tecidos transformados e não transformados (por exemplo, em plantas, uma origem transformada enxertada por um muda não transformada).

A presente invenção claramente estende-se a qualquer célula da planta ou plantas produzidas por quaisquer dos métodos descrito neste, e por todas partes das plantas e *propagules* destes. A presente invenção estende-se ainda para abranger o resultado de uma célula transfectada ou transformada

primária, tecido, órgão ou planta total que tem sido produzida por qualquer dos métodos mencionados acima, a necessidade apenas sendo o resultado exibe a mesma característica genotípica e/ou fenotípica como aquela produzida pela origem nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras que contém um ácido nucleico semelhante a CLE isolado como definido neste acima. As células hospedeiras preferidas de acordo com a invenção são células das plantas. As plantas hospedeiras para o ácido nucleicos ou o vetor usado no método de acordo com a invenção, a expressão cassete ou construto ou vetor são, em princípio, todas as plantas vantajosamente, que são capazes da sintetização dos polipeptídeos usados no método inventivo.

A invenção também estende-se a partes colhíveis de uma planta tal como, mas não limitado a sementes, folhas, frutos, flores, hastes, raízes, rizomas, tubérculos e bulbos. A invenção além disso diz respeito aos produtos derivados, preferivelmente derivado diretamente, da parte colhível de uma tal planta, tal como grânulos ou pós secos, óleo, gordura e ácidos graxos, amido ou proteínas.

Vantajosamente, o desempenho dos métodos de acordo com a presente invenção resulta em plantas que tem produção intensificada que diz respeito a traços, particularmente aumento da produção, mais particularmente semente aumentada na produção e/ou biomassa aumentada, que diz respeito as plantas de controle.

O desempenho dos métodos da invenção dando plantas que tem produção intensificada que diz respeito a traços. Em particular o desempenho dos métodos da invenção dão as plantas que tem produção aumentada, especialmente semente aumentada na produção que diz respeito as plantas de controle. Os termos "produção" e "produção da semente" são descritos em mais detalhes nas seção de "definições" neste.

A referência neste da produção intensificada que diz respeito a

traços é retirado por significar um aumento na biomassa (peso) de uma ou mais partes de uma planta, que pode incluir acima do solo (colhíveis) partes e/ou (colhíveis) partes abaixo trituradas. Em particular, tais partes colhíveis são sementes e/ou biomassa, e o desempenho dos métodos da invenção resultam em plantas tendo sementes aumentadas na produção relativa da produção da semente de plantas de controle.

Tomando o milho como um exemplo, um aumento da produção pode ser manifestada como um ou mais dos seguintes: aumento de diversas plantas estabelecidas por hectare ou acre, um aumento de diversas espigas por plantas, um aumento de diversas séries, diversas sementes por série, peso da semente, peso singular da semente, comprimento da espiga/diâmetro, aumento da taxa de enchimento da semente (que são diversas sementes enchedoras divididas pelo número total de sementes e multiplicadas por 100), entre outros. Tomando o arroz como um exemplo, um aumento da produção pode manifestar em si mesma como um aumento de um ou mais dos seguintes: diversas plantas por hectare ou acre, diversas panículas por plantas, diversas espigas por panícula, número de flores (flósculos) por panícula (que é expressado como uma razão de diversas sementes enchedoras em diversas panículas primárias), aumento da taxa de enchimento da semente (que são diversas sementes enchedoras divididas pelo número total de sementes e multiplicadas por 100), aumento no peso singular da semente, entre outros.

A presente invenção fornece um método para aumentar a produção, especialmente a produção da semente de plantas, que diz respeito as plantas de controle, que compreende métodos de modulação de expressão, preferivelmente expressão aumentada, na planta de um ácido nucleico codificando um polipeptídeo semelhante a CLE como definido neste.

Visto que as plantas transgênicas de acordo com a presente invenção tem produção aumentada, isto é provavelmente que estas plantas

exibem uma taxa de crescimento aumentada (durante pelo menos parte de seu ciclo de vida), relativa a taxa de crescimento de plantas de controle em estágios correspondentes em seu ciclo de vida.

A taxa de crescimento aumentada pode ser específica a uma ou mais partes de uma planta (incluindo sementes), ou pode ser completamente substancial na planta total. As plantas que tem uma taxa de crescimento aumentada pode ter um ciclo de vida curto. O ciclo de vida de uma planta pode ser retirada por significar o período necessário ao crescimento a partir de uma semente madura seca até o estágio onde as plantas tem produzido sementes maduras secas, similar ao material de partida. Este ciclo de vida pode ser influenciado pelos fatores tal como vigor precoce, taxa de crescimento, índice da verdura, período de florescência e rapidez da maturação da semente. O aumento na taxa de crescimento pode acontecer em um ou mais estágios no ciclo de vida de uma planta ou durante ou durante o ciclo de vida total na planta substancial. Aumento na taxa de crescimento durante o estágios precoces no ciclo de vida de uma planta pode refletir vigor intensificado. O aumento na taxa de crescimento pode alterar o ciclo de colheita de uma planta permitindo as plantas serem semeadas depois e/ou colhidas mais cedo do que seria possível de outra maneira (um efeito similar pode ser obtido com período de florescência precoce). Se a taxa de crescimento é suficientemente aumentada, isto permite uma semeadura adicional de sementes das mesmas espécies de plantas (por exemplo semeadura e colheita de plantas de arroz seguidas pela semeadura e colheita de plantas de arroz adicionais todas dentro de um período de crescimento convencional). Similarmente, se a taxa de crescimento é suficientemente aumentada, isto permite uma semeadura adicional de sementes de diferentes espécies de plantas (por exemplo a semeadura e colheita de plantas de milho seguidas pela, por exemplo, a semeadura e colheita opcional de soja, batata ou qualquer outra planta adequada). Os períodos de colheita adicionais a partir da

mesma fonte no caso de algumas plantas de colheita também pode ser possível. Alterando o ciclo de colheita de uma planta pode levar a um aumento na produção de biomassa anual por acre (devido a um aumento de diversos períodos (anuncia em um ano) que qualquer planta particular pode ser desenvolvida e colhida). Um aumento na taxa de crescimento também pode permitir para a cultivo de plantas transgênicas em uma área geográfica extensa do que suas contra-partes de tipo selvagem, visto que as limitações territoriais para crescimento de uma safra são freqüentemente determinadas pelas condições ambientais adversas no período do plantio (época precoce) ou no período da colheita (época atrasada). Tais condições adversas pode ser evitadas se o ciclo de colheita é encurtado. A taxa de crescimento pode ser determinada pelos vários parâmetros derivados a partir de curvas de crescimento, tais parâmetros pode ser: T-Mid (o período de retirada das plantas para atingir 50% de seu tamanho máximo) e T-90 (período de retirada da planta para atingir 90% de seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção dão as plantas que tem uma taxa de crescimento aumentada que diz respeito as plantas de controle. Portanto, de acordo com a presente invenção, aqui é fornecido um método para aumentar a taxa de crescimento das plantas, que compreende métodos de modulação de expressão, preferivelmente expressão aumentada, na planta de um ácido nucleico codificando um polipeptídeo semelhante a CLE como definido neste.

Um aumento na produção e/ou taxa de crescimento ocorre se a planta está sob condições sem-tensão ou se a planta está exposta a várias tensões comparadas as plantas de controle. Plantas tipicamente respondem a exposição por tensão pelo crescimento mais lentamente. Em condições de grave tensão, a planta ainda pode interromper crescimento completamente. Tensão branda por outro lado é definida neste como sendo qualquer tensão

que uma planta está exposta pelo qual não resulta na interrupção do crescimento completamente da planta sem a capacidade para recuperar o crescimento. Tensão branda no sentido da invenção conduz a uma redução no crescimento das plantas submetidas à tensão de menos do que 40%, 35% ou 30%, preferivelmente menos do que 25%, 20% ou 15%, mais preferivelmente menos do que 14%, 13%, 12%, 11% ou 10% ou menos em comparação ao controle da planta sob condições sem-tensão. Devido a melhora nas práticas agrícolas (irrigação, fertilização, tratamentos com pesticidas) graves tensões não são freqüentemente encontradas em safra de plantas cultivadas. Como uma conseqüência, o crescimento compromissado induzido pela tensão branda é freqüentemente uma característica indesejável para a agricultura. Tensões brandas são as tensões (ambientais) abióticas ou bióticas diárias que uma planta está exposta. Tensões abióticas pode ser devido à aridez ou excesso de água, tensão anaeróbica, tensão ao sal, toxicidade química, tensão oxidativa e temperaturas de calor, frio ou congeladas. A tensão abiótica pode ser uma tensão osmótica causada por um tensão à água (particularmente devido à aridez), tensão ao sal, tensão oxidativa ou uma tensão iônica. Tensões bióticas são tipicamente aquelas tensões causadas por patogênios, tal como bactéria, vírus, fungo e insetos.

Em particular, os métodos da presente invenção podem ser realizados sob condições sem-tensão ou sob condições de aridez branda para dar as plantas que tem produção aumentada que diz respeito as plantas de controle. Como relatadas em Wang et al. (Planta (2003) 218: 1-14), tensão abiótica conduz a uma série de mudanças morfológica, fisiológica, bioquímica e molecular que adversamente afetam o crescimento da planta e produtividade. Aridez, salinidade, temperaturas extremas e tensão oxidativa são conhecidas a serem interconectadas e podem induzir crescimento e dano celular através de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descreve um alto grau particularmente de "*cross talk*" entre

tensão da aridez e tensão de alta salinidade. Por exemplo, aridez e/ou salinização são manifestadas primariamente como tensão osmótica, resultando no rompimento de homeostase e distribuição de íon na célula. Tensão oxidativa, que freqüentemente acompanha temperatura alta ou baixa, salinidade ou tensão da aridez, pode causar desnaturização das proteínas funcionais e estruturais. Como uma conseqüência, estas diversas tensões ambientais freqüentemente ativam os caminhos de sinalização celular similares e respostas celulares, tal como a produção de proteínas de tensão, super-regulação de anti-oxidantes, acumulação de solutos compatíveis e detém o crescimento. O termo condições "sem-tensão" como usado neste são aquelas condições ambientais que permitem ótimo crescimento de plantas. As pessoas habilitadas na técnica estão atentas as condições normais do solo e condições climáticas para uma dado local.

O desempenho dos métodos da invenção dão crescimento de plantas sob condições sem-tensão ou produção aumentada de sob condições de aridez branda que diz respeito as plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, aqui é fornecido um método para aumentar a produção em plantas desenvolvidas sob condições sem-tensão ou sob branda condições de aridez, que compreende métodos da expressão aumentada na planta de um ácido nucleico codificando um polipeptídeo semelhante a CLE.

O desempenho dos métodos da invenção dão crescimento de plantas sob condições de deficiência de nutriente, particularmente sob condições de deficiência de nitrogênio, produção aumentada que diz respeito as plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, aqui é fornecido um método para aumentar a produção em plantas desenvolvidas sob condições de deficiência de nutriente, que compreende métodos de expressão aumentada na planta de um ácido nucleico codificando um polipeptídeo semelhante a CLE. Deficiência de

nutriente pode resultar a partir de uma falta de nutrientes tal como nitrogênio, fosfatos e outros compostos contendo fosforosos, potássio, cálcio, cádmio, magnésio, manganês, ferro e boro, entre outros.

A presente invenção também abrange o uso de ácidos nucleicos semelhante a CLE na alteração da produção que diz respeito a traços.

O ácido nucleico que codifica polipeptídeos semelhantes a CLE podem observar o uso nos programas de geração em que um marcador de DNA é identificado que pode ser geneticamente ligado a um gene semelhante a CLE. Os ácidos nucleicos/genes podem ser usados para definir um marcador molecular. Este marcador de DNA pode ser usado nOs programas de geração para selecionar as plantas que tem produção aumentada como definido neste acima nos métodos da invenção.

As variantes alélicas do ácido nucleico semelhante a CLE/gene também podem observar o uso em programas de geração auxiliados por marcador. Tais programas de geração as vezes requerem a introdução da variação alélica pelo tratamento mutagênico da plantas, usando por exemplo mutagênese EMS; alternativamente, o programa pode começar com uma coleção de variantes alélicas deste modo denominada origem "natural" não causada intencionalmente. A identificação de variantes alélicas então apresentam, por exemplo, por PCR. Esta é seguida por uma etapa para seleção de várias variantes alélicas da seqüência em questão e que dão a produção aumentada. A seleção é tipicamente realizada pelo monitoramento do desempenho desenvolvido de plantas contendo variantes alélicas diferentes da seqüência em questão. O desempenho desenvolvido pode ser monitorado em uma estufa para plantas ou no campo. Etapas opcionais adicionais incluem plantas de cruzamento em que a variante alélica superior foi identificada com uma outra planta. Esta pode ser usada, por exemplo, por fazer uma combinação de características fenotípicas de interesse.

Um ácido nucleico semelhante a CLE também pode ser usado como sondas para mapeamento geneticamente ou físico dos genes que estes são a parte de, e como marcadores por traços ligados aqueles genes. Tal informação pode ser útil na geração de plantas a fim de desenvolver linhas com fenotipos desejados. O tal uso de ácido nucleico semelhante a CLEs requer apenas uma seqüência de ácido nucleico de pelo menos 15 nucleotídeos no comprimento. Os ácidos nucleicos semelhantes a CLE podem ser usados como marcadores de poliformismo de comprimento do fragmento de restrição (RFLP). *Blots Southern* (Sambrook J, Fritsch EF e Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico de planta digerida com restrição pode ser fornecida com o ácido nucleico semelhante a CLE. A origem das fitas resultadas pode ser então submetida por análise genética usando programas de computador tal como MapMaker (Lander et al. (1987) Genomics 1: 174-181) a fim do construto de um mapa genético. Além disso, os ácidos nucleicos podem ser usados por fonte de *blots Southern* contendo restrição de DNA genômico tratado por endonuclease de uma série de indivíduos que representam a origem e a descendência de um cruzado genético definido. A segregação do poliformismo de DNA é notada e usada para calcular as posições do ácido nucleico semelhante a CLE no mapa genético previamente obtido usando esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

A produção e uso das sondas derivadas de gene de planta para o uso na cópia do mapa genético é descrito em Bernatzky e Tanksley (Plant Mol. Biol. Reporter 4: 37-41, 1986). Publicações numerosas descrevem a cópia do mapa genético dos clones de cDNA específico usando a metodologia resumida acima ou variações destas. Por exemplo, Populações inter-cruzadas F2, populações de suporte cruzado, populações unidas aleatoriamente, linhas isogênicas próximas, e outras séries de indivíduos que podem ser usado para o mapeamento. Tais metodologias são bem conhecidas aqueles habilitados na

técnica.

As sondas de ácido nucleico também pode ser usadas para mapeamento físico (isto é, disposição das seqüências em mapas físicos; ver Hoheisel et al. In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346, e referências citadas neste).

Em uma outra forma de realização, as sondas de ácido nucleico podem ser usadas em fluorescência direta em mapeamento de hibridização in situ (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora os métodos correntes do mapeamento FISH favorece o uso de clones amplos (diversos kb por diversos cem kb; ver Laan et al. (1995) Genome Res. 5:13-20), melhora na sensibilidade pode permitir o desempenho do mapeamento FISH usando sondas mais curtas.

Uma variedade de métodos com base na amplificação de ácido nucleico para o mapeamento genético e físico pode ser realizado usando os ácidos nucleicos. Exemplos incluem amplificação específica por alelo (Kazazian (1989) J. Lab. Clin. Med 11:95-96), poliformismo de fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332), ligação específica por alelo (Landegren et al. (1988) Science 241:1077-1080), reações de extensão por nucleotídeo (Sokolov (1990) Nucl. Acids Res. 18:3671), mapeamento híbrido de radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e mapeamento adequado (Dear e Cook (1989) Nucl. Acids Res. 17:6795-6807). Para estes métodos, a seqüência de um ácido nucleico é usada para designar e produzir pares iniciadores para o uso na reação de amplificação ou nas reações de extensão iniciadoras. O desenho de tais iniciadores é bem conhecidos por aqueles habilitados na técnica. Em métodos utilizados na Cópia de mapa genético com base em PCR, pode ser necessário para identificar as diferenças na seqüência de DNA entre a origem do cruzada do mapeamento na região que corresponde ao instante da seqüência de ácido nucleico. Este, entretanto, não é geralmente necessário para métodos de

mapeamento.

Os métodos de acordo com a presente invenção resultam nas plantas que tem produção alterada que diz respeito a traços, como descrito nestes acima. Estes traços também podem ser combinados com outros traços vantajosos economicamente, tal como trações de intensificação de produção adicional, a tolerância por outras tensões abióticas e bióticas, características arquitetural de vários traços de modificação e/ou características bioquímicas e/ou fisiológicas.

SYR

Surpreendentemente, tem sido agora observado a modulação da expressão na planta de um ácido nucleico codificando um polipeptídeo de SYR dando plantas, quando desenvolvidas sob condições de tensão abiótica, tendo produção intensificada que diz respeito a traços que diz respeito as plantas de controle. De acordo com uma primeira forma de realização, a presente invenção fornece um método para intensificar produção que diz respeito a traços em plantas desenvolvidas sob condições de tensão abiótica, que diz respeito as plantas de controle, que compreende a modulação da expressão na planta de um ácido nucleico codificando um polipeptídeo de SYR.

Um método preferido para modulação (preferivelmente, aumento) da expressão de um ácido nucleico codificando um polipeptídeo de SYR é pela introdução de uma expressão na planta por ácido nucleico codificando um polipeptídeo de SYR.

Qualquer referência depois disso a um "método da invenção útil para proteína" é retirado por significar um polipeptídeo de SYR como definido neste. Qualquer referência depois disso a um "ácido nucleico útil no método da invenção" é retirada por significar um ácido nucleico capaz de codificar um tal polipeptídeo de SYR. O ácido nucleico a ser introduzido na planta (e portanto útil no desempenho dos métodos da invenção) é qualquer

ácido nucleico que codifica o tipo de proteína que será descrita, depois também chamada "ácido nucleico SYR" ou "gene SYR".

O termo "proteína SYR ou homólogo deste" como definido neste refere-se a um polipeptídeo de cerca de 65 a cerca de 200 aminoácidos, que compreende (i) um domínio rico em leucina semelhante a um zíper de leucina na metade do terminal C das proteínas, que o domínio rico em leucina é (ii) precedido por um tripeptídeo com uma seqüência YFS (motivo conservado 1a, SEQ ID N°: 256), ou YFT (motivo conservado 1b, SEQ ID N°: 257), ou YFG (motivo conservado 1c, SEQ ID N°: 258) ou YLG (motivo conservado 1d, SEQ ID N°: 259), e (iii) seguido pelo motivo conservado 2 ((V/A/I)LAFMP(T/S), SEQ ID N°: 260). Preferivelmente, o motivo conservado 2 é (A/V)LAFMP(T/S), mais preferivelmente, o motivo conservado é VLAFMPT. A "proteína SYR ou homólogo deste" preferivelmente também tem um peptídeo de terminal C conservado codificando o motivo conservado 3 (SYL ou PYL, SEQ ID N°: 261). O domínio rico em leucina da proteína SYR ou deste homólogo é cerca de 38 a 48 aminoácidos longos, a partida imediatamente atrás do motivo conservado 1 e a interrupção imediata antes do motivo conservado 2, e que compreende pelo menos 30% de leucina. O domínio rico em Leu preferivelmente tem um motivo semelhante ao zíper do motivo de leucina (L-X_6-L-X_6-L-X_6-L, neste X_6 é uma seqüência de 6 aminoácidos consecutivos). Um exemplo preferido de uma proteína SYR é representado pela SEQ ID N°: 252, um resumo deste domínio é dado na figura 24. Será notado que o termo "proteína SYR ou homólogo deste" não abrange a proteína ARGOS de Arabidopsis thaliana (SEQ ID N°: 276).

Ainda preferivelmente, as proteínas SYR tem dois domínios de transmembrana, com o parte do terminal N e parte do terminal C das proteínas localizadas dentro e as partes entre os domínios da transmembrana localizados fora.

Alternativamente, o homólogo de uma proteína SYR tem um aumento a fim da preferência pelo menos 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, ou 99% da identidade da seqüência total ao aminoácido representado pela SEQ ID N°: 252, fornecido que a proteína de homólogo compreende os motivos conservados 1 (a, b, c ou d), 2 e 3, e o domínio rico em leucina como resumido acima. A identidade de seqüência total é determinada usando um algoritmo de alinhamento global, tal como o algoritmo Needleman Wunsch no programa GAP (GCG Wisconsin Package, Accelrys), preferivelmente com os parâmetros padrões.

O termo "domínio" e "motivo" é definido nas seção de "definições" neste. A base de dados especializada existe para a identificação dos domínios, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucl. Acids Res 30, 242-244, InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318, Prosite (Bucher and Bairoch (1994), uma regra de perfil generalizado pelos motivos das seqüências biomoleculares e esta função na interpretação da seqüência automática. (In) ISMB-94; Proceedings 2° International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004), ou Pfam (Bateman et al., Nucl. Acids Research 30(1): 276-280 (2002). Uma série de ferramentas para análise in silico das seqüências de proteína é disponível no servidor proteômico de ExPASY (Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucl. Acids

Res. 31:3784-3788(2003)). Os domínios também podem ser identidificados usando técnicas de rotina, tal como pelo alinhamento de seqüência.

Os métodos para o alinhamento das seqüências para comparação são bem conhecidos na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. GAP usa o algoritmo de Needleman e Wunsch ((1970) J Mol Biol 48: 443-453) para observar o global (isto é medindo-se as seqüências completas) alinhamento de duas seqüências que maximizam o número de pontos e minimiza o número de fendas. O algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-10) calcula o percentual da identidade da seqüência e realiza uma análise estatística de similaridade entre as duas seqüências. O software para a realização da análise BLAST é publicamente disponível através do National Centre for Biotechnology Information (NCBI). Os homólogos podem ser facilmente identificados usando, por exemplo, o algoritmo de alinhamento da seqüência múltipla ClustalW (versão 1.83), com os parâmetros de alinhamento em pares padrão, e um método de contagem em porcentagem. As porcentagens globais da similaridade e identidade também podem ser determinados usando um dos métodos disponíveis na embalagem do software MatGAT (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando proteína ou seqüências de DNA.). A edição do manual menor pode ser realizado para otimizar o alinhamento entre os motivos conservados, como seria apresentado por uma pessoa habilitada na técnica. Além disso, em vez do uso das seqüências de comprimento total para uma identificação de homólogos, domínios específicos também pode ser usados. A identidade de valores de seqüência pode ser determinado no ácido nucleico inteiro ou seqüência de aminoácido ou em domínios selecionados ou motivos conservados, usando Os programas mencionados acima usando o parâmetros padrões.

Os domínios da transmembrana são de cerca de 15 a 30

aminoácidos longos e são usualmente compostos de resíduos hidrofóbicos que formam um espiral alfa. Estes são usualmente preditos na base de hidrofobicidade (por exemplo Klein et al., Biochim. Biophys. Acta 815, 468, 1985; ou Sonnhammer et al., In J. Glasgow, T. Littlejohn, F. Major, R. Lathrop, D. Sankoff, e C. Sensen, editors, Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology, páginas 175-182, Menlo Park, CA, 1998. AAAI Press.).

Os exemplos de proteínas que quebram sob a definição de "polipeptídeo de SYR ou um homólogo deste" são dados na Tabela II da seção de exemplos e incluem seqüências de várias plantas monocotiledôneas, tal como arroz (SEQ ID N°: 252, SEQ ID N°: 262 e SEQ ID N°: 263), milho (SEQ ID N°: 264), trigo (SEQ ID N°: 265), cevada (SEQ ID N°: 266), cana de açúcar (SEQ ID N°: 267 e SEQ ID N°: 268), sorgo (SEQ ID N°: 269); e de plantas dicotiletiledônias tal como Arabidopsis (SEQ ID N°: 270 e SEQ ID N°: 271), uva (SEQ ID N°: 272), cítricas (SEQ ID N°: 273) ou tomate (SEQ ID N°: 274 e SEQ ID N°: 275). É considerado que o domínio rico em Leu é importante para a função das proteínas, visto que as proteínas com o domínio rico em Leu mas sem os motivos conservados 1 ou 2 podem ser úteis bem como nos métodos da presente invenção; exemplos de tais proteínas são dados na SEQ ID N°: 284 e 285.

Será entendido que o termo "polipeptídeo de SYR ou um homólogo deste" não é limitado a uma seqüência representada pela SEQ ID N°: 252 ou aos homólogos listados como SEQ ID N°: 262 à SEQ ID N°: 275, mas que qualquer polipeptídeo de cerca de 65 a cerca de 200 aminoácidos reunidos a critério que compreende um domínio rico em leucina como definido acima, precedido pelo motivo de tripeptídeo conservado 1 (a, b, c ou d) e seguido pelo o motivo conservado 2 e também preferivelmente pelo motivo conservado 3; ou tendo pelo menos 38% da seqüência da identidade a uma seqüência da SEQ ID N°: 252, pode ser adequado para o uso nos

métodos da invenção.

A atividade de uma proteína SYR ou homólogo deste pode ser submetido à ensaio pela expressão da proteína SYR ou homólogo deste sob controle de um promotor GOS2 em *Oryza sativa*, que resulta em plantas com biomassa aumentada e/ou produção da semente sem nenhuma atraso no período de florescência quando desenvolvida sob condições de deficiência de nitrogênio e comparadas como correspondente as plantas do tipo selvagem. Este aumento na produção da semente pode ser medido em diversos caminhos, por exemplo como um aumento do peso da semente total, número de sementes enchedoras ou peso singular da semente.

A presente invenção é ilustrada pela transformação das plantas com a seqüência de ácido nucleico representada pela SEQ ID N°: 251, que codifica a seqüência de polipeptídeo da SEQ ID N°: 252. Entretanto, o desempenho da invenção não é restrita aquelas seqüências; os métodos da invenção podem ser vantajosamente realizados usando qualquer SYR que codifica ácido nucleico ou polipeptídeo de SYR como definido neste.

Exemplos de ácido nucleicos que codificam polipeptídeos SYR são dados na Tabela II do Exemplo 38 neste. Tais ácidos nucleicos são úteis no desempenho dos métodos da invenção. As seqüências de aminoácidos dados na Tabela II do Exemplo 38 são seqüências de exemplos de ortólogos e parálogos do polipeptídeo de SYR representada pela SEQ ID N°: 252. Os termos "ortólogos" e "parálogos" sendo como definido neste. Ainda ortólogos e parálogos podem ser facilmente identificados pela realização de uma busca de blast recíproca denominada. Tipicamente, esta envolve o primeiro BLAST que envolve a explosão da seqüência em questão (por exemplo usando qualquer de uma seqüência listada na tabela II do Exemplo 38) contra qualquer base de dados de seqüência, tal como a base de dados NCBI disponível publicamente. BLASTN ou TBLASTX (usando valores de descuido padrão) são geralmente usados quando a partida de uma

seqüência de nucleotídeo, e BLASTP ou TBLASTN (usando valores de descuido padrão) quando a partir da seqüência da proteína. O resultado de BLAST pode ser opcionalmente filtrado. As seqüências de comprimento total dos resultados filtrados ou resultados não filtrados são então suporte estragado (segundo BLAST) contra as seqüências a partir de um organismo de que a seqüência em questão é derivada (onde a seqüência em questão é a SEQ ID N°: 251 ou SEQ ID N°: 252, o segundo BLAST seria portanto contra as seqüências de arroz). Os resultados do primeiro e segundo BLAST são então comparados. Um parálogo é identificado se um acerto de alta classificação do primeiro blast é da mesma espécie como da seqüência em questão é derivada, o BLAST baixo então idealmente resulta na seqüência em questão entre os acertos maiores; um ortólogo é identificado se um acerto de alta classificação no primeiro BLAST não é da mesma espécie como da seqüência em questão é derivada, e preferivelmente resulta no BLAST baixo na seqüência em questão sendo entre os acertos maiores.

Os acertos de alta classificação são aqueles tendo um valor E baixo. O valor E inferior, o mais significante da contagem (ou em outras palavras o inferior tem a chance que o acerto seja observado pela escolha). A computação do valor E é bem conhecidos na técnica. Além disso a comparação aos valores E também são contados pela identidade de porcentagem. A identidade de porcentagem refere-se à diversos nucleotídeos idênticos (ou aminoácidos) entre as seqüências de dois ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. No caso de grandes famílias, ClustalW pode ser usadas, seguidas pela árvore de ligação próxima, para ajudar a visualizar o agrupamento que diz respeito a um gene e para identificar ortólogos e parálogos.

As variantes de ácidos nucleicos também podem ser úteis na prática dos métodos da invenção. Exemplos de tais variantes incluem ácidos nucleicos que codificam homólogos e derivados de qualquer um das

seqüências de aminoácido dados na Tabela II do Exemplo 38, os termos "homólogo" e "derivado" sendo como definido neste. Também úteis nos métodos da invenção são os ácidos nucleicos que codificam homólogos e derivados de ortólogos ou parólogos de qualquer um das seqüências de aminoácido dados na Tabela II do Exemplo 38. Os homólogos e derivados úteis nos métodos da presente invenção tem substancialmente a mesma atividade biológica e funcional como as proteínas não modificadas de que são derivados.

As variantes adicionais de ácidos nucleicos úteis na prática dos métodos da invenção incluem porções de ácidos nucleicos que codificam polipeptídeos de SYR, ácidos nucleicos de hibridização por ácidos nucleicos que codificam polipeptídeos de SYR, variantes unidas de ácidos nucleicos que codificam polipeptídeos de SYR, variantes alélicas de ácidos nucleicos que codificam polipeptídeos de SYR e variantes de ácidos nucleicos que codificam polipeptídeos de SYR obtidos pela mistura do gene. Os termos seqüência de hibridização, variante unida, variante alélilca e a mistura do gene são como descritos neste.

Os ácidos nucleicos que codificam os polipeptídeos de SYR não necessitam ser os ácidos nucleicos de comprimento total, visto que o desempenho dos métodos da invenção não confiam no uso de seqüência de ácido nucleico de comprimento total. De acordo com a presente invenção, aqui é fornecido um método para intensificar a produção que diz respeito a traços em plantas, compreende a introdução de uma expressão na planta uma porção de qualquer uma da seqüência de ácido nucleico dado na Tabela II do Exemplo 38, ou uma porção de um ácido nucleico codificando um ortólogo, parálogo ou homólogo de qualquer das seqüências de aminoácido dados na Tabela II do Exemplo 38.

Uma porção de um ácido nucleico pode ser preparada, por exemplo, pela fabricação uma ou mais anulações ao ácido nucleico. As

porções podem ser usadas na forma isolada ou este pode ser fundido a outras seqüências codificados (ou não codificadas) a fim de, por exemplo, produzir uma proteína que combina com diversas atividades. Quando fundido a outras seqüências codificadoras, o polipeptídeo resultante produz na tradução podendo ser maior do que predito para a porção de proteína.

As porções úteis nos métodos da invenção, codificam um polipeptídeo de SYR como definido neste, e tem substancialmente a mesma atividade biológica como as seqüências de aminoácido dados na Tabela II do Exemplo 38. Preferivelmente, a porção é uma porção de qualquer um dos ácido nucleicos dados na Tabela II do Exemplo 38, ou é uma porção de um ácido nucleico codificando um ortólogo ou parálogo de qualquer uma das seqüências de aminoácido dados na Tabela II do Exemplo 38. Preferivelmente a porção é pelo menos 150, 200, 250, 300, 350, 400, 450, 500, 550, 600 nucleotídeos consecutivos no comprimento, os nucleotídeos consecutivos sendo de qualquer uma da seqüência de ácido nucleico dado na Tabela II do Exemplo 38, ou de um ácido nucleico codificando um ortólogo ou parálogo de qualquer uma das seqüências de aminoácido dados na Tabela II do Exemplo 38. Mais preferivelmente a porção é uma porção do ácido nucleico da SEQ ID N°: 251. Preferivelmente, as porções codificadas de um polipeptídeo de cerca de 65 a cerca de 200 aminoácidos, compreende um domínio rico em leucina como definido acima, precedido pelo motivo de tripeptídeo conservado 1 (a, b, c ou d) e seguido pelo motivo conservado 2 e também preferivelmente pelo motivo conservado 3; ou tendo pelo menos 38% da seqüência da identidade a uma seqüência da SEQ ID N°: 252.

Uma outra variante de ácido nucleico útil nos métodos da invenção é um ácido nucleico capaz de hibridização, sob condições de estringência reduzidas, preferivelmente sob condições estringentes, como um ácido nucleico codificando um polipeptídeo de SYR como definido neste, ou com uma porção como definido neste.

De acordo com a presente invenção, aqui é fornecido um método para intensificar a produção que diz respeito a traços em plantas, que compreende a introdução de uma expressão na planta do ácido nucleico capaz de hibridização a qualquer um dos ácido nucleicos dados na Tabela II do Exemplo 38, ou que compreende a introdução de uma expressão na planta a ácido nucleico capaz da hibridização por um ácido nucleico codificando um ortólogo, parálogo ou homólogo de qualquer da seqüência de ácido nucleico dados na Tabela II do Exemplo 38.

As seqüências de hibridização úteis nos métodos da invenção codificam um polipeptídeo de SYR como definido neste, e tem substancialmente a mesma atividade biológica como as seqüências de aminoácido dados na Tabela II do Exemplo 38. Preferivelmente, as seqüências de hibridização é capaz de hibridizar qualquer um dos ácidos nucleicos dados na Tabela II do Exemplo 38, ou por uma porção de qualquer destas seqüências, uma porção sendo como definido acima, ou neste as seqüências de hibridização são capazes da hibridização por um ácido nucleico codificando um ortólogo ou parálogo de qualquer uma das seqüências de aminoácido dados na Tabela II do Exemplo 38. Mais preferivelmente, as seqüências de hibridização é capaz da hibridização por a ácido nucleico como representado pela SEQ ID N°: 251 ou por uma porção deste.

Preferivelmente, as seqüências de hibridização codificadas por um polipeptídeo de cerca de 65 a cerca de 200 aminoácidos, compreende um domínio rico em leucina como definido acima, precedido pelo motivo de tripeptídeo conservado 1 (a, b, c ou d) e seguido pelo motivo conservado 2 e também preferivelmente pelo motivo conservado 3; ou tendo pelo menos 38% da seqüência da identidade a uma seqüência da SEQ ID N°: 252.

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante unida codificando um polipeptídeo de SYR como definido neste acima, uma variante unida sendo como definido neste.

De acordo com a presente invenção, aqui é fornecido um método para intensificar a produção que diz respeito a traços em plantas, que compreende a introdução de uma expressão na planta de uma variante unida de qualquer uma da seqüência de ácido nucleico dados na Tabela II do Exemplo 38, ou uma variante unida de um ácido nucleico codificando um ortólogo, parálogo ou homólogo de qualquer das seqüências de aminoácido dados na Tabela II do Exemplo 38.

As variantes unidas preferidas são variantes unidas de um ácido nucleico representada pela SEQ ID N°: 251, ou uma variante unida de um ácido nucleico codificando um ortólogo ou parálogo da SEQ ID N°: 252. Preferivelmente, a seqüência de aminoácido codificado pela variante unida é um polipeptídeo de cerca de 65 a cerca de 200 aminoácidos, que compreende um domínio rico em leucina como definido acima, precedido pelo motivo de tripeptídeo conservado 1 (a, b, c ou d) e seguido pelo motivo conservado 2 e também preferivelmente pelo motivo conservado 3; ou tendo pelo menos 38% da seqüência da identidade a uma seqüência da SEQ ID N°: 252.

Uma outra variante de ácido nucleico útil no desempenho dos métodos da invenção é uma variante alélica de um ácido nucleico codificando um polipeptídeo de SYR como definido neste acima, uma variante alélica sendo como definido neste.

De acordo com a presente invenção, aqui é fornecido um método para intensificar a produção que diz respeito a traços em plantas, que compreende a introdução de uma expressão na planta uma variante alélica de qualquer um dos ácido nucleico dados na Tabela II do Exemplo 38, ou que compreende a introdução de uma expressão na planta uma variante alélica de um ácido nucleico codificando um ortólogo, parálogo ou homólogo de qualquer das seqüências de aminoácido dados na Tabela II do Exemplo 38.

As variantes alélicas úteis nos métodos da presente invenção tem substancialmente a mesma atividade biológica como o polipeptídeo de

SYR da SEQ ID N°: 252 e qualquer dos aminoácidos descrito na tabela II do Exemplo 38. Variantes alélicas existem na natureza, e abrangem os métodos da presente invenção, é o uso destes alelos naturais. Preferivelmente, a variante alélica é uma variante alélica da SEQ ID N°: 251 ou uma variante alélica de um ácido nucleico codificando um ortólogo ou parálogo da SEQ ID N°: 252. Preferivelmente, a seqüência de aminoácido codificado pela variante alélica é um polipeptídeo de cerca de 65 a cerca de 200 aminoácidos, que compreende um domínio rico em leucina como definido acima, precedido pelo motivo de tripeptídeo conservado 1 (a, b, c ou d) e seguido pelo motivo conservado 2 e também preferivelmente pelo motivo conservado 3; ou tendo pelo menos 38% da seqüência da identidade a uma seqüência da SEQ ID N°: 252.

A mistura do gene ou evolução direcionada também pode ser usada como variantes geradas do ácido nucleico que codifica polipeptídeos de SYR como definido acima; o termo "a mistura do gene" sendo como definido neste.

De acordo com a presente invenção, aqui é fornecido um método para intensificar a produção que diz respeito a traços em plantas, que compreende a introdução de uma expressão na planta de uma variante de qualquer uma da seqüência de ácido nucleico dados na Tabela II do Exemplo 38, ou que compreende a introdução de uma expressão na planta de uma variante de um ácido nucleico codificando um ortólogo, parálogo ou homólogo de qualquer das seqüências de aminoácido dados na Tabela II do Exemplo 38, pelo qual a variante do ácido nucleico é obtida pela mistura do gene.

Preferivelmente, a seqüência de aminoácido codificado pela variante de ácido nucleico obtida pela mistura do gene é um polipeptídeo de cerca de 65 a cerca de 200 aminoácidos, que compreende um domínio rico em leucina como definido acima, precedido pelo motivo de tripeptídeo

conservado 1 (a, b, c ou d) e seguido pelo motivo conservado 2 e também preferivelmente pelo motivo conservado 3; ou tendo pelo menos 38% da seqüência da identidade a uma seqüência da SEQ ID N°: 252.

Além disso, variantes de ácidos nucleicos também podem ser obtidas pela mutagênese direcionado ao local. Diversos métodos são disponíveis para atingir mutagênese direcionado ao local, a mais comum sendo métodos com base em PCR (Current Protocols in Molecular Biology; Wiley Eds.).

Ácido nucleicos que codificam polipeptídeos de SYR podem ser derivados de qualquer fonte natural ou artificial. O ácido nucleico pode ser modificado a partir da forma natural em ambiente de composição e/ou genômico através da manipulação humana deliberada. Preferivelmente o polipeptídeo de SYR que codifica o ácido nucleico é a partir da planta, ainda preferivelmente a partir da planta monocotiledônia, mais preferivelmente a partir da família Poaceae, mais preferivelmente o ácido nucleico é de *Oryza sativa*.

O desempenho dos métodos da invenção dão as plantas que tem aumento da resistência da tensão abiótica (ou tolerância da tensão abiótica, pelo termos que são usados permutavelmente), efetuando a produção intensificada que diz respeito a traços comparando as plantas de controle quando desenvolvidas sob tensão abiótica. Em particular, o desempenho dos métodos da invenção dão as plantas que tem produção aumentada, especialmente semente aumentada na produção e biomassa aumentada que diz respeito as plantas de controle. Os termos "produção" e "produção da semente" são descritas em mais detalhes nas seção de "definições" neste.

A referência neste da produção intensificada que diz respeito a traços é retirada por significar um aumento na biomassa (peso) de uma ou mais partes de uma planta, que pode incluir acima do solo (colhíveis) partes e/ou (colhíveis) partes abaixo trituradas. Em particular, tais partes colhíveis

são sementes, e o desempenho dos métodos da invenção resulta em plantas tendo semente aumentada na produção relativa à produção da semente de plantas de controle.

Tomando o milho como um exemplo, um aumento da produção pode ser manifestada como um ou mais dos seguintes: aumento de diversas plantas estabelecidas por hectare ou acre, um aumento de diversas espigas por plantas, um aumento de diversas séries, diversas sementes por série, peso da semente, peso singular da semente, comprimento da espiga/diâmetro, aumento da taxa de enchimento da semente (que são diversas sementes enchedoras divididas pelo número total de sementes e multiplicadas por 100), entre outros. Tomando o arroz como um exemplo, um aumento da produção pode manifestar em si mesma como um aumento de um ou mais dos seguintes: diversas plantas por hectare ou acre, diversas panículas por plantas, diversas espigas por panícula, número de flores (flósculos) por panícula (que é expressado como uma razão de diversas sementes enchedoras em diversas panículas primárias), aumento da taxa de enchimento da semente (que são diversas sementes enchedoras divididas pelo número total de sementes e multiplicadas por 100), aumento no peso singular da semente, entre outros.

A presente invenção fornece um método para aumentar a resistência da tensão abiótica de plantas, resultando na produção aumentada, especialmente a produção da semente e/ou biomassa aumentada de plantas, que diz respeito as plantas de controle, quando desenvolvidas sob condições de tensão abiótica, que compreende métodos de modulação de expressão, preferivelmente expressão aumentada, na planta de um ácido nucleico codificando um polipeptídeo de SYR como definido neste.

Visto que as plantas transgênicas de acordo com a presente invenção tem produção aumentada, isto é provavelmente que estas plantas exibem uma taxa de crescimento aumentada (durante pelo menos parte de seu

ciclo de vida), relativa a taxa de crescimento de plantas de controle em estágios correspondentes em seu ciclo de vida. Além disso a capacidade da produção aumentada, uma eficiência aumentada de nutriente também pode entrar para contribuir com o aumento produção. É observado que as plantas de acordo com a presente invenção mostram uma eficiência maior na entrada do nutriente. A eficiência aumentada da entrada do nutriente permite melhor crescimento da planta, quando as plantas estão sob tensão.

A taxa de crescimento aumentada pode ser específica a uma ou mais partes de uma planta (incluindo sementes), ou pode ser completamente substancial na planta total. As plantas que tem uma taxa de crescimento aumentada pode ter um ciclo de vida curta. O ciclo de vida de uma planta pode ser retirada por significar o período necessário ao crescimento a partir de uma semente madura seca até o estágio onde as plantas tem produzido sementes maduras secas, similar ao material de partida. Este ciclo de vida pode ser influenciado pelos fatores tal como vigor precoce, taxa de crescimento, índice da verdura, período de florescência e rapidez da maturação da semente. O aumento na taxa de crescimento pode acontecer em um ou mais estágios no ciclo de vida de uma planta ou durante ou durante o ciclo de vida total na planta substancial. Aumento na taxa de crescimento durante o estágios precoces no ciclo de vida de uma planta pode refletir vigor intensificado. O aumento na taxa de crescimento pode alterar o ciclo de colheita de uma planta permitindo as plantas serem semeadas depois e/ou colhidas mais cedo do que seria possível de outra maneira (um efeito similar pode ser obtido com período de florescência precoce). Se a taxa de crescimento é suficientemente aumentada, isto permite uma semeadura adicional de sementes das mesmas espécies de plantas (por exemplo semeadura e colheita de plantas de arroz seguidas pela semeadura e colheita de plantas de arroz adicionais todas dentro de um período de crescimento convencional). Similarmente, se a taxa de crescimento é suficientemente

aumentada, isto permite uma semeadura adicional de sementes de diferentes espécies de plantas (por exemplo a semeadura e colheita de plantas de milho seguidas pela, por exemplo, a semeadura e colheita opcional de soja, batata ou qualquer outra planta adequada). Os períodos de colheita adicionais a partir da mesma fonte no caso de algumas plantas de colheita também pode ser possível. Alterando o ciclo de colheita de uma planta pode levar a um aumento na produção de biomassa anual por acre (devido a um aumento de diversos períodos (anuncia em um ano) que qualquer planta particular pode ser desenvolvida e colhida). Um aumento na taxa de crescimento também pode permitir para o cultivo de plantas transgênicas em uma área geográfica extensa do que suas contra-partes de tipo selvagem, visto que as limitações territoriais para crescimento de uma safra são freqüentemente determinadas pelas condições ambientais adversas no período do plantio (época precoce) ou no período da colheita (época atrasada). Tais condições adversas pode ser evitadas se o ciclo de colheita é encurtado. A taxa de crescimento pode ser determinada pelos vários parâmetros derivados a partir de curvas de crescimento, tais parâmetros pode ser: T-Mid (o período de retirada das plantas para atingir 50% de seu tamanho máximo) e T-90 (período de retirada da planta para atingir 90% de seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção dão as plantas que tem uma taxa de crescimento aumentada que diz respeito as plantas de controle quando desenvolvidas sob tensão abiótica condições. Portanto, de acordo com a presente invenção, aqui é fornecido um método para aumentar a taxa de crescimento de plantas sob condições de tensão abiótica, que compreende métodos de modulação de expressão, preferivelmente expressão aumentada, na planta de um ácido nucleico codificando um polipeptídeo de SYR como definido neste.

Um aumento na produção e/ou taxa de crescimento ocorre

quando as plantas estão expostas a várias tensões abióticas comparadas as plantas de controle. As plantas tipicamente respondem a exposição por tensão pelo crescimento mais lentamente. Em condições de grave tensão, a planta ainda pode interromper crescimento completamente. Tensão branda por outro lado é definida neste como sendo qualquer tensão que uma planta está exposta pelo qual não resulta na interrupção do crescimento completamente da planta sem a capacidade de recuperar o crescimento. Tensão branda no sentido da invenção conduz a uma redução no crescimento das plantas submetidas à tensão de menos do que 40%, 35% ou 30%, preferivelmente menos do que 25%, 20% ou 15%, mais preferivelmente menos do que 14%, 13%, 12%, 11% ou 10% ou menos em comparação ao controle da planta sob condições sem-tensão. Devido a melhora nas práticas agrícolas (irrigação, fertilização, tratamentos com pesticidas) graves tensões não são freqüentemente encontradas em safra de plantas cultivadas. Como uma conseqüência, o crescimento compromissado induzido pela tensão branda é freqüentemente uma característica indesejável para a agricultura. Tensões brandas são as tensões (ambientais) abióticas ou bióticas diárias que uma planta está exposta. Tensões abióticas pode ser devido à aridez ou excesso de água, tensão anaeróbica, tensão ao sal, toxicidade química, tensão oxidativa e temperaturas de calor, frio ou congeladas. A tensão abiótica pode ser uma tensão osmótica causada por um tensão à água (particularmente devido à aridez), tensão ao sal, tensão oxidativa ou uma tensão iônica. Em uma forma de realização particular, a tensão abiótica tem a disponibilidade reduzida de um ou mais nutrientes que necessitam serem assimilados pelas plantas pelo crescimento e crescimento. Tensões bióticas são tipicamente aquelas tensões causadas por patogênios, tal como bactéria, vírus, fungo e insetos.

Em particular, os métodos da presente invenção podem ser realizadas sob condições de tensão, preferivelmente sob condições de disponibilidade reduzida de um ou mais nutrientes, ou sob condições de

aridez branda para dar as plantas que tem produção aumentada que diz respeito as plantas de controle. Como relatadas em Wang et al. (Planta (2003) 218: 1-14), tensão abiótica conduz a uma série de mudanças morfológica, fisiológica, bioquímica e molecular que adversamente afetam crescimento da planta e produtividade. Aridez, salinidade, temperaturas extremas e tensão oxidativa são conhecidas a serem interconectadas e podem induzir crescimento e dano celular através de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descrevem um alto grau particularmente de "*cross talk*" entre tensão da aridez e tensão de alta salinidade. Por exemplo, aridez e/ou salinização são manifestadas primariamente como tensão osmótica, resultando no rompimento de homeostase e distribuição de íon na célula. Tensão oxidativa, que freqüentemente acompanha temperatura alta ou baixa, salinidade ou tensão da aridez, pode causar desnaturização das proteínas funcionais e estruturais. Como uma conseqüência, estas diversas tensões ambientais freqüentemente ativam os caminhos de sinalização celular similares e respostas celulares, tal como a produção de proteínas de tensão, super-regulação de anti-oxidantes, acumulação de solutos compatíveis e detenção do crescimento. O termo condições "sem-tensão" como usado neste são aquelas condições ambientais que permitem ótimo crescimento de plantas. As pessoas habilitadas na técnica estão atentas as condições normais do solo e condições climáticas para uma dado local.

O desempenho dos métodos da invenção dão crescimento de plantas sob condições de tensão abiótica ou produção aumentada de sob condições de aridez branda que diz respeito as plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, aqui é fornecido um método para aumentar a produção em plantas desenvolvidas sob condições de tensão abiótica ou sob condições de aridez branda, que compreendem métodos de expressão aumentada na planta

de um ácido nucleico codificando um polipeptídeo de SYR.

O desempenho dos métodos da invenção dão crescimento de plantas sob condições de deficiência de nutriente, particularmente sob condições de deficiência de nitrogênio, produção aumentada que diz respeito as plantas de controle desenvolvidas sob condições comparáveis. Portanto, de acordo com a presente invenção, aqui é fornecido um método para aumentar a produção em plantas desenvolvidas sob condições de deficiência de nutriente, que compreende métodos de expressão aumentada na planta de um ácido nucleico codificando um polipeptídeo de SYR. A deficiência de nutriente pode resultar a partir de uma falta de nutrientes tal como nitrogênio, fosfatos e outros compostos contendo fosforosos, potássio, cálcio, cádmio, magnésio, manganês, ferro e boro, entre outros.

A presente invenção abrange plantas ou partes destas (incluindo sementes) obtidas pelos métodos de acordo com a presente invenção. As plantas ou partes destas que compreende um transgene de ácido nucleico codificando um polipeptídeo de SYR como definido acima.

A invenção também fornece construtos genéticos e vetores para facilitar a introdução e/ou expressão em plantas do ácido nucleico que codifica os polipeptídeos de SYR. Os construtos do gene podem ser inseridos em vetores, que podem ser comercialmente disponíveis, adequados para a transformação em plantas e adequados para a expressão do gene de interesse nas células transformadas. A invenção também fornece o uso de um construto do gene como definido neste nos métodos da invenção.

Mais especificamente, a presente invenção fornece um construto que compreende:

(a) um ácido nucleico codificando um polipeptídeo de SYR como definido acima;

(b) uma ou mais seqüências de controle capaz de conduzir a expressão da seqüência de ácido nucleico de; e opcionalmente

(c) uma seqüência de terminação de transcrição.

Preferivelmente, o ácido nucleico codificando um polipeptídeo de SYR é como definido acima. O termo "seqüência de controle" e "seqüência de terminação" são como definidos neste.

As plantas são transformadas com um vetor que compreende qualquer um dos ácidos nucleicos descrito acima. A pessoa habilitada é bem cautelosa aos elementos genéticos que podem estar presentes no vetor a fim de transformar sucessivamente, selecionado e propagado células hospedeiras que contém uma seqüência de interesse. A seqüência de interesse é operavelmente ligada a um ou mais seqüências de controle (pelo menos a um promotor).

Vantajosamente, qualquer tipo de promotor, se o natural ou sintético, pode ser usado para conduzir a expressão da seqüência de ácido nucleico. Um promotor constitutivo é particularmente útil nos métodos. Ver a seção de "definições" neste pelas definições dos vários tipos de promotores.

Será claro que a aplicabilidade da presente invenção não é restrita ao polipeptídeo de SYR que codifica o ácido nucleico representado pela SEQ ID N°: 251, ou é a aplicabilidade da invenção restrita a expressão de um polipeptídeo de SYR que codifica o ácido nucleico quando conduzido por um promotor constitutivo.

O promotor constitutivo é preferivelmente um promotor GOS2, preferivelmente um promotor GOS2 a partir do arroz. Ainda preferivelmente o promotor constitutivo é representado pela seqüência de ácido nucleico substancialmente similar à SEQ ID N°: 255 ou SEQ ID N°: 58, mais preferivelmente o promotor constitutivo é como representado pela SEQ ID N°: 255 ou SEQ ID N°: 58. Ver tabela 2a na seção de "definições" neste pelos exemplos adicionais do promotor útil constitutivo.

Opcionalmente, uma ou mais seqüências terminadoras podem ser usadas no construto introduzido na planta. Elementos reguladores

adicionais podem incluir intensificadores de transcrição bem como as os de tradução. Aqueles habilitados na técnica serão informados das seqüências terminadoras e intensificadoras que podem ser adequadas para o uso no desempenho da invenção. Um seqüência íntron também pode ser adicionada à extremidade da região 5' não traduzida (UTR) ou na seqüência codificadora para aumentar a quantidade da mensagem madura que acumula no citosol, como descrito na seção de definições. Outras seqüências de controle (além disso promotor, intensificador, silenciador, seqüências de íntron, regiões 3'UTR e/ou 5'UTR) podem ser proteínas e/ou elementos estabilizadores de RNA. Tais seqüências serão conhecidas ou podem ser facilmente obtidas por uma pessoa habilitada na técnica.

Os construtos genéticos da invenção ainda podem incluir uma origem da seqüência de replicação que é requerida para manutenção e/ou replicação em um tipo celular específico. Um exemplo é então um construto genético é requerido a ser mantido em uma célula bacteriana como um elemento genético episssomal (por exemplo moléculas de plasmídio ou cosmídio). Origens preferidas da replicação incluem, mas não limitam a, o f1-ori e colE1.

Para a detecção de transferir bem sucedida da seqüência de ácido nucleico como usado os métodos da invenção e/ou seleção de plantas transgênicas compreende estes ácidos nucleicos, este é vantajoso para o uso genes marcadores (ou genes repórter). Portanto, o construto genético pode opcionalmente compreender um gene marcador selecionável. Marcadores selecionáveis são descritos em mais detalhes nas seção de "definições" neste.

É conhecido na integração estável ou transitória do ácido nucleico nas células das plantas, apenas a minoria das células tornar-se o DNA estranho e, se desejado, integra este no genoma, dependendo do vetor de expressão usado e as técnicas de transfecção usadas. Para identificar e selecionar estes integrantes, um gene que codifica por um marcador

selecionável (tal como um descrito acima) é usualmente introduzido nas células hospedeiras junto com o gene de interesse. Estes marcadores podem ser por exemplo usados nos mutantes em que estes genes não são funcionais, por exemplo, a anulação pelos métodos convencionais. Além disso, as moléculas de ácido nucleico codificando um marcador selecionável pode ser introduzido na célula hospedeira no mesmo vetor que compreende uma seqüência que codifica os polipeptídeos da invenção ou usado os métodos da invenção, ou em um outro vetor separado. As células que tem estado estáveis transfectadas com o ácido nucleico introduzido podem ser identificadas por exemplo pela seleção (por exemplo, células que tem o marcador selecionável integrado sobrevive considerando as outras células mortas). Os genes marcadores pode ser removidos ou taxados a partir da célula transgênica uma vez que estes não são necessariamente longos. As técnicas para remoção do gene marcador são conhecidos na técnica, técnicas úteis são descritas acima na seção de definições.

A invenção também fornece um método para a produção de plantas transgênicas tendo, quando desenvolvidas sob condições de tensão abiótica, produção intensificada que diz respeito a traços que diz respeito as plantas de controle, compreende introdução e expressão na planta de qualquer ácido nucleico codificando um polipeptídeo de SYR como definido neste acima.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo aumentado a produção intensificada que diz respeito a traços, particularmente produção (semente) aumentada e/ou biomassa aumentada, que compreende métodos:

(i) introdução de uma expressão na planta ou célula da planta de um polipeptídeo de SYR que codifica o ácido nucleico; e

(ii) cultivo da célula da planta sob condições que promovem o crescimento da planta e crescimento.

O ácido nucleico de (i) pode ser qualquer dos ácidos nucleicos capazes de codificar um polipeptídeo de SYR como definido neste.

O ácido nucleico pode ser introduzido diretamente na célula da planta ou na própria planta (incluindo introdução no tecido, órgão ou qualquer outra parte de uma planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é preferivelmente introduzido na planta pela transformação. O termo "transformação" é descrito em mais detalhes nas seção de "definições" neste.

As células das plantas geneticamente modificadas podem ser regeneradas por intermédio de todos os métodos com que o técnico habilitado esteja familiar. Os métodos adequados podem ser observados nas publicações mencionadas acima por S.D. Kung e R. Wu, Potrykus ou Höfgen e Willmitzer.

Geralmente após a transformação, as células das plantas ou disposições celulares são selecionadas na presença de um ou mais marcadores que são codificados pelos genes expressíveis em plantas co-transferidas com o gene de interesse, seguido pelo qual o material transformado é regenerado na planta total. Para plantas transformadas seletas, o material da planta obtido na transformação é, como uma regra, submetida às condições seletivas de modo que plantas transformadas podem ser distinguidas das plantas não transformadas. Por exemplo, as sementes obtidas na maneira descrita acima podem ser plantadas e, após um período de crescimento inicial, submetida a uma seleção adequada por pulverização. Uma possibilidade adicional consiste no crescimento das sementes, se apropriado após a esterilização, em placas de ágar usando um agente de seleção adequado de modo que apenas as sementes transformadas podem desenvolver-se nas plantas. Alternativamente, as plantas transformadas são avaliadas na presença de um marcador selecionável tal como um descrito acima.

Seguindo a regeneração e transferência do DNA,

putativamente as plantas transformadas também pode ser avaliadas, por exemplo usando a análise de Southern, na presença do gene de interesse, organização de número de cópias e/ou genômicas. Alternativamente ou adicionalmente, os níveis de expressão do DNA introduzida novamente pode ser monitorado usando a análise Northern e/ou Western, ambas as técnicas sendo bem conhecidas às pessoas comuns habilitadas na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de meios, tal como pela propagação clonal ou técnicas de geração clássica. Por exemplo, uma primeira geração de planta transformada (ou T1) pode ser tornada independente e geração de segundo homozigoto (ou T2) transformantes selecionados, e as plantas T2 podem então ser propagadas adicionais através de técnicas de geração clássica. Os organismos transformados gerados podem tornar-se uma variedade de formas. Por exemplo, estes podem ser quimeras de células transformadas e células não transformadas; transformadores de clone (por exemplo, todas as células transformadas que contém a expressão cassete); enxertados de tecidos transformados e não transformados (por exemplo, em plantas, uma origem transformada enxertada por um muda não transformada).

A presente invenção claramente estende-se a qualquer célula da planta ou plantas produzida por qualquer dos métodos descritos neste, e por todas partes das plantas e *propagules* deste. A presente invenção estende-se ainda para abranger o resultado de uma célula transfectada ou transformada primária, tecido, órgão ou planta total que tem sido produzida por qualquer um dos métodos mencionados acima, a necessidade apenas sendo o resultado que exibe a mesma característica genotípica e/ou fenotípica como aquela produzida pela origem nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras que contém um ácido nucleico isolado codificando um polipeptídeo de SYR como definido neste acima. As células hospedeiras preferidas de acordo com a invenção são

as células das plantas. As plantas hospedeiras para o ácido nucleicos ou o vetor usado no método de acordo com a invenção, a expressão cassete ou construto ou vetor são, em princípio, vantajosamente todas as plantas, que são capazes da sintetização dos polipeptídeos usados no método inventivo.

Os métodos da invenção são vantajosamente aplicáveis a qualquer planta. As plantas que são particularmente úteis nos métodos da invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular monocotiledôneas e plantas dicotiletiledônias incluindo legumes de alimentação e de forragem, plantas ornamentais, safra de alimentação, árvores ou arbustos. De acordo com uma forma de realização preferida da presente invenção, a planta é uma planta de safra. Exemplos de plantas de safra incluem soja, girassol, canola, alfafa, colza, algodão, tomate, batata e tabaco. Ainda preferivelmente, a planta é uma planta monocotiledônia. Exemplos de plantas monocotiledôneas incluem cana de açúcar. Mais preferivelmente a planta é uma cereal. Exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, triticale, sorgo e aveias.

A invenção também estende-se a partes colhíveis de uma planta tal como, mas não limitado a sementes, folhas, frutos, flores, hastes, rizomas, tubérculos e bulbos. A invenção além disso diz respeito aos produtos derivados, preferivelmente derivado diretamente, da parte colhível de uma tal planta, tal como grânulos ou pós secos, óleo, gordura e ácidos graxos, amido ou proteínas.

De acordo com uma característica preferida da invenção, a expressão modulada é a expressão aumentada. Os métodos para aumentar expressão do ácido nucleico ou genes, ou produtos do gene, são bem documentados na técnica e os Exemplos são fornecidos na seção de definições.

Como mencionado acima, um método preferido para modulação (preferivelmente, aumento) da expressão de um ácido nucleico

codificando um polipeptídeo de SYR é pela introdução de uma expressão na planta de um ácido nucleico codificando um polipeptídeo de SYR; entretanto os efeitos de realização do método, isto é produção de intensificação que diz respeito a traços também pode ser atingidos usando outras técnicas bem conhecidas, incluindo mas não limitado a etiqueta de ativação de T-DNA, TILLING, recombinação homóloga. A descrição destas técnicas é fornecida na seção de definições.

A presente invenção também abrange o uso de ácido nucleico que codifica polipeptídeos de SYR como descrito neste e o uso destes polipeptídeos de SYR em intensificar qualquer um dos mencionados acima na produção que diz respeito a traços em plantas quando desenvolvidas sob tensão abiótica condições.

Ácido nucleicos que codificam os polipeptídeo de SYR descrito neste, ou os polipeptídeos de SYR por si mesmo, podem observar o uso nOs programas de geração em que um marcador de DNA é identificado que pode ser geneticamente ligado por um polipeptídeo de SYR que codifica o gene. O ácido nucleico/genes, ou os polipeptídeos de SYR por si mesmo podem ser usados para definir um marcador molecular. Este marcador de DNA ou proteína pode ser usado nOs programas de geração para selecionar plantas que tem produção intensificada que diz respeito a traços como definido neste acima nos métodos da invenção.

As variantes alélicas de um polipeptídeo de SYR que codifica o ácido nucleico/gene também podem observar o uso em programas de geração auxiliados por marcador. Tais programas de geração as vezes requerem a introdução da variação alélica pelo tratamento mutagênico da plantas, usando por exemplo mutagênese EMS; alternativamente, Os programas podem começar com uma coleção de variantes alélicas deste modo denominada origem "natural" não causada intencionalmente. A identificação de variantes alélicas então apresentam, por exemplo, por PCR. Esta é seguida

por uma etapa para seleção de várias variantes alélicas da seqüência em questão e que dão a produção aumentada. A seleção é tipicamente realizada pelo monitoramento do desempenho desenvolvido de plantas contendo variantes alélicas diferentes da seqüência em questão. Desempenho desenvolvido pode ser monitorado em uma estufa para plantas ou no campo. Etapas opcionais adicionais incluem plantas de cruzamento em que a variante alélica superior foi identificada com uma outra planta. Esta pode ser usada, por exemplo, por fazer uma combinação de características fenotípicas de interesse.

Os ácidos nucleicos que codificam polipeptídeos de SYR também podem ser usados como sondas para o mapeamento geneticamente ou físico dos genes pelo qual estes são uma parte de, e como marcadores por traços ligados aqueles genes. Tal informação pode ser útil na geração de plantas a fim de desenvolver linhas com fenotipos desejados. O uso de tal polipeptídeo de SYR que codifica o ácido nucleico requer apenas uma seqüência de ácido nucleico de pelo menos 15 nucleotídeos no comprimento. O polipeptídeo de SYR que codifica ácido nucleico pode ser usado como marcadores de poliformismo de comprimento do fragmento de restrição (RFLP). *Blots* Southern (Sambrook J, Fritsch EF e Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico de planta digerida com restrição pode ser sondada com o SYR que codifica os ácidos nucleicos. As origens da fita de resultado podem ser então submetidas por análise genética usando programas de computador tal como MapMaker (Lander et al. (1987) Genomics 1: 174-181) a fim do construto de um mapa genético. Além disso, os ácidos nucleicos podem ser usados por sondas *blots* Southern contendo restrição de DNA genômico tratado por endonuclease de uma série de indivíduos que representam a origem e a descendência de um cruzado genético definido. A segregação do poliformismo de DNA é notado e usado para calcular as posições do polipeptídeo de SYR que codifica o ácido

nucleico no mapa genético previamente obtido usando esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

A produção e uso das sondas derivadas de gene de planta para o uso na cópia do mapa genético é descrito em Bernatzky e Tanksley (1986) Plant Mol. Biol. Reporter 4: 37-41. As publicações numerosas descrevem a cópia do mapa genético dos clones de cDNA específico usando a metodologia resumida acima ou variações destas. Por exemplo, as populações inter-cruzadas F2, populações de suporte cruzado, populações unidas aleatoriamente, linhas isogênicas próximas, e outras séries de indivíduos podem ser usadas para o mapeamento. Tais metodologias são bem conhecidas aqueles habilitados na técnica.

As sondas de ácido nucleico também podem ser usadas para mapeamento físico (isto é, disposição das seqüências em mapas físicos; ver Hoheisel et al. In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319-346, e referências citadas neste).

Em uma outra forma de realização, as sondas de ácido nucleico pode ser usadas em fluorescência direta em mapeamento de hibridização in situ (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora os métodos correntes do mapeamento FISH favorece o uso de clones amplos (diversos kb por diversos cem kb; ver Laan et al. (1995) Genome Res. 5:13-20), melhora na sensibilidade que pode permitir o desempenho do mapeamento FISH usando sondas mais curtas.

Uma variedade de métodos com base na amplificação de ácido nucleico para o mapeamento genético e físico pode ser realizado usando os ácidos nucleicos. Exemplos incluem amplificação específica por alelo (Kazazian (1989) J. Lab. Clin. Med 11:95-96), poliformismo de fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332), ligação específica por alelo (Landegren et al. (1988) Science 241:1077-1080), reações de extensão por nucleotídeo (Sokolov (1990) Nucl. Acids Res.

18:3671), mapeamento híbrido de radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e mapeamento adequado (Dear e Cook (1989) Nucl. Acids Res. 17:6795-6807). Para estes métodos, a seqüência de um ácido nucleico é usado para designar e produzir pares iniciadores para o uso na reação de amplificação ou nas reações de extensão iniciadoras. O desenho de tais iniciadores é bem conhecidos por aqueles habilitados na técnica. Em métodos utilizados a cópia de mapa genético com base em PCR, pode ser necessário para identificar as diferenças na seqüência de DNA entre a origem do mapeamento cruzado na região que corresponde ao instante da seqüência de ácido nucleico. Este, entretanto, não é geralmente necessário para os métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas que tem a produção intensificada que diz respeito a traços, como descrito nestes acima. Estes traços também podem ser combinados com outros traços vantajosos economicamente, tal como traços de intensificação de produção adicional, tolerância por outros tensões abióticas e bióticas, características arquitetural de vários traços de modificação e/ou características bioquímicas e/ou fisiológicas.

Descrição das figuras

ERLK

A Figura 1 dá um resumo do grupo das proteínas de quinase receptoras, classificadas de acordo com a sua região extracelular (Shiu e Bleecker, 2001). A linha vertical marcada como TM indica o domínio de transmembrana. Na esquerda, o nome do local ou nome MAtDB são fornecidos de proteínas representativas. Os repousos RLCK para a quinase citoplásmica semelhante ao receptor, o repouso RLK para quinase semelhante ao receptor. Os nomes dos domínios são dados de acordo com a base de dados SMART e Pfam.

A Figura 2 mostra a organização de domínio da proteína

ERLK usada na presente invenção (SEQ ID N°: 2): evidentemente indicada: domínio de complexidade inferior, sublinhado: domínio de transmembrana, sublinhados itálicos: domínio de quinase. A análise foi completa com SMART.

A Figura 3 dá um alinhamento múltiplo das proteínas listadas como SEQ ID N°: 2, SEQ ID N°: 14, SEQ ID N°: 16, SEQ ID N°: 18, SEQ ID N°: 20, SEQ ID N°: 22, SEQ ID N°: 24, SEQ ID N°: 26, SEQ ID N°: 28, SEQ ID N°: 30, SEQ ID N°: 34, SEQ ID N°: 38, SEQ ID N°: 40, SEQ ID N°: 42, SEQ ID N°: 44, SEQ ID N°: 46, SEQ ID N°: 48, SEQ ID N°: 50, SEQ ID N°:52, SEQ ID N°: 54 e SEQ ID N°: 56. O alinhamento foi feito com CLUSTAL W (1.83), matriz do peso: BLOSUM, penalidade aberta na fenda: 11, penalidade da extensão da fenda: 1.

A Figura 4 mostra um mapa do plasmídio binário p030, usado para aumentar a expressão em *Oryza sativa* de um *Arabidopsis* ERLK que codifica o ácido nucleico sob o controle de um promotor GOS2 (SEQ ID N°: 58).

FBXW

A Figura 5 é uma apresentação esquemática da estrutura dos polipeptídeos FBXW em plantas. A posição relativa das características diferentes é mostrada: o F-box (PFAM PF00646), o domínio WD40 (com sete indivíduos WD40 que repetem como em PFAM PF00400), e motivos 1 a 5 como representado respectivamente pela SEQ ID N°: 97 a 101.

A Figura 6 mostra uma alinhamento de seqüência múltipla de polipeptídeos FBXW de plantas usando CLUSTAL W (1.83) (em serviço GenomeNet no Kyoto University Bioinformatics Center), e valores defeituosos (Blosum 62 como matriz do peso, penalidade de abertura da fenda de 10; penalidade da extensão da fenda de 0,05). A caixa F e o WD40 repetidos são empacotados. Motivo 1 e Motivo 2 são ambos marcados por um suporte ondulado. Motivos 3, 4 e 5 são sublinhado por uma caixa preta.

A Figura 7 mostra um vetor binário p1017, para expressão aumentada em *Oryza sativa* de um ácido nucleico *Arabidopsis thaliana* codificando um polipeptídeo de FBXW sob o controle de um promotor GOS2 (SEQ ID N°: 58).

RANBP

A Figura 8 mostra um alinhamento de polipeptídeos RANBP como definido neste acima. As seqüências foram alinhadas usando um programa alignX do conjunto do vetor NTI (InforMax, Bethesda, MD). O alinhamento múltiplo foi completo com a penalidade aberta na fenda de 10 e uma extensão da fenda de 0,01. A edição do manual menor também foi realizado onde necessário à melhor posição de algumas regiões conservadas. Motivo II, III, IV, V, VI e VII são indicados.

A Figura 9 mostra um vetor binário p072, para expressão aumentada em *Oryza sativa* de um *Zea Mays* RANBP que codifica o ácido nucleico sob o controle de um promotor de prolamina (referência PRO0090 interna).

A Figura 10 mostra um vetor binário p074, para expressão aumentada em *Oryza sativa* de um *Arabidopsis thaliana* RANBP que codifica o ácido nucleico sob o controle de um promotor de prolamina (referência PRO0090 interna).

GLK

A Figura 11 dá um resumo gráfico de genes de milho e arroz GLK (Rossini et al., 2001). As linhas horizontais representam as regiões transcritas não traduzidas (UTRs). As caixas representam diferentes domínios da região codificadora como indicado. NLS é o sinal de localização nuclear predito, DBD é o domínio de ligação de DNA putativo, que compreende o domínio GARP. Os triângulos brancos indicam a posição dos íntrons presentes em todos os quatro genes; os triângulos escuros indicam a posição de íntron que não é observado no gene G2. Nota-se que embora o OsGLK1 é

predito para ter uma localização nuclear, não foi possível predizer com alta confidência da presença de uma seqüência NLS.

A Figura 12 mostra a organização de domínio da proteína GLK usada na presente invenção (SEQ ID N°: 157). O domínio GARP (evidente) tem algumas semelhanças ao domínio MYB (indicado em itálico, como identificado por SMART); o domínio GCT está sublinhado.

A Figura 13 dá um alinhamento múltiplo das proteínas listadas como SEQ ID N°: 157 (OsGLK1), SEQ ID N°: 169 (OsGLK2), SEQ ID N°: 171 (AtGLK1), SEQ ID N°: 173 (AtGLK2), SEQ ID N°: 175 (PpGLK1), SEQ ID N°: 177 (PpGLK2), SEQ ID N°: 179 (ZmGLK1), SEQ ID N°: 181 (ZmG2), SEQ ID N°: 183 (TaGLK1), SEQ ID N°: 185 (AcGLK1), SEQ ID N°: 189 (SbGLK1), SEQ ID N°: 193 (OsGLK1var). O alinhamento foi feito com CLUSTAL W (1.83), matriz do peso: BLOSUM, penalidade aberta na fenda: 10, penalidade da extensão da fenda: 0,05.

A Figura 14 mostra um mapa do plasmídio binário p045, usado para aumentar a expressão em *Oryza sativa* de um *Oryza sativa* GLK que codifica o ácido nucleico sob o controle de um promotor GOS2 (SEQ ID N°: 58).

REV ΔHDZip/START

A Figura 15 mostra um filograma de classe de polipeptídeos III HDZip. Estes são dois polipeptídeos REV de mocotiledônias (*Oryza sativa*) que agrupam com um simples polipeptídeo REV de dicotiledônia (*Arabidopsis thaliana*). O círculo indica o clado de polipeptídeos REV de que a seqüência de ácido nucleicos REV ΔHDZip/START pode ser útil no desempenho dos métodos da invenção. Após Floyd et al. (2006) Genetics 173(1): 373-88

A Figura 16. A produção de árvore de ligação próxima após um alinhamento da seqüência múltipla de todas as classes III de polipeptídeos HDZip de Arabidopsis *thaliana* (5 no total) e *Oryza sativa* (5 no total),

incluindo exemplos de polipeptídeo REV ortólogos e parálogos (ver Exemplo 27), usando CLUSTAL W (1.83) (em serviço GenomeNet no Kyoto University Bioinformatics Center), e valores defeituosos (Blosum 62 como matriz do peso, penalidade de abertura de fenda de 10; penalidade da extensão da fenda de 0,05). Os polipeptídeos da ramificação REV são indicados pelo suporte ondulado, e são separados de outras quatro classes III de polipeptídeos HDZip pela linha evidente. O círculo indica o ponto de ramificação REV.

A Figura 17 é uma representação esquemática de um polipeptídeo VER de comprimento total. Os polipeptídeos REV que compreende a partir de um Terminal N ao terminal C: (i) um domínio (HD) homeodomínio, para ligação de DNA; (ii) um zíper de leucina (Zip), para interação de proteína-proteína; (iii) um domínio START para ligação de lipídeo/esterol (compreende um local de ligação complementar miRNA, mir165/166), e (iv) uma região terminal C (CTR), de função não definida. Por exemplo, é um polipeptídeo REV de *Oryza sativa* como representado pela SEQ ID N°: 199, os aminoácidos HD expandidos de 27 a 87, o zíper de aminoácidos de leucina 91 a 127, o domínio de aminoácidos START 166 a 376 e os aminoácidos CTR 377 a 840.

A Figura 18 é um alinhamento da seqüência múltipla polipeptídeos REV de comprimento total de que a seqüências de ácidos nucleicos REV ΔHDZip/START são úteis no desempenho dos métodos de acordo com a invenção, usando CLUSTAL W (1.83) (em serviço GenomeNet no Kyoto University Bioinformatics Center), e valores defeituosos (Blosum 62 como matriz do peso, penalidade de abertura da fenda de 10; penalidade da extensão da fenda de 0,05). O homeodomínio, o zíper de leucina, o domínio START e o CTR são pesadamente encaixotados.

A Figura 19 é um alinhamento da seqüência múltipla dos CTR de polipeptídeos REV (ambos de comprimento total e polipeptídeos parciais,

como listado no exemplo 27), de que a seqüência de ácido nucleico REV ΔHDZip/STARTs é útil no desempenho dos métodos de acordo com a invenção, usando CLUSTAL W (1.83) (em serviço GenomeNet no Kyoto University Bioinformatics Center), e valores defeituosos (Blosum 62 como matriz do peso, penalidade de abertura de fenda de 10; penalidade da extensão da fenda de 0,05.

A Figura 20 representa os vetores binários p0443 e p0448 para a redução de uma expressão do gene REV endógeno em *Oryza sativa*, usando respectivamente a seqüência de ácido nucleico REV ΔHDZip/START como representado pela SEQ ID N°: 194 (codificando um CTR parcial de um polipeptídeo REV) e a seqüência de ácido nucleico REV como representado pela SEQ ID N°: 198 (que codifica o polipeptídeo VER inteiro). A seqüência de ácido nucleico é clonada como repetição invertida separada por uma região não codificadora (existe uma região de ligação de matriz parcial (MAR) de tabaco) com a mira de obter um mRNA com uma conformação de grampo de cabelo. Um promotor constitutivo (PRO0129) controla a expressão de ambas seqüências de ácidos nucleicos SEQ ID N°: 194 e SEQ ID N°: 198 nos dois plasmídios diferentes.

CLE

A Figura 21 mostra uma organização de domínio dos polipeptídeo semelhante a CLE da SEQ ID N°: 233: em itálicos: seqüência de sinal, o resíduo Arg conservado necessário para o procedimento proteolítico é evidentemente indicado (aqui Arg73), o domínio CLE está sublinhado.

A Figura 22. dá um alinhamento múltiplo de polipeptídeos semelhantes a CLE: um simples alinhamento de seqüência não abaixo indica a menos a substituição conservada, um cólon indica uma substituição conservada, um asterisco indica um resíduo conservado em todas as seqüências.

A Figura 23 representa o vetor binário p068 por silenciamento

do gene endógeno em *Oryza sativa*, preferencialmente usando a seqüência de ácido nucleico codificando um polipeptídeo semelhante a CLE, ou uma parte deste, usando um construto de grampo de cabelo sob o controle de um promotor específico de endosperma (PRO90).

SYR NUE

A Figura 24 dá um resumo dos motivos conservados presentes na SEQ ID N°: 252. O domínio rico em leucina está sublinhado, e os motivos conservados 1, 2 e 3 são evidentemente indicados e uma seqüência em itálico representa o local de glicosilação N putativa com o local de fosforilação C de quinase de proteína putativa.

A Figura 25 mostra um alinhamento múltiplo de várias proteínas SYR. Os asteriscos indicam os resíduos de aminoácido idênticos, os cólons representam as substituições altamente conservadas e os pontos que representam menos substituições conservadas. Com a informação da Figura 1, os vários domínios e motivos conservados na SEQ ID N°: 252 podem ser facilmente identificados nas outras proteínas SYR.

A Figura 26 mostra o vetor binário pGOS2::SYR para a transformação e a expressão em *Oryza sativa* de um ácido nucleico *Oryza sativa SYR* sob o controle de um promotor de arroz GOS2.

A Figura 27 detalha os exemplos das seqüências úteis no desempenho dos métodos de acordo com a presente invenção, ou úteis no isolamento de tais seqüências. As seqüências podem resultar a partir de assembléias públicas EST, com menos sequenciamento de qualidade. Como uma conseqüência, um pouco de substituições de ácido nucleico podem ser esperadas. Tanto a extremidade 5' quanto a extremidade 3' UTR também podem ser usados para a realização dos métodos da invenção. SEQ ID N°: 1 a SEQ ID N°: 58 relativo ao ERLK; SEQ ID N°: 58 a SEQ ID N°: 112 diz respeito a FBXWD40; SEQ ID N°: 113 a SEQ ID N°: 155 diz respeito a RANBP; SEQ ID N°: 156 a SEQ ID N°: 193 e SEQ ID N°: 58 diz respeito a

GLK; SEQ ID N°: 194 a SEQ ID N°: 231 e SEQ ID N°: 58 diz respeito a REV ΔHDZip/START; SEQ ID N°: 232 a SEQ ID N°: 250 diz respeito a CLE; SEQ ID N°: 58 e SEQ ID N°: 251 a SEQ ID N°: 292 diz respeito a SYR. SEQ ID N°: 276 representa a seqüência da proteína ARGOS (acessão GenBank AY305869).

Exemplos

A presente invenção será descrita com referência aos seguintes exemplos, que são pela maneira de ilustração sozinha. Os seguintes exemplos não são pretendidos definir completamente ou de outra amaneira limitar o escopo da invenção.

Manipulação de DNA: a não ser de outra maneira estado, as técnicas de DNA recombinantes são realizadas de acordo com os protocolos padrões descritos em (Sambrook (2001) Molecular Cloning: a laboratory manual, 3° Edition Cold Spring Harbor Laboratory Press, CSH, New York) ou em Volumes 1 e 2 de Ausubel et al. (1994), Current Protocols in Molecular Biology, Current Protocols. Materiais padrões e métodos para trabalhar a planta molecular são descritos em Plant Molecular Biology Labfax (1993) por R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (UK) e Blackwell Scientific Publications (UK).

Exemplo 1: Identificação de homólogos da proteína ERLK da SEQ ID N°: 2 em Arabidopsis, arroz e outras espécies de plantas.

As seqüências (cDNA de comprimento total, ESTs ou genômico) que diz respeito à seqüência de ácido nucleico usado os métodos da presente invenção foram identificados entre aqueles mantidos na base de dados de nucleotídeos Entrez no National Center for Biotechnology Information usando ferramentas de busca da seqüência de base de dados, tal como uma ferramenta de alinhamento de local básico (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucl. Acids Res. 25:3389-3402). Este programa é tipicamente usado para observar regiões de

similaridade local entre seqüências pela comparação com as seqüências de ácido nucleico ou polipeptídeo pela base de dados da seqüência e pelo cálculo da significância estatística dos pontos. Os polipeptídeo codificados pelo ácido nucleico da presente invenção como usado com o algoritmo TBLASTN, com falta de direção e o filtro para ignorar as seqüências de complexidade inferiores foi ajustado. A produção de uma análise foi vista pela comparação em forma de par, e classificada de acordo com a contagem de probabilidade (valor E), onde a contagem reflete a probabilidade que um alinhamento particular ocorra pela escolha (o valor E inferior, o mais significante do acerto). Além disso a comparação aos valores E também foram contadas pela identidade de porcentagem. A identidade de porcentagem refere-se à diversos nucleotídeos idênticos (ou aminoácidos) entre as seqüências de dois ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. Em alguns exemplos, os parâmetros padrões podem ser ajustados para modificar a severidade da busca.

Seqüências de arroz e seqüências EST de várias espécies de plantas também pode ser obtidas a partir de outras bases de dados, tal como KOME (Knowledge based Oryza Molecular biological Encyclopedia; Kikuchi et al., Science 301, 376-379, 2003), Sputnik (Rudd, S., Nucl. Acids Res., 33: D622 - D627, 2005) ou a base de dados ortólogos do gene eucariótico (EGO, hospedado pelo Institute for Genomic Research). Estas bases de dados são buscáveis com a ferramenta BLAST. SEQ ID N°: 11 a SEQ ID N°: 56 são ácidos nucleicos e a seqüência da proteínas de homólogos da SEQ ID N°: 2 e foram obtidas a partir da base de dados mencionadas acima usando a SEQ ID N°: 2 como uma seqüência em questão.

Tabela A: Seqüência de ácido nucleico que diz respeito à seqüência de ácido nucleico (SEQ ID N°: 1) úteis nos métodos da presente invenção, e o polipeptídeos correspondentes deduzidos.

Fonte da planta	Ácido nucleico da SEQ ID N°:	Proteína da SEQ ID N°:
Arabidopsis thaliana ERLK	1	2
Arabidopsis thaliana AAC64891	11	12
Arabidopsis thaliana NM_104355.2	13	14
Arabidopsis thaliana At3g58690	15	16
Arabidopsis thaliana At3g58690	17	18
Arabidopsis thaliana At4g02010	19	20
Arabidopsis thaliana At5g56890	21	22
Arabidopsis thaliana AT2G20300	23	24
Oryza sativa Osi093820.1	25	26
Oryza sativa Osi015947.1	27	28
Oryza sativa Osi003977.2	29	30
Oryza sativa Osi003977	31	32
Oryza sativa Osi000040.5	33	34
Oryza sativa Ozsa8316	35	36
Oryza sativa Osi001397.3	37	38
Oryza sativa Osi000142.1	39	40
Oryza sativa Osi009054.1	41	42
Oryza sativa Osi001078.4	43	44
Saccharum officinarum TC15181	45	46
Glycine max TC196912	47	48
Solanum tuberosum TC81885	49	50
Nicotiana tabacum coi2	51	52
Medicago truncatula ABE82646.1	53	54
Populus sp TC25047	55	56

Exemplo 2: Determinação da similaridade e identidade globais entre os domínios quinase das proteínas ERLK.

As porcentagens de similaridade e identidade entre os domínios quinase das proteínas ERLK foram determinados usando MatGAT (Ferramenta de alinhamento global da matriz) software (BMC Bioinformatics. 2003 4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando proteína ou seqüências de DNA. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). Software MatGAT que gera matrizes de similaridade/identidade por dNA ou seqüência da proteína sem necessidade do pré-alinhamento dos dados. Os programas realizam uma série de alinhamentos em formas de pares usando o algoritmo Myers e Miller do alinhamento global (com a penalidade aberta na fenda de 12, e a penalidade da extensão da fenda de 2), cálculos de similaridade e identidade usando por exemplo Blosum 62 (para polipeptídeos), e então colocando o resultado na matriz de distância. A similaridade de seqüência é

mostrada na metade funda da linha que divide e identidade de seqüência é mostrada na metade superior da linha que divide diagonal. A seqüência da SEQ ID N°: 2 é indicada como número 1 ma matriz.

Os resultados da análise do software são mostrados na Tabela C para similaridade e identidade globais nos domínios quinase dos polipeptídeos ERLK. Os domínios quinase foram delineados usando uma ferramenta SMART e as seqüências obtidas são listadas na Tabela B. A identidade de porcentagem é dado acima, a diagonal (em negrito) e a similaridade da porcentagem é dada abaixo, a diagonal (fonte normal). A identidade de porcentagem entre domínios quinase de parálogos e ortólogos ERLK da faixa CDS845 (SEQ ID N°: 2) entre 30% e 68,8%.

Tabela B: Seqüência dos domínios quinase como obtido na análise com SMART e usado na análise MATGAT:

CDS0845	FSEEKKIGNGDVYKGVLSDGTVAAIKKLHMFNDNASNQKHEERSFRLV QRSTSRLQCPYLVELLGYCADQNHRILIYEFMPNGTVEHHLHDHNFKNL KDRPQPLDWGARLRIALDCARALEFLHENTISTVIHRNFKCTNILLDQNN RAKVSDFGLAKTGSDKLNGEISTRVIGTTGYLAPEYASTGKLTTKSDVYS YGIVLLQLLTGRTPIDSRRPRGQDVLVSWALPRLTNREKISEMVDPTMK GQYSQKDLIQVAAIAAVCVQPEASYRPLMTDVVHSL
NP_175879	FSEEKKIGNGDVYKGVLSDGTVAAIKKLHMFNDNASNQKHEERSFRLEV DLLSRLQCPYLVELLGYCADQNHRILIYEFMPNGTVEHHLHDHNFKNLK DRPQPLDWGARLRIALDCARALEFLHENTISTVIHRNFKCTNILLDQNNR AKVSDFGLAKTGSDKLNGEISTRVIGTTGYLAPEYASTGKLTTKSDVYSY GIVLLQLLTGRTPIDSRRPRGQDVLVSWALPRLTNREKISEMVDPTMKG QYSQKDLIQVAAIAAVCVQPEASYRPLMTDVVHSL
AT3G58690	FSKSNVVGNGGFGLVYRGVLNDGRKVAIKLMDHAGKQGEEEFKMEVE LLSRLRSPYLLALLGYCSDNSHKLLVYEFMANGGLQEHLYLPNRSGSVP PRLDWETRMRIAVEAAKGLEYLHEQVSPPVIHRDFKSSNILLDRNFNAK VSDFGLAKVGSDKAGGHVSTRVLGTQGYVAPEYALTGHLTTKSDVYSY GVVLLELLTGRVPVDMKRATGEGVLVSWALPQLADRDKVVDIMDPTLE GQYSTKEVVQVAAIAAMCVQAEADYRPLMADVVQSL
At3g58690s	FSKSNVVGNGGFGLVYRGVLNDGRKVAIKLMDHAGKQGEEEFKMEVE LLSRLRSPYLLALLGYCSDNSHKLLVYEFMANGGLQEHLYLPNRSGSVP PRLDWETRMRIAVEAAKGLEYLHEQVSPPVIHRDFKSSNILLDRNFNAK VSDFGLAKVGSDKAGGHVSTRVLGTQGYVAPEYALTGHLTTKSDVYSY GVVLLELLTGRVPVDMKRATGEGVLVSWALPQLADRDKVVDIMDPTLE GQYSTKEVVQVAAIAAMCVQAEADYRPLMADVVQSL
AT4G02010	FESASILGEGGFGKVYRGILADGTAVAIKKLTSGGPQGDKEFQVEIDMLS RLHHRNLVKLVGYYSSRDSSQHLLCYELVPNGSLEAWLHGPLGLNCPL DWDTRMKIALDAARGLAYLHEDSQPSVIHRDFKASNILLENNFNAKVA DFGLAKQAPEGRGNHLSTRVMGTFGYVAPEYAMTGHLLVKSDVYSYG VVLLELLTGRKPVDMSQPSGQENLVTWTRPVLRDKDRLEELVDSRLEG KYPKEDFIRVCTIAAACVAPEASQRPTMGEVVQSL
At5g56890	FDESRVLGEGGFGRVYEGVFDDGTKVAVKVLKRDDQQGSREFLAEVEM

	LSRLHHRNLVNLIGICIEDRNRSLVYELIPNGSVESHLHGIDKASSPLDWD ARLKIALGAARGLAYLHEDSSPRVIHRDFKSSNILLENDFTPKVSDFGLA RNALDDEDNRHISTRVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLE LLTGRKPVDMSQPPGQENLVSWTRPFLTSAEGLAAIIDQSLGPEISFDSIA KVAAIASMCVQPEVSHRPFMGEVVQAL
AT2G20300	FSAKRVLGEGGFGRVYQGSMEDGTEVAVKLLTRDNQNRDREFIAEVEM LSRLHHRNLVKLIGICIEGRTRCLIYELVHNGSVESHLHEGTLDWDARLK IALGAARGLAYLHEDSNPRVIHRDFKASNVLLEDDFTPKVSDFGLAREA TEGSQHISTRVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLELLTGR RPVDMSQPSGEENLVTWARPLLANREGLEQLVDPALAGTYNFDDMAK VAAIASMCVHQEVSHRPFMGEVVQAL
Osi015947	FSECNVVGRGAYGVVFRGRLGDGTTAAIKRLKMDGRREGEREFRIEMG VAITAQVDLLSRMHSPYLVGLLGYCADQSHRLLVFEFMPNGSLKSHLHR RALAPAEQPPPLDWQTRLGIALDCARALEFLHEHSSPAVIHRDFKCSNIL LDHNYRARVSDFGMAKLGSNKANGQVAAITAMCIQTKADYRPLMTDV VQSL
Osi003977	FGRAHVVGQGSFGAVYRGVLPDGRKVAVKLMDRPGKQGEEEFEMEVE LLSRLRSPYLLGLIGHCSEGGHRLLVYEFMANGGLQEHLYPNGAFEKIET FSIYLVKQRPIFDKNIGHPICRRAHFSRQLISHKADIVGSCGGISKLDWPTR MRIALEAAKGLEYLHERVNPPVIHRDFKSSNILLDKDFRARVSDFGLAKL GSDRAGGHVSTRVLGTQGYVAPDYGVVLLELLTGRVPVDMKRPPGEGV LVNWALPMLTDREKVVQILDPALEGQYSLKDAVQVAAIAAMCVQQEA DYRPLMADVVQSL
Osi000040a	FDPSSMLGEGGFGRVFKGVLTDGTAVAIKKLTSGGHQGDKEFLVEVEM LSRLHHRNLVKLIGYYSNRTLGASRPLDWDTRMRIALDAARGLAYLHE DSQPCVIHRDFKASNILLEDDFHAKVSDFGLAKQAPEGRTNYLSTRVMG TFGYVAPEYAMTGHLLVKSDVYSYGVVLLELLTGRRPVDMSQPSGQEN LVTWLTQMFPDLSTKSLVSHQPLLAKLSGVEILICSILTTQARPILRDKDT LEELADPKLGGQYPKDDFVRVCTIAAACVSPEASQRPTMGEVVQSL
Osi000040b	FSNDLAIGVGGFGVVYRGVVDGDVKVAVKRSNPSSEQGITEFQTEVEML SKLRHRHLVSLIGFCEEDGEMVLVYDYMEHGTLREHLYHNGGKPTLSW RHRLDICIGAARGLHYLHTGESHVSTVVKGSFGYLDPEYYRRQQLTDKS DVYSFGVVLFEVLMARPALDPALPRDQVSLADYALACKRGGALPDVVD PAIRDQIAPECLAKFADTAEKCLSENGTERPTMGDVLWNL
Osi001397	FDNSRIIGEGGFGRVYEGILEDGERVAVKILKRDDQQGTREFLAEVEMLS RLHHRNLVKLIGICTEEHIRCLVYELVPNGSVESHLHGSDKGTAPLYWD ARLKIALGAARALAYLHEDSSPRVIHRDFKSSNILLEHDFTPKVSDFGLA RTAIGEGNEHISTRVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLELL TGRKPVDILRPPGQENLVAWACPFLTSRDGLETIIDPSLGNSILFDSIAKV AAIASMCVQPEVDQRPFMGEVVQAL
Osi000142	FDDSTVLGEGGFGCVYQGTLEDGTRVAVKVLKRYDGQGEREFLAEVEM LGRLHHRNLVKLLGICVEENARCLVYELIPNGSVESHLHGVDLETAPLD WNARMKIALGAARALAYLHEDSSPCVIHRDFKSSNILLEHDFTPKVSDF GLARTARGEGNQHISTRVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVL LELLTGRKPVDMSRPGGQENLVSWARPLLTNVVSLRQAVDPLLGPNVPL DNVAKAAAIASMCVQPEVAHRPSMGEVVQAL
Osi009054	FDSKRVLGQGGFGRVYHGTMDGGDEIAVKLLTREDRSGDREFIAEVEM LSRLHHRNLVKLIGICIEHNKRCLVYELIRNGSVESHLHGADKAKGMLN WDVRMKIALGAARGLAYLHEDSNPHVIHRDFKGSNILLEEDFTPKVTDF GLAREATNGIQPISTRVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLE LLSGRKPVCMSDTNGPQNLVTWARPLLCHKEGLERLIDPSLNGNFNFDD VAKVASIASMCVHNDPSQRPFMGEVVQAL
Osi001078	FSFNKIIGEGGYGRVYRGTIDDEVDVAVKLLTRKHQNRDREFIAEVEMLS RLHHRNLVKLIGICIERSTRCLVFELVPNGSVESHLHGSDKIYGPLDFDTR MKIALGAARGLAYLHEDANPHVIHRDFKASNVLLENDFTPKVADFGLA KEASEGMDHISTQVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLELL

	SGRKPVDMTQPPGSENLVTWARPLLTDRDGLQQLVDPSMPAASYGFEK LAKAAAIASMCVHVEASHRPFMGEVVQAL
TC15181	FGRAHMVGQGSFGAVYRGVLPDGRKVAVKLMDRPGKQGEEEFEMEVE LLSRLRSPYLLGLIGHCSEGGHRLLVYEFMANGGLQEHLYPNRGSCGGIS KLDWDTRMRIALEAAKGLEYLHERVNPPVIHRDFKSSNILLDKDFHARV SDFGLTKLGSDRAGGHVSTRVLGTQGYVAPEYALTGHLTTKSDVYSYG VVLLELLTGRVPVDMKRPPGEGVLVNWALPMLTDREKVVRILDPALEG QYSLKDAVQVAAIAAMCVQPEADYRPLMADVVQSL
TC196912	SKSNVIGHGGFGLVYRGVLNDGRKVAIKFMDQAGKQGEEEFKVEVELL SRLHSPYLLALLGYCSDSNHKLLVYEFMANGGLQEHLYPVSNSIITPVKL DWETRLRIALEAAKGLEYLHEHVSPPVIHRDFKSSNILLDKKFHAKVSDF GLAKLGPDRAGGHVSTRVLGTQGYVAPEYALTGHLTTKSDVYSYGVVL LELLTGRVPVDMKRPPGEGVLVSWALPLLTDREKVVKIMDPSLEGQYS MKEVVQVAAIAAMCVQPEADYRPLMADVVQSL
TC81885	EEEFKVEVELLCRLRSPYLLSLIGYCSESSHKLLVYEFMANGGLQEHLYPI KGSNNCCPKLDWKTRLRIALEAAKGLEYLHEHVNPPIIHRDLKSSNILLD KNFHAKVSDFGLAKLGSDKAGGHVSTRVLGTQGYVAPEYALTGHLTTK SDVYSYGVVLLELLTGRVPVDMKRSPGEGVLVSWALPRLTDREKVVEI MDPALEGQYSMKEVIQVAAIAAMCVQPEADYRPLMADVVQSL
ABB36644	FSLKRVLGEGGFGRVYHGILEDRTEVAVKVLTRDNQNGDREFIAEVEML SRLHHRNLVKLIGICSEERTRSLVYELVRNGSVESHLHGRDGRKEPLDW DVRLKIALGAARGLAYLHEDSNPRVIHRDFKASNVLLEDDFTPKVADFG LAREATEGSHHISTRVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLE LLSGRKPVDMSQPPGEENLVTWARPLLTTREGLEQLVDPSLAGSYDFDD MAKVAAIASMCVHPEVTQRPFMGEVVQAL
ABE82646	MLSRLHHRNLVKLIGICIEGRRRCLVYELVPNGSVESHLHGDDKNRGPL DWEARMKIALGAARGLAYLHEDSNPRVIHRDFKASNVLLEDDFTPKVS DFGLAREATEGSNHISTRVMGTFGYVAPEYAMTGHLLVKSDVYSYGVV LLELLTGRKPVDMSQPQGQENLVTWARALLTSREGLEQLVDPSLAGGY NFDDMAKVAAIASMCVHSEVTQRPFMGEVVQAL
Popsp	VTWNEGYGVVYGGTLSDGTVAAIKMLHRAGKQGERAFREVDLLSRLH SPYLVELLGYCADQNHRLLVFEFMPNGTLQHHLHHKQYRPLDWGTRLR IALDCARALEFLHELTIPAVIHRDFKCSNILLDQNFRAKVSDFGSAKMGS ERINARNSMCLPSTTGYLAPEHASTGKLTTKSDVYSYGVVLLQLLTGRK PVDTKQPSGEHVLVSWALPRLTNRDKVVEMVDPAMQDQYSKKDLIQV AAIAAVCVQPEADYRPLMTDVVQSL

Tabela C:

Sim\Ident	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. CDS0845		98,2	57,5	57,5	47,7	48,3	47,7	39,2	44,8	40,7	30,0	47,7	48,8	42,0	45,7	54,7	58,6	51,8	47,0	42,9	68,8
2. NP_175879	98,2		58,9	58,9	48,8	49,3	48,8	41,0	46,0	41,6	31,0	48,8	49,8	43,0	46,7	56,1	60,0	53,2	48,1	43,3	70,6
3. At3g58690	74,8	76,6		100,0	56,3	54,0	52,3	40,8	65,6	47,5	35,4	53,8	54,2	48,7	51,1	81,0	87,0	73,6	52,7	44,0	61,6
4. At3g58690s	74,8	76,6	100,0		56,3	54,0	52,3	40,8	65,6	47,5	35,4	53,8	54,2	48,7	51,1	81,0	87,0	73,6	52,7	44,0	61,6
5. At4g02010	66,0	67,7	73,6	73,6		64,0	64,9	33,2	42,6	68,4	32,6	63,1	63,2	61,2	65,0	54,5	56,1	49,1	67,4	58,0	52,5
6. At5g56890	64,5	66,3	70,8	70,8	75,7		76,4	33,7	44,9	56,1	37,2	83,3	78,9	73,1	71,7	56,5	56,8	48,2	78,2	66,9	50,2
7. At2g20300	64,9	66,7	70,8	70,8	76,4	84,7		30,4	44,1	58,2	37,7	75,5	74,1	77,7	78,1	54,9	53,1	47,3	87,5	74,0	49,8
8. Osi015947	50,4	52,1	53,4	53,4	47,8	48,0	47,4		35,8	26,9	27,3	33,8	34,0	29,7	30,3	37,6	40,1	33,1	30,8	24,5	46,0
9. Osi003977	63,5	65,1	76,9	76,9	60,3	59,0	59,3	48,2		38,7	28,9	44,1	43,8	41,9	42,7	78,9	66,1	59,0	44,4	38,2	49,1
10. Osi000040a	58,6	60,3	64,1	64,1	82,4	67,9	70,0	39,3	56,7		29,4	53,7	54,3	53,7	52,9	48,4	48,1	40,8	57,2	48,9	45,0
11.Osi000040b	46,1	47,9	51,6	51,6	48,6	49,8	51,5	43,2	42,0	42,8		38,0	35,3	36,0	37,7	36,5	35,0	29,6	38,0	29,9	30,8
12. Osi001397	64,2	66,0	69,7	69,7	75,0	89,5	84,7	46,7	58,3	66,2	50,4		80,7	71,5	72,7	56,3	56,3	48,0	78,5	67,2	50,7
13. Osi000142	64,5	66,3	67,1	67,1	76,1	86,5	82,5	48,2	57,7	68,3	48,2	87,6		71,2	69,8	54,5	56,0	46,9	74,8	65,0	51,8
14. Osi009054	62,1	63,8	69,0	69,0	73,9	80,7	88,6	44,7	57,7	67,9	50,5	82,1	79,6		72,6	52,0	50,2	44,0	78,8	67,0	44,9
15. Osi001078	65,6	67,4	70,4	70,4	76,4	82,2	89,4	46,0	59,3	69,3	49,3	83,2	83,2	86,1		53,6	52,5	46,4	79,6	68,2	50,5
16. TC15181	71,3	73,0	91,0	91,0	71,8	69,7	70,0	50,9	83,7	63,8	49,1	69,7	68,6	67,5	69,7		82,3	74,4	56,0	47,5	60,3
17. TC196912	74,1	75,9	92,4	92,4	75,0	71,4	70,7	53,3	77,5	64,8	50,4	70,3	68,5	68,8	72,1	91,0		77,2	54,9	46,4	63,8
18. TC81885	66,0	67,7	80,9	80,9	63,8	60,7	63,4	49,8	67,1	54,5	48,1	60,6	59,5	60,4	62,8	79,8	81,5		48,0	51,9	57,0
19. ABB36644	66,0	67,7	70,0	70,0	77,9	85,8	91,6	46,2	59,3	71,0	51,6	85,8	81,8	87,9	88,3	70,8	70,7	61,9		74,4	50,2
20. ABE82646	59,2	59,2	59,2	59,2	65,9	72,0	78,4	43,4	50,2	58,6	48,7	71,5	69,3	74,0	74,8	59,6	59,4	67,8	77,7		43,9
21. Popsp	78,7	80,5	76,5	76,5	67,4	65,1	68,3	55,4	63,8	59,0	49,4	66,4	65,7	64,8	68,6	74,0	77,5	70,8	67,4	59,2	

Exemplo 3: Clonagem da seqüência de ácido nucleico usado os métodos da invenção

Manipulação de DNA: a não ser de outra maneira estado, as técnicas de DNA recombinantes são realizadas de acordo com os protocolos padrões descritos em (Sambrook (2001) Molecular Cloning: a laboratory manual, 3° Edition Cold Spring Harbor Laboratory Press, CSH, New York) ou no Volumes 1 e 2 de Ausubel et al. (1994), Current Protocols in Molecular Biology, Current Protocols. Materiais padrões e métodos para trabalhar a planta molecular são descritos em Plant Molecular Biology Labfax (1993) por R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (UK) e Blackwell Scientific Publications (UK).

A seqüência de ácido nucleico usado os métodos da invenção foi amplificado por PCR usando, como padrão, um biblioteca de cDNA de muda de Arabidopsis thaliana feita por sob encomenda (em pCMV Sport 6,0; Invitrogen, Paisley, UK). O PCR foi realizado usando polimerase de DNA Hifi Taq em condições padrões, usando 200 ng de padrão em uma mistura de 50 µl de PCR. Os iniciadores usados foram prm2500 (SEQ ID N°: 3; sentido, códon iniciado em negrito:

5' ggggacaagtttgtacaaaaaagcaggcttcacaatggaaaacaaaagccatagc 3')

e prm2501 (SEQ ID N°: 4; reverso, complementar,:

5' ggggaccactttgtacaagaaagctgggtaaacaaaagagtgtcatggca 3'),

que inclui os locais AttB pela recombinação Gateway. O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizado, durante pelo qual o fragmento PCR recombina in vivo com o plasmídio pDONR201 para fornecer, de acordo com a tecnologia Gateway, uma "clone de entrada", p031. O plasmídio pDONR201 foi adquirido de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 4: Construção do vetor de expressão

O clone de entrada p031 foi subseqüentemente usado em uma reação LR com p05050, um vetor de destinação usado pela transformação da Oryza sativa. Este vetor contém como elementos funcionais dentro dos limitadores de DNA T: um marcador de planta selecionável; um marcador de expressão de cassete avaliável; e um cassete Gateway pretendido pela recombinação LR in vivo com a seqüência de ácido nucleico de interesse já clonado no clone de entrada. Um promotor constitutivo não viral de arroz, o promotor GOS2 (SEQ ID N°: 58) foi localizado a montante deste cassete Gateway.

Após a etapa de recombinação LR, o vetor resultante da expressão p030 (Figura 4) foi transformado em cepa de Agrobacterium LBA4044 e subseqüentemente por plantas Oryza sativa.

Exemplo 5: Transformação da planta

Transformação do arroz

A Agrobacterium que contém um vetor de expressão como usado para transformar as plantas Oryza sativa. As sementes maduras secas do arroz japônico que cultivam Nipponbare foram descascadas. A esterilização foi realizada pela incubação por um minuto em 70% de etanol, seguido por 20 minutos em 0,2% de $HgCl_2$, seguido por 6 vezes de lavagem de 15 minutos com água destilada estéril. As sementes estéreis foram então germinadas em um meio contendo 2,4-D (meio de indução de calos). Após a incubação no escuro por quatro semanas, calos derivados de escutelo embriogênico foram taxados e propagados no mesmo meio. Após duas semanas, os calos foram multiplicados e propagados pela sub-cultura no mesmo meio por outras duas semanas. As porções de calos embriogênicos foram sub-culturados em meio fresco por 3 dias antes do co-cultivo (pela atividade de divisão celular auxiliar).

A Cepa Agrobacterium LBA4404 que contém um vetor de expressão como usado para co-cultivo. Agrobacterium foi inoculado em meio

AB com os antibióticos apropriados e cultivadas por 3 dias a 28°C. As bactérias foram então colhidas e recolocadas em suspensão em meio de co-cultivo líquido para uma densidade (OD_{600}) de cerca de 1. A suspensão foi então transferida a uma placa de Petri e o caldo colocado em imersão na suspensão por 15 minutos. O tecidos dos calos foram então manchados a seco em um filtro de papel e transferido solidificado, meio de co-cultivo e incubado por 3 dias no escuro a 25°C. Os calos co-cultivados foram desenvolvidos em meio contendo 2,4-D por 4 semanas no escuro a 28°C na presença de um agente de seleção. Durante este período, desenvolveram-se ilhas de calos resistentes que desenvolvem-se rapidamente. Após a transferência deste material para um meio de regeneração e incubação na luz, o potencial embriogênico foi liberado e brotos desenvolvidos nas próximas quatro ou cinco semanas. Os brotos foram taxados a partir do calo e incubado por 2 a 3 semanas em um meio contendo auxina que foram transferidos ao solo. Os brotos endurecidos foram desenvolvidos sob alta umidade e dias curtos em uma estufa para plantas.

Aproximadamente 35 transformantes de arroz T0 independentes foram gerados por um construto. Os transformantes primários foram transferidos a partir de uma câmara de cultura de tecido para uma estufa para plantas. Após uma análise quantitativa de PCR para verificar o número de cópias do suplemento de DNA T, apenas plantas transgênicas de cópia simples que exibem a tolerância pelo agente de seleção que foi mantido para endurecer a semente T1. As sementes foram então endurecidas três a cinco meses após a transplante. O método produzido de transformantes de local simples em uma taxa de durante 50% (Aldemita e Hodges1996, Chan et al. 1993, Hiei et al. 1994).

Transformação do milho

A Transformação do milho (*Zea mays*) é realizado com uma modificação do método descrito por Ishida et al. (1996) Nature Biotech 14(6):

745-50. A transformação é dependente do genótipo em milho e apenas genótipos específicos são responsáveis pela transformação e regeneração. A linha inerente A188 (University of Minnesota) ou híbridas com A188 como uma origem são boas fontes de material doador para transformação, mas outros genótipos podem ser usados sucessivamente bem como. As espigas são endurecidas a partir da planta de milho aproximadamente 11 dias após a polinização (DAP) então o comprimento do embrião imaturo é cerca de 1 a 1,2 mm. Os Embriões imaturos são co-cultivados com *Agrobacterium tumefaciens* que contém um vetor de expressão, e plantas transgênicas são recuperadas através de organogênese. Embriões taxados são desenvolvidos em meio de indução de calos, então o meio de regeneração do milho, contendo o agente de seleção (por exemplo imidazolinona mas vários marcadores de seleção podem ser usadas). As placas de Petri são incubadas na luz a 25 °C por 2 a 3 semanas, ou até os brotos desenvolverem. Os brotos verdes são transferidos de cada embrião para meio de enraizamento do milho e incubado a 25 °C por 2 a 3 semanas, até as raízes desenvolverem. Os brotos enraizados são transplantados ao solo na estufa para plantas. As sementes T1 são produzidas a partir de plantas que exibem tolerância pelo agente de seleção e que contém uma cópia simples do suplemento de DNA T.

Transformação do Trigo

A transformação do trigo é realizado com o método descrito por Ishida et al. (1996) Nature Biotech 14(6): 745-50. O cultivar Bobwhite (disponível de CIMMYT, Mexico) é comumente usado na transformação. Embriões imaturos são co-cultivados com *Agrobacterium tumefaciens* que contém um vetor de expressão, e plantas transgênicas são recuperadas através de organogênese. Após a incubação com *Agrobacterium*, os embriões são desenvolvidos in vitro em meio de indução de calos, então o meio de regeneração, contendo o agente de seleção (por exemplo imidazolinona mas vários marcadores de seleção podem ser usados). As placas de Petri são

incubadas na luz a 25 °C por 2 a 3 semanas, ou até os brotos desenvolverem. Os brotos verdes são transferidos de cada embrião por meio enraizado e incubado a 25 °C por 2 a 3 semanas, até as raízes desenvolverem. Os brotos enraizados são transplantados ao solo na estufa para plantas. As sementes T1 são produzidas a partir de plantas que exibem tolerância pelo agente de seleção e que contém uma cópia simples do suplemento de DNA T.

Transformação de soja

A soja é transformada de acordo com uma modificação do método descrito no Texas A&M Patente U.S. 5.164.310. Diversas variedades de soja comerciais são responsáveis pela transformação deste método. O cultivar Jack (disponível da Illinois Seed foundation) é comumente usado para transformação. As sementes de soja são esterilizada em semeadura in vitro. A hipocotila, a pequena raiz e cotiledônia são taxadas a partir de mudas jovens com sete dias de idade. A epicotila e a cotiledônia remanescente ainda são desenvolvidas por nódulos auxiliares de crescimento. Estes nódulos auxiliares são taxados e incubados com *Agrobacterium tumefaciens* que contém um vetor de expressão. Após o tratamento de co-cultivo, as mudas são lavadas e transferidas para o meio de seleção. As raízes regeneradas são taxadas e colocadas em um meio de prolongamento de broto. Nenhum dos brotos são mais longos do que 1 cm e são colocados em meio enraizado até as raízes desenvolverem. Os brotos enraizados são transplantados ao solo na estufa para plantas. As sementes T1 são produzidas a partir de plantas que exibem a tolerância pelo agente de seleção e que contém uma cópia simples do suplemento de DNA T.

Transformação de Colza/Canola

Os pecíolos e hipocotilas cotiledonárias de mudas jovens de 5 a 6 dias de idade são usados como mudas para a cultura do tecido e transformadas de acordo com Babic et al. (1998, Plants Cell Rep 17: 183-188). O cultivar Westar comercial (Agriculture Canada) é a variedade padrão

usada para transformação, mas outras variedades também podem ser usadas. As sementes de canola são esterilizadas na superfície para uma semeadura in vitro. As mudas de pecíolo cotiledônio com a ligação de cotiledônias são taxadas a partir das mudas in vitro, e inoculadas com *Agrobacterium* (que contém um vetor de expressão) pela imersão da extremidade cortada das mudas de pecíolos na suspensão bacteriana. As mudas são então cultivadas por 2 dias em meio MSBAP-3 contendo 3 mg/l BAP, 3% de sacarose, 0,7% de Phytagar a 23 °C, 16 horas de luz. Após dois dias de co-cultivo com *Agrobacterium*, as mudas de pecíolos são transferidos para o meio MSBAP-3 contendo 3 mg/l BAP, cefotaxima, carbenicilina, ou timentina (300 mg/l) por 7 dias, e então cultivadas em meio MSBAP-3 com cefotaxima, carbenicilina, ou timentina e agente de seleção até a regeneração dos brotos. Quando os brotos são 5 a 10 mm no comprimento, estes são cortados e transferidos para o meio de prolongamento de broto (MSBAP-0,5, contendo 0,5 mg/l de BAP). Os brotos de cerca de 2 cm no comprimento são transferidos para o meio enraizado (MS0) para a indução da raiz. Os brotos enraizados são transplantados ao solo na estufa para plantas. As sementes T1 são produzidas a partir de plantas que exibem a tolerância pelo agente de seleção e que contém uma cópia simples do suplemento de DNA T.

Transformação da Alfafa

Um clone de regeneração da alfafa (*Medicago sativa*) é transformado usando o método de (McKersie et al., 1999 Plant Physiol 119: 839–847). Regeneração e transformação da alfafa é dependente do genótipo e portanto uma planta de regeneração é requerida. Métodos para obter as plantas de regeneração tem sido descrito. Por exemplo, estes podem ser selecionados a partir do cultivar Rangelander (Agriculture Canada) ou qualquer outra variedade de alfafa comercial como descrito por Brown DCW e A Atanassov (1985. Culture the organ, tissue, plant cell 4: 111-112). Alternativamente, a variedade RA3 (University of Wisconsin) foi selecionada

para o uso em cultura de tecido (Walker et al., 1978 Am J Bot 65:654-659). As mudas de pecíolos são co-cultivados com uma cultura durante a noite de *Agrobacterium tumefaciens* C58C1 pMP90 (McKersie et al., 1999 Plant Physiol 119: 839–847) ou LBA4404 que contém um vetor de expressão. As mudas são co-cultivadas por 3 dias no escuro no meio de indução SH contendo 288 mg/ L Pro, 53 mg/ L de tioprolina, 4,35 g/ L K2SO4, e 100 µm de acetosiringinona. As mudas são lavadas em meio Murashige-Skoog com metade da força (Murashige e Skoog, 1962) e colocado no mesmo meio de indução SH sem acetosiringinona mas com um agente de seleção adequado e antibiótico adequado para inibir a *Agrobacterium* desenvolvida. Após diversas semanas, os embriões somáticos são transferidos para o meio de crescimento BOi2Y contendo nenhum dos reguladores desenvolvidos, nenhum antibiótico, e 50 g/ L de sacarose. Os embriões somáticos são subseqüentemente germinados em meio Murashige-Skoog com metade da força. As mudas enraizadas foram transplantadas em potes e desenvolvidas em uma estufa para plantas. As sementes T1 são produzidas a partir de plantas que exibem a tolerância pelo agente de seleção e que contém uma cópia simples do suplemento de DNA T.

Exemplo 6: Procedimento de avaliação

6.1 Apresentação da avaliação

Aproximadamente 30 transformantes de arroz T0 independentes foram gerados. Os transformantes primários foram transferidos a partir de uma câmara de cultura de tecido para uma estufa para plantas para crescimento e endurecimento da semente T1. Sete eventos, pelo qual a produto T1 segregado 3:1 para a presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 mudas de T1 contendo o transgene (hetero e homo zigotos) e aproximadamente 10 mudas de T1 que precisam do transgene (nulizigotos) foram selecionados pelo monitoramento da expressão do marcador visual. As plantas transgênicas e o

nulizigotos correspondente foram desenvolvidos lado a lado em posições aleatórias. As condições da estufa para plantas foram de dias curtos (12 horas de luz), 28°C na luz e 22°C no escuro, e uma umidade relativa de 70%.

Quatro eventos T1 ainda foram avaliados na geração T2 seguindo os mesmos procedimentos de avaliação como para a geração T1 mas com mais indivíduos por evento. A partir do estágio da semeadura até o estágio da maturidade as plantas foram passadas diversas vezes através de uma cabine formadora de imagem digital. Em cada um ponto de tempo as imagens digitais (2048 x 1536 pixéis, 16 milhões de cores) foram tirada de cada planta de pelo menos 6 ângulos diferentes.

6.2 Análise estatística: Teste t e Teste F

Um dos dois fatores ANOVA (análise de variantes) como usado como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um Teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformados com o gene da presente invenção. O teste F foi realizado por verificar por um efeito do gene em todos os eventos de transformação e para verificar por um efeito total do gene, também conhecido como um efeito do gene global. O princípio para a significância por um efeito de gene verdadeiro global foi apresentado a 5% do nível de probabilidade para o teste F. Um ponto de valor significante no teste F a um efeito do gene, significa que não é apenas como mera presença ou posição do gene que está causando as diferenças nos fenotipos.

Exemplo 7: Resultados da avaliação

As plantas da área acima do solo (ou biomassa da folha) foi determinada pela contagem do número total de pixéis nas imagens digitais das partes acima do solo das plantas discriminadas a partir do fundamento. Este valor foi medido para as figuras retiradas no mesmo ponto dos ângulos diferentes e foi convertido a um valor de superfície física expressada no quadrado mm pela calibração. Os experimentos mostram que as plantas da

área acima do solo medido desta maneira correlacionada com a biomassa das partes das plantas acima trituradas. A área triturada acima é o ponto de tempo em que as plantas tem atingido esta biomassa de folha máxima.

As panículas primárias maduras foram endurecidas, contadas, empacotadas, rotulado com código de barra e então secadas por três dias em um forno a 37°C. As panículas foram então refugada e todas as sementes foram colhidas e contadas. As cascas enchedoras foram separadas a partir de um despejo usando dispositivo de sopro de ar. As cascas despejadas foram descartadas e a fração remanescente foi contada novamente. A casca enchedora foi pesada em uma balança analítica. O número de sementes enchedoras foi determinada pela contagem do número de cascas enchedoras que permaneceram após a etapa de separação. A produção total da semente foi medida pela pesagem de todas as cascas enchedoras endurecidas a partir de uma planta.

Como apresentado nas Tabelas D a F, a biomassa acima do solo, a produção da semente, o número de sementes enchedoras são aumentadas nas plantas transgênicas com aumento da expressão de um ácido nucleico codificando uma proteína ERLK, comparada as planta adequadas de controle. Os resultados a partir da geração T1 e T2 foram mostradas.

Tabela D mostra um aumento na biomassa acima do solo em porcentagem, bem como a relevância estatística deste aumento de acordo com um teste F, na geração T1 e T2 de arroz transgênico com aumento da expressão de um ácido nucleico codificando uma proteína ERLK.

Tabela D:

Biomassa acima do solo		
	% de diferença	**Valor P do teste F**
Geração T1	11	0,0137
Geração T2	12	0,0229

Tabela E mostra um aumento na produção total da semente (peso da semente total) em porcentagem, bem como a relevância estatística

deste aumento de acordo com um teste F, na geração T1 e T2 de arroz transgênico com aumento expressão de um ácido nucleico codificando uma proteína ERLK.

Tabela E:

Produção da semente		
	% da diferença	**Valor P do teste F**
Geração T1	18	0,0253
Geração T2	22	0,0062

Tabela F mostra um aumento no número de sementes enchedoras em porcentagem, bem como a relevância estatística deste aumento de acordo com um teste F, na geração T1 e T2 de arroz transgênico com aumento expressão de um ácido nucleico codificando uma proteína ERLK.

Tabela F:

Número de sementes enchedoras		
	% da diferença	**Valor P do teste F**
Geração T1	15	0,0042
Geração T2	18	0,0019

FBXW

Exemplo 8: Identificação das seqüências que diz respeito à seqüência de ácido nucleico usados nos métodos da invenção

As seqüências de (cDNA de comprimento total, ESTs ou genômico) que diz respeito à seqüência de ácido nucleico usado os métodos da presente invenção foram identificadas entre aquelas mantidos na base de dados de nucleotídeos Entrez no National Center for Biotechnology Information usando ferramentas de busca da seqüência de base de dados, tal como a ferramenta de alinhamento de local básico (BLAST) (Altschul *et al.* (1990) J. Mol. Biol. 215:403-410; e Altschul *et al.* (1997) Nucl. Acids Res. 25:3389-3402). Os programas são usados para observar as regiões de similaridade local entre seqüências pela comparação com as seqüências de ácido nucleico ou polipeptídeo pela base de dados da seqüência e pelo cálculo a significância estatística dos pontos. Os polipeptídeos codificados pelo ácido

nucleico da presente invenção como usado pelo algoritmo TBLASTN, com falta de direção e o filtro para ignorar as seqüências de complexidade inferiores apresentadas. A produção de uma análise foi vista pela comparação em forma de par, e classificada de acordo com a contagem de probabilidade (Valor E), onde a contagem reflete a probabilidade que um alinhamento particular ocorra pela escolha (o valor E inferior, o mais significante do acerto). Além disso a comparação aos valores E também foram contadas pela identidade de porcentagem. Identidade de porcentagem refere-se à diversos nucleotídeos idênticos (ou aminoácidos) entre as seqüências de dois ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. Em alguns exemplos, os parâmetros padrões pode ser ajustados para modificar a severidade da busca

A tabela abaixo fornece uma lista da seqüência de ácido nucleico que diz respeito à seqüência de ácido nucleico útil no desempenho dos métodos da presente invenção.

Tabela G: Seqüência de ácido nucleico que diz respeito à seqüência de ácido nucleico (SEQ ID N°: 59) usado os métodos da presente invenção, e o correspondente aos polipeptídeos deduzidos

nome	Ácido nucleico SEQ ID N°	Polipeptídeo SEQ ID N°	comprimento da seqüência	Número de acessão NCBI	organismo da fonte
Arath_FBXW (At5g21040)	59	60	Comprimento total	NM_122112.2	*Arabidopsis thaliana*
Orysa_FBXW	61	62	Comprimento total	AK111585	*Oryza sativa*
Medtr_FBXW	63	64	Comprimento total	CR931734 (unidos)	*Medicago trunculata*
Triae_FBXW	65	66	Comprimento total	CJ661176.1 CB307121.1 CJ553648.1	*Triticum aestivum*
Poptr_FBXW	67	68	Comprimento total	Proprietário	*Populus tremuloides*

Zeama_FBXW	69	70	Comprimento total	AC183938.1 (unidos)	*Zea mays*
Vitvi_FBXW	71	72	Parcial (3')	CF210354 CF413646 CF213082	*Vitis vinifera*
Senca_FBXW	73	74	Parcial (5')	DY662683.1	*Senecio cambrensis*
Helan_FBXW	75	76	Parcial (meio)	DY916708	*Helianthus annuus*
Eupes_FBXW	77	78	Parcial (meio)	DV129599	*Euphorbia esula*
Lyces_FBXW	79	80	Partial (meio)	BI931509	*Lycopersicon esculentum*
Aqufo_FBXW	81	82	Parcial (meio)	DT753991.1	*Aquilegia formosa x Aquilegia pubescens*
Goshi_FBXW	83	84	Parcial (3')	DT466472	*Gossypium hirsutum*
Sorbi_FBXW	85	86	Parcial (meio)	CF770159	*Sorgo bicolor*
Iponi_FBXW	87	88	Parcial (3')	BJ574759.1	*Ipomea nil*
Soltu_FBXW	89	90	Parcial (3')	CX161187	*Solanum tuberosum*
Zamfi_FBXW	91	92	Parcial (meio)	DY032229	*Zamia fischeri*
Peram_FBXW	93	94	Parcial (3')	CK756534	*Persea americana*
Glyma_FBXW	95	96	Parcial (3')	CD418593.1	*Glycine max*
Brara_FBXW	107	108	Comprimento total	AC189583	*Brassica rapa*
Sorbi_FBXW	109	110	Comprimento total	próximo do BI075893 CW484775 CF770159 CF770238	*Sorgo bicolor*
Vitvi_FBXW	111	112	Comprimento total	AM440865	*Vitis vinifera*

Exemplo 9: Determinação da similaridade e identidade globais entre polipeptídeos FBXW, e suas regiões conservadas como representado pela SEQ ID N°: 102 e SEQ ID N°: 103 (ambas compreendidas na SEQ ID N°: 60)

As porcentagens globais de similaridade e identidade entre comprimento total dos polipeptídeos FBXW foram determinados usando um dos métodos disponível na técnica, o software MatGAT (Ferramenta de alinhamento global da matriz) (BMC Bioinformatics. 2003 4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando proteína ou seqüências de DNA. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). O software MatGAT que gera matrizes de similaridade/identidade por dNA ou seqüência da proteínas sem necessidade do pré-alinhamento dos dados. Os programas realizam uma série de alinhamentos em formas de pares usando o algoritmo Myers e Miller de alinhamento global (com a penalidade aberta na fenda de 12, e a penalidade da extensão da fenda de 2), cálculos de similaridade e identidade usando por exemplo Blosum 62 (para polipeptídeos), e então colocando o resultado na matriz de distância. A similaridade da seqüência é mostrada na metade funda da linha que divide e a identidade de seqüência é mostrada na metade superior da linha que divide diagonal. A seqüência da SEQ ID N°: 60 é de Arabidopsis *thaliana* (código Arath_FBXW).

Os parâmetros usados na comparação foram:

Contagem da matriz: Blosum62

Primeira fenda: 12

Fenda estendida: 2

Os resultados da análise do software são mostrados na Tabela C a H para similaridade e identidade globais no comprimento total dos polipeptídeos FBXW. Identidade de porcentagem é dado acima a diagonal e a similaridade da porcentagem é dado abaixo a diagonal. Identidade de porcentagem

entre as faixas de parálogos e ortólogos FBXW entre 45 e 80%, refletem a relatividade da conservação da identidade de seqüência inferior entre estas.

Tabela H: Os resultados MatGAT para a similaridade e identidade globais no comprimento total dos polipeptídeos FBXW.

Similaridade e identidade globais no comprimento total dos polipeptídeos FBXW	1	2	3	4	5	6
1. Arath_FBXW		57,9	51,1	63,5	50,4	49,6
2. Medtr_FBXW	72		51,2	61,7	50,2	49,1
3. Orysa_FBXW	68,7	66,6		51,6	76,5	74,4
4. Poptr_FBXW	77,7	75,1	69,9		52,8	52,3
5. Triae_FBXW	67,7	65,7	87,7	69,6		72,4
6. Zeama_FBXW	67,2	65,2	83,7	67,2	82,6	

Os resultados da análise do software foram mostrados nas Tabelas I e J para similaridade e identidade globais nas regiões conservadas 1 (como representado pela SEQ ID N°: 102 que compreende a SEQ ID N°: 60) e 2 (SEQ ID N°: 103 que compreende a SEQ ID N°: 60) dos polipeptídeos FBXW. Identidade de porcentagem é dada acima a diagonal e a similaridade da porcentagem é dada abaixo a diagonal. Identidade de porcentagem dos parálogos e ortólogos FBXW dentro da região conservada 1 (como na SEQ ID N°: 102) e dentro da região conservada 2 (como na SEQ ID N°: 103) as faixas entre 65% e 100% (similaridade entre 85 e 100%).

Tabela I: Os resultados MatGAT para a similaridade e identidade globais na região conservada 1 (como na SEQ ID N°: 102) dos polipeptídeos FBXW.

Similaridade e identidade globais na região conservada 1 dos polipeptídeos FBXW	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Cons1_Iponi_FBXW		94,7	84,2	92,1	89,5	84,2	81,6	97,4	97,4	81,6	81,6	81,6	84,2
2. Cons1_Soltu_FBXW	97,4		81,6	86,8	89,5	84,2	86,8	97,4	92,1	76,3	76,3	76,3	78,9
3. Cons1_Peram_FBXW	94,7	97,4		78,9	81,6	76,3	78,9	84,2	84,2	81,6	78,9	73,7	78,9
4. Cons1_Glyma_FBXW	94,7	92,1	89,5		84,2	78,9	73,7	89,5	89,5	76,3	76,3	78,9	78,9
5. Cons1_Poptr_FBXW	100	97,4	94,7	94,7		84,2	78,9	86,8	89,5	78,9	78,9	76,3	81,6
6. Cons1_Vitvi_FBXW	100	97,4	97,4	94,7	100		76,3	81,6	81,6	78,9	78,9	73,7	81,6
7. Cons1_Aqufo_FBXW	94,7	97,4	97,4	89,5	94,7	94,7		84,2	81,6	78,9	73,7	71,1	76,3
8. Cons1_Goshi_FBXW	97,4	100	97,4	92,1	97,4	97,4	97,4		94,7	78,9	78,9	78,9	81,6
9. Cons1_Medtr_FBXW	97,4	94,7	92,1	92,1	97,4	97,4	94,7	94,7		81,6	81,6	81,6	84,2
10. Cons1_Orysa_FBXW	100	97,4	97,4	94,7	100	100	94,7	97,4	97,4		89,5	71,1	94,7
11. Cons1_Triae_FBXW	100	97,4	97,4	94,7	100	100	94,7	97,4	97,4	100		65,8	94,7
12. Cons1_Arath_FBXW	92,1	89,5	86,8	94,7	92,1	92,1	86,8	89,5	89,5	92,1	92,1		68,4
13. Cons1_Zeama_FBXW	100	97,4	97,4	94,7	100	100	94,7	97,4	97,4	100	100	92,1	

Tabela J: Os resultados MatGAT para a similaridade e identidade globais na região conservada 2 (como na SEQ ID N°: 103) dos polipeptídeos FBXW.

Similaridade e identidade globais na região conservada 2 dos polipeptídeos FBXW	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Cons2_Ipono_FBXW		92,3	84,6	78,5	81,5	83,1	84,6	83,1	84,6	72,3	73,8	84,6	69,2
2. Cons2_Soltu_FBXW	96,9		83,1	78,5	78,5	81,5	81,5	83,1	84,6	70,8	73,8	78,5	70,8
3. Cons2_Peram_FBXW	95,4	95,4		81,5	84,6	87,7	81,5	89,2	84,6	73,8	75,4	81,5	73,8
4. Cons2_Glyma_FBXW	92,3	92,3	93,8		80	81,5	78,5	83,1	87,7	70,8	72,3	78,5	67,7
5. Cons2_Poptr_FBXW	90,8	90,8	93,8	92,3		86,2	78,5	86,2	80	69,2	72,3	81,5	69,2
6. Cons2_Vitvi_FBXW	90,8	90,8	95,4	92,3	93,8		80	90,8	81,5	73,8	75,4	80	70,8
7. Cons2_Aqufo_FBXW	93,8	93,8	95,4	92,3	90,8	90,8		83,1	87,7	73,8	75,4	81,5	70,8
8. Cons2_Goshi_FBXW	92,3	92,3	96,9	93,8	93,8	95,4	95,4		87,7	80	80	80	75,4
9. Cons2_Medtr_FBXW	92,3	92,3	93,8	93,8	89,2	89,2	95,4	93,8		75,4	76,9	83,1	69,2
10. Cons2_Orysa_FBXW	90,8	87,7	92,3	87,7	89,2	89,2	89,2	93,8	87,7		86,2	69,2	84,6
11. Cons2_Triae_FBXW	89,2	86,2	90,8	89,2	87,7	87,7	89,2	93,8	87,7	96,9		69,2	83,1
12. Cons2_Arath_FBXW	93,8	92,3	92,3	90,8	89,2	89,2	90,8	90,8	89,2	89,2	89,2		66,2
13. Cons2_Zeama_FBXW	87,7	87,7	89,2	84,6	86,2	86,2	86,2	90,8	84,6	96,9	93,8	87,7	

Exemplo 10: Clonagem da seqüência de ácido nucleico usando os métodos da invenção

Manipulação de DNA: a não ser de outra maneira estado, as técnicas de DNA recombinantes são realizadas de acordo com os protocolos padrões descritos em (Sambrook (2001) Molecular Cloning: a laboratory manual, 3° Edition Cold Spring Harbor Laboratory Press, CSH, New York) ou nos Volumes 1 e 2 de Ausubel et al. (1994), Current Protocols in Molecular Biology, Current Protocols. Materiais padrões e métodos para trabalhar a planta molecular são descritos em Plant Molecular Biology Labfax (1993) por R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (UK) e Blackwell Scientific Publications (UK).

A seqüência de ácido nucleico usado os métodos da invenção foi amplificado por PCR usando como padrão, uma biblioteca de cDNA de tecido misturado *Arabidopsis thaliana* feito sob encomenda (em pCMV Sport 6,0; Invitrogen, Paisley, UK). PCR foi realizado usando polimerase de DNA Hifi Taq em condições padrões, usando 200 ng de padrão em uma mistura de 50 µl de PCR. Os iniciadores usados foram prm06999 (SEQ ID N°: 105; sentido, local AttB1 em caso inferior:

5'- *gggacaagtttgtacaaaaaagcaggcttaaaca*ATGAATCGTTT TTCTCGTTT 3')

e prm07000 (SEQ ID N°: 106; reverso, complementar, local AttB2 em caso inferior:

5' *ggggaccactttgtacaagaaagctgggt*ATCCAATCTTATCGC TTAGG3'),

que inclui os locais AttB pela recombinação Gateway. O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante pelo qual o fragmento PCR recombina in vivo com o plasmídio pDONR201 para fornecer, de acordo com a tecnologia Gateway, um "clone de entrada", p08433. O plasmídio pDONR201 foi adquirido de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 11: Construção do vetor de expressão

O clone de entrada p08433 foi subseqüentemente usado em uma reação LR com p00640, um vetor de destinação usado pela transformação da Oryza sativa. Este vetor contém como elementos funcionais dentro dos limitadores de DNA T: um marcador de planta selecionável; um marcador de expressão de cassete avaliável; e um cassete Gateway pretendido pela recombinação LR in vivo com a seqüência de ácido nucleico de interesse já clonado no clone de entrada. Um promotor de arroz GOS2 (SEQ ID N°: 58) para a expressão constitutiva (PRO0129) foi localizado a montante deste cassete Gateway.

Após a etapa de recombinação LR, o vetor resultante da expressão p15973 (Figura 7) foi transformado em Cepa *Agrobacterium* LBA4044 e subseqüentemente por plantas Oryza sativa. As plantas de arroz transformadas foram deixadas desenvolver-se e foram então examinadas para os parâmetros descritos abaixo.

Exemplo 12: Procedimento de avaliação

12.1 Apresentação da avaliação

Aproximadamente 35 transformantes de arroz T0

independentes foram gerados. Os transformantes primários foram transferidos a partir de uma câmara de cultura de tecido para uma estufa para plantas para crescimento e endurecimento da semente T1. Sete eventos, pelo qual a produto T1 segregado 3:1 para a presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 mudas de T1 contendo o transgene (hetero e homo zigotos) e aproximadamente 10 mudas de T1 que precisam do transgene (nulizigotos) foram selecionados pelo monitoramento expressão do marcador visual. As plantas transgênicas e o nulizigotos correspondentes foram desenvolvidos lado a lado em posições aleatórias. As condições da estufa para plantas foram de dias curtos (12 horas de luz), 28°C na luz e 22°C no escuro, e uma umidade relativa de 70%.

Quatro eventos T1 ainda foram avaliados na geração T2 seguindo os mesmos procedimentos de avaliação como para a geração T1 mas com mais indivíduos por evento. A partir do estágio da semeadura até o estágio da maturidade as plantas foram passadas diversas vezes através de uma cabine formadora de imagem digital. Em cada um ponto de tempo as imagens digitais (2048 x 1536 pixéis, 16 milhões de cores) foram tiradas de cada planta de pelo menos 6 ângulos diferentes.

12.2 Análise estatística: teste F

Um dos dois fatores ANOVA (análise de variantes) como usado como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um Teste F foi realizado em todos os parâmetros medido de todas as plantas de todos os eventos transformados com o gene da presente invenção. O teste F foi realizado por verificar por um efeito do gene em todas os eventos de transformação e para verificar por um efeito total do gene, também conhecido como um efeito do gene global. O princípio para a significância por um efeito de gene verdadeiro global foi apresentado a 5% nível de probabilidade para o teste F. Um ponto de valor significante no teste F a um efeito do gene, significa que não é apenas como mera presença ou

posição do gene que está causando as diferenças nos fenotipos.

Exemplo 13: Resultados da avaliação

As plantas da área acima do solo (ou biomassa da folha) foram determinadas pela contagem do número total de pixéis nas imagens digitais a partir das partes acima do solo das plantas discriminadas a partir do fundamento. Este valor foi medido para as figuras retiradas no mesmo ponto a partir de ângulos diferentes e foi convertido a um valor de superfície física expressada no quadrado mm pela calibração. Os experimentos mostram que as plantas da área acima do solo medidas desta maneira correlacionadas com a biomassa das partes das plantas acima trituradas. A área triturada acima é o ponto de tempo em que as plantas tem atingido esta biomassa de folha máxima.

As panículas primárias maduras foram endurecidas, contadas, empacotadas, rotulado com código de barra e então secadas por três dias em um forno a 37°C. As panículas foram então refugadas e todas as sementes foram colhidas e contadas. A casca enchedora foi separada a partir de um despejo usando dispositivo de sopro de ar. As cascas despejadas foram descartadas e a fração remanescente foi contada novamente. A casca enchedora foram pesada em uma balança analítica. O número de sementes enchedoras foi determinada pela contagem o número de cascas enchedoras que permaneceram após a etapa de separação. A produção total da semente foi medida pelo pesagem de todas as cascas enchedoras endurecidas a partir de uma planta. O número de semente total por plantas foi medida pela contagem do número de cascas endurecidas a partir de uma planta. Peso singular da semente (TKW) é extrapolado a partir do número de sementes enchedoras contadas e seu peso total. O índice colhido (HI) na presente invenção é definido como uma razão entre a produção total da semente e a área triturada acima (mm^2), multiplicado por um fator 10^6. O número total de flores por panícula como definido na presente invenção é a razão entre o

número total de sementes e o número de panículas primárias maduras. A taxa enchedora da semente como definido na presente invenção é uma proporção (expressada como uma%) de diversas sementes enchedoras no número total de sementes (ou flósculos).

Como apresentado nas Tabelas K a O, a biomassa acima do solo, o número de flores por panícula, a produção da semente, o número total de sementes, o número de sementes enchedoras, os peso singular da semente (TKW) e índice colhido são aumentados nas plantas transgênicas com aumento da expressão de um ácido nucleico codificando um polipeptídeo FBXW, comparado com as planta adequadas de controle. Os resultados a partir da geração T1 e T2 foram mostradas.

Tabela K mostra um número de eventos transgênicos com um aumento na produção total da semente (peso da semente total), as porcentagens deste aumento, bem como a relevância estatística deste aumento de acordo com um teste F.

Tabela K: Diversos eventos transgênicos com um aumento na produção total da semente, as porcentagens de aumento, e o valor P do teste F na geração T1 e T2 de arroz transgênico com aumento da expressão de um ácido nucleico codificando um polipeptídeo FBXW.

Produção total da semente			
	Diversos eventos mostrando um aumento	**% da diferença**	**Valor P do teste F**
Geração T1	6 fora de 7	21	0,001
Geração T2	3 fora de 4	17	0,0002

Tabela L mostra um número de eventos transgênicos com um aumento de diversas sementes enchedoras, as porcentagens deste aumento, bem como a relevância estatística deste aumento de acordo com um teste F.

Tabela L: Diversos eventos transgênicos com um aumento no número de sementes enchedoras, as porcentagens de aumento, e valor P do teste F na geração T1 e T2 de arroz transgênico com aumento expressão de um ácido nucleico codificando um polipeptídeo FBXW.

Número de sementes enchedoras			
	Diversos eventos mostrando um aumento	**% da diferença**	**Valor P do teste F**
Geração T1	6 fora de 7	20	0,0017
Geração T2	3 fora de 4	17	0,0002

Tabela M mostra um número de eventos transgênicos com um aumento no índice colhido, as porcentagens deste aumento, bem como a relevância estatística deste aumento de acordo com um teste F.

Tabela M: Diversos eventos transgênicos com um aumento no índice colhido, as porcentagens do aumento, e valor P do teste F na geração T1 e T2 de arroz transgênico com aumento expressão de um ácido nucleico codificando um polipeptídeo FBXW.

Índice colhido			
	Diversos eventos mostrando um aumento	**% da diferença**	**Valor P do teste F**
Geração T1	6 fora de 7	17	0,0009
Geração T2	3 fora de 4	15	<0,0001

Tabela N mostra um número de eventos transgênicos com um aumento no peso singular da semente (TKW), as porcentagens deste aumento, bem como a relevância estatística deste aumento de acordo com um teste F.

Tabela N: Diversos eventos transgênicos com um aumento no peso singular da semente (TKW), as porcentagens do aumento, e valor P do teste F na geração T1 e T2 de arroz transgênico com aumento expressão de um ácido nucleico codificando um polipeptídeo FBXW.

TKW			
	Diversos eventos mostrando um aumento	**% da diferença**	**Valor P do teste F**
Geração T1	5 fora de 7	2	0,0187
Geração T2	2 fora de 4	1	0,0022

Tabela O mostra um número de eventos transgênicos com um aumento na taxa enchedora da semente, as porcentagens deste aumento, bem como a relevância estatística deste aumento de acordo com um teste F.

Tabela O: Diversos eventos transgênicos com um aumento na taxa enchedora, as porcentagens do aumento, e valor P do teste F na geração T1 e T2 de arroz transgênico com aumento expressão de um ácido nucleico codificando um polipeptídeo FBXW.

Taxa da semente enchedora			
	Diversos eventos mostrando um aumento	**% da diferença**	**Valor P do teste F**
Geração T1	6 fora de 7	15	0,0005
Geração T2	3 fora de 4	9	<0,0001

RANBP

Exemplo 14: Identificação das seqüências que diz respeito à seqüência de ácido nucleico usado os métodos da invenção

As seqüências (cDNA de comprimento total, ESTs ou genômico) que diz respeito à seqüência de ácido nucleico usado os métodos da presente invenção foram identificados entre aqueles mantidos na base de dados de nucleotídeos Entrez no National Center for Biotechnology Information (NCBI) usando ferramentas de busca da seqüência de base de dados, tal como a ferramenta de alinhamento de local básico (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucl. Acids Res. 25:3389-3402). Os programas são usados para observar regiões de similaridade local entre seqüências pela comparação com as seqüências de ácido nucleico ou polipeptídeo pela base de dados da seqüência e pelo cálculo da significância estatística dos pontos. Por exemplo, os polipeptídeos codificados pelo ácido nucleico usados na presente invenção como usado pelo algoritmo TBLASTN, com falta de direção e o filtro para ignorar as seqüências de complexidade inferiores apresentadas. A produção de uma análise foi vista pela comparação em forma de par, e classificada de acordo com a contagem de probabilidade (Valor E), onde a contagem reflete a probabilidade que um alinhamento particular ocorra pela escolha (o valor E inferior, o mais significante do acerto). Além disso a comparação aos valores E também foram contados pela identidade da porcentagem. Identidade de porcentagem refere-se à diversos nucleotídeos idênticos (ou aminoácidos) entre as seqüências de dois ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. Em alguns exemplos, os parâmetros padrões pode ser ajustados para modificar a severidade da busca. Por exemplo o valor E pode ser aumentado para mostrar os pontos menos estringentes. Esta

maneira, os pontos exatos aproximadamente curtos pode ser identidificados.

Tabela P fornece uma lista da seqüência de ácido nucleico que diz respeito à seqüência de ácido nucleico usado os métodos da presente invenção.

Tabela P: Seqüência de ácido nucleico que diz respeito à seqüência de ácido nucleico (SEQ ID N°: 113) útil nos métodos da presente invenção, e os polipeptídeos correspondentes deduzidos.

Fonte da planta	Ácido nucleico SEQ ID N°:	Proteína SEQ ID N°:
Zea mays	113	114
Zea mays	113	115
Arabidopsis thaliana	116	117
Arabidopsis thaliana	116	118
Arabidopsis thaliana	119	120
Arabidopsis thaliana	121	122
Lycopersicon esculentum	123	124
Oryza sativa	125	126
Zea mays	127	128
Oryza sativa	129	130
Populus sp	131	132
Saccharum officinarum	133	134
Saccharum officinarum	135	136
Medicago	137	138

Em alguns exemplos, que diz respeito a seqüência que tem tentativamente ser reunida e publicada divulgada pelas instituições de busca, tal como The Institute for Genomic Research (TIGR). A base de dados do gene ortólogo eucariótico (EGO) pode ser usado para identificar um tal que diz respeito as seqüências, pela busca da palavra chave ou pelo uso do algoritmo BLAST com a seqüência de ácido nucleico ou polipeptídeo de interesse.

Exemplo 15: Clonagem de um Zea Mays RANBP que codifica a seqüência de ácido nucleico usado os métodos da invenção

A seqüência de ácido nucleico usado os métodos da invenção foi amplificada por PCR usando como padrão, uma biblioteca de cDNA de endosperma de milho. PCR foi realizado usando polimerase de DNA Hifi Taq em condições padrões, usando 200 ng de padrão em uma mistura de 50 µl de

PCR. Os iniciadores usados foram prm06703 (SEQ ID N°: 151; sentido, códon iniciado em negrito, AttB1 locais em itálicos:

5'-ggggacaagtttgtacaaaaaagcaggcttaaacaatggcggacaaggagc-3')

e prm06704 (SEQ ID N°: 152; reverso, complementar, AttB2 locais em itálicos:

5' ggggaccactttgtacaagaaagctgggtagtgcaacc acaccaactact 3'),

que inclui os locais AttB pela recombinação Gateway. O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante pelo qual o fragmento PCR recombina *in vivo* com o plasmídio pDONR201 para fornecer, de acordo com a tecnologia Gateway, um "clone de entrada". O plasmídio pDONR201 foi adquirido de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 16: Construção do vetor de expressão

O clone de entrada foi subseqüentemente usado em uma reação LR com um vetor de destinação usado pela transformação da Oryza sativa. Este vetor contém como elementos funcionais dentro dos limitadores de DNA T: um marcador de planta selecionável; um marcador de expressão de cassete avaliável; e um cassete Gateway pretendido pela recombinação LR in vivo com a seqüência de ácido nucleico de interesse já clonado no clone de entrada. Um promotor de prolamina (SEQ ID N°: 155) pela expressão específica de embrião (referência interna PRO90) foi localizado a montante deste cassete Gateway.

Após a etapa de recombinação LR, o vetor resultante da expressão p072 (Figura 9) foi transformado em Cepa *Agrobacterium* LBA4044 e subseqüentemente por plantas Oryza sativa. As plantas de arroz transformadas foram deixadas desenvolver-se e foram então examinadas para os parâmetros descritos abaixo.

Exemplo 17: Clonagem de Arabidopsis thaliana RANBP que codifica a

seqüência de ácido nucleico usado os métodos da invenção

A seqüência de ácido nucleico usado os métodos da invenção foi amplificada por PCR usando como padrão, uma biblioteca de cDNA de *Arabidopsis thaliana* (em pCMV Sport 6.0; Invitrogen, Paisley, UK). PCR foi realizado usando polimerase de DNA Hifi Taq em condições padrões, usando 200 ng de padrão em uma mistura de 50 µl de PCR. Os iniciadores usados foram 6491 (SEQ ID N°: 153; sentido, códon iniciado em negrito, AttB1 local em itálico:

5'- ggggacaagtttgtacaaaaaagcaggcttcacaatggcgagcattagcaac 3')

e 6492 (SEQ ID N°: 154; reverso, complementar, AttB2 local em itálico:

5' ggggaccactttgtacaagaaagctgggtgcatcttaagctgagggaac 3'),

que inclui os locais AttB pela recombinação Gateway. O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante pelo qual o fragmento PCR recombina in vivo com o plasmídio pDONR201 para fornecer, de acordo com a tecnologia Gateway, um "clone de entrada". O plasmídio pDONR201 foi adquirido de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 18: Construção do vetor de expressão

O clone de entrada foi subseqüentemente usado em uma reação LR com um vetor de destinação usado pela transformação da Oryza sativa. Este vetor contém como elementos funcionais dentro dos limitadores de DNA T: um marcador de planta selecionável; um marcador de expressão de cassete avaliável; e um cassete Gateway pretendido pela recombinação LR in vivo com a seqüência de ácido nucleico de interesse já clonado no clone de entrada. Um promotor de prolamina (SEQ ID N°: 155) pela expressão específica de embrião (referência interna PRO90) foi localizado a montante deste cassete Gateway.

Após a etapa de recombinação LR, o vetor resultante da

expressão p074 (Figura 10) foi transformado em Cepa *Agrobacterium* LBA4044 e subseqüentemente por plantas Oryza sativa. As plantas de arroz transformadas foram deixadas desenvolver-se e foram então examinadas para os parâmetros descritos abaixo.

Exemplo 19: Procedimento de avaliação

19.1 Apresentação da avaliação

Aproximadamente 30 transformantes de arroz T0 independentes foram gerados. Os transformantes primários foram transferidos a partir de uma câmara de cultura de tecido para uma estufa para plantas para crescimento e endurecimento da semente T1. Sete eventos, pelo qual a produto T1 segregado 3:1 para a presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 mudas de T1 contendo o transgene (hetero e homo zigotos) e aproximadamente 10 mudas de T1 que precisam do transgene (nulizigotos) foram selecionados pelo monitoramento expressão do marcador visual. As plantas transgênicas e o nulizigotos correspondente foram desenvolvidos lado a lado em posições aleatórias. As condições da estufa para plantas foram de dias curtos (12 horas de luz), 28°C na luz e 22°C no escuro, e uma umidade relativa de 70%.

Quatro eventos T1 ainda foram avaliados na geração T2 seguindo os mesmos procedimentos de avaliação como para a geração T1 mas com mais indivíduos por evento. A partir do estágio da semeadura até o estágio da maturidade as plantas foram passadas diversas vezes através de uma cabine formadora de imagem digital. Em cada um ponto de tempo as imagens digitais (2048 x 1536 pixéis, 16 milhões de cores) foram tirada de cada planta de pelo menos 6 ângulos diferentes.

19.2 Análise estatística: Teste t e Teste F

Um dos dois fatores ANOVA (análise de variantes) como usado como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um Teste F foi realizado em todos os parâmetros

medido de todas as plantas de todos os eventos transformados com o gene da presente invenção. O teste F foi realizado por verificar por um efeito do gene em todas os eventos de transformação e para verificar por um efeito total do gene, também conhecido como um efeito do gene global. O princípio para a significância por um efeito de gene verdadeiro global foi apresentado a 5% nível de probabilidade para o teste F. Um ponto de valor significante no teste F a um efeito do gene, significa que não é apenas como mera presença ou posição do gene que está causando as diferenças nos fenotipos.

Por verificar por um efeito do genes dentro de um evento, isto é, por um efeito específico de linha, um teste t foi realizado dentro de cada evento usando dados apresentados a partir das plantas transgênicas e o correspondente de plantas nulas. "plantas nulas" ou "segregadores nulos" ou "nulizigotos" são as plantas tratadas na mesma maneira como a planta transgênica, mais de que um transgene tem segregado. As plantas nulas também podem ser descritas como plantas negativas homozigotas transformadas. O princípio para a significância para o teste t é apresentado a 10% do nível de probabilidade. Os resultados por alguns eventos podem ser acima ou abaixo deste princípio. Este é baseado na hipótese que uma força do gene apenas tem um efeito em certas posições no genoma, e que a ocorrência deste efeito depende da posição e não é comum. Este tipo de efeito do gene também é denominado neste um "efeito do gene de linha". O valor P é obtido pela comparação com o valor t à distribuição t ou alternativamente, pela comparação com o valor F à distribuição F. O valor P então dá as probabilidades que a hipótese de nulo (isto é, que aqui não é nenhum efeito do transgene) seja correta.

Exemplo 20: Resultados da avaliação

As plantas de arroz transgênicas que expressam um RANBP de milho sob o controle de um promotor de prolamina dão um aumento na peso médio da semente, número de sementes enchedoras, índice colhido,

biomassa, taxa enchedora, peso singular da semente (TKW), área média da semente, comprimento médio da semente e altura média da semente, cada um que diz respeito as plantas de controle. Em particular, TKW foi aumentado na geração T1 e este aumento foi confirmado na geração de plantas T2. O aumento como observado a ser significante estatisticamente com um valor P a partir do teste F de >0,00001. Também notável foi o aumento na taxa enchedora comparada as plantas de controle, com o aumento na geração T1 sendo confirmada na geração de plantas T2. O aumento também foi observado a ser significante estatisticamente com um valor P a partir do teste F de 0,001.

Dado comparativo

As plantas de arroz transgênicas que expressam um RANBP de milho sob o controle de um promotor GOS2 constitutivo não dá uma diferença real na produção comparada as plantas de controle. Aqui não foi aumento no peso médio da semente, número de sementes enchedoras, índice colhido, biomassa, taxa enchedora ou peso singular da semente (TKW) em plantas transgênicas comparadas as plantas de controle.

As plantas de arroz transgênicas que expressam um RANBP *Arabidopsis thaliana* sob o controle de um promotor de prolamina dando um aumento na biomassa, peso médio da semente, número de sementes enchedoras, número de flores por panícula, índice colhido, taxa enchedora e peso singular da semente, cada um que diz respeito as plantas de controle.

GLK

Exemplo 21: Identificação de homólogos da proteína GLK da SEQ ID N°: 157 em Arabidopsis, arroz e outras espécies de plantas.

As seqüências (cDNA de comprimento total, ESTs ou genômico) que diz respeito à seqüência de ácido nucleico usado os métodos da presente invenção foram identificados entre aqueles mantidos na base de dados de nucleotídeos Entrez no National Center for Biotechnology

Information usando ferramentas de busca da seqüência de base de dados, tal como a ferramenta de alinhamento de local básico (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucl. Acids Res. 25:3389-3402). Este programa é tipicamente usado para observar regiões de similaridade local entre seqüências pela comparação com as seqüências de ácido nucleico ou polipeptídeo pela base de dados da seqüência e pelo cálculo a significância estatística dos pontos. Os polipeptídeo codificado pelo ácido nucleico da presente invenção como usado com o algoritmo TBLASTN, com falta de direção e o filtro para ignorar seqüências de complexidade inferiores foi ajustado. A produção de uma análise foi vista pela comparação em forma de par, e classificada de acordo com a contagem de probabilidade (Valor E), onde a contagem reflete a probabilidade que um alinhamento particular ocorre pela escolha (o valor E inferior, o mais significante do acerto). Além disso á comparação aos valores E também foram contados pela identidade de porcentagem. Identidade de porcentagem refere-se à diversos nucleotídeos idênticos (ou aminoácidos) entre as seqüências de dois ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. Em alguns exemplos, os parâmetros padrões pode ser ajustados para modificar a severidade da busca.

As seqüências de arroz e seqüências EST de várias espécies de plantas também pode ser obtidas a partir de outras base de dados, tal como KOME (conhecimento com base em Oryza Molecular biological Encyclopedia; Kikuchi et al., Science 301, 376-379, 2003), Sputnik (Rudd, S., Nucl. Acids Res., 33: D622 - D627, 2005) ou uma base de dados ortólogos do gene eucariótico (EGO, hospedado por Institute for Genomic Research). Estas bases de dados são buscáveis com a ferramenta BLAST. SEQ ID N°: 168 a SEQ ID N°: 193 são ácidos nucleicos e seqüência da proteínas de homólogos da SEQ ID N°: 157 e foram obtidas a partir da base de dados mencionadas acima usando SEQ ID N°: 157 como uma seqüência em questão.

Tabela Q: Seqüência de ácido nucleicos que diz respeito à seqüência de ácido nucleico (SEQ ID N°: 156) úteis nos métodos da presente invenção, e o polipeptídeos correspondentes deduzidos.

Fonte da planta	Ácido nucleico SEQ ID N°:	Proteína SEQ ID N°:
OsGLK	156	157
OSJNBa0086P08.18	168	169
Arabidopsis thaliana GLK1	170	171
Arabidopsis thaliana GLK2	172	173
Physcomitrella patens Glk1	174	175
Physcomitrella patens Glk2	176	177
Zea mays ZmGLK1	178	179
Zea mays ZmGLK2	180	181
Triticum aestivum TaGLK1	182	183
Allium cepa AcGLK1	184	185
Hordeum vulgare HvGLK2	186	187
Sorgo bicolor SbGLK1	188	189
Saccharum officinarum oGLK2	190	191
Oryza sativa OsGLK1	192	193

Exemplo 22: Determinação da similaridade e identidade globais entre proteínas GLK.

As porcentagens de similaridade e identidade entre o comprimento total da seqüência da proteína GLK e entre os domínios GARP ou GCT das proteínas GLK foram determinados usando software MatGAT (Ferramenta de alinhamento global da matriz) (BMC Bioinformatics. 2003 4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando proteína ou seqüências de DNA. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). O software MatGAT que gera matrizes de similaridade/identidade por DNA ou seqüência da proteínas sem necessidade do pré-alinhamento dos dados. Os programas realizam uma série de alinhamentos em formas de pares usando o algoritmo Myers e Miller de alinhamento global (com a penalidade aberta na fenda de 12, e a penalidade da extensão da fenda de 2), cálculos de similaridade e identidade usando por exemplo Blosum 62 (para polipeptídeos), e então colocando o resultado na matriz de distância. A similaridade de seqüência é mostrada na metade funda da linha que divide e a identidade da seqüência é mostrada na metade superior

da linha que divide diagonal.

Os domínios GARP e GCT foram delineados usando um alinhamento múltiplo e as seqüências obtidas são listadas nas Tabelas R e S. Os resultados da análise do software foram mostrados nas Tabelas T a V para similaridade e identidade de nenhum comprimento total da seqüência da proteínas e para os domínios GARP ou GCT dos polipeptídeos GLK. A seqüência da SEQ ID N°: 157 é indicada como número 6 (OsGLK1) nas matrizes. A identidade de porcentagem é dada acima a diagonal (em negrito) e a similaridade da porcentagem é dada abaixo a diagonal (fonte normal). A identidade de porcentagem entre seqüências de comprimento total de parálogos e ortólogos GLK da faixa da SEQ ID N°: 157 entre 30% e 98,7%. Estas porcentagens são consideravelmente maiores quando a seqüência do domínio GARP é usado em vez da seqüência de comprimento total.

Tabela R: Seqüência dos domínios GARP como obtido no alinhamento das seqüências de proteínas e usado a análise MATGAT:

ZmG2	KVKVDWTPELHRRFVQAVEQLGIDKAVPSRILEIMGTDCLTRHNI ASHLQKYRSHR
ZmGLK1	KAKVDWTPELHRRFVQAVEELGIDKAVPSRILEIMGIDSLTRHNIA SHLQKYRSHR
OsGLK2	KVKVDWTPELHRRFVQAVEQLGIDKAVPSRILELMGIECLTRHNI ASHLQKYRSHR
PpGLK2	KAKVDWTPELHRRFVHAVEQLGVEKAYPSRILELMGVQCLTRHN IASHLQKYRSHR
PpGlk1	KAKVDWTPELHRRFVHAVEQLGVEKAFPSRILELMGVQCLTRHNI ASHLQKYRSHR
OsGLK1	KAKVDWTPELHRRFVQAVEQLGIDKAVPSRILEIMGIDSLTRHNIA SHLQKYRSHR
AtGLK2	KPKVDWTPELHRKFVQAVEQLGVDKAVPSRILEIMNVKSLTRHN VASHLQKYRSHR
AtGLK1	KVKVDWTPELHRRFVEAVEQLGVDKAVPSRILELMGVHCLTRHN VASHLQKYRSHR
AcGLK1	KAKVDWTPELHRRFVQAVEQLGVDKAVPSRILELMGIDCLTRHNI ASHLQKYRSHR

Tabela S: Seqüências dos domínios GCT como obtido no alinhamento das seqüências de proteínas e usados na análise MATGAT:

ZmG2	KHLMAREAEAATWAQKRHMYAPPAPRTTTTTDAARPPWVVPTTIGFP PPRFCRPLHVWGHPP PHAAAAEAAAATPMLPVWPRHLAPPRHLAPWAHPTPVDPAFWHQQY SAARKWGPQAAAVTQG

	TPCVPLPRFPVPHPIYSRPAMVPPPPSTTKLAQLHLELQAHPSKESIDAAI GDVLVKPWLPL PLGLKPPSLDSVMSELHKQGVPKIPPAAATTTGATG
ZmGLK1	KHMLAREVEAATWTTHRRPMYAAPSGAVKRPDSNAWTVPTIGFPPPA GTPPRPVQHFGRPLH VWGHPSPTPAVESPRVPMWPRHLAPRAPPPPPWAPPPPADPASFWHHA YMRGPAAHMPDQVA VTPCVAVPMAAARFPAPHVRGSLPWPPPMYRPLVPPALAGKSQQDAL FQLQIQPSSESIDAA IGDVLTKPWLPLPLGLKPPSVDSVMGELQRQGVANVPQACG
OsGLK2	KHLMAREAEAASWTQKRQMYTAAAAAAAAVAAGGGPRKDAAAATA AVAPWVMPTIGFPPPHAA AMVPPPPHPPPFCRPPLHVWGHPTAGVEPTTAAAPPPPSPHAQPPLLPV WPRHLAPPPPPLP AAWAHGHQPAPVDPAAYWQQQYNLQRFPVPPVPGMVPHPMYRPIPPP SPPQGNKLAALQLQL DAHPSKESIDAAIGDVLVKPWLPLPLGLKPPSLDSVMSELHKQGIPKVP PAASGAAG
PpGLK2	RHLAAREAEAASWTHRRTYTQAPWPRSSRRDGLPYLVPIHTPHIQPRPS MAMAMQPQLQTPH HPISTPLKVWGYPTVDHSNVHMWQQPAVATPSYWQAADGSYWQHPA TGYDAFSARACYSHPM QRVPVTTTHAGLPIVAPGFPDESCYYGDDMLAGSMYLCNQSYDSEIGR AAGVAACSKPIETH LSKEVLDAAIGEALANPWTPPPLGLKPPSMEGVIAELQRQGINTVPPST C
PpGlk1	RHLAAREAEAASWTHRRAYTQMPWSRSSRRDGLPYLVPLHTPHIQPRP SMVMAMQPQLQTQH TPVSTPLKVWGYPTVDHSSVHMWQQPAVATPSYWQAPDGSYWQHPA TNYDAYSARACYPHPM RVSLGTTHAGSPMMAPGFPDESYYGEDVLAATMYLCNQSYDSELGRA AGVAACSKPPETHLS KEVLDAAIGEALANPWTPPPLGLKPPSMEGVIAELQRQGINTVPPSTC
OsGLK1	KHMIAREAEAASWTQRRQIYAAGGGAVAKRPESNAWTVPTIGFPPPPP PPPSPAPMQHFARP LHVWGHPTMDPSRVPVWPPRHLVPRGPAPPWVPPPPPSDPAFWHHPY MRGPAHVPTQGTPCM AMPMPAARFPAPPVPGVVPCPMYRPLTPPALTSKNQQDAQLQLQVQPS SESIDAAIGDVLSK PWLPLPLGLKPPSVDSVMGELQRQGVANVPPACG
AtGLK2	KHLLAREAEAASWNLRRHATVAVPGVGGGGKKPWTAPALGYPPHVA PMHHGHFRPLHVWGHP TWPKHKPNTPASAHRTYPMPAIAAAPASWPGHPPYWHQQPLYPQGYG MASSNHSSIGVPTRQ LGPTNPPIDIHPSNESIDAAIGDVISKPWLPLPLGLKPPSVDGVMTELQR QGVSNVPPLP
AtGLK1	KHLLAREAEAANWTRKRHIYGVDTGANLNGRTKNGWLAPAPTLGFPP PPPVAVAPPPVHHHH FRPLHVWGHPTVDQSIMPHVWPKHLPPPSTAMPNPPFWVSDSPYWHP MHNGTTPYLPTVATR FRAPPVAGIPHALPPHHTMYKPNLGFGGARPPVDLHPSKESVDAAIGD VLTRPWLPLPLGLN PPAVDGVMTELHRHGVSEVPPTASCA

Tabela T: Matriz de MATGAT do comprimento das seqüências totais.

	1	2	3	4	5	6	7	8	9
1. ZmG2		43,9	53,6	29,4	29,7	45,0	38,8	40,0	44,8
2. ZmGLK1	54,3		42,1	32,0	33,2	68,3	39,5	40,9	68,5
3. OsGLK2	63,2	54,8		32,9	31,3	46,9	37,6	40,5	46,9
4. PpGLK2	43,5	45,1	46,8		78,7	33,3	31,3	34,3	33,5
5. PpGlk1	42,9	44,4	43,5	85,1		33,6	30,9	32,9	34,2
6. OsGLK1	57,7	75,6	56,5	45,5	46,4		40,7	45,1	98,7
7. AtGLK2	49,5	49,9	47,0	42,2	41,9	50,8		46,1	40,3
8. AtGLK1	51,8	52,2	52,0	45,8	43,6	58,7	58,6		45,1
9. OsGLK1var	57,9	75,6	56,7	44,5	46,8	99,3	50,5	58,7	

Tabela U: Matriz de MATGAT dos domínios GARP.

	1	2	3	4	5	6	7	8	9
1. ZmG2		92,9	94,6	85,7	85,7	94,6	85,7	89,3	92,9
2. ZmGLK1	94,6		91,1	83,9	83,9	98,2	85,7	83,9	92,9
3. OsGLK2	98,2	96,4		87,5	87,5	92,9	83,9	91,1	94,6
4. PpGLK2	91,1	92,9	94,6		98,2	85,7	82,1	89,3	91,1
5. PpGlk1	91,1	92,9	94,6	100,0		85,7	82,1	89,3	91,1
6. OsGLK1	94,6	100,0	96,4	92,9	92,9		87,5	85,7	94,6
7. AtGLK2	91,1	94,6	94,6	91,1	91,1	94,6		87,5	85,7
8. AtGLK1	96,4	94,6	98,2	92,9	92,9	94,6	92,9		91,1
9. AcGLK1	96,4	98,2	98,2	94,6	94,6	98,2	92,9	96,4	

Tabela V: Matriz de MATGAT dos domínios GCT

	1	2	3	4	5	6	7	8
1. ZmG2		48,2	56,3	26,7	28,5	47,6	38,0	37,0
2. ZmGLK1	58,6		44,0	32,4	33,9	72,7	41,0	44,3
3. OsGLK2	65,0	57,6		33,6	32,0	51,6	40,6	42,5
4. PpGLK2	41,9	47,0	44,0		88,6	32,4	33,8	35,0
5. PpGlk1	42,3	45,7	42,4	93,2		35,2	32,6	34,5
6. OsGLK1	61,3	81,1	60,5	44,1	49,6		44,2	48,9
7. AtGLK2	48,6	48,5	48,6	43,2	43,2	52,3		47,3
8. AtGLK1	50,9	55,5	54,7	49,2	46,6	61,8	57,1	

Exemplo 23: Clonagem da seqüência de ácido nucleico usando os métodos da invenção

Manipulação de DNA: a não ser de outra maneira estado, as técnicas de DNA recombinantes são realizadas de acordo com os protocolos padrões descritos em (Sambrook (2001) Molecular Cloning: a laboratory manual, 3° Edition Cold Spring Harbor Laboratory Press, CSH, New York) ou nos Volumes 1 e 2 de Ausubel et al. (1994), Current Protocols in Molecular Biology, Current Protocols. Materiais padrões e métodos para trabalhar a planta molecular são descritos em Plant Molecular Biology Labfax (1993) por R.D.D. Croy, publicado por BIOS Scientific Publications Ltd

(UK) e Blackwell Scientific Publications (UK).

A seqüência de ácido nucleico usado nos métodos da invenção foi amplificado por PCR usando como padrão, a biblioteca de cDNA das mudas *Oryza sativa* feitas por encomenda (em pCMV Sport 6.0; Invitrogen, Paisley, UK). PCR foi realizado usando polimerase de DNA Hifi Taq em condições padrões, usando 200 ng de padrão em uma mistura de 50 µl de PCR. Os iniciadores usados foram prm2251 (SEQ ID N°: 158; sentido, códon iniciado em negrito:

5' ggggacaagtttgtacaaaaaagcaggcttcacaatgcttgccgtgtcgc 3')

e prm2252 (SEQ ID N°: 159; reverso, complementar,:

5' ggggaccactttgtacaagaaagctgggtaatatcatccacacgctgga 3'),

que inclui os locais AttB pela recombinação Gateway. O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizado, durante pelo qual o fragmento PCR recombina in vivo com o plasmídio pDONR201 para fornecer, de acordo com a tecnologia Gateway, um "clone de entrada", p034. O plasmídio pDONR201 foi adquirido de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 24: Construção do vetor de expressão

O clone de entrada p031 foi subseqüentemente usado em uma reação LR com p00640, um vetor de destinação usado pela transformação da Oryza sativa. Este vetor contém como elementos funcionais dentro dos limitadores de DNA T: um marcador de planta selecionável; um marcador de expressão de cassete avaliável ; e um cassete Gateway pretendido pela recombinação LR in vivo com a seqüência de ácido nucleico de interesse já clonado no clone de entrada. Um promotor constitutivo não viral de arroz, o promotor GOS2 (SEQ ID N°: 58) (PRO0129) foi localizado a montante deste cassete Gateway.

Após a etapa de recombinação LR, o vetor resultante da

expressão p045 (Figura 14) foi transformado em Cepa *Agrobacterium* LBA4044 e subseqüentemente por plantas Oryza sativa. As plantas de arroz transformadas foram deixadas desenvolver-se e foram então examinadas para os parâmetros descritos abaixo.

Exemplo 25: Procedimento de avaliação

25.1 Apresentação da avaliação

Aproximadamente 30 transformantes de arroz T0 independentes foram gerados. Os transformantes primários foram transferidos a partir de uma câmara de cultura de tecido para uma estufa para plantas para crescimento e endurecimento da semente T1. Quatro eventos, pelo qual um produto T1 segregado 3:1 para a presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 mudas de T1 contendo o transgene (hetero e homo zigotos) e aproximadamente 10 mudas de T1 que precisam do transgene (nulizigotos) foram selecionados pelo monitoramento da expressão do marcador visual. As plantas transgênicas e os nulizigotos correspondentes foram desenvolvidos lado a lado em posições aleatórias. As condições da estufa para plantas foram de dias curtos (12 horas de luz), 28°C na luz e 22°C no escuro, e uma umidade relativa de 70%.

Os quatro eventos T1 ainda foram avaliados na geração T2 seguindo os mesmos procedimentos de avaliação como para a geração T1 mas com mais indivíduos por evento. A partir do estágio da semeadura até o estágio da maturidade as plantas foram passadas diversas vezes através de uma cabine formadora de imagem digital. Em cada um ponto de tempo as imagens digitais (2048 x 1536 pixéis, 16 milhões de cores) foram tiradas de cada planta de pelo menos 6 ângulos diferentes.

25.2 Análise estatística: Teste t e Teste F

Um dos dois fatores ANOVA (análise de variantes) como usado como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um Teste F foi realizado em todos os parâmetros

medidos de todas as plantas de todos os eventos transformados com o gene da presente invenção. O Teste F foi realizado por verificar por um efeito do gene em todas os eventos de transformação e para verificar por um efeito total do gene, também conhecido como um efeito do gene global. O princípio para a significância por um efeito de gene verdadeiro global foi apresentado a 5% nível de probabilidade para o teste F. Um ponto de valor significante no teste F a um efeito do gene, significa que não é apenas como mera presença ou posição do gene que está causando as diferenças nos fenotipos.

Exemplo 26: Resultados da avaliação

As plantas da área acima do solo (ou biomassa da folha) foi determinada pela contagem do número total de pixéis nas imagens digitais a partir da parte acima do solo das plantas discriminadas a partir do fundamento. Este valor foi medido para as figuras retiradas no mesmo ponto a partir dos ângulos diferentes e foi convertido a um valor de superfície física expressada no quadrado mm pela calibração. Os experimentos mostram que as plantas da área acima do solo medidas desta maneira correlacionadas com a biomassa das partes das plantas acima trituradas. A área triturada acima é o ponto de tempo em que as plantas tem atingido esta biomassa de folha máxima.

As panículas primárias maduras foram endurecidas, contadas, empacotadas, *barcode-labeled* e então secadas por três dias em um forno a 37°C. As panículas foram então refugadas e todas as sementes foram colhidas e contadas. A casca enchedora foi separada a partir de um despejo usando dispositivo de sopro de ar. As cascas despejadas foram descartadas e a fração remanescente foi contada novamente. A casca enchedora foi pesada em uma balança analítica. O número de sementes enchedoras foi determinada pela contagem o número de cascas enchedoras que permaneceram após a etapa de separação. A produção total da semente foi medida pelo pesagem de todas as cascas enchedoras endurecidas a partir de uma planta.

Como apresentado nas Tabelas W a Y, a biomassa acima do solo, os produção da semente e o número de sementes enchedoras são aumentadas nas plantas transgênicas com aumento expressão de um ácido nucleico codificando uma proteína GLK, comparadas as planta adequadas de controle. Os resultados a partir da geração T1 e T2 foram mostradas.

Tabela W mostra um aumento na biomassa acima do solo em porcentagem, bem como a relevância estatística deste aumento de acordo com um teste F, na geração T1 e T2 de arroz transgênico com aumento expressão de um ácido nucleico codificando uma proteína GLK.

Tabela W:

Biomassa acima do solo		
	% da diferença	**Valor P do teste F**
Geração T1	12	0,0018
Geração T2	27	0,0000

Tabela X mostra um aumento na produção total da semente (peso da semente total) em porcentagem, bem como a relevância estatística deste aumento de acordo com um teste F, na geração T1 e T2 de arroz transgênico com aumento expressão de um ácido nucleico codificando uma proteína GLK.

Tabela X:

Produção da semente		
	% da diferença	**Valor P do teste F**
Geração T1	20	0,0048
Geração T2	20	0,0160

Tabela Y mostra um aumento no número de sementes enchedoras em porcentagem, bem como a relevância estatística deste aumento de acordo com um teste F, na geração T1 e T2 de arroz transgênico com aumento expressão de um ácido nucleico codificando uma proteína GLK.

Tabela Y:

Número de sementes enchedoras		
	% da diferença	**Valor P do teste F**
Geração T1	22	0,0021
Geração T2	22	0,0065

REV ΔHDZip/START

Exemplo 27: Identificação das seqüências que diz respeito à seqüência de ácido nucleico usando os métodos da invenção

As seqüências (cDNA de comprimento total, ESTs ou genômico) que diz respeito à seqüência de ácido nucleico usado nos métodos da presente invenção foram identificados entre aqueles mantidos na base de dados de nucleotídeos Entrez no National Center for Biotechnology Information usando ferramentas de busca da seqüência de base de dados, tal como a ferramenta de alinhamento de local básico (BLAST) (Altschul *et al.* (1990) J. Mol. Biol. 215:403-410; e Altschul *et al.* (1997) Nucl. Acids Res. 25:3389-3402). Os programas são usados para observar regiões de similaridade local entre seqüências pela comparação com as seqüências de ácido nucleico ou polipeptídeo pela base de dados da seqüência e pelo cálculo de uma significância estatística dos pontos. Os polipeptídeo codificado pela seqüência de ácido nucleico da presente invenção como usado pelo algoritmo TBLASTN, com falta de direção e o filtro para ignorar seqüências de complexidade inferiores apresentadas. A produção de uma análise foi vista pela comparação em forma de par, e classificada de acordo com a contagem de probabilidade (Valor E), onde a contagem reflita na probabilidade que um alinhamento particular ocorra pela escolha (o valor E inferior, o mais significante do acerto). Além disso a comparação aos valores E também foram contados pela identidade de porcentagem. Identidade de porcentagem refere-se à diversos nucleotídeos idênticos (ou aminoácidos) entre as seqüências de dois ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. Em alguns exemplos, os parâmetros padrões pode ser ajustados para modificar a severidade da busca

Tabela Z fornece uma lista da seqüência de ácido nucleico que diz respeito à seqüência de ácido nucleico usando os métodos da presente invenção.

Tabela Z: Seqüência de ácido nucleicos que diz respeito à seqüência de ácido nucleico (SEQ ID N°: 194) usando os métodos da presente invenção, e os polipeptídeos correspondentes deduzidos.

Nome	Organismo da sonda	Ácido nucleico SEQ ID N°:	Polipeptídeo SEQ ID N°:	Número da acessão de base de dados
Orysa_ REV CTR parcial	*Oryza sativa*	194	195	Parte de AK102830 (Os10g33960)
Orysa_ REV CTR	*Oryza sativa*	196	197	Parte de AK102830 (Os10g33960)
Orysa_ Rev	*Oryza sativa*	198	199	AK102830 (Os10g33960)
Orysa_ HOX10	*Oryza sativa*	200	201	AY425991.1
Arath_ REV	*Arabidopsis thaliana*	202	203	AF188994
Zeama_ HDIII RLD1 (folha enrolada)	*Zea mays*	204	205	AY501430.1
Poptr_HDIII	*Populus trichocarpa*	206	207	AY919617
Medtr_ HDIII	*Medicago trunculata*	208	209	Unido de AC138171.17
Sacof_ HDIII Parcial	*Saccharum officinarum*	210	211	próximo do CA125167.1 CA217027.1 CA241276.1 CA124509.1
Triae_ HDIII Parcial	*Triticum aestivum*	212	213	próximo do CD905903 BM135681.1 BQ578798.1 CJ565259.1
Horvu_ HDIII Parcial	*Hordeum vulgare*	214	215	coleção de BU996988.1 BJ452342.1 BJ459891.1
Phypr_ HDIII Parcial	*Phyllostachys praecox*	216	217	DQ013803
Orysa_REV	*Oryza sativa*	223		Variante da SEQ ID N°: 196
Brara Revoluta	*Brassica rapa*	224	225	AC189324.1
Ginbi Revoluta	*Ginkgo biloba*	226	227	DQ385525
Gosba Revoluta	*Gossypium barbadense*	228	229	AY966446.1
Lyces Revoluta	*Lycopersicon esculentum*	230	231	BT013577

Exemplo 28: Determinação da similaridade e identidade globais entre o CTR dos polipeptídeos REV.

As porcentagens globais de similaridade e identidade entre o

CTR dos polipeptídeos REV foram determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Ferramenta de alinhamento global da matriz) (BMC Bioinformatics. 2003 4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando proteína ou seqüências de DNA. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). O software MatGAT que gera matrizes de similaridade/identidade por DNA ou seqüência das proteínas sem necessidade do pré-alinhamento dos dados. Os programas realizam uma série de alinhamentos em formas de pares usando o algoritmo Myers e Miller de alinhamento global (com a penalidade aberta na fenda de 12, e a penalidade da extensão da fenda de 2), cálculos de similaridade e identidade usando por exemplo Blosum 62 (para polipeptídeos), e então colocando o resultado na matriz da distância. A similaridade de seqüência é mostrada na metade funda da linha que divide e a identidade da seqüência é mostrada na metade superior da linha que divide diagonal.

Os parâmetros usados na comparação foram:

Contagem da matriz: Blosum62

Primeira fenda: 12

Fenda estendida: 2

Os resultados da análise do software são mostrados na Tabela C AA para similaridade e identidade globais entre o CTR dos polipeptídeos REV. Identidade de porcentagem é dado acima a diagonal e a similaridade da porcentagem é dado abaixo a diagonal. Identidade de porcentagem entre o CTR das faixas de parálogos de polipeptídeo REV e ortólogos entre 30 e 70%, refletindo a identidade inferior da conservação da seqüência entre estes, fora dos domínios HDZip e START.

Tabela AA: Os resultados MatGAT para a similaridade e identidade globais entre o CTR dos polipeptídeo REV ortólogos e parálogos.

Similaridade e identidade globais no CTR de Polipeptídeo REV ortólogos e parálogos	1	2	3	4	5	6	7	8	9	10
1. CTR_Horvu_HDIII		61,1	57,2	81,1	88,2	89,7	60	79,6	96,5	79,8
2. CTR_Arath_REV	76,9		70	61,8	63	63,2	73,3	62,5	62,9	63,4
3. CTR_Medtr_REV	73,2	86,5		59,2	60,3	60,8	70,2	59,4	58,6	60,3
4. CTR_Orysa_HOX10	89	80,7	76,6		83,9	84,8	59,8	88,2	83	89,7
5. CTR_Orysa_REV	93,5	79,6	76,2	92,5		92,7	61,9	83,2	90,7	83,9
6. CTR_Phypr_HDIII	93,3	79,8	76,2	92,7	96,3		61,9	84,5	91,8	84,2
7. CTR_Poptr_HDIII	75,1	86,4	82,8	76,1	76,1	76,5		60,4	61,6	61,3
8. CTR_Sacof_HDIII	88,6	80,3	75,9	94,2	91,8	91,8	77,4		81,9	94,8
9. CTR_Triae_HDIII	97,4	79	74,9	90,8	95,7	95,1	76,3	90,5		81,8
10. CTR_Zeama_HDIII_LRD1	88,8	81,3	77,8	95,7	92,5	92,3	77,8	97	90,6	

Todos os polipeptídeos REV que compreendem um CTR tendo, um aumento a fim da preferência, pelo menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ou 98% da seqüência da identidade ao CTR de um polipeptídeo REV como representado pela SEQ ID N°: 197.

Exemplo 29: Clonagem do gene

Manipulação de DNA: a não ser de outra maneira estado, as técnicas de DNA recombinantes são realizadas de acordo com os protocolos padrões descritos em (Sambrook (2001) Molecular Cloning: a laboratory manual, 3° Edition, Cold Spring Harbor Laboratory Press, CSH, New York) ou nos Volumes 1 e 2 de Ausubel et al. (1994), Current Protocols in Molecular Biology, Current Protocols. Materiais padrões e métodos para trabalhar a planta molecular são descritos em Plant Molecular Biology Labfax (1993) by R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (UK) e Blackwell Scientific Publications (UK).

O gene de comprimento total *Oryza sativa* Orysa_REV da SEQ ID N°: 198 foi amplificado por PCR usando como padrão, uma biblioteca de cDNA *Oryza sativa* feita por encomenda (Invitrogen, Paisley, UK). Após a transcrição reversa de RNA extraído a partir de mudas, os cDNAs foram clonados em pCMV Sport 6.0. Após a extração do plasmídio, 200 ng de padrão como usado em uma mistura de 50 µl de PCR. Os iniciadores prm01983 (SEQ ID N°: 221; sentido, códon iniciado em negrito,

AttB1 local em itálico:

5'-*ggggacaagtttgtacaaaaaagcaggcttaaaca*atggcggcggcggtgg-3')

e prm01984 (SEQ ID N°: 222; reverso, complementar, AttB2 local em itálico:

5'-*ggggaccactttgtacaagaaagctgggt*ggattttgggtcacacgaaggacca -3'),

que inclui os locais AttB pela recombinação Gateway, foram usado pela amplificação de PCR. PCR foi realizado usando polimerase de DNA Hifi Taq em condições padrões. O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante pelo qual o fragmento PCR recombina com o plasmídio pDONR201 para fornecer, de acordo com a tecnologia Gateway, um "clone de entrada", p04562. O plasmídio pDONR201 foi adquirido de Invitrogen, como parte da tecnologia Gateway®.

O CTR parcial *Oryza sativa* (REV ΔHDZip/START) do gene Orysa_REV da SEQ ID N°: 194 foi amplificado por PCR como acima, com iniciadores prm03263 (SEQ ID N°: 219:

5'ggggacaagtttgtacaaaaaagcaggcttgtgctaaggcatccatgctac3')

e prm03264 (SEQ ID N°: 220:

5'ggggaccactttgtacaagaaagctgggtgcaccttccatgctacagcttg3').

Após a clonagem, o número do clone de entrada resultante foi p04436.

Exemplo 30: Construção do Vetor

O clone de entrada p04562 e p04436 foram subseqüentemente usados em uma reação LR com p01519, um vetor de destinação usado pela transformação da Oryza sativa para o construto de grampo de cabelo. Este vetor contém como elementos funcionais dentro dos limitadores de DNA T: um marcador de planta selecionável; um marcador de expressão de cassete avaliável; e dois cassetes Gateway clonados como repetição invertida e separados por um MAR (fragmento ao redor de 300 bp de uma região de ligação de matriz *Nicotiana tabacum*), pretendido pela recombinação LR in

vivo tal que a seqüência de interesse do clone de entrada é integrada tanto nas orientações de sentido quanto de anti-sentido para formar a estrutura secundária de grampo de cabelo. Um promotor GOS2 de arroz (SEQ ID N°: 58) para a expressão constitutiva do construto genético (PRO0129) foi localizado a montante destes cassetes Gateway.

Após a etapa de recombinação LR, o vetor resultante da expressão, p0443 (com o parcial CTR de Orysa_REV; SEQ ID N°: 194) e p0448 (com o comprimento total Orysa_REV; que compreende a SEQ ID N°: 198) (ver Figura 20) foram separadamente transformados na cepa *Agrobacterium* LBA4044, e subseqüentemente separadamente por plantas Oryza sativa. As plantas de arroz transformadas foram deixadas desenvolver-se e foram então examinadas para os parâmetros descritos no exemplo 31.

Exemplo 31: Procedimento de avaliação

31.1 Apresentação da avaliação

Aproximadamente 15 a 20 transformantes de arroz T0 independentes foram gerados. Os transformantes primários foram transferidos a partir de uma câmara de cultura de tecido para uma estufa para plantas para crescimento e endurecimento da semente T1. Cinco eventos para o construto de grampo de cabelo compreende o comprimento total Orysa_REV e seis eventos para o construto de grampo de cabelo compreendem o CTR parcial do Orysa_REV, pelo qual a produto T1 segregado 3:1 para a presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 mudas de T1 contendo o transgene (hetero e homozigotos) e aproximadamente 10 mudas de T1 que precisam do transgene (nulizigotos) foram selecionados pelo monitoramento da expressão do marcador visual. As plantas T1 selecionadas foram transferidas para uma estufa para plantas. Cada uma das plantas recebeu um único rótulo com código de barra para não ligar-se ambiguamente os dados dos fenotipos às plantas correspondentes. As plantas T1 selecionadas foram desenvolvidos em solo em potes de 10 cm de

diâmetro sob os ajustes ambientais seguintes: foto-período= 11,5 horas, intensidade da luz do dia = 30,000 lux ou mais, temperatura do período do dia = 28°C ou maior, temperatura do período noturno = 22°C, umidade relativa = 60 a 70%. Plantas transgênicas e o nulizigotos correspondentes (plantas de controle) foram desenvolvidas lado a lado em posições aleatórias.

Cinco eventos T1 ainda foram avaliados (se resultados positivos foram obtidas na primeira avaliação) na geração T2 seguindo os mesmos procedimentos de avaliação como para a geração T1 mas com mais indivíduos por evento. A partir do estágio da semeadura até o estágio da maturidade as plantas foram passadas diversas vezes através de uma cabine formadora de imagem digital. Em cada um ponto de tempo as imagens digitais (2048 x 1536 pixéis, 16 milhões de cores) foram tiradas de cada planta de pelo menos 6 ângulos diferentes.

31.2 Análise estatística: Teste F

Um dos dois fatores ANOVA (análise de variantes) como usado como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um Teste F foi realizado em todos os parâmetros medido de todas as plantas de todos os eventos transformados com o gene da presente invenção. O teste F foi realizado por verificar um efeito do gene em todos os eventos de transformação e para verificar um efeito total do gene, também conhecido como um efeito do gene global. O princípio para a significância por um efeito de gene verdadeiro global foi apresentado a 5% de nível de probabilidade para o teste F. Um ponto de valor significante no teste F a um efeito do gene, significa que não é apenas como mera presença ou posição do gene que está causando as diferenças nos fenotipos.

Exemplo 32: Resultados da avaliação

As plantas da área acima do solo (ou biomassa da folha) foi determinada pela contagem do número total de pixéis nas imagens digitais a partir da parte acima do solo das plantas discriminadas a partir do

fundamento. Este valor foi medido para as figuras retiradas no mesmo ponto a partir dos ângulos diferentes e foi convertido a um valor de superfície física expressada no quadrado mm pela calibração. Os experimentos mostram que as plantas da área acima do solo medidos desta maneira correlacionada com a biomassa das partes das plantas acima trituradas. A área triturada acima é o ponto de tempo em que as plantas tem atingido esta biomassa de folha máxima.

As partes das plantas abaixo trituradas (no caso essencialmente as raízes) foram determinada pelo crescimento das plantas em especial potes projetado com fundo transparentes permitindo a visualização das raízes. Uma câmara digital registrou as imagens através do fundo dos potes durante o crescimento da planta.

As características da raiz tal como total da área projetada (que pode ser da cor que diz respeito ao volume total da raiz), diâmetro médio e comprimento de raízes acima em certos princípios de espessadores (comprimento dos espessadores das raízes, ou comprimento do espessador da raiz) foram deduzidos das figuras usando o software apropriado.

As panículas primárias maduras foram endurecidas, contadas, empacotadas, rotuladas com código de barra e então secadas por três dias em um forno a 37°C. As panículas foram então refugadas e todas as sementes foram colhidas e contadas. A casca enchedora foi separada a partir de um despejo usando dispositivo de sopro de ar. As cascas despejadas foram descartadas e a fração remanescente foi contada novamente. A casca enchedora foi pesada em uma balança analítica. O número de sementes enchedoras foi determinada pela contagem do número de cascas enchedoras que permaneceram após a etapa de separação. A produção total da semente foi medida pela pesagem de todas as cascas enchedoras endurecidas a partir de uma planta. O número de semente total por plantas foi medido pela contagem do número de cascas endurecidas a partir de uma planta. Peso

singular da semente (TKW) é extrapolado a partir do número de sementes enchedoras contadas e seu peso total. Os parâmetros das sementes individuais (incluindo altura, comprimento, área, peso) foram medidos usando um dispositivo feito sob encomenda que consiste de dois principais componentes, um dispositivo de pesagem e imagem, ligados ao software pela análise da imagem. O índice colhido (HI) na presente invenção é definido como uma razão entre a produção total da semente e a área triturada acima (mm^2), multiplicado por um fator 10^6. O número total de flores por panícula como definido na presente invenção é a razão entre o número total de sementes e o número de panículas primárias maduras. A taxa enchedora da semente como definido na presente invenção é uma proporção (expressada como uma%) de diversas sementes enchedoras no número total de sementes (ou flósculos).

32.1 Medição da produção que diz respeito aos parâmetros para trasnformantes reduzidos com a expressão de um gene REV endógeno usando uma seqüência de ácido nucleico REV ΔHDZip/START como representado pela SEQ ID N°: 194

Como apresentado nas Tabelas BB a EE, a produção da semente, o número de sementes enchedoras, a taxa enchedora da semente e o índice colhido são aumentadas nas plantas transgênicas reduzindo a expressão de um gene REV endógeno usando uma seqüência de ácido nucleico REV ΔHDZip/START comparadas as plantas de controle. Os resultados a partir da geração T1 e T2 foram mostradas.

A Tabela BB mostra um número de eventos transgênicos com um aumento na produção total da semente (peso da semente total), as porcentagens deste aumento, bem como a relevância estatística deste aumento de acordo com um teste F.

Tabela BB: Diversos eventos transgênicos com um aumento na produção da semente, as porcentagens do aumento, e valor P do teste F na geração T1 e T2 de arroz transgênico com redução na expressão de um gene REV endógeno

usando uma seqüência de ácido nucleico REV ΔHDZip/START.

Semente peso			
	Diversos eventos mostrando um aumento	**% da diferença**	**Valor P do teste F**
Geração T1	4 fora de 6	16	0,058
Geração T2	3 fora de 4	4	0,0133

Tabela CC mostra um número de eventos transgênicos com um aumento de diversos de sementes enchedoras, as porcentagens deste aumento, bem como a relevância estatística deste aumento de acordo com um teste F.

Tabela CC: Diversos eventos transgênicos com um aumento no número de sementes enchedoras, as porcentagens do aumento, e valor P do teste F na geração T1 e T2 de arroz transgênico com redução na expressão de um gene REV endógeno usando uma seqüência de ácido nucleico REV ΔHDZip/START.

Número de sementes enchedoras			
	Diversos eventos mostrando um aumento	**% da diferença**	**Valor P do teste F**
Geração T1	4 fora de 6	14	0,0884
Geração T2	3 fora de 4	4	0,0156

Tabela DD mostra um número de eventos transgênicos com um aumento na taxa enchedora da semente, as porcentagens deste aumento, bem como a relevância estatística deste aumento de acordo com um teste F.

Tabela DD: Diversos eventos transgênicos com um aumento na taxa da semente enchedora, as porcentagens do aumento, e valor P do teste F na geração T1 e T2 de arroz transgênico com redução na expressão de um gene REV endógeno usando uma seqüência de ácido nucleico REV ΔHDZip/START.

Taxa da semente enchedora			
	Diversos eventos mostrando um aumento	**% da diferença**	**Valor P do teste F**
Geração T1	5 fora de 6	19	0,0002
Geração T2	4 fora de 4	22	<0,0001

A Tabela EE mostra um número de eventos transgênicos com um aumento no índice colhido, as porcentagens deste aumento, bem como a relevância estatística deste aumento de acordo com um teste F.

Tabela EE Diversos eventos transgênicos com um aumento no índice colhido, as porcentagens do aumento, e valor P do teste F na geração T1 e T2 de arroz transgênico com redução na expressão de um gene REV endógeno usando uma seqüência de ácido nucleico REV ΔHDZip/START.

Índice colhido			
	Diversos eventos mostrando um aumento	**% da diferença**	**Valor P do teste F**
Geração T1	4 fora de 6	14	0,0284
Geração T2	3 fora de 4	9	0,0187

Dois parâmetros adicionais foram medidos por apenas uma avaliação:

1) O comprimento médio da semente individual

2) O espessador médio da raiz

Como mostrado na tabela FF, tanto o comprimento médio da semente individual quanto o espessador médio da raiz foram significantemente aumentados em arroz transgênicos com a redução da expressão de um gene REV endógeno usando uma seqüência de ácido nucleico REV ΔHDZip/START comparada as plantas de controle.

Tabela FF: Diversos eventos transgênicos com um aumento no índice colhido e valor P do teste F na geração simples de arroz transgênico com a redução da expressão de um gene REV endógeno usando uma seqüência de ácido nucleico REV ΔHDZip/START.

	Diversos eventos mostrando um aumento	**Valor P do teste F**
Comprimento médio da semente individual (sementes T2)	5 fora de 6	0,009
Espessador médio de raiz (plantas T2)	4 fora de 4	<0,0001

32.2 Medição da produção que diz respeito aos parâmetros para os transformantes com a redução na expressão de um polipeptídeo REV endógeno usando um ácido nucleico que codifica o comprimento total do polipeptídeo Orysa_ REV, como representado pela SEQ ID N°: 198:

O mesmo procedimento de avaliação como descrito neste acima foi realizado para arroz transgênico tendo a redução da expressão de

um polipeptídeo REV endógeno usando uma seqüência de ácido nucleico que codifica o comprimento total de polipeptídeo Orysa_ REV. Todos os parâmetros medidos foram negativos levemente e significantemente, como mostrado na Tabela GG.

Tabela GG: Diversos eventos transgênicos com uma diminuição na biomassa acima do solo, produção total da semente, número de sementes enchedoras, número de flores por panícula, índice colhido, diversas panículas primárias e altura da planta, as porcentagens do aumento, e valor P do teste F na geração T1 de arroz transgênico com redução da expressão de um gene REV endógeno usando uma seqüência de ácido nucleico que codifica o comprimento total de polipeptídeo Orysa_ REV.

Parâmetro	Diversos eventos que mostram uma diminuição	% da diferença	Valor P do teste F
Biomassa acima do solo	5 fora de 5	-40	<0,0001
Produção total da semente	5 fora de 5	-65	<0,0001
Número de sementes enchedoras	5 fora de 5	-64	<0,0001
Número de flores por panícula	5 fora de 5	-39	<0,0001
Índice colhido	5 fora de 5	-53	<0,0001
Diversas panículas primárias	5 fora de 5	-39	<0,0001
Peso da planta	5 fora de 5	-13	0,0006

Pelas regiões conservadas incluindo nos construtos de grampo de cabelo (tal como os domínios HDZip e START) para redução da expressão REV endógena do gene, a expressão de outros genes de classe III HDZip também pode ser reduzido.

CLE

Exemplo 33: Identificação das seqüências que diz respeito à seqüência de ácido nucleico usando os métodos da invenção

As seqüências (cDNA de comprimento total, ESTs ou genômico) que diz respeito à seqüência de ácido nucleico usando os métodos da presente invenção foram identificados entre aqueles mantidos na base de dados de nucleotídeos Entrez no National Center for Biotechnology

Information usando ferramentas de busca da seqüência de base de dados, tal como a ferramenta de alinhamento de local básico (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucl. Acids Res. 25:3389-3402). Os programas são usados para observar regiões de similaridade local entre seqüências pela comparação com as seqüências de ácido nucleico ou polipeptídeo pela base de dados da seqüência e pelo cálculo de uma significância estatística dos pontos. Os polipeptídeo codificados pela seqüência de ácido nucleico da presente invenção como usado pelo algoritmo TBLASTN, com falta de direção e o filtro para ignorar seqüências de complexidade inferiores apresentadas. A produção de uma análise foi vista pela comparação em forma de par, e classificada de acordo com a contagem de probabilidade (Valor E), onde a contagem refletida da probabilidade que um alinhamento particular ocorra pela escolha (o valor E inferior, o mais significante do acerto). Além disso a comparação aos valores E também foram contados pela identidade da porcentagem. Identidade de porcentagem refere-se à diversos nucleotídeos idênticos (ou aminoácidos) entre as seqüências de dois ácidos nucleicos comparados (ou polipeptídeo) em um comprimento particular. Em alguns exemplos, os parâmetros padrões podem ser ajustados para modificar a severidade da busca

A Tabela HH fornece uma lista da seqüência de ácido nucleico que diz respeito à seqüência de ácido nucleico usando os métodos da presente invenção.

Tabela HH: Seqüência de ácido nucleico que diz respeito à seqüência de ácido nucleico (SEQ ID N°: 232) usado os métodos da presente invenção, e os polipeptídeos correspondentes deduzidos.

Fonte da planta	**Ácido nucleico SEQ ID N°:**	**Proteína SEQ ID N°:**
Saccharum officinarum semelhante a CLE	232	233
Populus trichocarpa x Populus deltóides semelhante a CLE	239	240
Oryza sativa semelhante a CLE	241	242
Saccharum officinarum semelhante a CLE	243	244
Arabidopsis thaliana semelhante a CLE	245	246
Brassica napus semelhante a CLE	247	248
Arabidopsis thaliana CLAVATA3	249	250

Exemplo 34: Clonagem do gene

O gene semelhante a CLE de cana de açúcar foi amplificado por PCR com iniciadores prm05843 (SEQ ID N°: 234; sentido, códon iniciado em negrito, AttB1 local em itálico:

5'-*ggggacaagtttgtacaaaaaagcaggcttaaaca*atgaggatgttcttccgg-3')

e prm05844 (SEQ ID N°: 235; reverso, complementar, AttB2 local em itálico: 5'-*ggggaccactttgtacaagaaagctgggt*cctctcatctgttgtggag-3'),

que inclui os locais AttB pela recombinação Gateway, foram usados pela amplificação de PCR. PCR foi realizado usando polimerase de DNA Hifi Taq em condições padrões. Os fragmentos de PCR também foram purificados usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante pelo qual o fragmento PCR recombina in vivo com o plasmídio pDONR201 para fornecer, de acordo com a tecnologia Gateway, um "clone de entrada", p066. O plasmídio pDONR201 foi adquirido de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 35: Construção do vetor

O clone de entrada p066 foi subseqüentemente usado em uma reação LR com p01519, um vetor de destinação usado pela transformação da Oryza sativa para o construto anti-sentido. Estes vetores contém como elementos funcionais dentro dos limitadores de DNA T: um marcador de planta selecionável; um marcador de expressão de cassete avaliável; e um cassete Gateway pretendido pela recombinação LR in vivo tal que a seqüência de interesse do clone de entrada é integrado na orientação de sentido ou anti-sentido. Um promotor de prolamina de arroz (SEQ ID N°: 236) pela expressão de semente específica (PRO090) foi localizado a montante deste cassete Gateway.

Após a etapa de recombinação LR, o vetor resultante da expressão, p068 (Figura 23) foi transformado em Cepa *Agrobacterium* LBA4044 e subseqüentemente por plantas Oryza sativa. As plantas de arroz transformadas foram deixadas desenvolver-se e foram então examinadas para

os parâmetros descritos no exemplo 36.

Exemplo 36: Avaliação dos métodos de plantas transformadas com orientação de anti-sentido semelhante a CLE

Aproximadamente 15 a 20 transformantes de arroz T0 independentes foram gerados. Os transformantes primários foram transferidos a partir de uma câmara de cultura de tecido para uma estufa para plantas para crescimento e endurecimento da semente T1. Seis eventos para que a progenia T1 segregada 3:1 para a presença/ausência do transgene, foram retidas. Para cada um destes eventos, aproximadamente 10 mudas de T1 contendo o transgene (hetero e homozigotos) e aproximadamente 10 mudas de T1 que precisam do transgene (nulizigotos) foram selecionados pelo monitoramento da expressão do marcador visual. As plantas T1 selecionadas foram transferidas para uma estufa para plantas. Cada uma das plantas recebeu um único rótulo com código de barras para não ligar-se ambiguamente aos dados dos fenotipos às plantas correspondentes. As plantas T1 selecionadas foram desenvolvidos em solo em potes de 10 cm de diâmetro sob os ajustes ambientais seguintes: foto-período = 11,5 horas, intensidade da luz do dia = 30,000 lux ou mais, temperatura do período do dia = 28°C, temperatura do período noturno = 22°C, umidade relativa = 60 a 70%. Plantas transgênicas e o nulizigotos correspondente foram desenvolvidos lado a lado em posições aleatórias. A partir do estágio da semeadura até o estágio da maturidade as plantas foram passadas diversas vezes através de uma cabine formadora de imagem digital. Em cada um ponto de tempo as imagens digitais (2048 x 1536 pixéis, 16 milhões de cores) foram tiradas de cada planta de pelo menos 6 ângulos diferentes.

As plantas da área acima do solo (ou biomassa da folha) foram determinadas pela contagem do número total de pixéis nas imagens digitais da parte acima do solo das plantas discriminadas a partir do fundamento. Este valor foi medido para as figuras retiradas no mesmo ponto a partir de ângulos diferentes e foi convertido a um valor de superfície física expressada no

quadrado mm pela calibração. Os experimentos mostram que as plantas da área acima do solo medido desta maneira correlacionada com a biomassa das partes das plantas acima trituradas. A Areamax é uma área triturada acima no ponto de tempo em que as plantas tem atingido esta biomassa de folha máxima.

As panículas primárias maduras foram endurecidas, empacotadas, rotuladas com código de barras e então secadas por três dias no forno a 37°C. As panículas foram então refugadas e todas as sementes colhidas. A casca enchedora foi separada a partir de um despejo usando dispositivo de sopro de ar. Após a separação, ambas porções de sementes foram então contadas usando uma máquina de contagem comercialmente disponível. As cascas despejadas foram descartadas. A casca enchedora foi pesada em uma balança analítica e a área de seção cruzada das sementes foi medida usando imagem digital. Este procedimento resultou nas seguintes semente apresentadas que diz respeito aos parâmetros:

As flores-por-panícula é um parâmetro estimado do número médio de flósculos por panícula na planta, derivado do número de sementes totais divididas pelo número das primeiras panículas. A mais alta panícula e todas as panículas que sobrepõem com a panícula mais alta então alinhada verticalmente, foram consideradas as primeiras panículas e foram contadas manualmente. O número de sementes enchedoras foi determinada pela contagem do número de cascas enchedoras que permaneceram após a etapa de separação. A produção total da semente (peso da semente total) foi medida pela pesagem de todas as cascas enchedoras endurecidas a partir de uma planta. O número de semente total por plantas foi medido pelo número de contagem de cascas endurecidas a partir de uma planta e correspondentes ao número de flósculos por plantas. Estes parâmetros foram derivados em uma maneira automática a partir das imagens digitais usando software da análise da imagem e foram analisadas estatisticamente. Os parâmetros da semente individual (incluindo altura, comprimento, área, peso) foram medidos usando um dispositivo feito por encomenda que consiste de dois principais componentes, um dispositivo de pesagem e de imagem, ligados ao software para análise da imagem. O

índice colhido na presente invenção é definido como uma razão entre a produção total da semente (g) e a área triturada acima (mm^2), multiplicado por um fator 10^6.

Um de dois fatores ANOVA (análise da variância) correlacionado ao não projeto balanceado como usado como modelo estatístico para a avaliação total das características fenotípicas das plantas. Um Teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformados pelo gene. O teste F foi realizado por verificar um efeito do gene em todos os eventos de transformação e para verificar um efeito total do gene, também denominado neste "efeito global do gene". Se o valor do teste F mostra que os dados são significantes, é concluído que aqui tem um efeito de "gene", medindo que não apenas a presença ou as posições do gene estão causando o efeito. O princípio para a significância por um efeito de gene verdadeiro global é apresentado a 5% do nível de probabilidade para o teste F.

Para verificar um efeito do gene dentro de um evento, isto é, um efeito específico de linha, um teste t foi realizado dentro de cada evento usando dados apresentados a partir das plantas transgênicas e o correspondente as plantas nulas. "Plantas nulas" ou "segregantes nulos" ou "nulizigotos" são as plantas tratadas na mesma maneira como a planta transgênica, mais de que um transgene tem segregado. As plantas nulas também podem ser descritas como as plantas negativas homozigotas transformadas. O princípio para a significância para o teste t é apresentado a 10% do nível de probabilidade. Os resultados por alguns eventos podem ser acima ou abaixo deste princípio. Este é baseado na hipótese que uma força do gene apenas tem um efeito em certas posições no genoma, e que a ocorrência deste efeito depende da posição e não é comum. Este tipo de efeito do gene também é denominado neste um "efeito do gene de linha". O valor P é obtido pela comparação com o valor t à distribuição t ou alternativamente, pela comparação com o Valor F à distribuição F. O valor P então dá as probabilidade que a hipótese de nulo (isto é, que aqui não é nenhum efeito do transgene) seja correta.

Exemplo 37: Medições da produção que diz respeito aos parâmetros para transformantes de construto de anti-sentido:

Na análise da semente como descrito acima, os inventores observaram que as plantas transformadas com o construto do gene semelhante a CLE anti-sentido tem uma alta produção da semente, expressada como número de sementes enchedoras, total peso de sementes, número total de sementes e índice colhido, comparadas como plantas que precisam do transgene semelhante a CLE. Em particular o peso da semente total e número da semente total foi significantemente aumentado na tanto na geração de plantas T1 quanto na T2.

SYR NUE

Exemplo 38: Identificação das seqüências que diz respeito a SEQ ID N°: 251 e SEQ ID N°: 252

As seqüências (cDNA de comprimento total, ESTs ou genômico) que diz respeito a SEQ ID N°: 251 e/ou seqüência da proteínas que diz respeito a SEQ ID N°: 252 foram identificadas entre aquelas mantidos na base de dados de nucleotídeos Entrez no National Center for Biotechnology Information (NCBI) usando ferramentas de busca da seqüência de base de dados, tal como a ferramenta de alinhamento de local básico (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucl. Acids Res. 25:3389-3402). Os programas são usados para observar regiões de similaridade local entre seqüências pela comparação com as seqüências de ácido nucleico ou polipeptídeo pela base de dados da seqüência e pelo cálculo da significância estatística dos pontos. O polipeptídeo codificado pela SEQ ID N°: 251 como usado pelo algoritmo TBLASTN, com falta de direção e o filtro para ignorar seqüências de complexidade inferiores apresentadas. A produção de uma análise foi vista pela comparação em forma de par, e classificada de acordo com a contagem de probabilidade (Valor E), onde uma contagem reflete nas probabilidades que um alinhamento particular ocorra pela escolha (o valor E inferior, o mais significante do acerto). Além disso a comparação aos valores E também foram contados pela identidade de porcentagem. Identidade de porcentagem refere-se à diversos nucleotídeos idênticos (ou aminoácidos) entre as seqüências de dois ácidos nucleicos comparados (ou

polipeptídeo) em um comprimento particular. Em alguns exemplos, os parâmetros padrões podem ser ajustados para modificar a severidade da busca.

Além das seqüências de ácido nucleico publicamente disponíveis em NCBI, a seqüência do proprietário com a base de dados também são buscados seguindo os mesmo procedimentos como descrito neste acima.

Tabela II fornece uma lista de ácido nucleico e seqüência da proteínas que diz respeito à seqüência de ácido nucleico como representado pela SEQ ID N°: 251 e a seqüência da proteína representada pela SEQ ID N°: 252.

Tabela II: Seqüência de ácido nucleico que diz respeito à seqüência de ácido nucleico (SEQ ID N°: 251) útil nos métodos da presente invenção, e o polipeptídeos correspondentes deduzidos.

Nome	Organismo da sonda	Polipeptídeo da SEQ ID N°:	Ácido nucleico SEQ ID N°:	Número de acessão da Base de dados	Condição
OsSYR	*Oryza sativa*	252	251	/	Comprimento total ou parcial
homólogo SYR de arroz 1	*Oryza sativa*	262	277	XP_472637	Comprimento total
homólogo SYR de arroz 2	*Oryza sativa*	263		AP008218	Comprimento total
homólogo SYR de milho	*Zea mays*	264	278	AY110705	Parcial
homólogo SYR de trigo	*Triticum aestivum*	265		/	Comprimento total
Cevada SYR homólogo	*Hordeum vulgare*	266	286	CB871444	Comprimento total
homólogo SYR de cana de açúcar 1	*Saccharum officinarum*	267	287	CA165713	Parcial
homólogo SYR de cana de açúcar 2	*Saccharum officinarum*	268	288	CA242805	Comprimento total
homólogo SYR de sorgo	*Sorgo bicolor*	269	289	CX611532	Comprimento total
homólogo AtSYR 1	*Arabidopsis thaliana*	270	290	NM_115853	Comprimento total
homólogo AtSYR 2	*Arabidopsis thaliana*	271	291	NM_180078	Comprimento total
homólogo SYR de uva	*Vitis vinifera*	272	279	CF404276	Comprimento total
homólogo SYR cítrico	*Citrus reticulata*	273	280	CF830612	Parcial
homólogo SYR de tomate 1	*Lycopersicon esculentum*	274	282	AI774560	Comprimento total
homólogo SYR de tomate 2	*Lycopersicon esculentum*	275	281	BG125370	Comprimento total

Exemplo 39: Alinhamento de seqüências de polipeptídeo relevantes

O AlignX do vetor NTI (Invitrogen) é baseado no algoritmo popular Clustal do alinhamento progressivo (Thompson *et al.* (1997) Nucl. Acids Res 25:4876-4882; Chenna *et al.* (2003). Nucl. Acids Res 31:3497-3500). Uma árvore filogenética pode ser construída usando um algoritmo de agrupamento unidos próximos. Os valores defeituosos são para a penalidade da fenda de 10, para a penalidade da extensão da fenda de 0,1 e a matriz selecionada do peso é Blosum 62 (se os polipeptídeos são alinhados).

O resultado do alinhamento da seqüência múltipla usando polipeptídeos relevantes em identificação de um útil no desempenho dos métodos da invenção é mostrado na figura 25. A repetição rica em leucina e os motivos conservados pode ser facilmente discriminados em várias seqüências.

Exemplo 40: Cálculo de identidade de porcentagem global entre seqüências de polipeptídeo úteis no desempenho dos métodos da invenção

As porcentagens globais de similaridade e identidade entre comprimento total das seqüências de polipeptídeo úteis no desempenho dos métodos da invenção foram determinadas usando um dos métodos disponível na técnica, o software MatGAT (Ferramenta de alinhamento global da matriz) (BMC Bioinformatics. 2003 4:29. MatGAT: uma aplicação que gera matrizes de similaridade/identidade usando proteína ou seqüências de DNA. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). O software MatGAT que gera matrizes de similaridade/identidade por DNA ou seqüência da proteína sem necessidade do pré-alinhamento dos dados. Os programas realizam uma série de alinhamentos em formas de pares usando o algoritmo Myers e Miller de alinhamento global (com a penalidade aberta na fenda de 12, e a penalidade da extensão da fenda de 2), cálculos de similaridade e identidade usando por exemplo Blosum 62 (para polipeptídeos), e então colocando o resultado na matriz de distância. A

similaridade de seqüência é mostrada na metade funda da linha que divide e a identidade da seqüência é mostrada na metade superior da linha que divide diagonal.

Os parâmetros usados na comparação foram:

Contagem da matriz: Blosum62

Primeira fenda: 12

Fenda estendida: 2

Os resultados da análise do software são mostradas na Tabela C JJ para similaridade e identidade globais no comprimento total das seqüências de polipeptídeo (excluindo as seqüências de polipeptídeos parciais). Identidade de porcentagem é dado acima a diagonal e a similaridade da porcentagem é dado abaixo a diagonal.

A identidade de porcentagem entre as seqüências de polipeptídeos úteis no desempenho os métodos da invenção podem estar como inferior a 27% da identidade de aminoácido comparada as SEQ ID N°: 252.

Tabela JJ: Os resultados MatGAT para a similaridade e identidade globais no comprimento total das seqüências do polipeptídeo.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. SEQ ID2		29,8	46,8	55,2	67,0	66,1	66,7	71,4	63,6	36,8	34,6	35,5	39,7	39,0	41,0	27,6	32,1
2. SEQ ID12	40,4		29,8	23,0	26,8	28,1	23,6	25,3	28,7	30,3	28,1	30,9	32,0	28,1	24,7	16,3	17,4
3. SEQ ID13	57,9	39,3		42,9	46,0	47,6	44,4	47,6	45,2	31,9	33,3	33,1	34,1	37,3	34,1	24,8	28,3
4. SEQ ID14	59,0	32,0	50,8		57,1	55,4	77,4	77,4	83,2	25,4	26,7	26,6	30,2	32,2	33,3	21,6	23,9
5. SEQ ID15	80,9	41,0	57,9	69,1		89,1	63,4	67,9	66,1	36,9	31,9	33,1	40,5	37,3	40,9	24,8	27,9
6. SEQ ID16	79,1	38,2	59,5	65,5	95,5		61,6	66,1	62,5	36,4	32,6	36,0	40,5	38,8	38,2	24,0	28,8
7. SEQ ID17	69,5	34,8	57,1	78,1	72,7	69,1		94,9	81,3	30,8	29,6	31,7	34,1	34,7	39,4	25,5	29,0
8. SEQ ID18	74,3	37,1	60,3	80,0	77,3	73,6	94,9		85,0	33,1	31,9	33,8	36,5	37,3	42,4	28,2	32,0
9. SEQ ID19	69,2	39,3	56,3	86,0	78,2	74,5	84,1	88,8		36,9	32,6	36,7	38,1	39,8	40,2	28,8	29,6
10. SEQ ID20	54,6	41,6	56,9	46,2	57,7	60,8	50,0	53,1	54,6		66,2	46,9	51,9	44,3	42,7	26,3	26,9
11. SEQ ID21	51,9	44,4	56,3	47,4	54,8	54,8	50,4	53,3	52,6	77,8		49,0	46,8	41,1	39,3	28,7	27,2
12. SEQ ID22	54,0	43,8	54,7	45,3	53,2	54,0	49,6	51,8	54,7	65,5	65,5		61,9	45,1	40,3	24,0	22,9
13. SEQ ID23	58,7	45,5	55,6	50,0	60,3	59,5	54,8	57,1	63,5	66,9	66,7	77,7		53,8	44,4	27,0	27,6
14. SEQ ID24	61,9	42,7	57,9	55,1	58,5	63,6	61,0	63,6	62,7	66,9	64,4	68,3	77,0		73,7	27,9	29,4
15. SEQ ID25	62,9	35,4	50,0	53,3	60,0	58,2	66,7	69,7	61,7	56,2	54,8	54,7	60,3	73,7		36,7	38,6
16. SEQ ID34	45,7	25,3	38,1	38,1	39,1	40,0	45,5	48,5	44,9	40,0	40,7	36,0	41,3	41,5	56,3		42,0
17. SEQ ID35	50,5	30,3	45,2	40,0	46,4	44,5	47,5	50,5	45,8	34,6	42,2	36,7	40,5	42,4	55,2	57,7	

Exemplo 41: Predição de topologia das seqüências de polipeptídeos úteis no desempenho dos métodos da invenção (transmembrana de localização subcelular ...)

O alvo P 1.1 prediz a posição sub-celular de proteínas eucarióticas. A indicação da posição é baseada na presença predita de qualquer das pré-seqüências do terminal N: peptídeo de passagem de cloroplasto (cTP), peptídeo de alvejamento mitocondrial (mTP) ou peptídeo de sinal de caminho secretor (SP). As contagens na predição final não são baseadas realmente nas probabilidades, e estas não são necessariamente adicionadas por uma. Entretanto, a posição com a maior contagem é mais semelhante de acordo com alvo P, e a conexão entre a contagens (a classe da confiança) pode ser uma indicação de que certas predições são. A faixa da classe de confiança (RC) de 1 a 5, onde 1 indica a predição mais forte. O alvo P é mantido ao servidor ao Technical University of Denmark.

Para uma seqüência predita que contém uma pré-seqüência de terminal N um local de clivagem potencial também pode ser predito.

Diversos parâmetros foram selecionados, tal como grupo de organismo (não planta ou planta), cortes apresentados (nenhum, cortes pré-definidos apresentados, ou cortes apresentados de uso específico), e o cálculo da predição dos locais de clivagem (sim ou não).

Os resultados da análise do alvo P 1.1 da seqüência do polipeptídeo como representado pela SEQ ID N°: 252 são representados na tabela KK. O grupo de organismo da "planta" foi selecionado, nenhum corte definido, e os comprimentos preditos do peptídeo de passagem requeridos. A localização sub-celular as seqüência de polipeptídeo como representado pela SEQ ID N°: 252 pode ser a mitocôndria; entretanto será notado que a classe de confiança é 5 (isto é a classe de confiança mais inferior).

Tabela KK: Análise do alvo P 1.1 da seqüência do polipeptídeo como representado pela SEQ ID N°: 252

Comprimento (AA)	105
Peptídeo de passagem cloroplástico	0,025
Peptídeo de passagem mitocondrial	0,552
Peptídeo de sinal de caminho secretor	0,009
Outros alvejamentos sub-celulares	0,416
Localização predita	Mitocôndrio
Classe de confiança	5

Dois domínios da transmembrana são identificados pelo programa TMHMM, hospedado no servidor do Center for Biological Sequence Analysis, Technical University of Denmark. Os probabilidade que o terminal N é localizado dentro é de 0,997. Os detalhes adicionais na orientação são dados na Tabela LL:

Tabela LL: Resultados de TMHMM 2.0

Orientação	Resíduo ínicio - final
Dentro	1 42
Tmhelix	43 65
Fora	66 74
Tmhelix	75 92
Dentro	93 105

Muitos outros algorítmos podem ser usados para realizar um tal análise, incluindo:

- Cloro P 1.1 hospedado no servidor de Technical University of Denmark;
- Protein Prowler Subcellular Localisation Predictor versão 1.2 hospedado no servidor de Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia;
- PENCE Proteome Analyst PA-GOSUB 2.5 hospedado no servidor de University of Alberta, Edmonton, Alberta, Canada;

Exemplo 42: Clonagem do gene

Manipulação de DNA: a não ser de outra maneira estado, as técnicas de DNA recombinantes são realizadas de acordo com os protocolos padrões descrito em (Sambrook (2001) Molecular Cloning: a laboratory manual, 3° Edition, Cold Spring Harbor Laboratory Press, CSH, New York)

nos Volumes 1 e 2 de Ausubel *et al.* (1994), Current Protocols in Molecular Biology, Current Protocols. Materiais padrões e métodos para trabalhar a planta molecular são descritos em Plant Molecular Biology Labfax (1993) por R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (UK) e Blackwell Scientific Publications (UK).

O gene *Oryza sativa SYR* foi amplificado por PCR usando como padrão, uma biblioteca de cDNA de muda *Oryza sativa* (Invitrogen, Paisley, UK). Após a transcrição reversa de RNA extraído das mudas, os cDNAs foram clonados em pCMV Sport 6.0. O tamanho médio inserido do banco foi 1,5 kb e diversos clones originais foi a fim de $1,59 \times 10^7$ cfu. O título original foi determinado a ser $9,6 \times 10^5$ cfu/ml após a primeira amplificação de 6×10^{11} cfu/ml. Após a extração de plasmídio, 200 ng de padrão como usado em uma mistura de 50 µl de PCR. Os iniciadores prm08170 (SEQ ID N°: 253; sentido, códon iniciado em negrito, AttB1 local em itálico:

5'-*ggggacaagtttgtacaaaaaagcaggcttaaaca*atggaaggtgtaggtgctagg-3')

e prm08171 (SEQ ID N°: 254; reverso, complementar, AttB2 local em itálico:

5'-*ggggaccactttgtacaagaaagctgggt*caaaaacaaaaataaattcccc-3'),

que inclui os locais AttB pela recombinação Gateway, foram usado pela amplificação PCR. PCR foi realizado usando polimerase de DNA Hifi Taq em condições padrões. Um fragmento de PCR do tamanho correto foi amplificado e purificado também usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante pelo qual o fragmento PCR recombina in vivo com o plasmídio pDONR201 para fornecer, de acordo com a tecnologia Gateway, um "clone de entrada", pSYR. O plasmídio pDONR201 foi adquirido de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 43: Construção do vetor

O clone de entrada pSYR foi subseqüentemente usado em uma

reação LR com um vetor de destinação usado pela transformação da Oryza sativa. Este vetor contém como elementos funcionais dentro dos limitadores de DNA T: um marcador de planta selecionável; um marcador de expressão de cassete avaliável; e um cassete Gateway pretendido pela recombinação LR in vivo com uma seqüência de interesse já clonado no clone de entrada. Um promotor de arroz GOS2 (SEQ ID N°: 58) para a expressão constitutiva foi localizado a montante deste cassete Gateway. Um construto de vetor similar foi preparado, mas com o promotor de proteína do grupo de maior mobilidade (HMGP, SEQ ID N°: 283 ou SEQ ID N°: 293) em vez do promotor GOS.

Após a etapa de recombinação LR, o vetor resultante da expressão, pGOS2::SYR (com o promotor GOS2) e pHMGP::SYR (com o promotor HMGP), ambos para a expressão constitutiva SYR (Figura 25) foram transformados em Cepa *Agrobacterium* LBA4044 e subseqüentemente por plantas Oryza sativa.

Exemplo 44: Avaliação dos métodos de plantas transformadas com SYR sob o controle do promotor de arroz GOS2 ou do promotor HMGP

44.1 Avaliação apresentada

Aproximadamente 15 a 20 transformantes de arroz T0 independentes foram gerados. Os transformantes primários foram transferidos a partir de uma câmara de cultura de tecido para uma estufa para plantas para crescimento e endurecimento da semente T1. Oito eventos, pelo qual um produto T1 segregado 3:1 para a presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 mudas de T1 contendo o transgene (hetero e homo zigotos) e aproximadamente 10 mudas de T1 que precisam do transgene (nulizigotos) foram selecionados pelo monitoramento expressão do marcador visual. As plantas T1 selecionadas foram transferidas para uma estufa para plantas. Cada planta recebeu um único rótulo com código de barras para não ligar-se ambiguamente aos dados dos fenotipos das plantas correspondentes. As plantas T1 selecionadas foram

desenvolvidas em solo em potes de 10 cm de diâmetro sob os ajustes ambientais seguintes: foto-período = 11,5 horas, intensidade da luz do dia = 30,000 lux ou mais, temperatura do período do dia = 28°C ou maior, temperatura do período noturno = 22°C, umidade relativa = 60 a 70%. Plantas transgênicas e o nulizigotos correspondente foram desenvolvidos lado a lado em posições aleatórias. A partir do estágio da semeadura até o estágio da maturidade as plantas foram passadas diversas vezes através de uma cabine formadora de imagem digital. Em cada um ponto de tempo as imagens digitais (2048 x 1536 pixéis, 16 milhões de cores) foram tiradas de cada planta de pelo menos 6 ângulos diferentes.

Avaliação da eficiência do uso de nitrogênio

As plantas de arroz a partir das sementes T2 foram desenvolvidas em potes com solo sob condições normais exceto pela solução de nutriente. Os potes foram lavados a partir da transplantação por maturação com uma solução específica de nutriente contendo teor nitrogênio reduzido (N), usualmente entre 7 a 8 vezes menos. O repouso do cultivo (maturação da planta, colheita da semente) foi o mesmo como foi para as plantas não desenvolvidas sob tensão abiótica. Os parâmetros de crescimento e produção são registrados detalhados para o crescimento sob condições normais.

44.2 Análise estatística: Teste F

Um dos dois fatores ANOVA (análise de variantes) como usado como um modelo estatístico para a avaliação total das características fenotípicas das plantas. Um teste F foi realizado em todos os parâmetros medido de todas as plantas de todos os eventos transformados com o gene da presente invenção. O teste F foi realizado por verificar por um efeito do gene em todos os eventos de transformação e para verificar um efeito total do gene, também conhecido como um efeito do gene global. O princípio para a significância um efeito de gene verdadeiro global foi apresentado a 5% do nível de probabilidade para o teste F. Os pontos de valor do teste F

significante a um efeito do gene, significa que não é apenas como mera presença ou posição do gene que está causando as diferenças nos fenotipos.

Porque dois experimentos com eventos sobrepostos foram realizados, uma análise combinada foi realizada. Esta é útil por verificar a consistência dos efeitos nos dois experimentos, e se esta é a causa, para acumular evidência a partir de ambos experimentos a fim de aumentar a confidência na conclusão. Os métodos usados foram um método de modelo misturado que torna-se em conta a estrutura multi-nível do dado (isto é experimento - evento - segregantes). O valor P foi obtido pela comparação com as distribuições quadradas chi para o teste de razão de probabilidade.

44.3 Parâmetros medidos

Biomassa que diz respeito as medidas dos parâmetros

As plantas da área acima do solo (ou biomassa da folha) foi determinada pela contagem do número total de pixéis nas imagens digitais da parte acima do solo das plantas discriminadas a partir do fundamento. Este valor foi medido para as figuras retiradas no mesmo ponto do ângulo diferente e foi convertido a um valor de superfície física expressada no quadrado mm pela calibração. Os experimentos mostram que as plantas da área acima do solo medido desta maneira correlacionada com a biomassa das partes das plantas acima trituradas. A área triturada acima é a área medida no mesmo ponto de tempo em que as plantas tem atingido esta biomassa de folha máxima. A biomassa de raiz aumentada é expressada como um aumento na biomassa da raiz total (medido como biomassa máxima de raízes observada durante a vida total de uma planta).

Medições de parâmetros que dizem respeito à sementes

As panículas primárias maduras foram endurecidas, contadas, empacotadas, rotuladas com código de barras e então secadas por três dias em um forno a 37°C. As panículas foram então refugadas e todas as sementes foram colhidas e contadas. A casca enchedora foram separadas a partir de um

despejo usando dispositivo de sopro de ar. As cascas despejadas foram descartadas e a fração remanescente foi contada novamente. A casca enchedora foi pesada em uma balança analítica. O número de sementes enchedoras foi determinada pela contagem o número de cascas enchedoras que permaneceram após a etapa de separação. A produção total da semente foi medida pela pesagem de todas as cascas enchedoras endurecidas a partir de uma planta. Peso singular da semente (TKW) é extrapolado a partir do número de sementes enchedoras contadas e seu peso total.

Exemplo 45: Medição da produção que diz respeito aos parâmetros para os transformantes pGOS2::SYR desenvolvidos sob condições de deficiência dos nutrientes:

Na análise das sementes como descrito acima, os inventores observaram que as plantas transformadas com o construto do gene SYR pGOS2:: e desenvolvidas sob deficiência de tensão de nutrientes, tem uma maior produção da semente, expressada como número de sementes enchedoras (aumento de mais do que 5%), total peso de sementes (aumento de mais do que 5%) e TKW (aumento de mais do que 2,5%), comparadas as plantas que necessitam do transgene SYR. Aqui também foi observado um aumento na biomassa do broto (mais do que 5%) e biomassa da raiz (diversas linhas mais do que 5%).

LISTAGEM DE SEQÜÊNCIAS

<110> CropDesign N.V.
<120> Plantas tendo características relacionadas com o crescimento intensificado e um método para fabricar as mesmas
<130> PF58342
<160> 293
<170> PatentIn version 3.3
<210> 1
<211> 1397
<212> DNA
<213> Arabidopsis thaliana
<400> 1

```
atggaaaaca aaagccatag caaccaccac cggtaccatc atctatcatc ccggagccac     60
gagcaaaacc agcgtggcct aatttcgata aatgccataa tcatcattgg catctccatt    120
atctccatat tcataatcct tgcaattctc ctaataatca ttttacttca tcgacttaaa    180
tccgcaagag tcaaagcaca agaattatct tgcaaagaaa gcttcaacaa catgaacaat    240
ggtggagctt ccaccaacta cagctacact tctagccctg atgacatcaa gagagattgt    300
ctatactcaa gaaacccaac atcgttcaga caattacccc cacagacaaa aagttgtagg    360
agaagccgag ccgaaggagt agaagtgtac acatacaagg aattagagat tgcaacaaat    420
aatttcagcg aggagaagaa aatcggaaat ggagatgtgt acaaaggagt actaagtgat    480
ggaactgttg cggccattaa aaagcttcat atgtttaatg ataatgctag taaccaaaag    540
catgaagaac ggtcctttag gcttgaggtt gatcttctta gtcgattaca atgcccatat    600
ttggtggaat tacttggata ttgtgcagat caaaaccata ggatcytgat atatgaattt    660
atgcctaatg gtactgtgga acatcatctc cacgatcata attttaagaa tctaaaggat    720
cgacctcaac cattagattg gggagctcgg cttaggatcg cccttgattg cgccagagcc    780
ctaragtttc ttcatgagaa tacaatctct actgttatcc accggaactt caagtgtacc    840
aacattttgt tggatcaaaa taatcgagct aaagtttcag atttcggatt agccaaaacc    900
ggatccgata aactaaacgg tgagatttca acaagggtta ttggcaccac aggatatctt    960
gctccagagt atgcctctac tggaaagctt actacaaaat cagatgttta tagttatggt   1020
attgtgttac tacaactctt aacgggtcgc acaccgatcg attcaagacg tccacgggga   1080
caagatgtcc ttgtctcttg ggctcttcca agactaacca atagagagaa aattagtgaa   1140
atggtggatc caaccatgaa aggtcaatac tcacaaaaag atcttattca ggtagcagcc   1200
atagcagcag tgtgtgtgca accagaagca agttatagac cgttgatgac ggatgttgtt   1260
cactcgctca ttccccttgt taaagctttc aacaaaagta ctgattcttc tcggtttcct   1320
agccgtagag aaagcttgtc atttgatgat attatgccat gacactcttt tgtttaccca   1380
gctttcttgt acaaagt                                                  1397
```

<210> 2
<211> 453
<212> PRT
<213> Arabidopsis thaliana
<220>
<221> INCERTO
<222> (216)..(216)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> INCERTO
<222> (262)..(262)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 2

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

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

<210> 3
<211> 55
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm2500
<400> 3
ggggacaagt ttgtacaaaa aagcaggctt cacaatggaa aacaaaagcc atagc 55

<210> 4
<211> 50
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm2501
<400> 4
ggggaccact ttgtacaaga aagctgggta aacaaaagag tgtcatggca 50

<210> 5

<211> 2193
<212> DNA
<213> Oryza sativa
<400> 5
aatccgaaaa gtttctgcac cgttttcacc ccctaactaa caatataggg aacgtgtgct 60
aaatataaaa tgagacctta tatatgtagc gctgataact agaactatgc aagaaaaact 120
catccaccta ctttagtggc aatcgggcta aataaaaaag agtcgctaca ctagtttcgt 180
tttccttagt aattaagtgg gaaaatgaaa tcattattgc ttagaatata cgttcacatc 240
tctgtcatga agttaaatta ttcgaggtag ccataattgt catcaaactc ttcttgaata 300
aaaaaatctt tctagctgaa ctcaatgggt aaagagagag attttttttа aaaaaataga 360
atgaagatat tctgaacgta ttggcaaaga tttaaacata taattatata attttatagt 420
ttgtgcattc gtcatatcgc acatcattaa ggacatgtct tactccatcc caatttttat 480
ttagtaatta aagacaattg acttatttтt attatttatc ttttttcgat tagatgcaag 540
gtacttacgc acacactttg tgctcatgtg catgtgtgag tgcacctcct caatacacgt 600
tcaactagca acacatctct aatatcactc gcctatttaa tacatttagg tagcaatatc 660
tgaattcaag cactccacca tcaccagacc acttttaata atatctaaaa tacaaaaaat 720
aattttacag aatagcatga aaagtatgaa acgaactatt taggtttttc acatacaaaa 780
aaaaaaagaa ttttgctcgt gcgcgagcgc caatctccca tattgggcac acaggcaaca 840
acagagtggc tgcccacaga acaacccaca aaaaacgatg atctaacgga ggacagcaag 900
tccgcaacaa ccttttaaca gcaggctttg cggccaggag agaggaggag aggcaaagaa 960
aaccaagcat cctcctcctc ccatctataa attcctcccc ccttttcccc tctctatata 1020
ggaggcatcc aagccaagaa gagggagagc accaaggaca cgcgactagc agaagccgag 1080
cgaccgcctt cttcgatcca tatcttccgg tcgagttctt ggtcgatctc ttccctcctc 1140
cacctcctcc tcacagggta tgtgcccttc ggttgttctt ggatttattg ttctaggttg 1200
tgtagtacgg gcgttgatgt taggaaaggg gatctgtatc tgtgatgatt cctgttcttg 1260
gatttgggat agaggggttc ttgatgttgc atgttatcgg ttcggtttga ttagtagtat 1320
ggttttcaat cgtctggaga gctctatgga aatgaaatgg tttagggtac ggaatcttgc 1380
gattttgtga gtacctttтg tttgaggtaa aatcagagca ccggtgattt tgcttggtgt 1440
aataaaagta cggttgtttg gtcctcgatt ctggtagtga tgcttctcga tttgacgaag 1500
ctatcctttg tttattccct attgaacaaa aataatccaa ctttgaagac ggtcccgttg 1560
atgagattga atgattgatt cttaagcctg tccaaaattt cgcagctggc ttgtttagat 1620
acagtagtcc ccatcacgaa attcatggaa acagttataa tcctcaggaa caggggattc 1680
cctgttcttc cgatttgctt tagtcccaga attttttttc ccaaatatct taaaaagtca 1740
ctttctggtt cagttcaatg aattgattgc tacaaataat gcttttatag cgttatccta 1800
gctgtagttc agttaatagg taatacccct atagtttagt caggagaaga acttatccga 1860
tttctgatct ccatttttaa ttatatgaaa tgaactgtag cataagcagt attcatttgg 1920
attatttttt ttattagctc tcaccccttc attattctga gctgaaagtc tggcatgaac 1980
tgtcctcaat tttgtttca aattcacatc gattatctat gcattatcct cttgtatcta 2040
cctgtagaag tttctttttg gttattcctt gactgcttga ttacagaaag aaatttatga 2100
agctgtaatc gggatagtta tactgcttgt tcttatgatt catttccttt gtgcagttct 2160
tggtgtagct tgccactttc accagcaaag ttc 2193
<210> 6
<211> 15
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 1
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir = "Leu"
<220>
<221> VARIANTE
<222> (3)..(3)
<223> /substituir = "Gly" /substituir = "Arg"
<220>
<221> VARIANTE
<222> (5)..(5)
<223> /substituir = "Met"
<220>
<221> VARIANTE
<222> (6)..(6)
<223> /substituir = "Arg" /substituir = "Gln"

<220>
<221> VARIANTE
<222> (7)..(7)
<223> /substituir = "Ser" /substituir = "Cys"
<220>
<221> VARIANTE
<222> (8)..(8)
<223> /substituir = "Pro"
<220>
<221> VARIANTE
<222> (9)..(9)
<223> /substituir = "Tyr"
<220>
<221> VARIANTE
<222> (11)..(11)
<223> /substituir = "Val"
<220>
<221> INCERTO
<222> (12)..(12)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente, preferivelmente um de Lys, Asn, Ala, Ser, Gly ou Glu
<220>
<221> VARIANTE
<222> (14)..(14)
<223> /substituir = "Leu" /substituir = "Val"
<400> 6

Met	Leu	Ser	Arg	Leu	His	His	Arg	Asn	Leu	Leu	Xaa	Leu	Ile	Gly
1				5				10					15	

<210> 7
<211> 11
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 2
<220>
<221> VARIANTE
<222> (2)..(2)
<223> /substituir = "Tyr" /substituir = "Asn"
<220>
<221> VARIANTE
<222> (3)..(3)
<223> /substituir = "Phe"
<220>
<221> INCERTO
<222> (4)..(4)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente, preferivelmente um de Asp, Asn, Glu, Lys, Pro, Gln ou Gly
<220>
<221> VARIANTE
<222> (5)..(5)
<223> /substituir = "Val" /substituir = "Thr"
<220>
<221> VARIANTE
<222> (7)..(7)
<223> /substituir = "Met"
<220>
<221> VARIANTE
<222> (8)..(8)
<223> /substituir = "Arg" /substituir = "Gly"
<220>
<221> VARIANTE
<222> (11)..(11)
<223> /substituir = "Val"

<400> 7
Leu Tyr Trp Xaa Ala Arg Leu Leu Ile Ala Leu
1 5 10
<210> 8
<211> 9
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 3
<220>
<221> VARIANTE
<222> (2)..(2)
<223> /substituir = "Lys"
<220>
<221> VARIANTE
<222> (3)..(3)
<223> /substituir = "Gly"
<220>
<221> VARIANTE
<222> (5)..(5)
<223> /substituir = "Glu"
<220>
<221> VARIANTE
<222> (6)..(6)
<223> /substituir = "Phe"
<400> 8
Ala Arg Ala Leu Ala Tyr Leu His Glu
1 5
<210> 9
<211> 14
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 4
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir = "Ile"
<220>
<221> VARIANTE
<222> (5)..(5)
<223> /substituir = "Asn"
<220>
<221> VARIANTE
<222> (6)..(6)
<223> /substituir = "Leu"
<220>
<221> VARIANTE
<222> (6)..(6)
<223> aminoácido não é preferivelmente Ile, Val ou Met
<220>
<221> VARIANTE
<222> (8)..(8)
<223> /substituir = "Ala" /substituir = "Gly" /substituir = "Cys"
<220>
<221> VARIANTE
<222> (9)..(9)
<223> /substituir = "Thr"
<220>
<221> VARIANTE
<222> (11)..(11)
<223> /substituir = "Val"
<220>

<221> VARIANTE
<222> (14)..(14)
<223> /substituir = "Asp"
<400> 9
Val Ile His Arg Asp Phe Lys Ser Ser Asn Ile Leu Leu Glu
1 5 10
<210> 10
<211> 7
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 5
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir = "Arg"
<220>
<221> VARIANTE
<222> (3)..(3)
<223> /substituir = "Ala" /substituir = "Thr"
<220>
<221> VARIANTE
<222> (7)..(7)
<223> /substituir = "Met" /substituir = "Ser"
<400> 10
Lys Val Ser Asp Phe Gly Leu
1 5
<210> 11
<211> 1524
<212> DNA
<213> Arabidopsis thaliana
<400> 11
atggaaaaca aaagccatag caaccaccac cggtaccatc atctatcatc ccggagccac 60
gagcaaaacc agcgtggcct aatttcgata aatgccataa tcatcattgg catctccatt 120
atctccatat tcataatcct tgcaattctc ctaataatca ttttacttca tcgacttaaa 180
tccgcaagag tcaaagcaca agaattatct tgcaaagaaa gcttcaacaa catgaacaat 240
ggtggagctt ccaccaacta cagctacact tctagccctg gtaaaacacc taaacattca 300
ttccataatt ttgtaaagaa tcataagttt ttaattcaaa atcaatcagc aatgagtgga 360
gattctcata tctttatatc agtattgaat tacttccaca tccaatgtgt atttcttatt 420
cgaactttaa tgttcttgtt agatgacatc aagagagatt gtctatactc aagaaaccca 480
acatcgttca gacaattacc cccacagaca aaaagttgta ggagaagccg agccgaagga 540
gtagaagtgt acacatacaa ggaattagag attgcaacaa ataatttcag cgaggagaag 600
aaaatcggaa atggagatgt gtacaaagga gtactaagtg atggaactgt tgcggccatt 660
aaaaagcttc atatgtttaa tgataatgct agtaaccaaa agcatgaaga acggtccttt 720
aggcttgagg ttgatcttct tagtcgatta caatgcccat atttggtgga attacttgga 780
tattgtgcag atcaaaacca taggatcctg atatatgaat ttatgcctaa tggtactgtg 840
gaacatcatc tccacgatca taattttaag aatctaaagg atcgacctca accattagat 900
tggggagctc ggcttaggat cgcccttgat tgcgccagag ccctagagtt tcttcatgag 960
aatacaatct ctactgttat ccaccggaac ttcaagtgta ccaacatttt gttggatcaa 1020
aataatcgag ctaaagtttc agatttcgga ttagccaaaa ccggatccga taaactaaac 1080
ggtgagattt caacaagggt tattggcacc acaggatatc ttgctccaga gtatgcctct 1140
actggaaagc ttactacaaa atcagatgtt tatagttatg gtattgtgtt actacaactc 1200
ttaacgggtc gcacaccgat cgattcaaga cgtccacggg gacaagatgt ccttgtctct 1260
tgggctcttc caagactaac caatagagag aaaattagtg aaatggtgga tccaaccatg 1320
aaaggtcaat actcacaaaa agatcttatt caggtagcag ccatagcagc agtgtgtgtg 1380
caaccagaag caagttatag accgttgatg acggatgttg ttcactcgct cattcccctt 1440
gttaaagctt tcaacaaaag tactgattct tctcggtttc ctagccgtag agaaagcttg 1500
tcatttgatg atattatgcc atga 1524
<210> 12
<211> 507
<212> PRT
<213> Arabidopsis thaliana
<400> 12

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

500 505

<210> 13
<211> 1482
<212> DNA
<213> Arabidopsis thaliana
<400> 13

caaaaatgaa caacaccaaa atggaaaaca aaagccatag caaccaccac cggtaccatc 60
atctatcatc ccggagccac gagcaaaacc agcgtggcct aatttcgata aatgccataa 120
tcatcattgg catctccatt atctccatat tcataatcct tgcaattctc ctaataatca 180
ttttacttca tcgacttaaa tccgcaagag tcaaagcaca agaattatct tgcaaagaaa 240
gcttcaacaa catgaacaat ggtggagctt ccaccaacta cagctacact tctagccctg 300
atgacatcaa gagagattgt ctatactcaa gaaacccaac atcgttcaga caattacccc 360
cacagacaaa aagttgtagg agaagccgag ccgaaggagt agaagtgtac acatacaagg 420
aattagagat tgcaacaaat aatttcagcg aggagaagaa aatcggaaat ggagatgtgt 480
acaaaggagt actaagtgat ggaactgttg cggccattaa aaagcttcat atgtttaatg 540
ataatgctag taaccaaaag catgaagaac ggtcctttag gcttgaggtt gatcttctta 600
gtcgattaca atgcccatat ttggtggaat tacttggata ttgtgcagat caaaaccata 660
ggatcctgat atatgaattt atgcctaatg gtactgtgga acatcatctc cacgatcata 720
attttaagaa tctaaaggat cgacctcaac cattagattg gggagctcgg cttaggatcg 780
cccttgattg cgccagagcc ctagagtttc ttcatgagaa tacaatctct actgttatcc 840
accggaactt caagtgtacc aacattttgt tggatcaaaa taatcgagct aaagtttcag 900
atttcggatt agccaaaacc ggatccgata aactaaacgg tgagatttca acaagggtta 960
ttggcaccac aggatatctt gctccagagt atgcctctac tggaaagctt actacaaaat 1020
cagatgttta tagttatggt attgtgttac tacaactctt aacgggtcgc acaccgatcg 1080
attcaagacg tccacgggga caagatgtcc ttgtctcttg ggctcttcca agactaacca 1140
atagagagaa aattagtgaa atggtggatc caaccatgaa aggtcaatac tcacaaaaag 1200
atcttattca ggtagcagcc atagcagcag tgtgtgtgca accagaagca agttatagac 1260
cgttgatgac ggatgttgtt cactcgctca ttccccttgt taaagctttc aacaaaagta 1320
ctgattcttc tcggtttcct agccgtagag aaagcttgtc atttgatgat attatgccat 1380
gacactcttt tgtttaatat gtattaattt ctctctttta agttgagatt ctcgttgtta 1440
ttgtatattt gtataggtaa atgaatccat ttatcttgta ct 1482

<210> 14
<211> 458
<212> PRT
<213> Arabidopsis thaliana
<400> 14

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

```
   210                 215                 220
Phe Met Pro Asn Gly Thr Val Glu His His Leu His Asp His Asn Phe
225                 230                 235                 240
Lys Asn Leu Lys Asp Arg Pro Gln Pro Leu Asp Trp Gly Ala Arg Leu
                245                 250                 255
Arg Ile Ala Leu Asp Cys Ala Arg Ala Leu Glu Phe Leu His Glu Asn
            260                 265                 270
Thr Ile Ser Thr Val Ile His Arg Asn Phe Lys Cys Thr Asn Ile Leu
        275                 280                 285
Leu Asp Gln Asn Asn Arg Ala Lys Val Ser Asp Phe Gly Leu Ala Lys
   290                 295                 300
Thr Gly Ser Asp Lys Leu Asn Gly Glu Ile Ser Thr Arg Val Ile Gly
305                 310                 315                 320
Thr Thr Gly Tyr Leu Ala Pro Glu Tyr Ala Ser Thr Gly Lys Leu Thr
                325                 330                 335
Thr Lys Ser Asp Val Tyr Ser Tyr Gly Ile Val Leu Leu Gln Leu Leu
            340                 345                 350
Thr Gly Arg Thr Pro Ile Asp Ser Arg Arg Pro Arg Gly Gln Asp Val
        355                 360                 365
Leu Val Ser Trp Ala Leu Pro Arg Leu Thr Asn Arg Glu Lys Ile Ser
   370                 375                 380
Glu Met Val Asp Pro Thr Met Lys Gly Gln Tyr Ser Gln Lys Asp Leu
385                 390                 395                 400
Ile Gln Val Ala Ala Ile Ala Ala Val Cys Val Gln Pro Glu Ala Ser
                405                 410                 415
Tyr Arg Pro Leu Met Thr Asp Val Val His Ser Leu Ile Pro Leu Val
            420                 425                 430
Lys Ala Phe Asn Lys Ser Thr Asp Ser Ser Arg Phe Pro Ser Arg Arg
        435                 440                 445
Glu Ser Leu Ser Phe Asp Asp Ile Met Pro
   450                 455
<210> 15
<211> 1516
<212> DNA
<213> Arabidopsis thaliana
<400> 15
caaaaaccaa aaactaattt tcacgaaatc cgaagcttaa aaattattcc tctactgttt     60
caatttcatc tcttcttcct cttgagttgg caatttctag ggttttctca ttttcctttt    120
tttccatttc tggaaactgg taaaaggctt tttctttgtc tgtgggggag aaacaaaaat    180
ggagacagac gaagcatacc aaaagaaaga gcgagctgca ttagtagcca tcgttgtcct    240
tgcttgtctt gctttatctt ctttgttcgt cgcctttagc tactactgct atattcgtaa    300
caaggtttct aagcgtcaca gaattagcaa gaggtttgat tgtgaggaaa aaggcgattg    360
tcagaaagta caagatgtta ctgaaaatgg gttgcaaatt ttcaccttta agcaattgca    420
ttcagcaact ggtggattta gcaagtctaa tgtggttggg aatggtgggt ttggcttggt    480
ttatcgtggt gtgcttaacg atggcagaaa agttgctatc aagctcatgg atcatgcagg    540
aaagcaaggg gaagaagaat tcaagatgga ggttgagtta ctgagccgtc tgcgttcacc    600
atacttgttg gcacttcttg gttattgctc agacaatagt cataagctgc ttgtgtatga    660
gttcatggca aatggtggtc tgcaggaaca cctgtatctt cccaacaggt ctggttctgt    720
cccaccaaga ttagattggg aaactaggat gagaatagct gtggaagctg caaaaggctt    780
ggaatatctc catgaacaag tatcaccccc ggtgattcat agagacttta agagcagcaa    840
tattctgctg gacagaaact tcaatgccaa agtttcagat tttggattgg ctaaagtcgg    900
atcagataaa gctggtggac acgtttctac ccgtgtatta ggcacacaag gatatgtagc    960
tcctgagtac gctttaactg gtcacttgac aacaaagtcg gatgtctata gctacggtgt   1020
tgtcctgtta gagttactaa ccggtagagt tccagtagat atgaagagag ctacaggaga   1080
aggtgttctt gtctcttggg ctttgcctca attggctgat agagacaaag tagtagacat   1140
aatggatcca acactcgaag gacaatactc tacgaaagaa gttgttcaag ttgcagcaat   1200
agcagcaatg tgtgttcaag cagaagctga ctacagaccc ttaatggcag atgtggtgca   1260
gtcactggtt ccattagtga gaaaccgtag gtcagcttca aagcttagtg gctgctctag   1320
tagctttagc ttggctcggt ctcctaactc tccgggtaaa gcaagcatag gttctcaatg   1380
aataaatcaa tcagagaaga cgaaatgtga ggcttttatt atgtaacggt atttactaaa   1440
acagacaatg atgtaaacac ttgagtgatc aaacatggag tcatagagat aatcatcaaa   1500
taaacttttt cgaatc                                                   1516
<210> 16
```

<211> 400
<212> PRT
<213> Arabidopsis thaliana
<400> 16

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

<210> 17
<211> 1161
<212> DNA
<213> Arabidopsis thaliana
<400> 17

```
atggagacag acgaagcata ccaaaagaaa gagcgagctg cattagtagc catcgttgtc     60
cttgcttgtc ttgctttatc ttctttgttc gtcgccttta gctactactg ctatattcgt    120
aacaaggaaa aaggcgattg tcagaaagta caagatgtta ctgaaaatgg gttgcaaatt    180
ttcaccttta agcaattgca ttcagcaact ggtggattta gcaagtctaa tgtggttggg    240
```

```
aatggtgggt ttggcttggt ttatcgtggt gtgcttaacg atggcagaaa agttgctatc    300
aagctcatgg atcatgcagg aaagcaaggg gaagaagaat tcaagatgga ggttgagtta    360
ctgagccgtc tgcgttcacc atacttgttg gcacttcttg gttattgctc agacaatagt    420
cataagctgc ttgtgtatga gttcatggca aatggtggtc tgcaggaaca cctgtatctt    480
cccaacaggt ctggttctgt cccaccaaga ttagattggg aaactaggat gagaatagct    540
gtggaagctg caaaaggctt ggaatatctc catgaacaag tatcaccccc ggtgattcat    600
agagacttta agagcagcaa tattctgctg gacagaaact tcaatgccaa agtttcagat    660
tttggattgg ctaaagtcgg atcagataaa gctggtggac acgtttctac ccgtgtatta    720
ggcacacaag gatatgtagc tcctgagtac gctttaactg gtcacttgac aacaaagtcg    780
gatgtctata gctacggtgt tgtcctgtta gagttactaa ccggtagagt tccagtagat    840
atgaagagag ctacaggaga aggtgttctt gtctcttggg ctttgcctca attggctgat    900
agagacaaag tagtagacat aatggatcca acactcgaag gacaatactc tacgaaagaa    960
gttgttcaag ttgcagcaat agcagcaatg tgtgttcaag cagaagctga ctacagaccc   1020
ttaatggcag atgtggtgca gtcactggtt ccattagtga gaaaccgtag gtcagcttca   1080
aagcttagtg gctgctctag tagctttagc ttggctcggt ctcctaactc tccgggtaaa   1140
gcaagcatag gttctcaatg a                                             1161
<210> 18
<211> 386
<212> PRT
<213> Arabidopsis thaliana
<400> 18
Met Glu Thr Asp Glu Ala Tyr Gln Lys Lys Glu Arg Ala Ala Leu Val
1               5                  10                  15
Ala Ile Val Val Leu Ala Cys Leu Ala Leu Ser Ser Leu Phe Val Ala
            20                  25                  30
Phe Ser Tyr Tyr Cys Tyr Ile Arg Asn Lys Glu Lys Gly Asp Cys Gln
        35                  40                  45
Lys Val Gln Asp Val Thr Glu Asn Gly Leu Gln Ile Phe Thr Phe Lys
    50                  55                  60
Gln Leu His Ser Ala Thr Gly Gly Phe Ser Lys Ser Asn Val Val Gly
65                  70                  75                  80
Asn Gly Gly Phe Gly Leu Val Tyr Arg Gly Val Leu Asn Asp Gly Arg
                85                  90                  95
Lys Val Ala Ile Lys Leu Met Asp His Ala Gly Lys Gln Gly Glu Glu
            100                 105                 110
Glu Phe Lys Met Glu Val Glu Leu Leu Ser Arg Leu Arg Ser Pro Tyr
        115                 120                 125
Leu Leu Ala Leu Leu Gly Tyr Cys Ser Asp Asn Ser His Lys Leu Leu
    130                 135                 140
Val Tyr Glu Phe Met Ala Asn Gly Gly Leu Gln Glu His Leu Tyr Leu
145                 150                 155                 160
Pro Asn Arg Ser Gly Ser Val Pro Pro Arg Leu Asp Trp Glu Thr Arg
                165                 170                 175
Met Arg Ile Ala Val Glu Ala Ala Lys Gly Leu Glu Tyr Leu His Glu
            180                 185                 190
Gln Val Ser Pro Pro Val Ile His Arg Asp Phe Lys Ser Ser Asn Ile
        195                 200                 205
Leu Leu Asp Arg Asn Phe Asn Ala Lys Val Ser Asp Phe Gly Leu Ala
    210                 215                 220
Lys Val Gly Ser Asp Lys Ala Gly Gly His Val Ser Thr Arg Val Leu
225                 230                 235                 240
Gly Thr Gln Gly Tyr Val Ala Pro Glu Tyr Ala Leu Thr Gly His Leu
                245                 250                 255
Thr Thr Lys Ser Asp Val Tyr Ser Tyr Gly Val Val Leu Leu Glu Leu
            260                 265                 270
Leu Thr Gly Arg Val Pro Val Asp Met Lys Arg Ala Thr Gly Glu Gly
        275                 280                 285
Val Leu Val Ser Trp Ala Leu Pro Gln Leu Ala Asp Arg Asp Lys Val
    290                 295                 300
Val Asp Ile Met Asp Pro Thr Leu Glu Gly Gln Tyr Ser Thr Lys Glu
305                 310                 315                 320
Val Val Gln Val Ala Ala Ile Ala Ala Met Cys Val Gln Ala Glu Ala
                325                 330                 335
```

Asp Tyr Arg Pro Leu Met Ala Asp Val Val Gln Ser Leu Val Pro Leu
340 345 350
Val Arg Asn Arg Arg Ser Ala Ser Lys Leu Ser Gly Cys Ser Ser Ser
355 360 365
Phe Ser Leu Ala Arg Ser Pro Asn Ser Pro Gly Lys Ala Ser Ile Gly
370 375 380
Ser Gln
385

<210> 19
<211> 2627
<212> DNA
<213> Arabidopsis thaliana
<400> 19

aacggaatat ttttttctg ggagctgaaa atcgttggtg acgatggcgg attctctttc 60
tagggtttac ggaactctgt ttagtggtgt taatatatct ctcttctcgg atccaccgtt 120
gaaaaccccc tgaagagtgt tttttttcc gggtttaagg tcttttacgt ttcgtttttt 180
cttgagattt cgtcttccgc tgtgagtttc tgttctcgcg gcgagttttt gtgtagatct 240
gagggttttc aaggcaagag aaagtgtcgg tgatgtcgat tttctcactc gcgagctctt 300
ctgattcaca aagagatgac ctaataatgg cgttgaaggc ggtggttatt gtatattgtg 360
tggtttctct cgtcagtgtt caattagctg atgctcaaca tgaaggactg ccagtttcac 420
caacgttatc accttcaact tcaccagtta tcactgatct gcccctacca gctgaatttc 480
cgcggtttca cagaaagtat ttcgcaccac aacaagcaga agcacctcag cattctcccc 540
cttatagtcg tttggtcgct tctgatcatc cccctaccag ctcacatttc tccaaacctt 600
ccatgaaaag gaatgctcag tctcctggag ccggcttggc tgatattgct ccagcacaat 660
ctagcaatgg tgttcttcct gatgccttaa ctcagccacc tttgtcgccc tccatttcaa 720
attgttgcaa atcagatatg gtgcttaaac gaagaagtat tggttgccac tgcgtgtatc 780
ctataaaact ggacatcctt ctcttgaatg tttcagaaac tcctagttgg aacatgttct 840
tgaacgaatt tgctacccag cttggtctcc tacctcacca aatcgagctg attaacttct 900
atgtgctaag cttatcaagg atgaacatat cgatggatat cacccctcat tctggaatta 960
gtttctcagc tagtcaggca tccgcaataa actcttccct tatcagccac aagattcaat 1020
ttagccctac tttggtggga gattacaaac ttctaaacct tacttggttt gaggcccctg 1080
caccttcgca agcacctcta gtggcttctt cacctcataa agcaccatca caaggatcct 1140
cagcaactac gtcagtaaga tctccaggga aaaagaggca tcccaatctt attcttatct 1200
tttctatagc cgctggtgtg cttatacttg ccataatcac tgtacttgtt atttgttccc 1260
gcgcactccg agaagagaaa gctccagatc ctcacaaaga agctgtaaaa ccaaggaacc 1320
tggacgctgg ttcatttggg ggatctcttc ctcacccagc aagtacacgg tttctgtcat 1380
atgaagaact caaagaggca actagcaatt ttgaatctgc tagcattcta ggagaaggtg 1440
ggtttggcaa ggtttacaga ggcatcttag ccgatggtac tgctgtagcg attaagaagc 1500
tcacaagtgg tgggccacaa ggtgataaag aattccaggt ggagattgat atgcttagcc 1560
gtcttcatca tcgtaatctt gtgaaacttg tgggttacta tagtagtcga gattcttctc 1620
agcacctact ttgttatgag cttgttccaa atggcagcct cgaggcttgg ctccatgggc 1680
ctctcgggtt gaactgtcct cttgattggg acaccagaat gaagattgca cttgatgctg 1740
caagaggact tgcatacctt catgaagact cgcaaccctc cgttatacac agagatttta 1800
aagcctctaa tatactcctt gaaaacaact tcaacgccaa agttgcagat tttggcctag 1860
ccaaacaagc tcctgaaggc aggggtaatc acttatctac tcgtgttatg ggcacatttg 1920
gatatgttgc gcctgaatat gcaatgacgg gacacctact cgtcaagagt gatgtttata 1980
gttacggtgt ggtccttctc gaattgttaa ctggtagaaa acctgtggat atgtcacaac 2040
cttcaggcca agaaaatctc gtcacttgga caaggccagt tttaagagac aaagaccggt 2100
tagaagaact agtcgattca agacttgaag gaaaataccc gaaagaagat ttcataagag 2160
tatgcacaat cgctgcagct tgtgttgcac ctgaagctag ccagagacca acgatgggcg 2220
aagtggttca gtcacttaaa atggttcaac gggtggttga gtatcaagac ccggttttaa 2280
acacttcaaa taaagctcgt cctaaccgga gacaatcatc agctacgttc gagtcagaag 2340
taacctcttc tatgttctct tctggtcctt attctggtct aagcgctttt gatcatgaaa 2400
atattacacg aacaactgtt ttctcagaag atcttcacga aggccgatga tagaagccag 2460
ggttttcttc tttttatttg tttttctccc acttacggtg agaaaatttc caccagagaa 2520
gcttttgagt ttgggacatt attacagctc tttggatttt ggattctcct ttattggagg 2580
atagaatttt gtatatattt ttgcgttaat taattaatat ttgccac 2627

<210> 20
<211> 725
<212> PRT
<213> Arabidopsis thaliana
<400> 20

Met Ser Ile Phe Ser Leu Ala Ser Ser Ser Asp Ser Gln Arg Asp Asp

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

```
Asn Ile Leu Leu Glu Asn Asn Phe Asn Ala Lys Val Ala Asp Phe Gly
        515                 520                 525
Leu Ala Lys Gln Ala Pro Glu Gly Arg Gly Asn His Leu Ser Thr Arg
    530                 535                 540
Val Met Gly Thr Phe Gly Tyr Val Ala Pro Glu Tyr Ala Met Thr Gly
545                 550                 555                 560
His Leu Leu Val Lys Ser Asp Val Tyr Ser Tyr Gly Val Val Leu Leu
                565                 570                 575
Glu Leu Leu Thr Gly Arg Lys Pro Val Asp Met Ser Gln Pro Ser Gly
            580                 585                 590
Gln Glu Asn Leu Val Thr Trp Thr Arg Pro Val Leu Arg Asp Lys Asp
        595                 600                 605
Arg Leu Glu Glu Leu Val Asp Ser Arg Leu Glu Gly Lys Tyr Pro Lys
    610                 615                 620
Glu Asp Phe Ile Arg Val Cys Thr Ile Ala Ala Ala Cys Val Ala Pro
625                 630                 635                 640
Glu Ala Ser Gln Arg Pro Thr Met Gly Glu Val Val Gln Ser Leu Lys
                645                 650                 655
Met Val Gln Arg Val Val Glu Tyr Gln Asp Pro Val Leu Asn Thr Ser
            660                 665                 670
Asn Lys Ala Arg Pro Asn Arg Arg Gln Ser Ser Ala Thr Phe Glu Ser
        675                 680                 685
Glu Val Thr Ser Ser Met Phe Ser Ser Gly Pro Tyr Ser Gly Leu Ser
    690                 695                 700
Ala Phe Asp His Glu Asn Ile Thr Arg Thr Thr Val Phe Ser Glu Asp
705                 710                 715                 720
Leu His Glu Gly Arg
                725
<210> 21
<211> 3732
<212> DNA
<213> Arabidopsis thaliana
<400> 21
cttcttcttc tgccacttcg attgctttaa acttggagat taagattaca catcaaataa     60
ttcttctttg tcttcctcgt gatcatcgag gaagcttgat ctcttctatg gaatggagaa    120
tttgactctc ctccttagga tctgtcttgt ttcgtcggtt ttagttgctg cttcttcctc    180
aggatcggaa ttgttatctc cgttatcatc accaccgtct ccattacccg aaacttcaaa    240
agggtttggt caagcaccga gtaattcacc tgaatctcac aaatccgaca atgtgccacc    300
gtctaaagct tcttctcaac cgtctcttcc tcctttagct gatttggcag ctccaccacc    360
gtcagattcc gtcggaggta aagcaccggc cggtgtgcct gttgtctttg ttcctaatgc    420
tccagctcca gctacaatac ccgttaagga tttgcccgta gcttcgcctc cggttcttca    480
acccattaca ccgattgctt caccacctcg attcattcca ggagatgcac caaaggagcc    540
tcctttctca ggaagagtta ctccagcccc ggtttcatca cctgtttcgg atattcctcc    600
aattccatca gtggctctgc ctccgcctac accgagcaat gtacctccta gaaatgcttc    660
taacaatcac aagcctcctc ctatcgaaaa gtctattgct cctgttgcat caccaccaac    720
aatatcaatt gacattgcac caccagtaca tcccgtcata cctaagttaa caccctcgag    780
ctcacctgta cctacctcga ctccaactaa gggatctccg agaagaaatc cgccaactac    840
ccacccagtt ttccccatag aatctcctgc tgtctcacca gatcatgctg caaatccagt    900
aaaacatccg ccccccagcg ataacggaga tgatagtaaa agtccgggtg ctgcaccggc    960
aaatgaaact gctaaaccct tacctgtctt cccacataaa gcttctcctc cttcgattgc   1020
tccctcagcc ccaaaattta acagacactc tcatcataca tctccatcta ctactccacc   1080
gccagattct actcctagta atgtacacca ccatccttca tcaccatctc ctcctcctct   1140
atcttcacac catcaacatc atcaagaacg aaagaaaatc gcagatagtc cagctccttc   1200
accactgcca ccgcatctca tatctccaaa gaagtcaaac cgtaaaggtt caatgactcc   1260
tcctccacaa tcgcaccatg ctccttctcc tcccattcct gattctttga tttctcctgc   1320
tcacgctcca gtctctttct ctatgaaacg catctccccg gctctagcac cttctccaac   1380
ccaagtgttt cctctcaggt cctcttcaag accttcaaaa tctcggaaat tccctctagg   1440
tccacctctt caagcatttc ctcctcctcc ccctaattca gattgttctt caactatttg   1500
tttggaaccg tacacaaaca cacctccggg atctccatgt ggctgtgtct ggccaattca   1560
agttgagcta cgccttagca tggcgctcta cgacttcttc ccaatggttt cagaatttgc   1620
tcgggaaatc agcgccggag ttttcatgaa acagagccaa gtccgtataa tgggagctaa   1680
cgcggccagc gaacaacctg agaaatccat tgttctaatc gatttagtac cgcttggaga   1740
taaattcgat aacatgactg caatgctgac ttaccagaga ttctggagta aaaaagtcta   1800
```

tatagatgaa ccaatctttg gcggatacga cgtgatttac gtgcgttatc ctggtttacc 1860
cgcttccccg ccaacttctg gtatgaccat tatagatcaa ggaccgtact ctggtaataa 1920
taacggaagg gcggtcaaac cgctcggagt tgatgttcca aggaagccgc gcaagaaaga 1980
gctcaatggt ggaagtattg ctgtgattgt tttatccgca gctgctttta tcggtttatg 2040
tttcgttatc gtttggttct tggttttccg gcgacaaaga gatcgacgac gtctttcgaa 2100
aagaacacca cttgctcgcc cttcactgcc ttccctctcg aaaccatcag gctcggcgag 2160
atctttaact ggaagccgat ttagctcaac ttcattgtca tttgaatcca gcattgctcc 2220
gtttactctc tctgcaaaga ctttcaccgc aagtgaaata atgaaagcta caaataactt 2280
tgatgaatcg agggttcttg gtgaaggcgg atttggacga gtttacgaag gtgttttcga 2340
tgacggaact aaagtagcag ttaaggtatt gaagagagat gaccagcaag gtagcaggga 2400
gttcttagcc gaagtggaaa tgcttagtcg tcttcatcac agaaacctcg tgaatttaat 2460
tggtatttgt atcgaagatc gtaaccgttc cctggtttat gaactcatac caaatggcag 2520
cgttgaatct cacctccacg ggattgataa agcatcttcg cctttagatt gggatgctcg 2580
gttgaagata gctcttggtg cggcccgtgg tttagcgtat ttacacgaag actcgagccc 2640
gcgagttata cacagagact tcaagtccag caacatcttg ctggaaaacg atttcacacc 2700
taaagtatct gactttggat tggctcgtaa cgcccttgat gatgaagata acagacatat 2760
atcgacacgc gttatgggaa cgtttgggta tgtggcaccg gaatacgcaa tgacagggca 2820
tcttttggtg aagagtgatg tgtatagtta cggagttgtg cttctcgagc tacttacagg 2880
gcggaaacca gtggatatgt ctcaaccgcc cggacaagag aacttagttt catggactcg 2940
gccttttctt acgagtgcag aagggcttgc agctattata gatcagtctt taggacccga 3000
gatctcattc gatagcatag ctaaagtcgc ggcaatagct tctatgtgcg ttcaaccaga 3060
agtatcacac cgtcctttca tgggagaagt agtgcaagct ttgaaacttg tcagcaacga 3120
atgtgatgaa gctaaagaac tcaattcttt aacttctatc tctaaagacg attttagaga 3180
tgacactcaa gctgagagct cttgtggtga cagctcagca aggatggctc ggtatccatt 3240
gctaccaaat tacgactctg agcctgatac agagagagga ttatctacat cggaaatgta 3300
cacggggtca gggaggttcg agagacagtc taactcgggt cccttgacct cgggtcgagg 3360
caagagtttc tggcaaaaga tgaggagatt atcgaccgga agcttgagtg agcatggaac 3420
accaacggtc atgttacgat ccggttctcg ctaaacaatc ttttgcctac agaacactga 3480
agagtttttg attgcccgtt taagttaaga gactgaaagg aaagaaggtg ccttctccta 3540
caagcaaaac tctttgcctc atggaaaccg gttaagataa aaccaggagt gatcatttcc 3600
cggttcaaaa tttttagttg aatagatctg tttatctgag aagaagaaac aagagattct 3660
gttaaatcta atttgaatgt acatatcacg ttttaaagta acaggcttaa gcattaagta 3720
tcatcatcct tc 3732

<210> 22
<211> 1113
<212> PRT
<213> Arabidopsis thaliana
<400> 22

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

```
       195                 200                 205
Ala Pro Pro Val His Pro Val Ile Pro Lys Leu Thr Pro Ser Ser Ser
   210                 215                 220
Pro Val Pro Thr Ser Thr Pro Thr Lys Gly Ser Pro Arg Arg Asn Pro
225                 230                 235                 240
Pro Thr Thr His Pro Val Phe Pro Ile Glu Ser Pro Ala Val Ser Pro
                245                 250                 255
Asp His Ala Ala Asn Pro Val Lys His Pro Pro Pro Ser Asp Asn Gly
            260                 265                 270
Asp Asp Ser Lys Ser Pro Gly Ala Ala Pro Ala Asn Glu Thr Ala Lys
        275                 280                 285
Pro Leu Pro Val Phe Pro His Lys Ala Ser Pro Pro Ser Ile Ala Pro
   290                 295                 300
Ser Ala Pro Lys Phe Asn Arg His Ser His His Thr Ser Pro Ser Thr
305                 310                 315                 320
Thr Pro Pro Pro Asp Ser Thr Pro Ser Asn Val His His His Pro Ser
                325                 330                 335
Ser Pro Ser Pro Pro Pro Leu Ser Ser His His Gln His His Gln Glu
            340                 345                 350
Arg Lys Lys Ile Ala Asp Ser Pro Ala Pro Ser Pro Leu Pro Pro His
        355                 360                 365
Leu Ile Ser Pro Lys Lys Ser Asn Arg Lys Gly Ser Met Thr Pro Pro
   370                 375                 380
Pro Gln Ser His His Ala Pro Ser Pro Pro Ile Pro Asp Ser Leu Ile
385                 390                 395                 400
Ser Pro Ala His Ala Pro Val Ser Phe Ser Met Lys Arg Ile Ser Pro
                405                 410                 415
Ala Leu Ala Pro Ser Pro Thr Gln Val Phe Pro Leu Arg Ser Ser Ser
            420                 425                 430
Arg Pro Ser Lys Ser Arg Lys Phe Pro Leu Gly Pro Pro Leu Gln Ala
        435                 440                 445
Phe Pro Pro Pro Pro Pro Asn Ser Asp Cys Ser Ser Thr Ile Cys Leu
   450                 455                 460
Glu Pro Tyr Thr Asn Thr Pro Pro Gly Ser Pro Cys Gly Cys Val Trp
465                 470                 475                 480
Pro Ile Gln Val Glu Leu Arg Leu Ser Met Ala Leu Tyr Asp Phe Phe
                485                 490                 495
Pro Met Val Ser Glu Phe Ala Arg Glu Ile Ser Ala Gly Val Phe Met
            500                 505                 510
Lys Gln Ser Gln Val Arg Ile Met Gly Ala Asn Ala Ala Ser Glu Gln
        515                 520                 525
Pro Glu Lys Ser Ile Val Leu Ile Asp Leu Val Pro Leu Gly Asp Lys
   530                 535                 540
Phe Asp Asn Met Thr Ala Met Leu Thr Tyr Gln Arg Phe Trp Ser Lys
545                 550                 555                 560
Lys Val Tyr Ile Asp Glu Pro Ile Phe Gly Gly Tyr Asp Val Ile Tyr
                565                 570                 575
Val Arg Tyr Pro Gly Leu Pro Ala Ser Pro Pro Thr Ser Gly Met Thr
            580                 585                 590
Ile Ile Asp Gln Gly Pro Tyr Ser Gly Asn Asn Asn Gly Arg Ala Val
        595                 600                 605
Lys Pro Leu Gly Val Asp Val Pro Arg Lys Pro Arg Lys Lys Glu Leu
   610                 615                 620
Asn Gly Gly Ser Ile Ala Val Ile Val Leu Ser Ala Ala Ala Phe Ile
625                 630                 635                 640
Gly Leu Cys Phe Val Ile Val Trp Phe Leu Val Phe Arg Arg Gln Arg
                645                 650                 655
Asp Arg Arg Arg Leu Ser Lys Arg Thr Pro Leu Ala Arg Pro Ser Leu
            660                 665                 670
Pro Ser Leu Ser Lys Pro Ser Gly Ser Ala Arg Ser Leu Thr Gly Ser
        675                 680                 685
Arg Phe Ser Ser Thr Ser Leu Ser Phe Glu Ser Ser Ile Ala Pro Phe
   690                 695                 700
```

```
Thr Leu Ser Ala Lys Thr Phe Thr Ala Ser Glu Ile Met Lys Ala Thr
705                 710                 715                 720
Asn Asn Phe Asp Glu Ser Arg Val Leu Gly Glu Gly Gly Phe Gly Arg
                725                 730                 735
Val Tyr Glu Gly Val Phe Asp Asp Gly Thr Lys Val Ala Val Lys Val
            740                 745                 750
Leu Lys Arg Asp Asp Gln Gln Gly Ser Arg Glu Phe Leu Ala Glu Val
        755                 760                 765
Glu Met Leu Ser Arg Leu His His Arg Asn Leu Val Asn Leu Ile Gly
    770                 775                 780
Ile Cys Ile Glu Asp Arg Asn Arg Ser Leu Val Tyr Glu Leu Ile Pro
785                 790                 795                 800
Asn Gly Ser Val Glu Ser His Leu His Gly Ile Asp Lys Ala Ser Ser
                805                 810                 815
Pro Leu Asp Trp Asp Ala Arg Leu Lys Ile Ala Leu Gly Ala Ala Arg
            820                 825                 830
Gly Leu Ala Tyr Leu His Glu Asp Ser Ser Pro Arg Val Ile His Arg
        835                 840                 845
Asp Phe Lys Ser Ser Asn Ile Leu Leu Glu Asn Asp Phe Thr Pro Lys
    850                 855                 860
Val Ser Asp Phe Gly Leu Ala Arg Asn Ala Leu Asp Asp Glu Asp Asn
865                 870                 875                 880
Arg His Ile Ser Thr Arg Val Met Gly Thr Phe Gly Tyr Val Ala Pro
                885                 890                 895
Glu Tyr Ala Met Thr Gly His Leu Leu Val Lys Ser Asp Val Tyr Ser
            900                 905                 910
Tyr Gly Val Val Leu Leu Glu Leu Leu Thr Gly Arg Lys Pro Val Asp
        915                 920                 925
Met Ser Gln Pro Pro Gly Gln Glu Asn Leu Val Ser Trp Thr Arg Pro
    930                 935                 940
Phe Leu Thr Ser Ala Glu Gly Leu Ala Ala Ile Ile Asp Gln Ser Leu
945                 950                 955                 960
Gly Pro Glu Ile Ser Phe Asp Ser Ile Ala Lys Val Ala Ala Ile Ala
                965                 970                 975
Ser Met Cys Val Gln Pro Glu Val Ser His Arg Pro Phe Met Gly Glu
            980                 985                 990
Val Val Gln Ala Leu Lys Leu Val  Ser Asn Glu Cys Asp  Glu Ala Lys
        995                1000                1005
Glu Leu  Asn Ser Leu Thr Ser  Ile Ser Lys Asp Asp  Phe Arg Asp
   1010                1015                1020
Asp Thr  Gln Ala Glu Ser Ser  Cys Gly Asp Ser Ser  Ala Arg Met
   1025                1030                1035
Ala Arg  Tyr Pro Leu Leu Pro  Asn Tyr Asp Ser Glu  Pro Asp Thr
   1040                1045                1050
Glu Arg  Gly Leu Ser Thr Ser  Glu Met Tyr Thr Gly  Ser Gly Arg
   1055                1060                1065
Phe Glu  Arg Gln Ser Asn Ser  Gly Pro Leu Thr Ser  Gly Arg Gly
   1070                1075                1080
Lys Ser  Phe Trp Gln Lys Met  Arg Arg Leu Ser Thr  Gly Ser Leu
   1085                1090                1095
Ser Glu  His Gly Thr Pro Thr  Val Met Leu Arg Ser  Gly Ser Arg
   1100                1105                1110
<210> 23
<211> 3162
<212> DNA
<213> Arabidopsis thaliana
<400> 23
ttctggtttc gaattcactt tactctcttc gcttccaatc tctctctacc actctgatac     60
cttcgccgga gaagaagctc tcgttctctc tatcgtcgga gaagctctcg ttctctgctt    120
ccctgctctg tacagatctc gagccgcatt tcgtggatct gtcaaaatcg cgttaaaacc    180
tactcgcttc tctctaaccc tagggttttt gaatttttt cgggacgagg gagaaatcaa    240
tgtctgtgaa gctttgagga agctatatgg ttacaccttt ggttctggtt agtcttcgtt    300
ttgatcggcg gagagtgttc gccgaggctt ggtgtctccg ttattgacta atggtggttt    360
```

ttgatttgag ataagagatg cggaactttg cgatgcttct gctgttaatt cttcttcttc 420
actctctcgc ctcgtttcct atatgctttg ctcggttatt tccgatgagt ttgccattta 480
cccgatcaaa ggcgcatcaa atgcatttct ttcatcctta tcttaatcca tcagttgctc 540
ctacaccttc accagctttt tcccccaacc caagtcgcat tccaccacta aggcacaagg 600
gtcatcaccg tcatcgtcgc tggcatttaa gacgcaatgc cactgcagtg tcaccttcgt 660
cacatgactg tcagcagaca tgtgtggagc ctctcacttc aactcccttt ggttcacctt 720
gtggttgtgt cttccccatg aaagttcagc ttctactaag tgtcgcacct ttttctattt 780
tccccgtgac caacgagtta gaaattgagg ttgctgctgg aacatacttg gaacaaagcc 840
aagtgaaaat aatgggtgcc agtgccgaca gtgaaaatca agggaaaaca gtggtggata 900
ttaaccttgt tccacttggg gaaaagttcg acaacactac agcaacactt atttatcaga 960
gattccggca caagaaagta cctctaaatg agactgtttt tggtgattat gaggttacgc 1020
acataagtta tccaggaatt ccttcttctt cgccaaatgg agatgttact ggagatgctc 1080
caggaggcct tccaattccg atcaatgcaa caacttttgc caacaaaagc cagggaatag 1140
gttttagaac gattgcaatt attgcgttgt caggttttgt cctcattctg gttttggttg 1200
gagctatttc tataatcgtg aaatggaaga aaattgggaa gtcatctaat gctgttggtc 1260
cggctttggc tccatcgatt aacaaaaggc ctggtgctgg atctatgttt tccagtagcg 1320
ccagaagctc aggctcagat tctttgatgt cttccatggc tacatgtgct ttatcagtta 1380
agaccttcac actctccgag cttgagaagg ctactgatag attcagtgcc aagagggttt 1440
tgggcgaggg cggatttggt cgtgtttatc aggggagtat ggaagatgga actgaggttg 1500
cagtgaaact gctaactagg gacaatcaga atcgagaccg tgaatttatt gccgaagtcg 1560
agatgctaag ccgtttgcac caccgcaatc ttgtgaaact gattggaata tgcattgaag 1620
gtcgtacacg ttgcttgatt tatgagctcg tccacaatgg gagtgttgag tcccacttac 1680
acgaaggaac gcttgattgg gatgcgcggt tgaagattgc acttggagca gcgagaggac 1740
tggcttatct ccatgaagat tcaaatcccc gagtaatcca ccgtgatttc aaggctagca 1800
atgttctcct agaagatgac tttaccccaa aagtctcaga ctttggactg gcaagagaag 1860
caaccgaagg aagtcaacat atttcaactc gagtcatggg gacctttggg tacgtggctc 1920
ctgagtatgc aatgacgggg catctccttg ttaaaagcga tgtctatagt tatggtgtag 1980
ttctactaga gcttctaacc ggtagaagac cagtagacat gtctcaacct tcgggagagg 2040
aaaaccttgt gacatgggcg agacccttac tagccaacag agagggcctg gagcaactgg 2100
tggaccctgc attggctgga acctacaact ttgatgacat ggcgaaagtg gctgcgattg 2160
cctccatgtg cgtccaccaa gaggtctcac acagaccatt catgggtgaa gttgtgcagg 2220
cactaaaact tatatacaac gatgcagatg agacctgtgg tgattattgc agccaaaagg 2280
actcatcagt accagactct gctgacttca aaggcgatct ggctccttcg gatagcagct 2340
ggtggaattt gacacctcgg ttaaggtatg gtcaagcttc gtcgttcata acaatggatt 2400
atagctcagg accactagag gacatggaga accggcctca ctctgcctct agcattccaa 2460
gagtaggagg tcttattcta ccaaacagat caggtccgtt aagacccatg cgtagtagaa 2520
gaaacttctt tagactgaga ggaagcatga gcgaacacgg tggaccgtcg tcgtctagac 2580
atctctggtc cggcaatggt gactggctct gagaaaccag aaaatgtaca gtgaagaaaa 2640
gggaaaagga ggctcacaaa ctcttgagct gcatggttag gttctggtta atccggccag 2700
gttcggttca tctgggtaaa gatttccggt ttgatttctt ttaggagctg tttgttgtta 2760
tgattatcct caatagagat ttaattgttt gtttttgcga ggatgcaaaa gcacagggag 2820
ggttgtattt tgaaggacgc ctgtatttag ttctctcctt ttctctctca cccacaaaaa 2880
aacattttta tatattttat tttagtgaag ttagtaatta cctaattttt tttagttttt 2940
caagttccta agaaatttgg cttttggtca tcgttgttaa gtttccggtt ttctgaattt 3000
acatccgaaa tatttttttgg ttaaataaaa ttctggcttg agagtaattg gctggtgaat 3060
ggtaacaagt cacttggtca aagctgttgt aattcttttc taatttctag tatcttctgt 3120
agattggaca caatgacatg gattgttgat ctttaaagtc gt 3162

<210> 24
<211> 744
<212> PRT
<213> Arabidopsis thaliana
<400> 24

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

```
                85                  90                  95
Asp Cys Gln Gln Thr Cys Val Glu Pro Leu Thr Ser Thr Pro Phe Gly
            100                 105                 110
Ser Pro Cys Gly Cys Val Phe Pro Met Lys Val Gln Leu Leu Leu Ser
        115                 120                 125
Val Ala Pro Phe Ser Ile Phe Pro Val Thr Asn Glu Leu Glu Ile Glu
    130                 135                 140
Val Ala Ala Gly Thr Tyr Leu Glu Gln Ser Gln Val Lys Ile Met Gly
145                 150                 155                 160
Ala Ser Ala Asp Ser Glu Asn Gln Gly Lys Thr Val Val Asp Ile Asn
                165                 170                 175
Leu Val Pro Leu Gly Glu Lys Phe Asp Asn Thr Thr Ala Thr Leu Ile
            180                 185                 190
Tyr Gln Arg Phe Arg His Lys Lys Val Pro Leu Asn Glu Thr Val Phe
        195                 200                 205
Gly Asp Tyr Glu Val Thr His Ile Ser Tyr Pro Gly Ile Pro Ser Ser
    210                 215                 220
Ser Pro Asn Gly Asp Val Thr Gly Asp Ala Pro Gly Gly Leu Pro Ile
225                 230                 235                 240
Pro Ile Asn Ala Thr Thr Phe Ala Asn Lys Ser Gln Gly Ile Gly Phe
                245                 250                 255
Arg Thr Ile Ala Ile Ile Ala Leu Ser Gly Phe Val Leu Ile Leu Val
            260                 265                 270
Leu Val Gly Ala Ile Ser Ile Ile Val Lys Trp Lys Lys Ile Gly Lys
        275                 280                 285
Ser Ser Asn Ala Val Gly Pro Ala Leu Ala Pro Ser Ile Asn Lys Arg
    290                 295                 300
Pro Gly Ala Gly Ser Met Phe Ser Ser Ser Ala Arg Ser Ser Gly Ser
305                 310                 315                 320
Asp Ser Leu Met Ser Ser Met Ala Thr Cys Ala Leu Ser Val Lys Thr
                325                 330                 335
Phe Thr Leu Ser Glu Leu Glu Lys Ala Thr Asp Arg Phe Ser Ala Lys
            340                 345                 350
Arg Val Leu Gly Glu Gly Gly Phe Gly Arg Val Tyr Gln Gly Ser Met
        355                 360                 365
Glu Asp Gly Thr Glu Val Ala Val Lys Leu Leu Thr Arg Asp Asn Gln
    370                 375                 380
Asn Arg Asp Arg Glu Phe Ile Ala Glu Val Glu Met Leu Ser Arg Leu
385                 390                 395                 400
His His Arg Asn Leu Val Lys Leu Ile Gly Ile Cys Ile Glu Gly Arg
                405                 410                 415
Thr Arg Cys Leu Ile Tyr Glu Leu Val His Asn Gly Ser Val Glu Ser
            420                 425                 430
His Leu His Glu Gly Thr Leu Asp Trp Asp Ala Arg Leu Lys Ile Ala
        435                 440                 445
Leu Gly Ala Ala Arg Gly Leu Ala Tyr Leu His Glu Asp Ser Asn Pro
    450                 455                 460
Arg Val Ile His Arg Asp Phe Lys Ala Ser Asn Val Leu Leu Glu Asp
465                 470                 475                 480
Asp Phe Thr Pro Lys Val Ser Asp Phe Gly Leu Ala Arg Glu Ala Thr
                485                 490                 495
Glu Gly Ser Gln His Ile Ser Thr Arg Val Met Gly Thr Phe Gly Tyr
            500                 505                 510
Val Ala Pro Glu Tyr Ala Met Thr Gly His Leu Leu Val Lys Ser Asp
        515                 520                 525
Val Tyr Ser Tyr Gly Val Val Leu Leu Glu Leu Leu Thr Gly Arg Arg
    530                 535                 540
Pro Val Asp Met Ser Gln Pro Ser Gly Glu Glu Asn Leu Val Thr Trp
545                 550                 555                 560
Ala Arg Pro Leu Leu Ala Asn Arg Glu Gly Leu Glu Gln Leu Val Asp
                565                 570                 575
Pro Ala Leu Ala Gly Thr Tyr Asn Phe Asp Asp Met Ala Lys Val Ala
            580                 585                 590
```

Ala Ile Ala Ser Met Cys Val His Gln Glu Val Ser His Arg Pro Phe
595 600 605
Met Gly Glu Val Val Gln Ala Leu Lys Leu Ile Tyr Asn Asp Ala Asp
610 615 620
Glu Thr Cys Gly Asp Tyr Cys Ser Gln Lys Asp Ser Ser Val Pro Asp
625 630 635 640
Ser Ala Asp Phe Lys Gly Asp Leu Ala Pro Ser Asp Ser Ser Trp Trp
645 650 655
Asn Leu Thr Pro Arg Leu Arg Tyr Gly Gln Ala Ser Ser Phe Ile Thr
660 665 670
Met Asp Tyr Ser Ser Gly Pro Leu Glu Asp Met Glu Asn Arg Pro His
675 680 685
Ser Ala Ser Ser Ile Pro Arg Val Gly Gly Leu Ile Leu Pro Asn Arg
690 695 700
Ser Gly Pro Leu Arg Pro Met Arg Ser Arg Arg Asn Phe Phe Arg Leu
705 710 715 720
Arg Gly Ser Met Ser Glu His Gly Gly Pro Ser Ser Ser Arg His Leu
725 730 735
Trp Ser Gly Asn Gly Asp Trp Leu
740

<210> 25
<211> 369
<212> DNA
<213> Oryza sativa
<400> 25
atgccaacta gtagagtcga cctgctgagc cggatgcact cgccgtacct ggtggggttg 60
ctgggctact gcgccgacca gagccaccgt ctgctggtgt tcgagttcat gcccaacggc 120
agcctcaaga gccacctcca ccggcgcgcg ctggcgccgg cggagcagcc gccgccgctg 180
gactggcaga cgcggctggg catcgcgctg gactgcgccc gcgcgctgga gttcctccac 240
gagcacagct cccccgccgt gatccaccgc gacttcaagt gcagcaacat cctcctcgac 300
cacaactacc gcgcccgcgt ctccgacttc ggcatgggca agctcggctc caacaaggcc 360
aatggccag 369

<210> 26
<211> 123
<212> PRT
<213> Oryza sativa
<400> 26
Met Pro Thr Ser Arg Val Asp Leu Leu Ser Arg Met His Ser Pro Tyr
1 5 10 15
Leu Val Gly Leu Leu Gly Tyr Cys Ala Asp Gln Ser His Arg Leu Leu
20 25 30
Val Phe Glu Phe Met Pro Asn Gly Ser Leu Lys Ser His Leu His Arg
35 40 45
Arg Ala Leu Ala Pro Ala Glu Gln Pro Pro Pro Leu Asp Trp Gln Thr
50 55 60
Arg Leu Gly Ile Ala Leu Asp Cys Ala Arg Ala Leu Glu Phe Leu His
65 70 75 80
Glu His Ser Ser Pro Ala Val Ile His Arg Asp Phe Lys Cys Ser Asn
85 90 95
Ile Leu Leu Asp His Asn Tyr Arg Ala Arg Val Ser Asp Phe Gly Met
100 105 110
Gly Lys Leu Gly Ser Asn Lys Ala Asn Gly Gln
115 120

<210> 27
<211> 1233
<212> DNA
<213> Oryza sativa
<400> 27
atggcggatg cttgggcggc gtcccctcca tctcctacgg ctccttcgcc tttcgccgcg 60
gacgacctcg tcgacgcgcg gctcgcgccg tggccgccgt tcgccccgtg gccggcgccc 120
gggcagctcc accacaaccg ccgcggcggc gggcacccga acccgctctt caccatcctg 180
ccggcctcgg cgctggcaat aggggtcgtg ctgcttgtcg ccgtggtcgt catcctggtg 240
atgacgcgtc ggtggaagcc cggggcggtg gacggcgccg gcggcgccag ctgcaacggc 300

```
gacaagcctg gcggcgcccc cgcgtccagc tgcggcagca gcgtccgcgg ctacaacaac    360
tccaggtact acgccgccgc cgccgccggg tgcatatatg gcggaaggct gggtttctcg    420
gtgcagccgc ggaaccgcgg cgcgcaagtg ttcacgtacc gggagctgga gagcgcgacg    480
gacgggttca gcgagtgcaa cgtggtgggc cgcggcgcgt acggcgtggt cttccgcggc    540
cggctcggcg acggcaccac ggccgccatc aagcggctca agatggacgg ccggcgcgag    600
ggcgagcgcg agttccgcat cgagatgggc gttgctatta ctgctcaggt cgacctgctg    660
agccggatgc actcgccgta cctggtgggg ttgctgggct actgcgccga ccagagccac    720
cgtctgctgg ttttcgagtt catgcccaac ggcagcctca agagccacct ccaccggcgc    780
gcgctggcgc cggcggagca gccgccgccg ctggactggc agacgcggct gggcatcgcg    840
ctggactgcg cccgcgcgct ggagttcctc cacgagcaca gctcccccgc cgtgatccac    900
cgcgacttca agtgcagcaa catcctcctc gaccacaact accgcgcccg cgtctccgac    960
ttcggcatgg ccaagctcgg ctccaacaag gccaatggcc aggtggccgc catcacggcg   1020
atgtgcatac agacgaaggc ggactaccgg ccgctgatga cggacgtggt gcagtcgctg   1080
atccccatcg tgaagagccc actcatgtcc tgcacctcca cgccgctgag gcccgcgcac   1140
gggcaccacc acgtcgtcta catgagcccg agcaggggct cttccaacgg cggtgccctg   1200
gaaacgcgat gcgtcatgca cggcttggat tga                                1233
<210> 28
<211> 410
<212> PRT
<213> Oryza sativa
<400> 28
Met Ala Asp Ala Trp Ala Ala Ser Pro Pro Ser Pro Thr Ala Pro Ser
1               5                  10                  15
Pro Phe Ala Ala Asp Asp Leu Val Asp Ala Arg Leu Ala Pro Trp Pro
            20                  25                  30
Pro Phe Ala Pro Trp Pro Ala Pro Gly Gln Leu His His Asn Arg Arg
        35                  40                  45
Gly Gly Gly His Pro Asn Pro Leu Phe Thr Ile Leu Pro Ala Ser Ala
    50                  55                  60
Leu Ala Ile Gly Val Val Leu Leu Val Ala Val Val Val Ile Leu Val
65                  70                  75                  80
Met Thr Arg Arg Trp Lys Pro Gly Ala Val Asp Gly Ala Gly Gly Ala
                85                  90                  95
Ser Cys Asn Gly Asp Lys Pro Gly Gly Ala Pro Ala Ser Ser Cys Gly
            100                 105                 110
Ser Ser Val Arg Gly Tyr Asn Asn Ser Arg Tyr Tyr Ala Ala Ala Ala
        115                 120                 125
Ala Gly Cys Ile Tyr Gly Gly Arg Leu Gly Phe Ser Val Gln Pro Arg
    130                 135                 140
Asn Arg Gly Ala Gln Val Phe Thr Tyr Arg Glu Leu Glu Ser Ala Thr
145                 150                 155                 160
Asp Gly Phe Ser Glu Cys Asn Val Val Gly Arg Gly Ala Tyr Gly Val
                165                 170                 175
Val Phe Arg Gly Arg Leu Gly Asp Gly Thr Thr Ala Ala Ile Lys Arg
            180                 185                 190
Leu Lys Met Asp Gly Arg Arg Glu Gly Glu Arg Glu Phe Arg Ile Glu
        195                 200                 205
Met Gly Val Ala Ile Thr Ala Gln Val Asp Leu Leu Ser Arg Met His
    210                 215                 220
Ser Pro Tyr Leu Val Gly Leu Leu Gly Tyr Cys Ala Asp Gln Ser His
225                 230                 235                 240
Arg Leu Leu Val Phe Glu Phe Met Pro Asn Gly Ser Leu Lys Ser His
                245                 250                 255
Leu His Arg Arg Ala Leu Ala Pro Ala Glu Gln Pro Pro Pro Leu Asp
            260                 265                 270
Trp Gln Thr Arg Leu Gly Ile Ala Leu Asp Cys Ala Arg Ala Leu Glu
        275                 280                 285
Phe Leu His Glu His Ser Ser Pro Ala Val Ile His Arg Asp Phe Lys
    290                 295                 300
Cys Ser Asn Ile Leu Leu Asp His Asn Tyr Arg Ala Arg Val Ser Asp
305                 310                 315                 320
Phe Gly Met Ala Lys Leu Gly Ser Asn Lys Ala Asn Gly Gln Val Ala
                325                 330                 335
```

```
Ala Ile Thr Ala Met Cys Ile Gln Thr Lys Ala Asp Tyr Arg Pro Leu
          340                 345                 350
Met Thr Asp Val Val Gln Ser Leu Ile Pro Ile Val Lys Ser Pro Leu
       355                 360                 365
Met Ser Cys Thr Ser Thr Pro Leu Arg Pro Ala His Gly His His His
   370                 375                 380
Val Val Tyr Met Ser Pro Ser Arg Gly Ser Ser Asn Gly Gly Ala Leu
385                 390                 395                 400
Glu Thr Arg Cys Val Met His Gly Leu Asp
                405                 410
<210> 29
<211> 1365
<212> DNA
<213> Oryza sativa
<400> 29
atgtcgagcg gcggcggcgg cggcgacgac tacaagcggg aggagtcggt ggcactcatg     60
gtcatcgtct ccctcgcggc gctctccctc ctctccctcg tcgccgcctt cgcctactac    120
tgctacatca cccgcaaggt ctcccgccgc ctccactcgc tcgacctccc caagcaccac    180
cgccgctcct cctcctcctc gcctcctccc atgcccccgc cgctgcctcc ccctcctcct    240
tccgcgaatg ccccgacgct cgggaaggag agcccgtcca gcaactccgc cagcgacggc    300
gcggcggcgg cggtcgtggt cggcggggag agggggccg tccaggtgtt cagctaccgc    360
cagctgcacg ccgccacggg cgggttcggg cgggcgcacg tggtggggca ggggagcttc    420
ggggccgtct accgcggggt gctccccgat gggaggaagg tcgcggtcaa gctcatggat    480
cgccccggca agcaggggga agaggagttc gagatggagg tggagctgct gagtcggctt    540
cgatcacctt atcttttggg cttgattggc cattgttcag agggtggaca ccggctgctg    600
gtctatgaat tcatggccaa tgggggtctc caggagcatc tctacccaaa tggagctttt    660
gaaaagattg aaacattttc gatctacttg gtgaaacaac gcccgatttt tgacaaaaat    720
atcgggcacc cgatttgtag gcgcgcccat ttttctcgtc aactgatatc ccacaaagct    780
gacattgtag gttcctgtgg tggtatctca aaattggact ggcctactag aatgagaata    840
gctcttgaag cagcaaaagg tctggaatat cttcacgagc gtgttaatcc acctgttata    900
cacagagact tcaagagcag caacattcta ctggacaagg acttccgtgc aagagtgtca    960
gattttggtt tggctaagct tggatctgat agagctggtg gtcatgtatc aactcgagtt   1020
ctaggcacac aaggttatgt tgcaccagat tacggtgttg tacttctgga gttgcttact   1080
ggcagggtgc cagttgatat gaagaggcct ccaggcgaag gtgttctagt taactgggca   1140
ttgcctatgc tcacagaccg agagaaggtt gtgcagatct tggatcctgc actggagggt   1200
caatattcat tgaaagatgc tgtccaagtg gctgccattg ctgccatgtg tgttcaacaa   1260
gaggctgact acaggccact aatggctgat gtggttcaat cgttggttcc gcttgtcaag   1320
aatcgttcta ctccaaagac gtgcaaccct agtgtccaag cgtag                   1365
<210> 30
<211> 454
<212> PRT
<213> Oryza sativa
<400> 30
Met Ser Ser Gly Gly Gly Gly Gly Asp Asp Tyr Lys Arg Glu Glu Ser
1               5                   10                  15
Val Ala Leu Met Val Ile Val Ser Leu Ala Ala Leu Ser Leu Leu Ser
          20                  25                  30
Leu Val Ala Ala Phe Ala Tyr Tyr Cys Tyr Ile Thr Arg Lys Val Ser
       35                  40                  45
Arg Arg Leu His Ser Leu Asp Leu Pro Lys His His Arg Arg Ser Ser
   50                  55                  60
Ser Ser Ser Pro Pro Pro Met Pro Pro Pro Leu Pro Pro Pro Pro Pro
65                  70                  75                  80
Ser Ala Asn Ala Pro Thr Leu Gly Lys Glu Ser Pro Ser Ser Asn Ser
                85                  90                  95
Ala Ser Asp Gly Ala Ala Ala Ala Val Val Val Gly Gly Glu Arg Gly
          100                 105                 110
Ala Val Gln Val Phe Ser Tyr Arg Gln Leu His Ala Ala Thr Gly Gly
       115                 120                 125
Phe Gly Arg Ala His Val Val Gly Gln Gly Ser Phe Gly Ala Val Tyr
   130                 135                 140
Arg Gly Val Leu Pro Asp Gly Arg Lys Val Ala Val Lys Leu Met Asp
145                 150                 155                 160
```

```
Arg Pro Gly Lys Gln Gly Glu Glu Glu Phe Glu Met Glu Val Glu Leu
                165                 170                 175
Leu Ser Arg Leu Arg Ser Pro Tyr Leu Leu Gly Leu Ile Gly His Cys
            180                 185                 190
Ser Glu Gly Gly His Arg Leu Leu Val Tyr Glu Phe Met Ala Asn Gly
        195                 200                 205
Gly Leu Gln Glu His Leu Tyr Pro Asn Gly Ala Phe Glu Lys Ile Glu
    210                 215                 220
Thr Phe Ser Ile Tyr Leu Val Lys Gln Arg Pro Ile Phe Asp Lys Asn
225                 230                 235                 240
Ile Gly His Pro Ile Cys Arg Arg Ala His Phe Ser Arg Gln Leu Ile
                245                 250                 255
Ser His Lys Ala Asp Ile Val Gly Ser Cys Gly Gly Ile Ser Lys Leu
            260                 265                 270
Asp Trp Pro Thr Arg Met Arg Ile Ala Leu Glu Ala Ala Lys Gly Leu
        275                 280                 285
Glu Tyr Leu His Glu Arg Val Asn Pro Pro Val Ile His Arg Asp Phe
    290                 295                 300
Lys Ser Ser Asn Ile Leu Leu Asp Lys Asp Phe Arg Ala Arg Val Ser
305                 310                 315                 320
Asp Phe Gly Leu Ala Lys Leu Gly Ser Asp Arg Ala Gly Gly His Val
                325                 330                 335
Ser Thr Arg Val Leu Gly Thr Gln Gly Tyr Val Ala Pro Asp Tyr Gly
            340                 345                 350
Val Val Leu Leu Glu Leu Leu Thr Gly Arg Val Pro Val Asp Met Lys
        355                 360                 365
Arg Pro Pro Gly Glu Gly Val Leu Val Asn Trp Ala Leu Pro Met Leu
    370                 375                 380
Thr Asp Arg Glu Lys Val Val Gln Ile Leu Asp Pro Ala Leu Glu Gly
385                 390                 395                 400
Gln Tyr Ser Leu Lys Asp Ala Val Gln Val Ala Ala Ile Ala Ala Met
                405                 410                 415
Cys Val Gln Gln Glu Ala Asp Tyr Arg Pro Leu Met Ala Asp Val Val
            420                 425                 430
Gln Ser Leu Val Pro Leu Val Lys Asn Arg Ser Thr Pro Lys Thr Cys
        435                 440                 445
Asn Pro Ser Val Gln Ala
    450
<210> 31
<211> 1923
<212> DNA
<213> Oryza sativa
<400> 31
ggcccccaaa cccctcaaaa ccctcgtcgt cttctcgcga tcgctattcc ctcctccatt     60
cctcctcctc cacctccgag acctagggtt ccaatgtcga gcggcggcgg cggcggcgac    120
gactacaagc gggaggagtc ggtggcactc atggtcatcg tctccctcgc ggcgctctcc    180
ctcctctccc tcgtcgccgc cttcgcctac tactgctaca tcacccgcaa ggtctcccgc    240
cgcctccact cgctcgacct ccccaagcac caccgccgct cctcctcctc ctcgcctcct    300
cccatgcccc cgccgctgcc tccccctcct ccttccgcga atgccccgac gctcgggaag    360
gagagcccgt ctagcaactc cgccagcgac ggcgcggcgg cggcggtcgt ggtcggcggg    420
gagaggggcg ccgtccaggt gttcagctac cgccagctgc acgccgccac gggcgggttc    480
gggcgggcgc acgtggtggg gcaggggagc ttcggggccg tctaccgcgg ggtgctcccc    540
gatgggagga aggtcgcggt caagctcatg gatcgccccg gcaagcaggg ggaagaggag    600
ttcgagatgg aggtggagct gctgagtcgg cttcgatcac cttatctttt gggcttgatt    660
ggccattgtt cagagggtgg acaccggctg ctggtctatg aattcatggc caatgggggt    720
ctccaggagc atctctaccc aaatggaggt tcctgtggtg gtatctcaaa attggactgg    780
cctactagaa tgagaatagc tcttgaagca gcaaaaggtc tggaatatct tcacgagcgt    840
gttaatccac ctgttataca cagagacttc aagagcagca acattctact ggacaaggac    900
ttccgtgcaa gagtgtcaga ttttggtttg gctaagcttg gatctgatag agctggtggt    960
catgtatcaa ctcgagttct aggcacacaa ggttatgttg caccagaata tgctttgacc   1020
gggcatttga cgaccaaatc cgatgtctat agttacggtg ttgtacttct ggagttgctt   1080
actggcaggg tgccagttga tatgaagagg cctccaggcg aaggtgttct agttaactgg   1140
gcattgccta tgctcacaga ccgagagaag gttgtgcaga tcttggatcc tgcactggag   1200
```

ggtcaatatt cattgaaaga tgctgtccaa gtggctgcca ttgctgccat gtgtgtttaa 1260
caagaggctg actacaggcc actaatggct gatgtggttc aatcgttggt tccgcttgtc 1320
aagaatcgtt ctactccaaa gacgtgcaac cctagtgtcc aagcgtagaa gccacttgaa 1380
ggggaagttg aatgctcagt ttccaggttg gtgcttgact tttctggaaa ttttcatcta 1440
catctaacaa ctactgctga gtttatgaga tcagtatgct ccataactta attctcttcc 1500
tctgccagtg aattcctgat tttgcgaggt gaatcgtaag cagttagatt tagaaatcaa 1560
gcattaattt agttcgatgg accagaaagc tagtttcagt ttgaagggaa gtcctctagt 1620
catgtgcaat tagtgatcat gttaggtggt ttggcttagt acttatggtt gtgccctgtg 1680
ggaatgatca gagagtgttg tttccagctg gcagaattgc aatgtactgc tgaattttct 1740
tttggcttgt caattatgtt aggtgtcttt ggagatccat gttcgttgtc ttaactcttg 1800
tgctaatacg ttggtctcat cagcttggcc tataaacatc gccgatgtat ctttttgtat 1860
atgtatatag taggattctg gaacttataa aaacaaacat ttgccagttg agcattttca 1920
tgc 1923

<210> 32
<211> 419
<212> PRT
<213> Oryza sativa
<400> 32

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

```
       355                 360                 365
Lys Arg Pro Pro Gly Glu Gly Val Leu Val Asn Trp Ala Leu Pro Met
   370                 375                 380
Leu Thr Asp Arg Glu Lys Val Val Gln Ile Leu Asp Pro Ala Leu Glu
385                 390                 395                 400
Gly Gln Tyr Ser Leu Lys Asp Ala Val Gln Val Ala Ala Ile Ala Ala
                405                 410                 415
Met Cys Val
```

<210> 33
<211> 3912
<212> DNA
<213> Oryza sativa
<400> 33

```
atggcgcgga tacggcgcgc gagctccgga cgagacatgt cagataattt ggtgcagcat     60
aacagacgct cagaagcaga aatttctact catgtagatg ctgcgcctcc tgatgcagcc    120
tcaaatacta gtgcagctcc ttctggatta gtacaacctc cagtatctcc tcacaatgca    180
tgttgttcac ataacatggt acaaaagaga gggagccaag attgccattg tgtttaccca    240
gtaagagttg agctttttct ccgaaatgtt tcactgacgt caaattggag cgatgaattt    300
ctgggggaac ttgcttctca actcagtctg agggttactc agtttgagat tgtaaatttc    360
tatgttgttg gggcttctgg gctaaacatc acaatgtata tagctcccca tacaggaatt    420
agtttctcag ccgaccaagt taccgcaatg aattattcac ttagtcagca tacagttcag    480
atcaaccctg tactagttgg tgactacaat cttcttaatt taacctggtt caggccactg    540
gttctagctc ctgtttattt tagtgttaag gtagagaaca ctaaaaggat gtttggtgat    600
atcaatatga tacttcctat gaatgatata aacattattt ctttatattt acagccctct    660
ccaacattca caatatcacc taaaccctct ccatctcaag catctacagt accaagacac    720
agtgcggata ccagcaatga aaaacacatg agcttgatta ctatcatttg catatttatt    780
ggtgctctaa ttgctgtctt ggtgattgcc atgtttattt gcttctgcaa attaaggaaa    840
gggaaaagga aggttcctcc agttgagaca cccaagcaaa gaacaccaga tgcggtttct    900
gcagtggact cacttcctcg cccaacaagt acaaggttcc ttgcatatga tgaactgaaa    960
gaagcaacca acaactttga tccgtcaagt atgcttggag aaggaggttt tggccgtgtc   1020
tttaaagggg tactaactga tggcactgcc gttgctataa agaagcttac tagtggaggg   1080
caccaaggag ataaggaatt tctagttgaa gtggagatgc tgagcaggtt gcaccaccga   1140
aaccttgtga aactcattgg ttactatagt aaccgtacat tgggagctag ccgccccttg   1200
gactgggaca ctagaatgag gatagcacta gatgctgcca ggggattggc ataccttcat   1260
gaggattctc agccttgcgt aatccacagg gatttcaagg cttctaatat attgcttgag   1320
gatgactttc atgctaaagt gtctgatttt ggtttggcca aacaggcacc ggaaggtcgc   1380
acaaattatc tctcaacacg tgttatggga acattcgggt atgtagcacc agagtatgca   1440
atgacaggac acttgcttgt aaaaagcgat gtatatagct atggagttgt tctacttgaa   1500
ctacttactg ggaggaggcc tgtggatatg tcacaacctt ctggccagga aaatctagtg   1560
acatggctta cacaaatgtt tcctgatctt tccactaaaa gccttgtgtc acatcaacct   1620
ttgctggcca agttgtcagg tgtagagata cttatttgct caattttgac tacgcaggca   1680
cgaccaattt tacgagataa agatacgcta gaagaactag ctgatcccaa gcttggtggc   1740
cagtatccaa aagatgattt tgtgcgagta tgcacaattg cagccgcctg tgtttcaccg   1800
gaggcaagcc aaaggccaac aatgggcgag gtggtgcagt cactgaagat ggtccaacgc   1860
tctgagttcc aggaatccat accaacacct ccagcacgac ccaatgcctc agccccttg    1920
agactgagaa catatcgaga acggctttct ctgaagatct tcacgaaggg cggtaacagt   1980
ggtgaacaag acccggccac catggtcggc ttcgccgacc agaccgccga cgtctccgcc   2040
gccgcgttcc agaccatgta ccgcctcaac gtcggcggcg cctacatccc gccgtccaac   2100
gactccggcc tcacgcggcc gtggtacgac gacacgccgt tcgtccaggg ccccctgcgc   2160
ggcctcgtct acaacgccgg cccgcacttc cacatcaagt accccagcga cgccgccgag   2220
tacgccgcgc cgccggaggt gtacctcggc ggccgctcca tgggcaggga ccagcgcctc   2280
aaccagaact ccaacctcac ctggagcctc cacgtggagt gcaacttcac ctacgtcgtg   2340
cgcctccatt tctgcgagct ccagctcatc cacggcaacc agcgggtgtt cgacatctac   2400
atcaacaacc ggacggcgca gacggacgtg gacgtgctgg agatggcgac ggagcgcggc   2460
gtgccggtgt acaaggacta cgcggtgagg ctgagcaacg acaccgccga cgagcatctg   2520
tgggtggcgg tgcacccctc ggtgatgctc cgcccgcagt tctacgacgc catcctcaac   2580
ggcctcgagg tgttcaaggt gaacaacacc ggcggcagcc tggcgtcgcc ggaccccgtc   2640
ccgtacaagc tgctcgcgga gaaggagctc ggctggggcg ggccgccgga gttctccacc   2700
gacaacccgg ccaacatggc gtcggtgatg ggcggcacgg cgggcggcgc ggcggcggcc   2760
gggatcgtgg cggccatctg cgtggtggtg tacagcaaca agaggagcaa gaagctgggc   2820
ggcggcggcg ccgactcgca cacctccgcg tggctgccgc tgtaccactc ccacaccagc   2880
ggcaagtcgt cggggcacat cacggcgaac atcgccggca tgtgccgcca cttctcgttc   2940
```

```
gcggagatca aggcggcgac caagaacttc tccaacgacc tggccatcgg cgtgggcggg   3000
ttcggcgtgg tgtaccgcgg cgtggtggac ggcgacgtga aggtggcggt gaagcggtcc   3060
aacccgtcgt cggagcaagg cataaccgag ttccagacgg aggtggagat gctctccaag   3120
ctccgccacc gccacctcgt ctccctcatc ggcttctgcg aggaggacgg cgagatggtg   3180
ctcgtctacg actacatgga gcacggcacc ctgcgcgagc acctctacca caacggcggc   3240
aagcccacgc tgtcgtggcg ccaccgcctc gacatctgca tcggcgccgc ccgcggcctc   3300
cactacctcc acaccggcga atcgcacgtc agcaccgtcg tcaagggcag cttcggctac   3360
ctcgacccgg agtactaccg ccggcagcag ctcaccgaca agtccgacgt ctactccttc   3420
ggcgtcgtgc tgttcgaggt gctcatggcg cgcccggcgc tcgacccggc gctgccgcgc   3480
gaccaggtca gcctcgccga ctacgcgctc gcctgcaagc gcggcggcgc gctgccggac   3540
gtcgtcgacc cggcgatcag ggaccagatc gcgcccgagt gcctcgccaa gttcgccgac   3600
acggcggaga agtgcctctc cgagaacggc accgagcgcc ccaccatggg cgacgtgctc   3660
tggaacctcg agagcgcgat gcacttccag gacgcgttcg acgccgccgc cggccgcccc   3720
gtccccgcgc tcgacgccgc cgccggcagc agcagccacc tcgacgacgg gagcacggcc   3780
agcatcaaca cgctggccac gtcgtccacg tcgcacccgc acgagccgtg cgtcgacgtt   3840
gtcttggagc ccgacgacgt cgtcgcggag cgcgccacct tctcgcagct cgtccagccc   3900
accggccggt ga                                                      3912
<210> 34
<211> 1303
<212> PRT
<213> Oryza sativa
<400> 34
Met Ala Arg Ile Arg Arg Ala Ser Ser Gly Arg Asp Met Ser Asp Asn
1               5                  10                  15
Leu Val Gln His Asn Arg Arg Ser Glu Ala Glu Ile Ser Thr His Val
            20                  25                  30
Asp Ala Ala Pro Pro Asp Ala Ala Ser Asn Thr Ser Ala Ala Pro Ser
        35                  40                  45
Gly Leu Val Gln Pro Pro Val Ser Pro His Asn Ala Cys Cys Ser His
    50                  55                  60
Asn Met Val Gln Lys Arg Gly Ser Gln Asp Cys His Cys Val Tyr Pro
65                  70                  75                  80
Val Arg Val Glu Leu Phe Leu Arg Asn Val Ser Leu Thr Ser Asn Trp
                85                  90                  95
Ser Asp Glu Phe Leu Gly Glu Leu Ala Ser Gln Leu Ser Leu Arg Val
            100                 105                 110
Thr Gln Phe Glu Ile Val Asn Phe Tyr Val Val Gly Ala Ser Gly Leu
        115                 120                 125
Asn Ile Thr Met Tyr Ile Ala Pro His Thr Gly Ile Ser Phe Ser Ala
    130                 135                 140
Asp Gln Val Thr Ala Met Asn Tyr Ser Leu Ser Gln His Thr Val Gln
145                 150                 155                 160
Ile Asn Pro Val Leu Val Gly Asp Tyr Asn Leu Leu Asn Leu Thr Trp
                165                 170                 175
Phe Arg Pro Leu Val Leu Ala Pro Val Tyr Phe Ser Val Lys Val Glu
            180                 185                 190
Asn Thr Lys Arg Met Phe Gly Asp Ile Asn Met Ile Leu Pro Met Asn
        195                 200                 205
Asp Ile Asn Ile Ile Ser Leu Tyr Leu Gln Pro Ser Pro Thr Phe Thr
    210                 215                 220
Ile Ser Pro Lys Pro Ser Pro Ser Gln Ala Ser Thr Val Pro Arg His
225                 230                 235                 240
Ser Ala Asp Thr Ser Asn Glu Lys His Met Ser Leu Ile Thr Ile Ile
                245                 250                 255
Cys Ile Phe Ile Gly Ala Leu Ile Ala Val Leu Val Ile Ala Met Phe
            260                 265                 270
Ile Cys Phe Cys Lys Leu Arg Lys Gly Lys Arg Lys Val Pro Pro Val
        275                 280                 285
Glu Thr Pro Lys Gln Arg Thr Pro Asp Ala Val Ser Ala Val Asp Ser
    290                 295                 300
Leu Pro Arg Pro Thr Ser Thr Arg Phe Leu Ala Tyr Asp Glu Leu Lys
305                 310                 315                 320
Glu Ala Thr Asn Asn Phe Asp Pro Ser Ser Met Leu Gly Glu Gly Gly
```

```
                325                 330                 335
Phe Gly Arg Val Phe Lys Gly Val Leu Thr Asp Gly Thr Ala Val Ala
            340                 345                 350
Ile Lys Lys Leu Thr Ser Gly Gly His Gln Gly Asp Lys Glu Phe Leu
        355                 360                 365
Val Glu Val Glu Met Leu Ser Arg Leu His His Arg Asn Leu Val Lys
    370                 375                 380
Leu Ile Gly Tyr Tyr Ser Asn Arg Thr Leu Gly Ala Ser Arg Pro Leu
385                 390                 395                 400
Asp Trp Asp Thr Arg Met Arg Ile Ala Leu Asp Ala Ala Arg Gly Leu
                405                 410                 415
Ala Tyr Leu His Glu Asp Ser Gln Pro Cys Val Ile His Arg Asp Phe
            420                 425                 430
Lys Ala Ser Asn Ile Leu Leu Glu Asp Asp Phe His Ala Lys Val Ser
        435                 440                 445
Asp Phe Gly Leu Ala Lys Gln Ala Pro Glu Gly Arg Thr Asn Tyr Leu
    450                 455                 460
Ser Thr Arg Val Met Gly Thr Phe Gly Tyr Val Ala Pro Glu Tyr Ala
465                 470                 475                 480
Met Thr Gly His Leu Leu Val Lys Ser Asp Val Tyr Ser Tyr Gly Val
                485                 490                 495
Val Leu Leu Glu Leu Leu Thr Gly Arg Arg Pro Val Asp Met Ser Gln
            500                 505                 510
Pro Ser Gly Gln Glu Asn Leu Val Thr Trp Leu Thr Gln Met Phe Pro
        515                 520                 525
Asp Leu Ser Thr Lys Ser Leu Val Ser His Gln Pro Leu Leu Ala Lys
    530                 535                 540
Leu Ser Gly Val Glu Ile Leu Ile Cys Ser Ile Leu Thr Thr Gln Ala
545                 550                 555                 560
Arg Pro Ile Leu Arg Asp Lys Asp Thr Leu Glu Glu Leu Ala Asp Pro
                565                 570                 575
Lys Leu Gly Gly Gln Tyr Pro Lys Asp Asp Phe Val Arg Val Cys Thr
            580                 585                 590
Ile Ala Ala Ala Cys Val Ser Pro Glu Ala Ser Gln Arg Pro Thr Met
        595                 600                 605
Gly Glu Val Val Gln Ser Leu Lys Met Val Gln Arg Ser Glu Phe Gln
    610                 615                 620
Glu Ser Ile Pro Thr Pro Pro Ala Arg Pro Asn Ala Ser Ala Pro Leu
625                 630                 635                 640
Arg Leu Arg Thr Tyr Arg Glu Arg Leu Ser Leu Lys Ile Phe Thr Lys
                645                 650                 655
Gly Gly Asn Ser Gly Glu Gln Asp Pro Ala Thr Met Val Gly Phe Ala
            660                 665                 670
Asp Gln Thr Ala Asp Val Ser Ala Ala Ala Phe Gln Thr Met Tyr Arg
        675                 680                 685
Leu Asn Val Gly Gly Ala Tyr Ile Pro Pro Ser Asn Asp Ser Gly Leu
    690                 695                 700
Thr Arg Pro Trp Tyr Asp Asp Thr Pro Phe Val Gln Gly Pro Leu Arg
705                 710                 715                 720
Gly Leu Val Tyr Asn Ala Gly Pro His Phe His Ile Lys Tyr Pro Ser
                725                 730                 735
Asp Ala Ala Glu Tyr Ala Ala Pro Pro Glu Val Tyr Leu Gly Gly Arg
            740                 745                 750
Ser Met Gly Arg Asp Gln Arg Leu Asn Gln Asn Ser Asn Leu Thr Trp
        755                 760                 765
Ser Leu His Val Glu Cys Asn Phe Thr Tyr Val Val Arg Leu His Phe
    770                 775                 780
Cys Glu Leu Gln Leu Ile His Gly Asn Gln Arg Val Phe Asp Ile Tyr
785                 790                 795                 800
Ile Asn Asn Arg Thr Ala Gln Thr Asp Val Asp Val Leu Glu Met Ala
                805                 810                 815
Thr Glu Arg Gly Val Pro Val Tyr Lys Asp Tyr Ala Val Arg Leu Ser
            820                 825                 830
```

```
Asn Asp Thr Ala Asp Glu His Leu Trp Val Ala Val His Pro Ser Val
        835                 840                 845
Met Leu Arg Pro Gln Phe Tyr Asp Ala Ile Leu Asn Gly Leu Glu Val
    850                 855                 860
Phe Lys Val Asn Asn Thr Gly Gly Ser Leu Ala Ser Pro Asp Pro Val
865                 870                 875                 880
Pro Tyr Lys Leu Leu Ala Glu Lys Glu Leu Gly Trp Gly Gly Pro Pro
                885                 890                 895
Glu Phe Ser Thr Asp Asn Pro Ala Asn Met Ala Ser Val Met Gly Gly
            900                 905                 910
Thr Ala Gly Gly Ala Ala Ala Ala Gly Ile Val Ala Ala Ile Cys Val
        915                 920                 925
Val Val Tyr Ser Asn Lys Arg Ser Lys Lys Leu Gly Gly Gly Gly Ala
    930                 935                 940
Asp Ser His Thr Ser Ala Trp Leu Pro Leu Tyr His Ser His Thr Ser
945                 950                 955                 960
Gly Lys Ser Ser Gly His Ile Thr Ala Asn Ile Ala Gly Met Cys Arg
                965                 970                 975
His Phe Ser Phe Ala Glu Ile Lys Ala Ala Thr Lys Asn Phe Ser Asn
            980                 985                 990
Asp Leu Ala Ile Gly Val Gly Gly  Phe Gly Val Val Tyr  Arg Gly Val
        995                 1000                1005
Val Asp  Gly Asp Val Lys Val  Ala Val Lys Arg Ser  Asn Pro Ser
   1010                1015                1020
Ser Glu  Gln Gly Ile Thr Glu  Phe Gln Thr Glu Val  Glu Met Leu
   1025                1030                1035
Ser Lys  Leu Arg His Arg His  Leu Val Ser Leu Ile  Gly Phe Cys
   1040                1045                1050
Glu Glu  Asp Gly Glu Met Val  Leu Val Tyr Asp Tyr  Met Glu His
   1055                1060                1065
Gly Thr  Leu Arg Glu His Leu  Tyr His Asn Gly Gly  Lys Pro Thr
   1070                1075                1080
Leu Ser  Trp Arg His Arg Leu  Asp Ile Cys Ile Gly  Ala Ala Arg
   1085                1090                1095
Gly Leu  His Tyr Leu His Thr  Gly Glu Ser His Val  Ser Thr Val
   1100                1105                1110
Val Lys  Gly Ser Phe Gly Tyr  Leu Asp Pro Glu Tyr  Tyr Arg Arg
   1115                1120                1125
Gln Gln  Leu Thr Asp Lys Ser  Asp Val Tyr Ser Phe  Gly Val Val
   1130                1135                1140
Leu Phe  Glu Val Leu Met Ala  Arg Pro Ala Leu Asp  Pro Ala Leu
   1145                1150                1155
Pro Arg  Asp Gln Val Ser Leu  Ala Asp Tyr Ala Leu  Ala Cys Lys
   1160                1165                1170
Arg Gly  Gly Ala Leu Pro Asp  Val Val Asp Pro Ala  Ile Arg Asp
   1175                1180                1185
Gln Ile  Ala Pro Glu Cys Leu  Ala Lys Phe Ala Asp  Thr Ala Glu
   1190                1195                1200
Lys Cys  Leu Ser Glu Asn Gly  Thr Glu Arg Pro Thr  Met Gly Asp
   1205                1210                1215
Val Leu  Trp Asn Leu Glu Ser  Ala Met His Phe Gln  Asp Ala Phe
   1220                1225                1230
Asp Ala  Ala Ala Gly Arg Pro  Val Pro Ala Leu Asp  Ala Ala Ala
   1235                1240                1245
Gly Ser  Ser Ser His Leu Asp  Asp Gly Ser Thr Ala  Ser Ile Asn
   1250                1255                1260
Thr Leu  Ala Thr Ser Ser Thr  Ser His Pro His Glu  Pro Cys Val
   1265                1270                1275
Asp Val  Val Leu Glu Pro Asp  Asp Val Val Ala Glu  Arg Ala Thr
   1280                1285                1290
Phe Ser  Gln Leu Val Gln Pro  Thr Gly Arg
   1295                1300
<210> 35
```

<211> 2463
<212> DNA
<213> Oryza sativa
<400> 35

```
atggcgcgga tacggcgcgc gagctccgga cgagctcaac tctcaggccg tgagaatttg     60
ggtttggtgt cgcgggaatg gtggtcatac tactatgtag gattcagtgc tctctgctgc    120
cgtggcgata ctggatatct ggaatctgat cgaagaacct ctgatctaag tttgagtcac    180
tacaaatgga tttgtgatat tgattatcaa attcaggcca atggatcctc ctttacaaga    240
aaatggtcat tactgaactc ttgttactgg aatgatgaga attcctgcat tgctgggttt    300
gcggaaaggc ttttggaaga tgctcgaagg aaaaccttct taaatttttt ggctgtggta    360
tttactctgg aattgattac cagaatcaat ttgattgccc aaacatcagc tctaaatgtg    420
gttgcccctg ctatatctcc atctcaaagt tggagaccag ttcgctccat gttgtcaaaa    480
gctaaagttg acattagtat ttcagtcagc gaacaaagac gaaagaagct atattcatca    540
ccagctactc tatctgtgca ccctccaatg tcagcgccaa gctatagctc catttctgac    600
atgtcagata atttggtgca gcataacaga cgctcagaag cagaaatttc tactcatgta    660
gatgctgcgc ctcctgatgc agcctcaaat actagtgcag ctccttctgg attagtacaa    720
cctccagtat ctcctcacaa tgcatgttgt tcacataaca tggtacaaaa gagagggagc    780
caagattgcc attgtgttta cccagtaaga gttgagcttt ttctccgaaa tgtttcactg    840
acgtcaaatt ggagcgatga atttctgggg gaacttgctt ctcaactcag tctgagggtt    900
actcagtttg agattgtaaa tttctatgtt gttggggctt ctgggctaaa catcacaatg    960
tatatagctc cccatacagg aattagtttc tcagccgacc aagttaccgc aatgaattat   1020
tcacttagtc agcatacagt tcagatcaac cctgtactag ttggtgacta caatcttctt   1080
aatttaacct ggttcaggcc actggttcta gctcctgctc caacattcac aatatcacct   1140
aaaccctctc catctcaagc atctacagta ccaagacaca gtgcggatac cagcaatgaa   1200
aaacacatga gcttgattac tatcatttgc atatttattg gtgctctaat tgctgtcttg   1260
gtgattgcca tgtttatttg cttctgcaaa ttaaggaaag ggaaaaggaa ggttcctcca   1320
gttgagacac ccaagcaaag aacaccagat gcggtttctg cagtggactc acttcctcgc   1380
ccaacaagta caaggttcct tgcatatgat gaactgaaag aagcaaccaa caactttgat   1440
ccgtcaagta tgcttggaga aggaggtttt ggccgtgtct ttaaaggggt actaactgat   1500
ggcactgccg ttgctataaa gaagcttact agtggagggc accaaggaga taaggaattt   1560
ctagttgaag tggagatgct gagcaggttg caccaccgaa accttgtgaa actcattggt   1620
tactatagta accgtgagtc atcacagaac ctactgtgct atgagcttgt tcctaatgga   1680
agcctagagg cttggcttca tggtacattg ggagctagcc gccccttgga ctgggacact   1740
agaatgagga tagcactaga tgctgccagg ggattggcat accttcatga ggattctcag   1800
ccttgcgtaa tccacaggga tttcaaggct tctaatatat tgcttgagga tgactttcat   1860
gctaaagtgt ctgattttgg tttggccaaa caggcaccgg aaggttgcac aaattatctc   1920
tcaacacgtg ttatgggaac attcgggtat gtagcaccag agtatgcaat gacaggacac   1980
ttgcttgtaa aaagcgatgt atatagctat ggagttgttc tacttgaact acttactggg   2040
aggaggcctg tggatatgtc acaaccttct ggccaggaaa atctagtgac atgggcacga   2100
ccaattttac gagataaaga tacgctagaa gaactagctg atcccaagct tggtggccag   2160
tatccaaaag atgattttgt gcgagtatgc acaattgcag ccgcctgtgt ttcaccggag   2220
gcaagccaaa ggccaacaat gggcgaggtg gtgcagtcac tgaagatggt ccaacgctct   2280
gagttccagg aatccatacc aacacctcca gcacgaccca atgttaggca gtcctcaacc   2340
acctatgaat ctgatggcac ttcatccatg ttctcttctg gacccttttc aggcctcagc   2400
ccctttgaga ctgagaacat atcgagaacg gctttctctg aagatcttca cgaagggcgg   2460
taa                                                                2463
```

<210> 36
<211> 820
<212> PRT
<213> Oryza sativa
<400> 36

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

```
Ile Ala Gly Phe Ala Glu Arg Leu Leu Glu Asp Ala Arg Arg Lys Thr
          100                 105                 110
Phe Leu Asn Phe Leu Ala Val Val Phe Thr Leu Glu Leu Ile Thr Arg
       115                 120                 125
Ile Asn Leu Ile Ala Gln Thr Ser Ala Leu Asn Val Val Ala Pro Ala
   130                 135                 140
Ile Ser Pro Ser Gln Ser Trp Arg Pro Val Arg Ser Met Leu Ser Lys
145                 150                 155                 160
Ala Lys Val Asp Ile Ser Ile Ser Val Ser Glu Gln Arg Arg Lys Lys
                165                 170                 175
Leu Tyr Ser Ser Pro Ala Thr Leu Ser Val His Pro Pro Met Ser Ala
          180                 185                 190
Pro Ser Tyr Ser Ser Ile Ser Asp Met Ser Asp Asn Leu Val Gln His
       195                 200                 205
Asn Arg Arg Ser Glu Ala Glu Ile Ser Thr His Val Asp Ala Ala Pro
   210                 215                 220
Pro Asp Ala Ala Ser Asn Thr Ser Ala Ala Pro Ser Gly Leu Val Gln
225                 230                 235                 240
Pro Pro Val Ser Pro His Asn Ala Cys Cys Ser His Asn Met Val Gln
                245                 250                 255
Lys Arg Gly Ser Gln Asp Cys His Cys Val Tyr Pro Val Arg Val Glu
          260                 265                 270
Leu Phe Leu Arg Asn Val Ser Leu Thr Ser Asn Trp Ser Asp Glu Phe
       275                 280                 285
Leu Gly Glu Leu Ala Ser Gln Leu Ser Leu Arg Val Thr Gln Phe Glu
   290                 295                 300
Ile Val Asn Phe Tyr Val Val Gly Ala Ser Gly Leu Asn Ile Thr Met
305                 310                 315                 320
Tyr Ile Ala Pro His Thr Gly Ile Ser Phe Ser Ala Asp Gln Val Thr
                325                 330                 335
Ala Met Asn Tyr Ser Leu Ser Gln His Thr Val Gln Ile Asn Pro Val
          340                 345                 350
Leu Val Gly Asp Tyr Asn Leu Leu Asn Leu Thr Trp Phe Arg Pro Leu
       355                 360                 365
Val Leu Ala Pro Ala Pro Thr Phe Thr Ile Ser Pro Lys Pro Ser Pro
   370                 375                 380
Ser Gln Ala Ser Thr Val Pro Arg His Ser Ala Asp Thr Ser Asn Glu
385                 390                 395                 400
Lys His Met Ser Leu Ile Thr Ile Ile Cys Ile Phe Ile Gly Ala Leu
                405                 410                 415
Ile Ala Val Leu Val Ile Ala Met Phe Ile Cys Phe Cys Lys Leu Arg
          420                 425                 430
Lys Gly Lys Arg Lys Val Pro Pro Val Glu Thr Pro Lys Gln Arg Thr
       435                 440                 445
Pro Asp Ala Val Ser Ala Val Asp Ser Leu Pro Arg Pro Thr Ser Thr
   450                 455                 460
Arg Phe Leu Ala Tyr Asp Glu Leu Lys Glu Ala Thr Asn Asn Phe Asp
465                 470                 475                 480
Pro Ser Ser Met Leu Gly Glu Gly Gly Phe Gly Arg Val Phe Lys Gly
                485                 490                 495
Val Leu Thr Asp Gly Thr Ala Val Ala Ile Lys Lys Leu Thr Ser Gly
          500                 505                 510
Gly His Gln Gly Asp Lys Glu Phe Leu Val Glu Val Glu Met Leu Ser
       515                 520                 525
Arg Leu His His Arg Asn Leu Val Lys Leu Ile Gly Tyr Tyr Ser Asn
   530                 535                 540
Arg Glu Ser Ser Gln Asn Leu Leu Cys Tyr Glu Leu Val Pro Asn Gly
545                 550                 555                 560
Ser Leu Glu Ala Trp Leu His Gly Thr Leu Gly Ala Ser Arg Pro Leu
                565                 570                 575
Asp Trp Asp Thr Arg Met Arg Ile Ala Leu Asp Ala Ala Arg Gly Leu
          580                 585                 590
Ala Tyr Leu His Glu Asp Ser Gln Pro Cys Val Ile His Arg Asp Phe
```

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

<210> 37
<211> 2670
<212> DNA
<213> Oryza sativa
<400> 37

atggggcggc gtggaggagg aggacgagcg ggaggcgcgt ggattggtgg ctgtgtactt 60
gcgctcgccg ccaccttggt cgtctgcgtg ctcgtgatcc gtggagctgg aggactaaat 120
gtaaaaccac ctgcagcaac tagcagacgt gttaacatac agcaggaact atatgcacca 180
agcccaacga tcagccacag aggtcttgct attcctccac ttccaacaac ttcatcccca 240
gttttcccac ctcccatcag atcttatcct ccacttcctg caaataagcc ataccccaat 300
agtccagttg tggcacctcc aaagggaagg catcaccatt ctttgcctgt aaataatacc 360
cgtgtaaaag gacctgccta ttctccttca aattctccca gtatccatag aaaacatggc 420
attccagttg ctgcacctcc aaagcaacat tccagcaatt taccaccttc acatcatagg 480
ccccacaaag gatcctttcc tgtcataagc cctactccac ataaagctga taatgcttct 540
gcgacgaaac atggacgttc tggcttacac catagtcctg cacctgcacc tgtaggcttg 600
cctccatctg aaggaaatgc acgaggaaat cctgcgtatg caccacgtca tccccatgaa 660
tatcactcgc cttcaaattc tccagaacct ggtctaccac ctgtgaatcc gcctgacagt 720
catgcattta aaaagcccaa gagtttggca ccagcacccc aatcatttcc gccaccgcct 780
ctgagttcat attgcatggc gttaaattgt caagatcctc tgaccaatag tctccctgga 840
actacatgtc tttgtgtgtg gccaataaaa gttgagcttc gtttaggtat agctttgtac 900
acattctttg cattggtatc agaacttgca caagatattg catctggagt gctcatgaaa 960
caaagtcaag ttcgtgtaat gggagcaaat gctgcaactg aggacccaga gaaaacagtt 1020
gtccttattg atcttgtgcc actgggagaa aaatttgaca aggcaacagc acttttagta 1080
tttgaaaggt tttggcacaa gcaggtcaac ataaactcta tgcattttgg aaactatgat 1140
gtgttatatg ttacctacca aggtcttcct ccgtcgccac caacagctcc cgggatgaac 1200
aatggtctta gcaatgttaa tgatccaagg ttgcatccac ttgctgttga tgtgggaaac 1260
cacagagaaa caaaaagcag gggcataatt gttataattg ttctttcaag tgtctttgct 1320
tttatttttat gttctggagc tgcgttggta atttgtttca agattagaaa ccgcaaccat 1380
ttaactgaag agtcacctat gccaccgaag cctgcaggtc ctgggtctgc agttgttggg 1440
agcaggctag gaagcagacc tatttcagca tctccatctt tcagctcaag catagtgaca 1500
tataaaggga cagcgaaaac atttagcttg attgagatgg aaagagctac acaaagattt 1560
gacaactcca gaattattgg cgagggtggt tttggacgtg tttatgaagg tattcttgag 1620
gatggagaac gggttgctgt caaaattctt aagagggacg accagcaagg tacacgggaa 1680
tttttagctg aggttgagat gcttagccga ttgcatcaca gaaacctggt taagttgata 1740

ggtatatgca cagaagaaca tatccggtgt cttgtgtatg agcttgttcc aaatggtagt 1800
gtggaatctc acttgcacgg atcagataag ggaactgctc cactttattg ggatgctagg 1860
cttaaaattg cacttggtgc tgcacgtgct cttgcttatt tgcatgaaga ttctagtccc 1920
cgtgtcatac accgtgactt caagtcaagt aacatcttat tggaacatga cttcaccccc 1980
aaggtgtcag actttggcct agcaagaaca gccataggtg aggggaatga gcatatttca 2040
actcgtgtta tgggaacttt cgggtatgtt gctcctgaat acgcaatgac tgggcatctt 2100
ctagtaaaga gtgatgtcta cagctatggt gtcgtcctcc ttgagcttct taccggaagg 2160
aaaccggtgg atattttgag acctccaggg caagaaaatt tagttgcatg ggcttgtcct 2220
tttcttacta gcagagatgg cttggaaaca ataattgatc catcacttgg aaatagcatc 2280
ctatttgaca gcattgcaaa agtagcagct attgcttcta tgtgcgtgca gcctgaggtg 2340
gatcagcgcc catttatggg tgaagttgtt caagctttga agttggtatg cgatgaaggc 2400
agcgagttca atgaatcaag aagtttcagc caagatttac atattcagga ttctgggatt 2460
ataagtagag caagcctgga tgtggacgta gaacctgttg tatctgctga gctgttcaat 2520
gcatcagcac attatgacac tcttgatgca tctggttctt ttcgacgata ttcaagttca 2580
ggtcctctta gggtaggtag aaccgggcac aataggggta gctcaagcga gcactgtggt 2640
acacaaagat tcaggataga ttcagagtag 2670

<210> 38
<211> 889
<212> PRT
<213> Oryza sativa
<400> 38

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

```
Glu Lys Thr Val Val Leu Ile Asp Leu Val Pro Leu Gly Glu Lys Phe
            340                 345                 350
Asp Lys Ala Thr Ala Leu Leu Val Phe Glu Arg Phe Trp His Lys Gln
        355                 360                 365
Val Asn Ile Asn Ser Met His Phe Gly Asn Tyr Asp Val Leu Tyr Val
    370                 375                 380
Thr Tyr Gln Gly Leu Pro Pro Ser Pro Pro Thr Ala Pro Gly Met Asn
385                 390                 395                 400
Asn Gly Leu Ser Asn Val Asn Asp Pro Arg Leu His Pro Leu Ala Val
                405                 410                 415
Asp Val Gly Asn His Arg Glu Thr Lys Ser Arg Gly Ile Ile Val Ile
            420                 425                 430
Ile Val Leu Ser Ser Val Phe Ala Phe Ile Leu Cys Ser Gly Ala Ala
        435                 440                 445
Leu Val Ile Cys Phe Lys Ile Arg Asn Arg Asn His Leu Thr Glu Glu
    450                 455                 460
Ser Pro Met Pro Pro Lys Pro Ala Gly Pro Gly Ser Ala Val Val Gly
465                 470                 475                 480
Ser Arg Leu Gly Ser Arg Pro Ile Ser Ala Ser Pro Ser Phe Ser Ser
                485                 490                 495
Ser Ile Val Thr Tyr Lys Gly Thr Ala Lys Thr Phe Ser Leu Ile Glu
            500                 505                 510
Met Glu Arg Ala Thr Gln Arg Phe Asp Asn Ser Arg Ile Ile Gly Glu
        515                 520                 525
Gly Gly Phe Gly Arg Val Tyr Glu Gly Ile Leu Glu Asp Gly Glu Arg
    530                 535                 540
Val Ala Val Lys Ile Leu Lys Arg Asp Asp Gln Gln Gly Thr Arg Glu
545                 550                 555                 560
Phe Leu Ala Glu Val Glu Met Leu Ser Arg Leu His His Arg Asn Leu
                565                 570                 575
Val Lys Leu Ile Gly Ile Cys Thr Glu Glu His Ile Arg Cys Leu Val
            580                 585                 590
Tyr Glu Leu Val Pro Asn Gly Ser Val Glu Ser His Leu His Gly Ser
        595                 600                 605
Asp Lys Gly Thr Ala Pro Leu Tyr Trp Asp Ala Arg Leu Lys Ile Ala
    610                 615                 620
Leu Gly Ala Ala Arg Ala Leu Ala Tyr Leu His Glu Asp Ser Ser Pro
625                 630                 635                 640
Arg Val Ile His Arg Asp Phe Lys Ser Ser Asn Ile Leu Leu Glu His
                645                 650                 655
Asp Phe Thr Pro Lys Val Ser Asp Phe Gly Leu Ala Arg Thr Ala Ile
            660                 665                 670
Gly Glu Gly Asn Glu His Ile Ser Thr Arg Val Met Gly Thr Phe Gly
        675                 680                 685
Tyr Val Ala Pro Glu Tyr Ala Met Thr Gly His Leu Leu Val Lys Ser
    690                 695                 700
Asp Val Tyr Ser Tyr Gly Val Val Leu Leu Glu Leu Leu Thr Gly Arg
705                 710                 715                 720
Lys Pro Val Asp Ile Leu Arg Pro Pro Gly Gln Glu Asn Leu Val Ala
                725                 730                 735
Trp Ala Cys Pro Phe Leu Thr Ser Arg Asp Gly Leu Glu Thr Ile Ile
            740                 745                 750
Asp Pro Ser Leu Gly Asn Ser Ile Leu Phe Asp Ser Ile Ala Lys Val
        755                 760                 765
Ala Ala Ile Ala Ser Met Cys Val Gln Pro Glu Val Asp Gln Arg Pro
    770                 775                 780
Phe Met Gly Glu Val Val Gln Ala Leu Lys Leu Val Cys Asp Glu Gly
785                 790                 795                 800
Ser Glu Phe Asn Glu Ser Arg Ser Phe Ser Gln Asp Leu His Ile Gln
                805                 810                 815
Asp Ser Gly Ile Ile Ser Arg Ala Ser Leu Asp Val Asp Val Glu Pro
            820                 825                 830
Val Val Ser Ala Glu Leu Phe Asn Ala Ser Ala His Tyr Asp Thr Leu
```

835 840 845
Asp Ala Ser Gly Ser Phe Arg Arg Tyr Ser Ser Ser Gly Pro Leu Arg
850 855 860
Val Gly Arg Thr Gly His Asn Arg Gly Ser Ser Ser Glu His Cys Gly
865 870 875 880
Thr Gln Arg Phe Arg Ile Asp Ser Glu
885

<210> 39
<211> 2913
<212> DNA
<213> Oryza sativa
<400> 39
atgcggtcgc tcgcttgttg cgcggtgctt cttctcgctt cggttcttgg atctacaggt 60
actgatctgg gtccttctcc tgtggtcgct aactcgccag atgctcaaga tcaaacttct 120
agccctcctg aacctacaat cgccctcggt ccagtaactc ttccaacagg ttgctctgaa 180
tttctaatgt ggtgtggcat tatggctctc aagatcataa ccagcattga caattgtcct 240
tacaattgct ttgtctcagc accctctgct ccatcagcaa gccctcctgt ggcgaagggc 300
gctgtcagcc cagctgttcc cactcgacca caaaatgcgc ctactccagt aacacccccct 360
aaagaatata atgcgcctcc tccagttgag cttacgcctc ctgctcccac tcatgtggcg 420
ccagtagctc ccccgcaagc tgcagttgaa aatcctgcac cagtgcttcc tgggacacct 480
gctttgttgc cttctgttca ggctcctgct ccatctgtgg ccaggaatcc taatctaccg 540
atagtgcaac ctccttctgt gaacaatcca ccaagcggac caattggatc agggaatggc 600
gtccctccgt atccaccacc tcaaagaagt ttacctgcca ttcctccttc cacttcagga 660
gttccgcgag aaagtgtaaa acctccagtt gcaccaccta ttattgcaca agcaccacgt 720
cagcaagcac tggcaccaag tagtgaccat agcaatggaa actcagtgcc cccagcaaat 780
acatcacctc cccataagaa tagtcatatt ccacgtgcac tgccaccaaa ggaatccagt 840
agtcaaactg gcacagctca caaaccgcca attagagggc ctcatatttc tccaaccatg 900
ccaccaatcc ctccccaacc agggccgaaa gcaccatcag cccatcctat ttgggcatta 960
cctccaccac cgcctaatct agattgcaat tcattagcat gcccagagcc tctaacagat 1020
ccacctgcgg gagccccatg tgtttgtgtt ctaccaatca aagttggagt ccgcttaagt 1080
gttgatctgt attcattctt tccattggta tctgattttg cggaagaagt gtcatctggg 1140
gtaaatatgg cacagagaca ggttcgtgtt atgggtgcaa atgtagctgg tgatcagcct 1200
gacaagacag tagttcttgt tgacctagta ccaatgcaag tgaagtttga caatgctact 1260
gctttttaa cgtttgaaaa cttatggagc aaaaaaattt ccctaaaacc atcagttttt 1320
ggggactacg agattctgta tgttgtatat ccagggcttc ctccttctcc accttcagca 1380
ccggaaagtg ttggtgacgg ggcatttgga aacaacagaa atgcaagggc aatgaagcct 1440
ctcggggttg atgtaggcag gcccaaaaaa agagtcaacg gcagcctaat tgctattgcc 1500
gtcttgtcta ctgttatagc attgattatt tgcactctag ctgcatggtt gctgataatc 1560
agattcaggg gttcggatgg cttggctcaa agatttccac atagcgcact tccaaaattt 1620
tccagatcat ctggtacagg tcaaacactg ttagctggtc gctatagttc accatcaggt 1680
ccatcaggtt cattgggctc aagtatcgca acatatgcag gacaggcaaa gacattcaaa 1740
tttgctgaga ttgagaaggc cacaaacagc tttgatgatt caacagtact tggagagggt 1800
ggcttcggat gtgtctacca gggcacgctt gaagatggaa ccagggttgc agtaaaggtt 1860
ctgaagaggt atgatggcca aggcgagaga gagttcttgg ctgaggttga gatgctggga 1920
cgcttgcatc accgaaattt ggtcaagttg ttaggcatat gtgtagagga gaacgcaaga 1980
tgtctggttt atgaacttat tccgaatggc agtgtagaat cccatttgca tggagtggat 2040
cttgagacag caccactcga ttggaatgcc cgcatgaaga tagccttggg ggctgcacga 2100
gctcttgcat atttacacga agattcaagc ccttgtgtga ttcatcgtga ttttaagtca 2160
agcaacatcc tattggagca tgatttcaca ccgaaagtgt ctgatttcgg actggccagg 2220
actgcaaggg gggaggggaa ccagcacatc tccactcgtg tcatgggaac atttggatat 2280
gttgcaccgg agtatgccat gacaggacat ctccttgtca agagcgatgt gtacagctat 2340
ggagtagtgt tgcttgagct cctcaccggt agaaagccag tggacatgtc taggcccggg 2400
gggcaagaaa acctagtttc atgggctcgc ccgcttctga ccaatgtggt aagcctacgt 2460
caagctgttg atccacttct tggccctaac gtacccctgg acaatgtcgc aaaagcagct 2520
gccatcgcat caatgtgtgt acaacctgag gttgcgcatc gcccgagcat gggtgaagta 2580
gtgcaggcct taaaacttgt ttgcagtgac ggtgatgagg gcctcggatc aggaagtttc 2640
agccaggaat tggcggcaca agctgcagca atctatgatg taactggcat ggaagctgag 2700
agggtgctgt tatctgagat gtttggctca acccctgtct tcactcccgc tgcggattca 2760
ggttctttcc gaaagcagtc cagctcaggt ccactgatga caggcaagaa cagaaagttc 2820
tggcagagat tgcgaagcct atcaagaggg agtatgagtg agcatggtgc ctcaccagat 2880
tttgagacgc gctcgcagtg tagtaatagg tga 2913
<210> 40
<211> 970

<212> PRT
<213> Oryza sativa
<400> 40

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

```
Leu Gly Val Asp Val Gly Arg Pro Lys Lys Arg Val Asn Gly Ser Leu
                485                 490                 495
Ile Ala Ile Ala Val Leu Ser Thr Val Ile Ala Leu Ile Ile Cys Thr
            500                 505                 510
Leu Ala Ala Trp Leu Leu Ile Ile Arg Phe Arg Gly Ser Asp Gly Leu
        515                 520                 525
Ala Gln Arg Phe Pro His Ser Ala Leu Pro Lys Phe Ser Arg Ser Ser
    530                 535                 540
Gly Thr Gly Gln Thr Leu Leu Ala Gly Arg Tyr Ser Ser Pro Ser Gly
545                 550                 555                 560
Pro Ser Gly Ser Leu Gly Ser Ser Ile Ala Thr Tyr Ala Gly Gln Ala
                565                 570                 575
Lys Thr Phe Lys Phe Ala Glu Ile Glu Lys Ala Thr Asn Ser Phe Asp
            580                 585                 590
Asp Ser Thr Val Leu Gly Glu Gly Gly Phe Gly Cys Val Tyr Gln Gly
        595                 600                 605
Thr Leu Glu Asp Gly Thr Arg Val Ala Val Lys Val Leu Lys Arg Tyr
    610                 615                 620
Asp Gly Gln Gly Glu Arg Glu Phe Leu Ala Glu Val Glu Met Leu Gly
625                 630                 635                 640
Arg Leu His His Arg Asn Leu Val Lys Leu Leu Gly Ile Cys Val Glu
                645                 650                 655
Glu Asn Ala Arg Cys Leu Val Tyr Glu Leu Ile Pro Asn Gly Ser Val
            660                 665                 670
Glu Ser His Leu His Gly Val Asp Leu Glu Thr Ala Pro Leu Asp Trp
        675                 680                 685
Asn Ala Arg Met Lys Ile Ala Leu Gly Ala Ala Arg Ala Leu Ala Tyr
    690                 695                 700
Leu His Glu Asp Ser Ser Pro Cys Val Ile His Arg Asp Phe Lys Ser
705                 710                 715                 720
Ser Asn Ile Leu Leu Glu His Asp Phe Thr Pro Lys Val Ser Asp Phe
                725                 730                 735
Gly Leu Ala Arg Thr Ala Arg Gly Glu Gly Asn Gln His Ile Ser Thr
            740                 745                 750
Arg Val Met Gly Thr Phe Gly Tyr Val Ala Pro Glu Tyr Ala Met Thr
        755                 760                 765
Gly His Leu Leu Val Lys Ser Asp Val Tyr Ser Tyr Gly Val Val Leu
    770                 775                 780
Leu Glu Leu Leu Thr Gly Arg Lys Pro Val Asp Met Ser Arg Pro Gly
785                 790                 795                 800
Gly Gln Glu Asn Leu Val Ser Trp Ala Arg Pro Leu Leu Thr Asn Val
                805                 810                 815
Val Ser Leu Arg Gln Ala Val Asp Pro Leu Leu Gly Pro Asn Val Pro
            820                 825                 830
Leu Asp Asn Val Ala Lys Ala Ala Ala Ile Ala Ser Met Cys Val Gln
        835                 840                 845
Pro Glu Val Ala His Arg Pro Ser Met Gly Glu Val Val Gln Ala Leu
    850                 855                 860
Lys Leu Val Cys Ser Asp Gly Asp Glu Gly Leu Gly Ser Gly Ser Phe
865                 870                 875                 880
Ser Gln Glu Leu Ala Ala Gln Ala Ala Ala Ile Tyr Asp Val Thr Gly
                885                 890                 895
Met Glu Ala Glu Arg Val Leu Leu Ser Glu Met Phe Gly Ser Thr Pro
            900                 905                 910
Val Phe Thr Pro Ala Ala Asp Ser Gly Ser Phe Arg Lys Gln Ser Ser
        915                 920                 925
Ser Gly Pro Leu Met Thr Gly Lys Asn Arg Lys Phe Trp Gln Arg Leu
    930                 935                 940
Arg Ser Leu Ser Arg Gly Ser Met Ser Glu His Gly Ala Ser Pro Asp
945                 950                 955                 960
Phe Glu Thr Arg Ser Gln Cys Ser Asn Arg
                965                 970
<210> 41
```

<211> 1317
<212> DNA
<213> Oryza sativa
<400> 41
ggcagatcaa ctttgtctgt tagcagggtg agctctgcat cggcgtcaat gctctcaaca 60
gtggcaactt gcacaacatc agtaaagaca ttttcacttt cccagcttga aaaggccact 120
gatggttttg attcaaaaag agtgttgggt caaggtgggt ttggacgcgt ctaccatggg 180
accatggatg gcggagacga gattgctgtt aaattgctca caagagaaga caggagtggc 240
gatcgagagt tcattgcgga ggttgaaatg ctcagtcgac tacatcaccg taatcttgtc 300
aaattgattg gcatttgcat tgagcacaac aaaaggtgtc ttgtttatga gcttatacgc 360
aatggaagtg ttgaatctca ccttcatggt gctgataagg ctaagggcat gctgaactgg 420
gatgttagga tgaaaattgc tctgggtgca gctagaggct tagcatatct acatgaagac 480
tcaaaccctc atgttataca tcgtgacttc aagggcagca atatattgtt ggaggaagat 540
ttcactccta aggttactga ttttggttta gcaagggagg caactaatgg aattcaaccc 600
atttctacta gggtaatggg tacttttggg tacgttgctc cagaatatgc catgaccggg 660
catcttcttg tgaagagtga tgtctacagt tacggtgttg tcttgctgga gcttttatca 720
ggaaggaagc ccgtatgcat gtctgatacc aatggtcctc agaaccttgt gacttgggct 780
cgtcctctgc tatgccataa ggagggtttg gagaggttga tcgacccttc tctgaatggc 840
aatttcaact ttgatgatgt agctaaagta gcatccatcg cttccatgtg tgttcacaat 900
gatccgtcac agaggccatt catgggtgaa gtagtgcagg cactgaagct tatttacaat 960
gatgcggagg cggcttgcga tgattcttac agccacaggg actcatcgtg tgaccagtac 1020
gatgactacc atggggccct ggcccttgac agcggtagtg gtagctggtg gaatagaagc 1080
tcaaatcctt ctgggttttt tgacaaccgg aacccctac cagttattac catggaatac 1140
agctctggca ggatcgaagg agcgcgtgac ccccgctttg cattgtcaac aggtggtcat 1200
gcacaatcac cagccctgca gaatagatca ggacccatcc ggatgaagaa aaagcttgcc 1260
tcattttacc ggtctagagg tagcttcagc gagcatggcc agctgcctcg ccattga 1317
<210> 42
<211> 438
<212> PRT
<213> Oryza sativa
<400> 42
Gly Arg Ser Thr Leu Ser Val Ser Arg Val Ser Ser Ala Ser Ala Ser
1 5 10 15
Met Leu Ser Thr Val Ala Thr Cys Thr Thr Ser Val Lys Thr Phe Ser
20 25 30
Leu Ser Gln Leu Glu Lys Ala Thr Asp Gly Phe Asp Ser Lys Arg Val
35 40 45
Leu Gly Gln Gly Gly Phe Gly Arg Val Tyr His Gly Thr Met Asp Gly
50 55 60
Gly Asp Glu Ile Ala Val Lys Leu Leu Thr Arg Glu Asp Arg Ser Gly
65 70 75 80
Asp Arg Glu Phe Ile Ala Glu Val Glu Met Leu Ser Arg Leu His His
85 90 95
Arg Asn Leu Val Lys Leu Ile Gly Ile Cys Ile Glu His Asn Lys Arg
100 105 110
Cys Leu Val Tyr Glu Leu Ile Arg Asn Gly Ser Val Glu Ser His Leu
115 120 125
His Gly Ala Asp Lys Ala Lys Gly Met Leu Asn Trp Asp Val Arg Met
130 135 140
Lys Ile Ala Leu Gly Ala Ala Arg Gly Leu Ala Tyr Leu His Glu Asp
145 150 155 160
Ser Asn Pro His Val Ile His Arg Asp Phe Lys Gly Ser Asn Ile Leu
165 170 175
Leu Glu Glu Asp Phe Thr Pro Lys Val Thr Asp Phe Gly Leu Ala Arg
180 185 190
Glu Ala Thr Asn Gly Ile Gln Pro Ile Ser Thr Arg Val Met Gly Thr
195 200 205
Phe Gly Tyr Val Ala Pro Glu Tyr Ala Met Thr Gly His Leu Leu Val
210 215 220
Lys Ser Asp Val Tyr Ser Tyr Gly Val Val Leu Leu Glu Leu Leu Ser
225 230 235 240
Gly Arg Lys Pro Val Cys Met Ser Asp Thr Asn Gly Pro Gln Asn Leu
245 250 255

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

<210> 43
<211> 2172
<212> DNA
<213> Oryza sativa
<400> 43

atgccgccgc ggcgagcggc gctgctcctc ctcgctctcg ccgtgtatgt gccactggga 60
acggcaagct caacaactat cgcgtcgtac ttacttgggt tatggagtag agcacatcgg 120
cattctcttc ctgcccctgc tccagctcca gctccagctc cagcccctga aactcaccga 180
cctggtatta gacacccagt gcctaggcat cacaggaaaa ggcctcatgt tgccccacca 240
ctaccaccac catcatcatc agaaagacaa gtagcaccat atcagttgtt tccaagaatt 300
gatgagttag aaattgagat tgcagcagga acatttctga aacagagtca agttcggata 360
atgggtgctg ggagtagtct tcaagatcct gaaaaaacta ctgtcaccgt tgatttggtg 420
ccactaggtc agaagtttga taggacttca gccttactga ctagcaacag atttttgcaa 480
aagaaggtgc caataaactc gtccatattt ggtgattata acgtaatata tgttcattat 540
cctggcctgc cttctttggt ccctagtgtt ccaggatcct tgggcccaat aagcagtagc 600
caatacccct ttagtgcaaa tgtacacaat agaagacatc agaagattaa ctctaaaagt 660
gttgccataa ttgcattatc agctgttgtg cttgtattaa tgagctttgg catttgcatc 720
atatggaaat ataaaggatt tgaaaagtct cgtggcactg gtcgtgtttc gaattcgtca 780
gccacaagaa aaacaggtat gaggtcatca ttctccagta tgaccagctc gacggcatct 840
tttgtttcaa caattgcaac atgcccacct acagttaaga cattctctat ttctgagctt 900
gaaaaggcca cagaaaactt cagcttcaac aaaataatag gtgaaggagg gtatggccgt 960
gtttatcgag gtacaattga tgatgaagtt gatgttgcag tcaaattgct tacaagaaaa 1020
catcaaaaca gagatcgtga gtttattgcc gaggttgaga tgctaagtcg tttgcatcat 1080
cgtaatcttg tcaagctgat cggtatatgc attgaacgga gcacaaggtg cttggtgttt 1140
gagcttgttc caaacggaag tgtggagtct catctgcacg gttctgataa aatatacggc 1200
ccacttgatt ttgatactag aatgaaaata gcccttggtg cagcaagggg tctggcctac 1260
ctacatgagg atgccaatcc ccatgttata caccgtgatt tcaaggccag caacgttctt 1320
ttggaaaatg atttcactcc caaggttgcg gatttcgggt tggcaaagga agcctcagaa 1380
ggaatggacc atatttctac tcaggtcatg gggacttttg ggtacgttgc cccggagtat 1440
gcgatgaccg ggcatctcct ggtgaaaagc gacgtgtaca gctacggtgt cgtgctgctg 1500
gagctcctct cggggaggaa gccggtggac atgacgcagc cgccggggtc ggagaacctc 1560
gtcacctggg cgcgcccccc cctcaccgac cgggacggcc tccagcagct ggtcgacccg 1620
tcgatgccgg cggcgagcta cggcttcgag aagctggcca aggcggcggc catcgcgtcc 1680
atgtgcgtgc acgtcgaggc gtcgcaccgg ccgttcatgg gcgaggtcgt gcaggccctg 1740
aagctcatct acaacggcaa caacgacgac acctgcacca gcggctcgtt cggcggcggc 1800
ggcggcgagg agtacgagga cgaggaggcg tcgtcgccgt ggaacaaccg ctcgtggagc 1860
cacgacttcg ccgcgacgcc gccgccggcg tcgcgcaggc tggcgttccc gcgggctccg 1920
gcccgcccga cgacgatgga ctacagctcg gacccggccg acggggcggc ggggacgtcg 1980
tcggcgtcgg cgcggcggca gcggtcgacg tcgagcctgg tgctggacaa gatcgagtcg 2040

ctggcggcgt acgactggtc cggcccgctg agggccagca gggggaggaa cttctacagg 2100
ctgaggggaa gcatgagcga gcatggcggc catccttccg aagattgctc catggagggc 2160
tactggatgt ag 2172

<210> 44
<211> 723
<212> PRT
<213> Oryza sativa
<400> 44

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

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

<210> 45
<211> 1257
<212> DNA
<213> Saccharum officinarum
<400> 45

```
atgacggacg gcggcgacga gtacaagcgc gaggagtcgg tgaccctgct ggtcatcgtc     60
tccctcgccg cgctctcgct cctctccctc atcgcggcct tcgcctacta ctgctacatt    120
acgcgcaagg tctcccgccg cctccactcg ctccatctcc ccaagcgttc cagctccgct    180
cctccgcctc ctccccgggc cccgccgcgg cagcagcagg ggaaggacag cccgtccagc    240
aactccgcca gcgacggggg tggggcggcg gcggcgatgt cggtggtggt tgctggggag    300
aggggcgtgc aggtgttcag ctaccggcag ctccacgcgg ccacgggcgg gttcggtcgg    360
gcgcacatgg ttgggcaggg cagcttcggc gccgtgtacc gcggggtgct ccccgacggg    420
aggaaggtcg cggtcaagct catggaccgc ccagggaagc agggcgaaga ggagttcgag    480
atggaggtgg agttgctgag taggctccga tcgccgtatt tattgggcct gattggccat    540
tgttcggaag gtggacaccg tctactggta tatgagttca tggccaatgg gggtctccag    600
gagcatctct atccaaacag aggttcctgt ggtggaatct caaaattgga ctgggacaca    660
agaatgagaa ttgcacttga agcagccaaa ggtttggaat atcttcatga gcgtgttaat    720
ccacctgtta tacatagaga cttcaagagt agcaatattc tcctggacaa ggacttccat    780
gcaagagttt cagactttgg tctaaccaaa cttggatctg atagagctgg tggtcatgtc    840
tcaactcgag ttttgggcac acaaggctat gttgcaccag aatacgcatt gactgggcat    900
ttgactacca aatcagatgt gtacagttac ggtgttgtgc ttttggaact gcttactggc    960
agagtaccag ttgatatgaa gaggcctcct ggagaaggtg ttttagttaa ctgggcattg   1020
cctatgctca ctgaccgaga aaaagtcgtg cggatcttgg atccagcatt ggaaggtcaa   1080
tattccttga aagatgctgt ccaagtggct gctattgctg ccatgtgtgt tcaaccagag   1140
gctgactata ggccattaat ggctgatgtt gtccaatcgc tggttccact tgtcaagaac   1200
cgttctaatc agaaagcttg caaccccaat gtgcaatcat cgaagccact cgactag      1257
```

<210> 46
<211> 418
<212> PRT
<213> Saccharum officinarum
<400> 46

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

<210> 47
<211> 921
<212> DNA
<213> Glycine max

<400> 47
agcaagtcaa atgtcattgg ccatggtgga tttggacttg tttaccgggg agtgctcaac 60
gatggcagga aagttgcaat caaattcatg gatcaggcag gaaagcaagg agaagaagaa 120
tttaaagtag aggtggaact gctaagtcgg ttgcattctc catatctgct ggcattgctt 180
gggtactgct ctgatagtaa tcataaattg ttggtgtatg aatttatggc aaatggtgga 240
ctgcaggaac atctctatcc tgtcagcaat tctattatta cgcctgtaaa attggattgg 300
gaaacacgac taagaatagc acttgaagct gctaagggtt tggaatatct tcatgagcat 360
gtcagtcccc cagtgattca cagagatttt aagagtagca acatcctctt ggacaagaaa 420
tttcatgcca aggtttctga ttttggattg gcgaagcttg gacctgatag agctggtgga 480
catgtttcaa ctcgggtttt gggcacccag ggatatgttg cccctgaata tgcactaaca 540
gggcatttaa caacaaaatc agatgtgtac agttatggtg ttgtactttt ggagctgctc 600
actggccggg tgcctgttga tatgaagaga cctccagggg aaggtgttct tgtttcttgg 660
gctctgcccc ttttgactga tagagaaaag gttgtaaaga ttatggatcc ttcattggag 720
ggacaatact ctatgaaaga ggttgtccag gtggctgcaa ttgctgcaat gtgtgtgcaa 780
ccagaggcag attatcggcc tctcatggct gatgttgtgc agtcattggt tccactggtg 840
aagactcaaa ggtccccttc aaaggtagga agctcctcta gtttcaattc ccctaagttg 900
tcccctggcc caacatttta a 921
<210> 48
<211> 306
<212> PRT
<213> Glycine max
<400> 48
Ser Lys Ser Asn Val Ile Gly His Gly Gly Phe Gly Leu Val Tyr Arg
1 5 10 15
Gly Val Leu Asn Asp Gly Arg Lys Val Ala Ile Lys Phe Met Asp Gln
20 25 30
Ala Gly Lys Gln Gly Glu Glu Glu Phe Lys Val Glu Val Glu Leu Leu
35 40 45
Ser Arg Leu His Ser Pro Tyr Leu Leu Ala Leu Leu Gly Tyr Cys Ser
50 55 60
Asp Ser Asn His Lys Leu Leu Val Tyr Glu Phe Met Ala Asn Gly Gly
65 70 75 80
Leu Gln Glu His Leu Tyr Pro Val Ser Asn Ser Ile Ile Thr Pro Val
85 90 95
Lys Leu Asp Trp Glu Thr Arg Leu Arg Ile Ala Leu Glu Ala Ala Lys
100 105 110
Gly Leu Glu Tyr Leu His Glu His Val Ser Pro Pro Val Ile His Arg
115 120 125
Asp Phe Lys Ser Ser Asn Ile Leu Leu Asp Lys Lys Phe His Ala Lys
130 135 140
Val Ser Asp Phe Gly Leu Ala Lys Leu Gly Pro Asp Arg Ala Gly Gly
145 150 155 160
His Val Ser Thr Arg Val Leu Gly Thr Gln Gly Tyr Val Ala Pro Glu
165 170 175
Tyr Ala Leu Thr Gly His Leu Thr Thr Lys Ser Asp Val Tyr Ser Tyr
180 185 190
Gly Val Val Leu Leu Glu Leu Leu Thr Gly Arg Val Pro Val Asp Met
195 200 205
Lys Arg Pro Pro Gly Glu Gly Val Leu Val Ser Trp Ala Leu Pro Leu
210 215 220
Leu Thr Asp Arg Glu Lys Val Val Lys Ile Met Asp Pro Ser Leu Glu
225 230 235 240
Gly Gln Tyr Ser Met Lys Glu Val Val Gln Val Ala Ala Ile Ala Ala
245 250 255
Met Cys Val Gln Pro Glu Ala Asp Tyr Arg Pro Leu Met Ala Asp Val
260 265 270
Val Gln Ser Leu Val Pro Leu Val Lys Thr Gln Arg Ser Pro Ser Lys
275 280 285
Val Gly Ser Ser Ser Ser Phe Asn Ser Pro Lys Leu Ser Pro Gly Pro
290 295 300
Thr Phe
305
<210> 49

<211> 1140
<212> DNA
<213> Solanum tuberosum
<400> 49
acagggtgag gaagaattca aagtggaggt ggagttgcta tgccgcttgc gctcaccgta 60
tttgctgtca ttgattggct attgttcaga aagcagccat aaattgcttg tttatgagtt 120
catggcaaat ggtggtttac aggagcactt gtatccaatt aaaggttcca ataattgttg 180
tccgaagttg gactggaaga cccggttaag aatagctttg gaggctgcta aaggtttgga 240
atatctacat gagcatgtca atcccccaat tatacataga gacctcaaaa gtagcaatat 300
tcttttggac aaaaacttcc atgccaaagt ttctgacttt ggattggcca agcttggatc 360
tgataaagct ggtggccatg tctctacccg cgttttgggg acacagggat atgttgctcc 420
agaatatgca ttaacaggac acttgaccac caaatcagat gtttacagtt atggggttgt 480
cctcctagag ttgttgactg gcagagttcc agttgatatg aagagatcac ctggcgaagg 540
tgttcttgtt tcttgggcat tgccccgtct cactgatagg gaaaaagtcg tagagataat 600
ggatccagct ctggagggtc agtattcaat gaaagaagtt attcaagtag ctgctattgc 660
tgcaatgtgt gtacagcccg aggctgatta caggccactg atggcagacg ttgtgcagtc 720
gttggttcca ctggtgaaac aaccgcgacc aactgtgaag cctggtagct cgtctagctt 780
tcacgccaca cagtcccccc ccccccatgt tacacaatct cctaaggctt agacttttt 840
caagcgaaat gggcatatca tatctcttga tccttaattc tagtccaact tctataacta 900
gtggccattt agtttgtgtt tggactatct agcttatgga acccccttc atttagtaac 960
tttaatctaa caaattttga tggagcaagc cagctgatgt aatattttct catgctaaac 1020
taggtggagc aaaaacagtt tctctgtatg taacatgtca aagctcccat ctaatatggg 1080
gattgtattt gagtattctt ggttagaagc agattcaaga tttgaagtta ctatatgata 1140
<210> 50
<211> 276
<212> PRT
<213> Solanum tuberosum
<400> 50
Gln Gly Glu Glu Glu Phe Lys Val Glu Val Glu Leu Leu Cys Arg Leu
1 5 10 15
Arg Ser Pro Tyr Leu Leu Ser Leu Ile Gly Tyr Cys Ser Glu Ser Ser
20 25 30
His Lys Leu Leu Val Tyr Glu Phe Met Ala Asn Gly Gly Leu Gln Glu
35 40 45
His Leu Tyr Pro Ile Lys Gly Ser Asn Asn Cys Cys Pro Lys Leu Asp
50 55 60
Trp Lys Thr Arg Leu Arg Ile Ala Leu Glu Ala Ala Lys Gly Leu Glu
65 70 75 80
Tyr Leu His Glu His Val Asn Pro Pro Ile Ile His Arg Asp Leu Lys
85 90 95
Ser Ser Asn Ile Leu Leu Asp Lys Asn Phe His Ala Lys Val Ser Asp
100 105 110
Phe Gly Leu Ala Lys Leu Gly Ser Asp Lys Ala Gly Gly His Val Ser
115 120 125
Thr Arg Val Leu Gly Thr Gln Gly Tyr Val Ala Pro Glu Tyr Ala Leu
130 135 140
Thr Gly His Leu Thr Thr Lys Ser Asp Val Tyr Ser Tyr Gly Val Val
145 150 155 160
Leu Leu Glu Leu Leu Thr Gly Arg Val Pro Val Asp Met Lys Arg Ser
165 170 175
Pro Gly Glu Gly Val Leu Val Ser Trp Ala Leu Pro Arg Leu Thr Asp
180 185 190
Arg Glu Lys Val Val Glu Ile Met Asp Pro Ala Leu Glu Gly Gln Tyr
195 200 205
Ser Met Lys Glu Val Ile Gln Val Ala Ala Ile Ala Ala Met Cys Val
210 215 220
Gln Pro Glu Ala Asp Tyr Arg Pro Leu Met Ala Asp Val Val Gln Ser
225 230 235 240
Leu Val Pro Leu Val Lys Gln Pro Arg Pro Thr Val Lys Pro Gly Ser
245 250 255
Ser Ser Ser Phe His Ala Thr Gln Ser Pro Ser Pro His Val Thr Gln
260 265 270
Ser Pro Lys Ala

275
<210> 51
<211> 1995
<212> DNA
<213> Nicotiana tabacum
<400> 51
taagatacca cagctatgct gacttatgaa agattttgga aaaagaaggt gcctctcaat 60
agaacgatgt ttggggatta tgaagtgatg cacattatct atccagggct gccttcttct 120
cctccatctg gcattgattc tggtaatggt ccaactggaa gtgctgtcaa tcaacaattc 180
cctattactg ctgattttgt aaacaagagt cagagaatga gtccccgagt cattttctc 240
attgcttcat cagcattagt actgttggtg gtatgctgtg gtgcactagt tgtcctctta 300
aaatgcagga ggactggccg accgtcaaat gctgttggtc cagtgtttac accatctatg 360
cacaagagat ctggtaaggg aattgggtcg acaatatcaa gtagcccgac tagttcaaca 420
tcgatttccc tcatttctgc tatgcctgct tctatccttt ctgtcaaaac atttacgctt 480
gcggagcttg agagggcaac tgacaagttt agtttaaaaa gggttttagg tgaaggagga 540
tttggacgtg tttatcatgg tatcttagaa gacagaacag aagttgccgt gaaagtactg 600
actagggata accaaaatgg agatcgtgaa ttcattgctg aagttgagat gctaagccga 660
ttgcatcacc gtaacctggt gaaattaatt ggtatatgca gtgaagagcg gactcgcagc 720
ttggtatatg aacttgttcg gaacggtagt gtggagtctc atttgcatgg aagagacggg 780
agaaaagagc cacttgattg ggatgtgagg ttgaaaattg ctcttggtgc tgcgagagga 840
ctagcttacc ttcatgaaga ttctaatcct cgtgtaattc atcgtgattt taaagccagc 900
aatgttttgt tagaagatga cttcacgccc aaggttgcag attttgggtt agcaagggaa 960
gcaaccgaag gaagtcacca catatctaca agagtcatgg gaacttttgg gtatgttgct 1020
cctgaatatg caatgacggg acacctactt gttaaaagtg atgtgtatag ttatggagtt 1080
gtattattgg agcttctctc cggaagaaaa cctgtggaca tgtctcaacc tcctggagaa 1140
gaaaacctgg taacttgggc gcgacctctt ctgaccacta gagaaggttt ggagcaactt 1200
gtggatcctt ctttggctgg aagctatgac tttgatgata tggcaaaggt ggctgccatt 1260
gcttcaatgt gcgttcaccc ggaggtgact caaaggccat ttatgggaga agtggtgcaa 1320
gctctcaaac taatttataa tgacaatgat gaaacttgtg ctgatggatg tagccagaag 1380
gagtcttctc tgccagattc agatttcaaa ggtgtccctt ccgatagcag ttggtggaat 1440
gctggtgggg ttacaccaag attaacatat ggacaagcct ccactttcat gaccatggat 1500
tacagttctg gtccgcttga agagttcgaa aacagaccgt tttcagcttc aagttttaat 1560
cttggtggag gggcgggttt aacaataagc cacggtaaca gatcaggtcc tttgaggact 1620
gtaaggagta aacctgctct ctataggtta aggggcagta tgagtgaaca cggtgcactt 1680
cttccaagac atgattggag ggatggcacc aattatgatg cttcttttta gagtgatttg 1740
tcaaatgtac agtgccgaag gccgtttcta tttgggaaaa aagatggttc tccggtgtag 1800
attaggagtt tgaaattgcc gctgtatccc tgagagagtg gactaagcct tctaattttg 1860
gtaatcttac attcatttta atagacagga aagataaaca ggacactcct tgttgtatat 1920
gttgtcaaat taacttttc tttagtccaa tattggattg gactactagt ctttctaaaa 1980
aaaaaaaaaa aaaaa 1995
<210> 52
<211> 571
<212> PRT
<213> Nicotiana tabacum
<400> 52
Met Leu Thr Tyr Glu Arg Phe Trp Lys Lys Lys Val Pro Leu Asn Arg
1 5 10 15
Thr Met Phe Gly Asp Tyr Glu Val Met His Ile Ile Tyr Pro Gly Leu
20 25 30
Pro Ser Ser Pro Pro Ser Gly Ile Asp Ser Gly Asn Gly Pro Thr Gly
35 40 45
Ser Ala Val Asn Gln Gln Phe Pro Ile Thr Ala Asp Phe Val Asn Lys
50 55 60
Ser Gln Arg Met Ser Pro Arg Val Ile Phe Leu Ile Ala Ser Ser Ala
65 70 75 80
Leu Val Leu Leu Val Val Cys Cys Gly Ala Leu Val Val Leu Leu Lys
85 90 95
Cys Arg Arg Thr Gly Arg Pro Ser Asn Ala Val Gly Pro Val Phe Thr
100 105 110
Pro Ser Met His Lys Arg Ser Gly Lys Gly Ile Gly Ser Thr Ile Ser
115 120 125
Ser Ser Pro Thr Ser Ser Thr Ser Ile Ser Leu Ile Ser Ala Met Pro
130 135 140

```
Ala Ser Ile Leu Ser Val Lys Thr Phe Thr Leu Ala Glu Leu Glu Arg
145                 150                 155                 160
Ala Thr Asp Lys Phe Ser Leu Lys Arg Val Leu Gly Glu Gly Gly Phe
                165                 170                 175
Gly Arg Val Tyr His Gly Ile Leu Glu Asp Arg Thr Glu Val Ala Val
            180                 185                 190
Lys Val Leu Thr Arg Asp Asn Gln Asn Gly Asp Arg Glu Phe Ile Ala
        195                 200                 205
Glu Val Glu Met Leu Ser Arg Leu His His Arg Asn Leu Val Lys Leu
    210                 215                 220
Ile Gly Ile Cys Ser Glu Glu Arg Thr Arg Ser Leu Val Tyr Glu Leu
225                 230                 235                 240
Val Arg Asn Gly Ser Val Glu Ser His Leu His Gly Arg Asp Gly Arg
                245                 250                 255
Lys Glu Pro Leu Asp Trp Asp Val Arg Leu Lys Ile Ala Leu Gly Ala
            260                 265                 270
Ala Arg Gly Leu Ala Tyr Leu His Glu Asp Ser Asn Pro Arg Val Ile
        275                 280                 285
His Arg Asp Phe Lys Ala Ser Asn Val Leu Leu Glu Asp Asp Phe Thr
    290                 295                 300
Pro Lys Val Ala Asp Phe Gly Leu Ala Arg Glu Ala Thr Glu Gly Ser
305                 310                 315                 320
His His Ile Ser Thr Arg Val Met Gly Thr Phe Gly Tyr Val Ala Pro
                325                 330                 335
Glu Tyr Ala Met Thr Gly His Leu Leu Val Lys Ser Asp Val Tyr Ser
            340                 345                 350
Tyr Gly Val Val Leu Leu Glu Leu Leu Ser Gly Arg Lys Pro Val Asp
        355                 360                 365
Met Ser Gln Pro Pro Gly Glu Glu Asn Leu Val Thr Trp Ala Arg Pro
    370                 375                 380
Leu Leu Thr Thr Arg Glu Gly Leu Glu Gln Leu Val Asp Pro Ser Leu
385                 390                 395                 400
Ala Gly Ser Tyr Asp Phe Asp Asp Met Ala Lys Val Ala Ala Ile Ala
                405                 410                 415
Ser Met Cys Val His Pro Glu Val Thr Gln Arg Pro Phe Met Gly Glu
            420                 425                 430
Val Val Gln Ala Leu Lys Leu Ile Tyr Asn Asp Asn Asp Glu Thr Cys
        435                 440                 445
Ala Asp Gly Cys Ser Gln Lys Glu Ser Ser Leu Pro Asp Ser Asp Phe
    450                 455                 460
Lys Gly Val Pro Ser Asp Ser Ser Trp Trp Asn Ala Gly Gly Val Thr
465                 470                 475                 480
Pro Arg Leu Thr Tyr Gly Gln Ala Ser Thr Phe Met Thr Met Asp Tyr
                485                 490                 495
Ser Ser Gly Pro Leu Glu Glu Phe Glu Asn Arg Pro Phe Ser Ala Ser
            500                 505                 510
Ser Phe Asn Leu Gly Gly Gly Ala Gly Leu Thr Ile Ser His Gly Asn
        515                 520                 525
Arg Ser Gly Pro Leu Arg Thr Val Arg Ser Lys Pro Ala Leu Tyr Arg
    530                 535                 540
Leu Arg Gly Ser Met Ser Glu His Gly Ala Leu Leu Pro Arg His Asp
545                 550                 555                 560
Trp Arg Asp Gly Thr Asn Tyr Asp Ala Ser Phe
                565                 570
<210> 53
<211> 1059
<212> DNA
<213> Medicago truncatula
<400> 53
atgttaagtc gtttgcatca tcgcaatcta gtgaaactta tcggtatatg cattgaaggg     60
cgcaggcgtt gtctggtata cgagcttgtt cctaatggca gtgttgagtc ccatctgcat    120
ggtgatgaca agaatagggg acctctagat tgggaagcac gtatgaaaat tgcccttgga    180
gctgcgagag gattggcgta tcttcacgag gattccaatc cccgtgtaat tcatcgagac    240
```

```
ttcaaagcta gcaatgtgtt attagaagac gactttaccc ctaaggtttc tgatttcggt    300
ttagcaagag aagcaactga ggggagtaat catatttcta cacgggtgat ggggactttt    360
gggtacgttg ccccagaata tgcaatgacg ggccatttac tggttaaaag tgatgtttat    420
agttacggtg ttgtgcttct tgaacttctc acaggcagaa aaccagtgga tatgtctcaa    480
cctcagggac aggagaatct tgttacctgg gcacgggcac tgttgactag tagagaaggt    540
ttagagcagc tagtggatcc atctttggct ggaggttaca actttgatga catggcaaag    600
gtagcagcta ttgcgtcaat gtgtgttcac tccgaggtga cacagagacc ttttatggga    660
gaagttgtgc aggctcttaa attaatatac aatgacacag acgagactgg tggagattat    720
tgtagtcaga aggactcctc tgctcaggag tctgatttca gaggtgagct tgccccttct    780
gatagcagtt ggtggaatgg tggaggttta actcctcgat taacttatgg acaagcatct    840
tcctttatca caatggagta cagttctggt ccacttgaag atatggaaaa cagaccgttt    900
tcaacttcaa gctttaatgg agatgagcta tctttaccaa ttaggcatgg gaataggtca    960
ggtcccttaa gaacgacccg aagcaagtta tccttatata gattttcagg aagtcggagt   1020
gagcatgggg aagtttcgtc taagcgaaat tggatttga                          1059
<210> 54
<211> 352
<212> PRT
<213> Medicago truncatula
<400> 54
Met Leu Ser Arg Leu His His Arg Asn Leu Val Lys Leu Ile Gly Ile
1               5                  10                  15
Cys Ile Glu Gly Arg Arg Arg Cys Leu Val Tyr Glu Leu Val Pro Asn
            20                  25                  30
Gly Ser Val Glu Ser His Leu His Gly Asp Asp Lys Asn Arg Gly Pro
        35                  40                  45
Leu Asp Trp Glu Ala Arg Met Lys Ile Ala Leu Gly Ala Ala Arg Gly
    50                  55                  60
Leu Ala Tyr Leu His Glu Asp Ser Asn Pro Arg Val Ile His Arg Asp
65                  70                  75                  80
Phe Lys Ala Ser Asn Val Leu Leu Glu Asp Asp Phe Thr Pro Lys Val
                85                  90                  95
Ser Asp Phe Gly Leu Ala Arg Glu Ala Thr Glu Gly Ser Asn His Ile
            100                 105                 110
Ser Thr Arg Val Met Gly Thr Phe Gly Tyr Val Ala Pro Glu Tyr Ala
        115                 120                 125
Met Thr Gly His Leu Leu Val Lys Ser Asp Val Tyr Ser Tyr Gly Val
    130                 135                 140
Val Leu Leu Glu Leu Leu Thr Gly Arg Lys Pro Val Asp Met Ser Gln
145                 150                 155                 160
Pro Gln Gly Gln Glu Asn Leu Val Thr Trp Ala Arg Ala Leu Leu Thr
                165                 170                 175
Ser Arg Glu Gly Leu Glu Gln Leu Val Asp Pro Ser Leu Ala Gly Gly
            180                 185                 190
Tyr Asn Phe Asp Asp Met Ala Lys Val Ala Ala Ile Ala Ser Met Cys
        195                 200                 205
Val His Ser Glu Val Thr Gln Arg Pro Phe Met Gly Glu Val Val Gln
    210                 215                 220
Ala Leu Lys Leu Ile Tyr Asn Asp Thr Asp Glu Thr Gly Gly Asp Tyr
225                 230                 235                 240
Cys Ser Gln Lys Asp Ser Ser Ala Gln Glu Ser Asp Phe Arg Gly Glu
                245                 250                 255
Leu Ala Pro Ser Asp Ser Ser Trp Trp Asn Gly Gly Gly Leu Thr Pro
            260                 265                 270
Arg Leu Thr Tyr Gly Gln Ala Ser Ser Phe Ile Thr Met Glu Tyr Ser
        275                 280                 285
Ser Gly Pro Leu Glu Asp Met Glu Asn Arg Pro Phe Ser Thr Ser Ser
    290                 295                 300
Phe Asn Gly Asp Glu Leu Ser Leu Pro Ile Arg His Gly Asn Arg Ser
305                 310                 315                 320
Gly Pro Leu Arg Thr Thr Arg Ser Lys Leu Ser Leu Tyr Arg Phe Ser
                325                 330                 335
Gly Ser Arg Ser Glu His Gly Glu Val Ser Ser Lys Arg Asn Trp Ile
            340                 345                 350
```

<210> 55
<211> 983
<212> DNA
<213> Populus sp.
<400> 55
atgggccata agcctccaga caaatacaaa ggagtccagg ttttcacata caaggagctt 60
gaaatcgcca caaacaaatt cagtgaagca aatgtgacat ggaatgaagg gtatggagtg 120
gtgtatggag gtaccctaag tgatggaact gtggcagcaa ttaagatgct ccatagagca 180
gggaagcaag gagagcgtgc atttaggata gaggtggatt tgctaagcag gttgcactca 240
ccatacttgg tggagctact tggttattgt gctgaccaga accacaggct cctagtattt 300
gaatttatgc ctaatgggac tcttcaacac catctccatc acaagcaata tcgaccgttg 360
gattggggga cacgattgag aattgcccct gattgtgcta gggccctgga gttccttcat 420
gagctcacaa tcccagcagt aatccatcgt gacttcaagt gtagtaacat cttactagac 480
caaaattttc gggctaaggt ttcagatttc ggatcggcta agatgggctc ggaaaggatc 540
aatgctcgaa attcgatgtg tttgccgagt accactggtt atctagcacc agagcatgct 600
tcaacaggta agctcaccac aaaatcagat gtatacagct atggagttgt tcttttacag 660
ctcttaactg gtcgtaaacc tgttgatacc aagcagccct ctggtgaaca tgtccttgtc 720
tcctgggctc ttccaaggct aaccaacaga gataaggtag tggaaatggt tgatccagcc 780
atgcaagacc agtattcaaa gaaggacttg atccaggtgg ctgctattgc agctgtgtgt 840
gtgcaaccag aagcagatta taggcctctc atgacagatg ttgtgcagtc tctaatccct 900
ctggtcaaga acctctcttc tgtatcttcc tcttgttcct ctaaatttat gaatcagatg 960
ctgtgcagtc caaggcctat gta 983
<210> 56
<211> 327
<212> PRT
<213> Populus sp.
<400> 56
Met Gly His Lys Pro Pro Asp Lys Tyr Lys Gly Val Gln Val Phe Thr
1 5 10 15
Tyr Lys Glu Leu Glu Ile Ala Thr Asn Lys Phe Ser Glu Ala Asn Val
20 25 30
Thr Trp Asn Glu Gly Tyr Gly Val Val Tyr Gly Gly Thr Leu Ser Asp
35 40 45
Gly Thr Val Ala Ala Ile Lys Met Leu His Arg Ala Gly Lys Gln Gly
50 55 60
Glu Arg Ala Phe Arg Ile Glu Val Asp Leu Leu Ser Arg Leu His Ser
65 70 75 80
Pro Tyr Leu Val Glu Leu Leu Gly Tyr Cys Ala Asp Gln Asn His Arg
85 90 95
Leu Leu Val Phe Glu Phe Met Pro Asn Gly Thr Leu Gln His His Leu
100 105 110
His His Lys Gln Tyr Arg Pro Leu Asp Trp Gly Thr Arg Leu Arg Ile
115 120 125
Ala Leu Asp Cys Ala Arg Ala Leu Glu Phe Leu His Glu Leu Thr Ile
130 135 140
Pro Ala Val Ile His Arg Asp Phe Lys Cys Ser Asn Ile Leu Leu Asp
145 150 155 160
Gln Asn Phe Arg Ala Lys Val Ser Asp Phe Gly Ser Ala Lys Met Gly
165 170 175
Ser Glu Arg Ile Asn Ala Arg Asn Ser Met Cys Leu Pro Ser Thr Thr
180 185 190
Gly Tyr Leu Ala Pro Glu His Ala Ser Thr Gly Lys Leu Thr Thr Lys
195 200 205
Ser Asp Val Tyr Ser Tyr Gly Val Val Leu Leu Gln Leu Leu Thr Gly
210 215 220
Arg Lys Pro Val Asp Thr Lys Gln Pro Ser Gly Glu His Val Leu Val
225 230 235 240
Ser Trp Ala Leu Pro Arg Leu Thr Asn Arg Asp Lys Val Val Glu Met
245 250 255
Val Asp Pro Ala Met Gln Asp Gln Tyr Ser Lys Lys Asp Leu Ile Gln
260 265 270
Val Ala Ala Ile Ala Ala Val Cys Val Gln Pro Glu Ala Asp Tyr Arg
275 280 285

```
Pro Leu Met Thr Asp Val Val Gln Ser Leu Ile Pro Leu Val Lys Asn
   290                 295                 300
Leu Ser Ser Val Ser Ser Ser Cys Ser Ser Lys Phe Met Asn Gln Met
305                 310                 315                 320
Leu Cys Ser Pro Arg Pro Met
                325
<210> 57
<211> 282
<212> PRT
<213> Arabidopsis thaliana
<220>
<221> INCERTO
<222> (75)..(75)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> INCERTO
<222> (121)..(121)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 57
Phe Ser Glu Glu Lys Lys Ile Gly Asn Gly Asp Val Tyr Lys Gly Val
1               5                  10                  15
Leu Ser Asp Gly Thr Val Ala Ala Ile Lys Lys Leu His Met Phe Asn
            20                  25                  30
Asp Asn Ala Ser Asn Gln Lys His Glu Glu Arg Ser Phe Arg Leu Glu
        35                  40                  45
Val Asp Leu Leu Ser Arg Leu Gln Cys Pro Tyr Leu Val Glu Leu Leu
    50                  55                  60
Gly Tyr Cys Ala Asp Gln Asn His Arg Ile Xaa Ile Tyr Glu Phe Met
65                  70                  75                  80
Pro Asn Gly Thr Val Glu His His Leu His Asp His Asn Phe Lys Asn
                85                  90                  95
Leu Lys Asp Arg Pro Gln Pro Leu Asp Trp Gly Ala Arg Leu Arg Ile
            100                 105                 110
Ala Leu Asp Cys Ala Arg Ala Leu Xaa Phe Leu His Glu Asn Thr Ile
        115                 120                 125
Ser Thr Val Ile His Arg Asn Phe Lys Cys Thr Asn Ile Leu Leu Asp
    130                 135                 140
Gln Asn Asn Arg Ala Lys Val Ser Asp Phe Gly Leu Ala Lys Thr Gly
145                 150                 155                 160
Ser Asp Lys Leu Asn Gly Glu Ile Ser Thr Arg Val Ile Gly Thr Thr
                165                 170                 175
Gly Tyr Leu Ala Pro Glu Tyr Ala Ser Thr Gly Lys Leu Thr Thr Lys
            180                 185                 190
Ser Asp Val Tyr Ser Tyr Gly Ile Val Leu Leu Gln Leu Leu Thr Gly
        195                 200                 205
Arg Thr Pro Ile Asp Ser Arg Arg Pro Arg Gly Gln Asp Val Leu Val
    210                 215                 220
Ser Trp Ala Leu Pro Arg Leu Thr Asn Arg Glu Lys Ile Ser Glu Met
225                 230                 235                 240
Val Asp Pro Thr Met Lys Gly Gln Tyr Ser Gln Lys Asp Leu Ile Gln
                245                 250                 255
Val Ala Ala Ile Ala Ala Val Cys Val Gln Pro Glu Ala Ser Tyr Arg
            260                 265                 270
Pro Leu Met Thr Asp Val Val His Ser Leu
        275                 280
<210> 58
<211> 2194
<212> DNA
<213> Oryza sativa
<400> 58
aatccgaaaa gtttctgcac cgttttcacc ccctaactaa caatataggg aacgtgtgct     60
aaatataaaa tgagacctta tatatgtagc gctgataact agaactatgc aagaaaaact    120
catccaccta ctttagtggc aatcgggcta aataaaaaag agtcgctaca ctagtttcgt    180
```

tttccttagt aattaagtgg gaaaatgaaa tcattattgc ttagaatata cgttcacatc 240
tctgtcatga agttaaatta ttcgaggtag ccataattgt catcaaactc ttcttgaata 300
aaaaaatctt tctagctgaa ctcaatgggt aaagagagag attttttttа aaaaaataga 360
atgaagatat tctgaacgta ttggcaaaga tttaaacata taattatata attttatagt 420
ttgtgcattc gtcatatcgc acatcattaa ggacatgtct tactccatcc caatttttat 480
ttagtaatta aagacaattg acttatttтt attatttatc ttttttcgat tagatgcaag 540
gtacttacgc acacactttg tgctcatgtg catgtgtgag tgcacctcct caatacacgt 600
tcaactagca acacatctct aatatcactc gcctatttaa tacatttagg tagcaatatc 660
tgaattcaag cactccacca tcaccagacc acttttaata atatctaaaa tacaaaaaat 720
aattttacag aatagcatga aaagtatgaa acgaactatt taggtttttc acatacaaaa 780
aaaaaaagaa ttttgctcgt gcgcgagcgc caatctccca tattgggcac acaggcaaca 840
acagagtggc tgcccacaga acaacccaca aaaaacgatg atctaacgga ggacagcaag 900
tccgcaacaa ccttttaaca gcaggctttg cggccaggag agaggaggag aggcaaagaa 960
aaccaagcat cctccttctc ccatctataa attcctcccc ccttttcccc tctctatata 1020
ggaggcatcc aagccaagaa gagggagagc accaaggaca cgcgactagc agaagccgag 1080
cgaccgcctt ctcgatccat atcttccggt cgagttcttg gtcgatctct tccctcctcc 1140
acctcctcct cacagggtat gtgcctccct tcggttgttc ttggatttat tgttctaggt 1200
tgtgtagtac gggcgttgat gttaggaaag gggatctgta tctgtgatga ttcctgttct 1260
tggatttggg atagaggggt tcttgatgtt gcatgttatc ggttcggttt gattagtagt 1320
atggttttca atcgtctgga gagctctatg gaaatgaaat ggtttaggga tcggaatctt 1380
gcgattttgt gagtaccttt tgtttgaggt aaaatcagag caccggtgat tttgcttggt 1440
gtaataaagt acggttgttt ggtcctcgat tctggtagtg atgcttctcg atttgacgaa 1500
gctatccttt gtttattccc tattgaacaa aaataatcca actttgaaga cggtcccgtt 1560
gatgagattg aatgattgat tcttaagcct gtccaaaatt tcgcagctgg cttgtttaga 1620
tacagtagtc cccatcacga aattcatgga aacagttata atcctcagga acaggggatt 1680
ccctgttctt ccgatttgct ttagtcccag aatttttttt cccaaatatc ttaaaaagtc 1740
actttctggt tcagttcaat gaattgattg ctacaaataa tgcttttata gcgttatcct 1800
agctgtagtt cagttaatag gtaatacccc tatagtttag tcaggagaag aacttatccg 1860
atttctgatc tccatttttа attatatgaa atgaactgta gcataagcag tattcatttg 1920
gattattttt tttattagct ctcacccctt cattattctg agctgaaagt ctggcatgaa 1980
ctgtcctcaa ttttgttttc aaattcacat cgattatcta tgcattatcc tcttgtatct 2040
acctgtagaa gtttcttttt ggttattcct tgactgcttg attacagaaa gaaatttatg 2100
aagctgtaat cgggatagtt atactgcttg ttcttatgat tcatttcctt tgtgcagttc 2160
ttggtgtagc ttgccacttt caccagcaaa gttc 2194

<210> 59
<211> 1620
<212> DNA
<213> Arabidopsis thaliana
<400> 59

atggaatttg agtgccaaga gagacttgaa gctgcgaacg atcaaagaag tgagtctggt 60
cttttgaata ccgatttgag aataggtaat gatgggtctg ttgagattcc aaatgtgaag 120
ttgtgttgtc agaaaaagaa gggaacttta gtgcccagtg gctcgaaaca gttactctct 180
gacaaggatt tgtctactac cattattgat ttacctcagg cattgatatc tgagattctt 240
aattgccttg acccaaaaga gttgggtcta gtgtcttgtg tttcaactta tctgcatagg 300
ttagcatctg agcatcacgc ttggaaggag ttctatcgtg agagatgggg actccctgta 360
gttttcggag ctgcaagttc agggctttct gatgagagat catggaaaga tctgtttgtt 420
gaaagggaat ttaggagtag gacctttтta ggacgttata gtattgatac tctgtatggt 480
cacactgaag cagtccggac tgtttttctc ttagcttctg ccaagctcgt tttcacttct 540
ggatatgatt ccattgttcg gatgtgggac atggaagaag gcttgtctat tgcggcatct 600
aaacctctcg gttgtacgat tagggcactt gctgctgata caaagctgtt ggtagctggc 660
ggtactgatg gattcattca ttgctggaaa tcactggatg gtctccggaa cctgtttgat 720
ctcacaggtt ttcagaaaga gaagactgag tttcggctat ggggacatga gggtcctatt 780
acatctcttg ccctggacat gacaagtatc tttagcggtt catgggatat gtctgttcgt 840
atatgggatc gatcttcaat gaagtgtgtc aaaaccttga ggcatagtga ttgggtttgg 900
ggacttgcac ctcatgaaac caccctcgct agcacgtcag gttctgatgt atatatttgg 960
gacgttagca gtgagactcc attggcgatc atccctgatg ctcatgaggg taccacttac 1020
tctttggcaa gaagccatac tggggatttc ctgtttactg gtggagaaga tggagggatt 1080
aaaatgtttg agatcagaag atacggtagt gaaacaagcg ttgtacttat atctcaatgg 1140
atgcctcaca ctagtcctgt ctactctctt tcttttgagt ttccttggct tgtctcagca 1200
tccggtgacg gtaaactcgc acttatcgac gttaggaaac tgttaaaaac taaccgttgt 1260
gcctattcca aaaggatttc ttcgagtact gtggaaccgc cacaacgaat gctgcacggg 1320
ttcgggtcaa acttgttctc tgtggacgtg ggttatgacc gtatagtgtg tggaggtgaa 1380
gaaggcacag tccggatatg gaacttcaca caagcattgg agatagagcg aaggacccga 1440

gctctgaaag gaatgaggca tgagaatagg atgagacgaa ggagaatgca aatggagatg 1500
aacgcaaaga atggaaggcc tgaccaatgc tccattgctg ctcataagaa ccccatcaat 1560
ggagaaagga accgagcctg gcatagtaaa cgaagagcaa gcgggaaggc gaaggcctaa 1620

<210> 60
<211> 539
<212> PRT
<213> Arabidopsis thaliana
<400> 60

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

```
       435                 440                 445
Asp Val Gly Tyr Asp Arg Ile Val Cys Gly Gly Glu Glu Gly Thr Val
   450                 455                 460
Arg Ile Trp Asn Phe Thr Gln Ala Leu Glu Ile Glu Arg Arg Thr Arg
465                 470                 475                 480
Ala Leu Lys Gly Met Arg His Glu Asn Arg Met Arg Arg Arg Arg Met
                485                 490                 495
Gln Met Glu Met Asn Ala Lys Asn Gly Arg Pro Asp Gln Cys Ser Ile
            500                 505                 510
Ala Ala His Lys Asn Pro Ile Asn Gly Glu Arg Asn Arg Ala Trp His
        515                 520                 525
Ser Lys Arg Arg Ala Ser Gly Lys Ala Lys Ala
    530                 535
<210> 61
<211> 1689
<212> DNA
<213> Oryza sativa
<400> 61
atggactttg attgcaaaac agcaagagga gactcttcat ctgtgaatcg ttcatgcatt     60
gtcactgaag gcactgtcgt ccaggcaaag cccgtttctc acaacgggaa agctaaacac    120
tggaatagcc tcagtacatt gaataaccag aagtgcagtt atgaattact ttccgaccca    180
aagaaaaatg ttgaaacaag tgatggtgag accgcctcaa agtgtgactc atggtgcttc    240
actgatttgc catccgcatt ggtctgtgaa gtgcttgaac acctcgatcc aaaggaactt    300
ggcattgtat cttgcgtctc caccctactg cataccctag ccacagatca tcaggggtgg    360
aagaagttct actgtgaaag gtgggggatt cctactcctc cagtcaccct caatgggccc    420
ttggttccag gaggaacttc agattggaag tcttggaaaa cattatttgt ggagcgggag    480
tttcgaagta aatcattcat gggaagattc agtgtggatg ttctccgtga ccacagtgag    540
gatgtacgca ctgtgttcct tctagcatca gtaaatctga tatttactgg tggtaatgac    600
tctgtgatcc gaatgtggga cttggaggaa gggcttttga ttgataagtc ccgcccactt    660
tgttgcacca tccgggccat tgcagctgac actaggcttt tggtcactgc aggaaccaat    720
gcctttattc attgttggag ggctgttgaa ggtaattctt acccttttcca catctctggg    780
aatggcactg accagagtcc tgagtttcgc ctttgggggc atgaaggacc tgtgacttgt    840
cttgccttgg attcattgag gattttcagt ggtagctggg atatgactgt ccgtgtttgg    900
gacagatccg aaatgaaatg tgttcagaag ttcatgcatg cagattgggt ttggagcgtg    960
gcacctcatg gaaatactgt tgccagtaca gctggtaggg atgcctatgt atgggatatc   1020
aggagtggtg agttggaaaa tgttatttcc aatgcccatt atggtaatgc attttcttta   1080
gctcgaacac acctagctga tgtgctgttt accggaggag aggatggggc aattcgcctg   1140
ttcaatgttt ctgaggtctc tgatgatgaa gatattaagc cagctgctac ttgggttcca   1200
cataccggcc ctgttcattc cctcgctttt gagtatccat ggcttgtctc agcctctagt   1260
gatggcaggg ttgcactgat tgatttgagg aagcttctga ccccaagaaa gtcatcaaaa   1320
caaccattca gggttaagaa ctttgatcca agttccattg aacctccaca gagaatgctt   1380
catggctttg ggtgcgatct tttttctgtc gccattggtg cagacaggat tgtctgtgga   1440
ggcgaggatg gcgctgtcaa agtctggaac ttctcagaag cactggagat tgagaagagg   1500
gcacaagctc taagaagtat gaggcaggag aaccgcatga ggcgaaaaaa ggcacaagta   1560
gagatgaatg caaatggtag aaggtctgac caatgtggct caatagccat gaaaagaaac   1620
caactgaagg gtgataagag tgtcacttgg cacagcaagc gtgccatcaa tgataaggtc   1680
aagtcttag                                                           1689
<210> 62
<211> 562
<212> PRT
<213> Oryza sativa
<400> 62
Met Asp Phe Asp Cys Lys Thr Ala Arg Gly Asp Ser Ser Ser Val Asn
1               5                   10                  15
Arg Ser Cys Ile Val Thr Glu Gly Thr Val Val Gln Ala Lys Pro Val
            20                  25                  30
Ser His Asn Gly Lys Ala Lys His Trp Asn Ser Leu Ser Thr Leu Asn
        35                  40                  45
Asn Gln Lys Cys Ser Tyr Glu Leu Leu Ser Asp Pro Lys Lys Asn Val
    50                  55                  60
Glu Thr Ser Asp Gly Glu Thr Ala Ser Lys Cys Asp Ser Trp Cys Phe
65                  70                  75                  80
Thr Asp Leu Pro Ser Ala Leu Val Cys Glu Val Leu Glu His Leu Asp
```

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

<210> 63
<211> 1728

<212> DNA
<213> Medicago trunculata
<400> 63
atggcatttg attgtccaca gagtattaag gtttctcaaa atttggtaga aaatccaggt 60
ataagcatag acatagttga attgaatcca aaacccaatt ccaaagcaat ttcaatttca 120
ttccctgatt ttgttacaaa taccaaatct gaatctggaa aaggtgagat ttttaagggg 180
aaatccaagc tgaattctaa gaaaggaatc aaagcagcct ctgctggtgc tttgaatcaa 240
tcatcagttc ctggtgatgt ggttattggt aattgtaggt caattaccga cctgcctccg 300
gtgttgatat ctgagattct gaattgtctt gatcccaaag aacttggaat tgtttcatgt 360
gtgtcgttga ttctgcgtag tcttgcttct gagcatcatg cttggaagga attctattgt 420
gaaaggtggg ggcttccagc agttcctgag gctgtcctgg attcggattc tggggttggg 480
gattcggtta aggatgaaaa gtcatggaag gatattttcg tggaaagaga ttataggagt 540
aagactttta tgggaaggta tagtatggat gtgttgtatg gacatactga ggcggttcgg 600
actgttttct tgttggcttc tacaaagctt atatttacat ctgggtatga cactgttgtg 660
aggatgtgga acatggaaag tgggttgtcg gttgcatcgt ctaaaccact tggctgcacc 720
attcgtgcgg ttgcagctga tacaaggctg ctagttgctg gtggtactga tgggttcatt 780
cattgttgga gggctgttga aggtttgcca catctgtttg agcttagaaa ctcacaacaa 840
aataagaatg aggttcgact ttggggtcat gatggtccgg ttacttcact tgctttagat 900
ttgacaagga tttatagcgg ctcttgggac acgactgttc gtgtttggga ccgtcactcg 960
atgaaatgca ctgtcgtgtt gaggcatagt gactgggttt ggggacttgt ccctcatgat 1020
actacagttg tcagcacatc aggttcaaat gtctatgttt gggatacaaa tagtggtaat 1080
ttggcaactg ttgttctaaa tgctcatgtt ggtaatactt atgctctggc aagaagccat 1140
actggggatt tcatttttac tgggggtgaa gatggttcaa ttcacatgta cgagattgtt 1200
gacggtagtt atgtgaccga agctctgcat gttgctacat gggatcccca ctctggtcct 1260
gtgtattccc ttgcctttga gtttccttgg cttgtttctg catcaagtga cggcaaattg 1320
gctctaattg acgtgaggaa gctgctgcgc aggagcaagc gtgccattgg aaagagagct 1380
acgaaggcaa agtattcggg cgaagttaat gtagagcctc cacagaggat gttgcatgga 1440
tttaagagta atcttttctc tgtaggtata ggagctgatc gaattgtttg cggaggtgaa 1500
gaaggtgtcg ttaggatttg gaatttcaca gaagctttgg aaattgaaag cagagttcgt 1560
gcattgagag gaatgcgatt agagaacaga atgagacgac gtaagaacca aacagagctc 1620
aacagtaaag gcggtaagag tgatcaatgt tcagctgcgg ccaagaagag ttcagtgact 1680
tgtatttggc cgagtaaacg tgggatgggt ggaaagacga aggcatag 1728
<210> 64
<211> 575
<212> PRT
<213> Medicago trunculata
<400> 64
Met Ala Phe Asp Cys Pro Gln Ser Ile Lys Val Ser Gln Asn Leu Val
1 5 10 15
Glu Asn Pro Gly Ile Ser Ile Asp Ile Val Glu Leu Asn Pro Lys Pro
20 25 30
Asn Ser Lys Ala Ile Ser Ile Ser Phe Pro Asp Phe Val Thr Asn Thr
35 40 45
Lys Ser Glu Ser Gly Lys Gly Glu Ile Phe Lys Gly Lys Ser Lys Leu
50 55 60
Asn Ser Lys Lys Gly Ile Lys Ala Ala Ser Ala Gly Ala Leu Asn Gln
65 70 75 80
Ser Ser Val Pro Gly Asp Val Val Ile Gly Asn Cys Arg Ser Ile Thr
85 90 95
Asp Leu Pro Pro Val Leu Ile Ser Glu Ile Leu Asn Cys Leu Asp Pro
100 105 110
Lys Glu Leu Gly Ile Val Ser Cys Val Ser Leu Ile Leu Arg Ser Leu
115 120 125
Ala Ser Glu His His Ala Trp Lys Glu Phe Tyr Cys Glu Arg Trp Gly
130 135 140
Leu Pro Ala Val Pro Glu Ala Val Leu Asp Ser Asp Ser Gly Val Gly
145 150 155 160
Asp Ser Val Lys Asp Glu Lys Ser Trp Lys Asp Ile Phe Val Glu Arg
165 170 175
Asp Tyr Arg Ser Lys Thr Phe Met Gly Arg Tyr Ser Met Asp Val Leu
180 185 190
Tyr Gly His Thr Glu Ala Val Arg Thr Val Phe Leu Leu Ala Ser Thr
195 200 205

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

<210> 65
<211> 1683
<212> DNA
<213> Triticum aestivum
<400> 65

```
atggactttg attgcaataa ggcaggagag tcttcagcca agcattgttc tagcatttgc     60
aacgagggca cactcatcca ggcaaacacc ttaattcatt gtggaaaggc taaaaagtgg    120
aatagcctca acaaattcaa caacccggag tcaagtcatg gatcacttcc aagggttaat    180
gaccccaagg aagatggtga aacaggaaat gatgcaactg cctctgagtg tagcgttatg    240
tgcttcactg atctgccgtc tgcattggtc tgtgaagtcc ttgcgcgcct tgatccaaag    300
gggcttgggg ttgtatcttg tgtctccaca gttctgcaga ccctagccac agatcatcag    360
ggatggaaga aattctactg tgagaggtgg ggacttccaa atgctcccat cggaccccta    420
gttccaggtg gaaccccaga tgggaggtcg tggaaagcat tgtttgtgga ccgggagttt    480
caaagcagat cattcatggg aagatttagt gtggatgttc tccgtggtca caatgaggat    540
gtacgcactg tgtaccttct ggcatcagca aatctgatat tcactggtgg ccatgattct    600
gtggttcgga tgtggaatat ggaggaaggg ctactaattg atgagtcccg cccatttggt    660
tgcactatcc gggcaattgc agctgacagt aggcttttgg taactggagg atccaaagcc    720
```

```
ttcattcagt gttggagggc tattgagggg gcctcgcacc tttctcacat ttctgggaat    780
ggtactaacc agaattctga atttcgccta tgggggcatg aagggcctgt gacttgtctt    840
tccttggatt caacaaggat ttacagtggt tcttgggata tgactgttcg tgtttgggac    900
agagccgaga tgaagtgtgt gcaaaagttc atgcatgctg actgggtttt ggccctggct    960
cctcatggaa atactgttgc cagtacagct ggtagagacg catatgtgtg ggatatcgga   1020
agtggcgagt taacaactat aatttccaat gcccatgttg gtaatgcata ttctgtagct   1080
cgaacacacc tggcaggtgt gctgtttact ggaggagagg acggggcaat tcgcttgttc   1140
gatgtttctg agatatcgga tgatgagaat attaaaccag ctgccacttg gttgccgcat   1200
tctggccctg ttcattctct tgcttttgag tacccatggc ttgtttctgc ctctagtgat   1260
ggcaggattg cactaattga tttgaggaag atcctgaccc ccctaaactc atcaaagcgc   1320
cgtttcaggg ttaagccctt tgatccaagt accgtagagc ctccacagcg aatgcttcat   1380
ggctttggat gctatctttt ctccgtcggc atcggtgcag atagaatcat atgtggaggc   1440
gaggatggct ctgtcagggt ctggaacttc tcggaagcac tggagattga gaagaaggca   1500
caggctttaa ggagtctgag acaggagaac cgcatgaggc ggaggaaggc gcaagtggag   1560
atgaatgcaa atggtagaag ggttgaccat tgctcagtag ccatgaaaag gaacccattg   1620
aagggtgata agagtgtcac ttggcaaatc aagcgtccca tcagtgacaa ggtcaagtct   1680
tag                                                                 1683
<210> 66
<211> 560
<212> PRT
<213> Triticum aestivum
<400> 66
Met Asp Phe Asp Cys Asn Lys Ala Gly Glu Ser Ser Ala Lys His Cys
1               5                  10                  15
Ser Ser Ile Cys Asn Glu Gly Thr Leu Ile Gln Ala Asn Thr Leu Ile
            20                  25                  30
His Cys Gly Lys Ala Lys Lys Trp Asn Ser Leu Asn Lys Phe Asn Asn
        35                  40                  45
Pro Glu Ser Ser His Gly Ser Leu Pro Arg Val Asn Asp Pro Lys Glu
    50                  55                  60
Asp Gly Glu Thr Gly Asn Asp Ala Thr Ala Ser Glu Cys Ser Val Met
65                  70                  75                  80
Cys Phe Thr Asp Leu Pro Ser Ala Leu Val Cys Glu Val Leu Ala Arg
                85                  90                  95
Leu Asp Pro Lys Gly Leu Gly Val Val Ser Cys Val Ser Thr Val Leu
            100                 105                 110
Gln Thr Leu Ala Thr Asp His Gln Gly Trp Lys Lys Phe Tyr Cys Glu
        115                 120                 125
Arg Trp Gly Leu Pro Asn Ala Pro Ile Gly Pro Leu Val Pro Gly Gly
    130                 135                 140
Thr Pro Asp Gly Arg Ser Trp Lys Ala Leu Phe Val Asp Arg Glu Phe
145                 150                 155                 160
Gln Ser Arg Ser Phe Met Gly Arg Phe Ser Val Asp Val Leu Arg Gly
                165                 170                 175
His Asn Glu Asp Val Arg Thr Val Tyr Leu Leu Ala Ser Ala Asn Leu
            180                 185                 190
Ile Phe Thr Gly Gly His Asp Ser Val Val Arg Met Trp Asn Met Glu
        195                 200                 205
Glu Gly Leu Leu Ile Asp Glu Ser Arg Pro Phe Gly Cys Thr Ile Arg
    210                 215                 220
Ala Ile Ala Ala Asp Ser Arg Leu Leu Val Thr Gly Gly Ser Lys Ala
225                 230                 235                 240
Phe Ile Gln Cys Trp Arg Ala Ile Glu Gly Ala Ser His Leu Ser His
                245                 250                 255
Ile Ser Gly Asn Gly Thr Asn Gln Asn Ser Glu Phe Arg Leu Trp Gly
            260                 265                 270
His Glu Gly Pro Val Thr Cys Leu Ser Leu Asp Ser Thr Arg Ile Tyr
        275                 280                 285
Ser Gly Ser Trp Asp Met Thr Val Arg Val Trp Asp Arg Ala Glu Met
    290                 295                 300
Lys Cys Val Gln Lys Phe Met His Ala Asp Trp Val Leu Ala Leu Ala
305                 310                 315                 320
Pro His Gly Asn Thr Val Ala Ser Thr Ala Gly Arg Asp Ala Tyr Val
```

```
                325                 330                 335
Trp Asp Ile Gly Ser Gly Glu Leu Thr Thr Ile Ile Ser Asn Ala His
            340                 345                 350
Val Gly Asn Ala Tyr Ser Val Ala Arg Thr His Leu Ala Gly Val Leu
        355                 360                 365
Phe Thr Gly Gly Glu Asp Gly Ala Ile Arg Leu Phe Asp Val Ser Glu
    370                 375                 380
Ile Ser Asp Asp Glu Asn Ile Lys Pro Ala Ala Thr Trp Leu Pro His
385                 390                 395                 400
Ser Gly Pro Val His Ser Leu Ala Phe Glu Tyr Pro Trp Leu Val Ser
                405                 410                 415
Ala Ser Ser Asp Gly Arg Ile Ala Leu Ile Asp Leu Arg Lys Ile Leu
            420                 425                 430
Thr Pro Leu Asn Ser Ser Lys Arg Arg Phe Arg Val Lys Pro Phe Asp
        435                 440                 445
Pro Ser Thr Val Glu Pro Pro Gln Arg Met Leu His Gly Phe Gly Cys
    450                 455                 460
Tyr Leu Phe Ser Val Gly Ile Gly Ala Asp Arg Ile Ile Cys Gly Gly
465                 470                 475                 480
Glu Asp Gly Ser Val Arg Val Trp Asn Phe Ser Glu Ala Leu Glu Ile
                485                 490                 495
Glu Lys Lys Ala Gln Ala Leu Arg Ser Leu Arg Gln Glu Asn Arg Met
            500                 505                 510
Arg Arg Arg Lys Ala Gln Val Glu Met Asn Ala Asn Gly Arg Arg Val
        515                 520                 525
Asp His Cys Ser Val Ala Met Lys Arg Asn Pro Leu Lys Gly Asp Lys
    530                 535                 540
Ser Val Thr Trp Gln Ile Lys Arg Pro Ile Ser Asp Lys Val Lys Ser
545                 550                 555                 560
<210> 67
<211> 1632
<212> DNA
<213> Populus tremuloides
<400> 67
atggcatttg aatgccaaaa gagcactgag gtgtcgaaaa gtcttgccat ttgtgttgaa     60
aaaagtaatt ccaatacgga acatcttgaa atttcgaagt ctcctttaga ctgttaccct    120
aagaagggga ttttaagcag tttgagttca aaacagcttt cttgtaatga tgttttgctc    180
aattgtcgcc ggtcaattac cgaccttcct ccggcattga tttctgagat tcttagttgc    240
cttgatcctc aagagcttgg tatagtttca tgtgtatcga gagtacttaa tagtcttgca    300
accgagaata gtgtttggaa ggaagtatat ggtgagagat ggggactccc tatagttcct    360
gcaccattgg gtgcaggggt ttcaaatgag aagtcatgga aggaactgtt tgtggagagg    420
gagtacagga gtaagatttt tttgagtcgt tatagcattg atactttgca tgggcatact    480
gaagcagtta gaacagtttc tctattggct tctgccaagc tcatttttac ctctgggtat    540
gatatgattg tgcgaatgtg ggacatggaa gacgggttgt atattgcatc atcacggcct    600
cttggttgca ctatacgagc agttgcagcg gacacgaagc tgttggttgc tggtggtact    660
gatggtttta ttcaaggctg gagggcagtt gagggtctca agcacttgtt tgaccttaaa    720
ggttctgagg tgccaaacac tgaatttaga atttgggagc atgagggacc tataacctct    780
cttgccttgg accccacaag gatttatagt gggtcatggg acatgactgt tcgtatatgg    840
gatcgttcct cacttgaatg cataaagatc ttgaggcatg gtgactgggt gtggagcctt    900
gttccgcatg atactacagt tgctagcaca tcaggttcag acgtgtatgt ttgggacact    960
aatagtggga ctctgctaac tgttattcat agtgctcatg taggtaacac ttattctttg   1020
gctcgaagcc atacagagga tttcattttt acaggaggag aagatggggc tatgcacatg   1080
tttgagatca ctggtcctaa acctgaggct aatgttttta aggttgcaac ttggatgcct   1140
cactcagggc ctgtttattc ccttgcattt gagtttccat ggcttgtttc agcttctagt   1200
gatgggaggc tgtcactagt tgatgtgaga aaactactaa ggaccaacag acgttctcta   1260
cgaaagaatg tttcaagggt taccgataag gaccgtaaca atgttgaacc acctcagagg   1320
atgttacatg ggtttgggcc caatttgttt tcagtagatg ttggtgctga ccgtattgta   1380
tgtggaggag aggaaggtgt tgttaggatc tggaacttct caaaggcgct ggaaagggag   1440
cggacagctc gtgccatgag aggaatacgg ttggagaaca ggatgaggcg tcgtagactt   1500
caaatagaga tgagcagtaa gggtggtagg actgatcaat gctctgttgc agccaagaag   1560
aacccaatgc atgtagacaa gagtggtgtc tggcgcaata aacatgggat gagtggcaag   1620
ctgaaagcat ag                                                       1632
<210> 68
```

<211> 543
<212> PRT
<213> Populus tremuloides
<400> 68

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

```
465                 470                 475                 480
Arg Thr Ala Arg Ala Met Arg Gly Ile Arg Leu Glu Asn Arg Met Arg
                485                 490                 495
Arg Arg Arg Leu Gln Ile Glu Met Ser Ser Lys Gly Gly Arg Thr Asp
            500                 505                 510
Gln Cys Ser Val Ala Ala Lys Lys Asn Pro Met His Val Asp Lys Ser
        515                 520                 525
Gly Val Trp Arg Asn Lys His Gly Met Ser Gly Lys Leu Lys Ala
    530                 535                 540
<210> 69
<211> 1695
<212> DNA
<213> Zea mays
<400> 69
atggcctttg actgcaacaa ggaaagagga gtttcttcgc ctgacagttg ctcccgcatt     60
tgcaccgagg gcactctcgt ccaggcaaat cccttatctc actactggaa gcctaaacgt    120
ttggagaact tcaacaagtt ggagaaccag aagtccagtt atgaatctgc tccgagtggc    180
tctgatccaa aacaggatgc tgaagcgagt gatgaggcaa ctgcctcact gtgtggcttc    240
aggtgcttca ctgatctgcc agcttcgttg gtctgcgagg tccttgcacg tctcgatgcg    300
aaggaccttg gaattgtatc ctgtgtctcc accctcctgc acgccttagc cacagatcat    360
cagggatgga agaagttata ctgcgaaagg tgggggcttc cagacctccc caccaccctg    420
aatgggccct tgctgccggg tgtccccctа gatgggaagt tttggaaaac atttttcgtg    480
gagcgggagt ttaggagcaa atccttcatg ggaagattca atctggatat tctccgtggc    540
cacaatgagg atgtgcgtgc tgtgtccctt ttggtatcag caaatctgat attcacaggc    600
ggccgcgatt ctgtggttag gatgtggaag atggaggaag ggttattgat tgatacatcc    660
cgtgcacttg gtggcaccat ccgggcgatt gcagctgact cgaggctttt agtgagtgga    720
ggaactaatg cctatattca gtgttggagg gccgttgaag gcaatggtca attatttcac    780
atctctggaa gtggtaccga ccagaatgcg gagtttcgcc tctggggtca tgaaggtcct    840
gtgacttgtc ttgccctgga ttcactgagg atttatagtg gttcttggga catgactgtt    900
cgtgtttggg acagagacca gatggagtgt gtgcaaaagt tcatgcatgc agactgggtt    960
tgggatcttg ctctgcgtgg aaatactgtt gccagtactg ctggtagaga tgcatatgta   1020
tgggatatca gggctggcga gttaacaagc ttaatttcca atgcccatgt tggtagcgca   1080
tattcaattg cttggacaca cctgccagat gtgctgttta ctggtggaga ggatggtgct   1140
attcgattgt tcaatgtttt tgatgtctgt gatgatgagg gtgacattaa accagtggca   1200
acttgggtgc cacattcagg tcctgttcat tctcttgctt ttgagtatcc atggcttgtg   1260
tcagcctcga gtgatggcag gattgcactt attgatttga ggaagcttct gtcctccaaa   1320
agttcatcaa agggtctgcg tgttaagaga gttgatggaa gcgcaataga gcctccacag   1380
aggatgcttc atggcgttgg atgtgatctc ttctccatcg ctattggtgc agataggatt   1440
gtatgtgccg gtgaggatgg ttctgttaga gtctggaact tctcaaaagc attggagatt   1500
gagaggaggg cacaggctct aaggagtttg aggcaggaga accgcatcag gcggaggaaa   1560
tcacaagcgg agatgaatac aaatactaga aggcctgacc agtgttcaat agccatgaaa   1620
aagaaccaac tgaagggtga taagagcgca acttggcaca acaagcgtgc cattaatgag   1680
aaggccaagt cttag                                                    1695
<210> 70
<211> 564
<212> PRT
<213> Zea mays
<400> 70
Met Ala Phe Asp Cys Asn Lys Glu Arg Gly Val Ser Ser Pro Asp Ser
1               5                   10                  15
Cys Ser Arg Ile Cys Thr Glu Gly Thr Leu Val Gln Ala Asn Pro Leu
            20                  25                  30
Ser His Tyr Trp Lys Pro Lys Arg Leu Glu Asn Phe Asn Lys Leu Glu
        35                  40                  45
Asn Gln Lys Ser Ser Tyr Glu Ser Ala Pro Ser Gly Ser Asp Pro Lys
    50                  55                  60
Gln Asp Ala Glu Ala Ser Asp Glu Ala Thr Ala Ser Leu Cys Gly Phe
65                  70                  75                  80
Arg Cys Phe Thr Asp Leu Pro Ala Ser Leu Val Cys Glu Val Leu Ala
                85                  90                  95
Arg Leu Asp Ala Lys Asp Leu Gly Ile Val Ser Cys Val Ser Thr Leu
            100                 105                 110
Leu His Ala Leu Ala Thr Asp His Gln Gly Trp Lys Lys Leu Tyr Cys
```

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

<210> 71
<211> 1278
<212> DNA
<213> Vitis vinifera
<400> 71
tggggagctc cggttgtttc ggggttttca gatgagaaat catggaggga attgtttgtg 60

```
gagagagagt ttaggagcaa aacctttagg ggacgattta gtattgatgt tctgtatggt    120
cgcactgagg cggttcgagc tgttttcctt ttggcctctg caaagctcat tttcacttct    180
gggtatgatt ccattgttcg aatgtgggat atggaagaag ggttgtcaat tgcgtgttca    240
cggcctctcg gttgcacaat aagagcagtg gctgcggata agaaactgtt gcttgcgggg    300
ggtagtgatg ggtttattca ttgttggaga gctgtagagg ggctctcttg cttgtttgac    360
cttgtgggtt ctcaaaacct gagtactgag ttccgaattt gggaacatga aggtcctgtg    420
acttgtcttg ctttggatat taagagaatt tatagtggct catgggacat gactgtgcgc    480
atatgggacc gttcttcgtt taaggttgta aaggtcttga ggcacacaga ctgggtttgg    540
ggccttgttc ctcgtgatac tacagttgct agcacttctg gctcagatgt gtatgtttgg    600
gacgctgata gtgggactct gctgacgata atctctaatg ctcatgtggg taatgcttat    660
gctttggcac ggagccacac aggggatttt ctattcactg ggggagaaga tggggcaata    720
catatgtttg aggtcgtgag tgattgcatg gaaaggaatg tattggaggt ttcaacgtgg    780
attcctcatt ctggtcctgt tcattccctg gcatttgagt ttccctggct tgtttcagct    840
tcagctgatg ggaggatgtc actgattgat gtgagaaagc ttttgcaaac ttgcaagcct    900
ttcttaggga agaatgtttc caaggttagg cacagggacc acaaaagtgt ggagccccct    960
cagaggatgt tacatgggtt tggctgcaat ttattctcag tggacattgg tgcagatcgt   1020
attgtctgtg gaggtgagaa aggtgtggtt agaatttgga acttctcaca agctttagaa   1080
gcagagcaga gggcccgtgc tttaagagga atacgtttag aaaccaggat gaggcggcgt   1140
aggcttcaaa tagagctgac tagtaagggt agtcgaactg atcaatgttc agttgcagcc   1200
agtaagaatc ctattaatgg tgataggagt ggtgtgtggc ccaataagcg gggaatgagt   1260
ggccggatga aagcatag                                                 1278
<210> 72
<211> 425
<212> PRT
<213> Vitis vinifera
<400> 72
Trp Gly Ala Pro Val Val Ser Gly Phe Ser Asp Glu Lys Ser Trp Arg
1               5                  10                  15
Glu Leu Phe Val Glu Arg Glu Phe Arg Ser Lys Thr Phe Arg Gly Arg
            20                  25                  30
Phe Ser Ile Asp Val Leu Tyr Gly Arg Thr Glu Ala Val Arg Ala Val
        35                  40                  45
Phe Leu Leu Ala Ser Ala Lys Leu Ile Phe Thr Ser Gly Tyr Asp Ser
    50                  55                  60
Ile Val Arg Met Trp Asp Met Glu Glu Gly Leu Ser Ile Ala Cys Ser
65                  70                  75                  80
Arg Pro Leu Gly Cys Thr Ile Arg Ala Val Ala Ala Asp Lys Lys Leu
                85                  90                  95
Leu Leu Ala Gly Gly Ser Asp Gly Phe Ile His Cys Trp Arg Ala Val
            100                 105                 110
Glu Gly Leu Ser Cys Leu Phe Asp Leu Val Gly Ser Gln Asn Leu Ser
        115                 120                 125
Thr Glu Phe Arg Ile Trp Glu His Glu Gly Pro Val Thr Cys Leu Ala
    130                 135                 140
Leu Asp Ile Lys Arg Ile Tyr Ser Gly Ser Trp Asp Met Thr Val Arg
145                 150                 155                 160
Ile Trp Asp Arg Ser Ser Phe Lys Val Val Lys Val Leu Arg His Thr
                165                 170                 175
Asp Trp Val Trp Gly Leu Val Pro Arg Asp Thr Thr Val Ala Ser Thr
            180                 185                 190
Ser Gly Ser Asp Val Tyr Val Trp Asp Ala Asp Ser Gly Thr Leu Leu
        195                 200                 205
Thr Ile Ile Ser Asn Ala His Val Gly Asn Ala Tyr Ala Leu Ala Arg
    210                 215                 220
Ser His Thr Gly Asp Phe Leu Phe Thr Gly Gly Glu Asp Gly Ala Ile
225                 230                 235                 240
His Met Phe Glu Val Val Ser Asp Cys Met Glu Arg Asn Val Leu Glu
                245                 250                 255
Val Ser Thr Trp Ile Pro His Ser Gly Pro Val His Ser Leu Ala Phe
            260                 265                 270
Glu Phe Pro Trp Leu Val Ser Ala Ser Ala Asp Gly Arg Met Ser Leu
        275                 280                 285
Ile Asp Val Arg Lys Leu Leu Gln Thr Cys Lys Pro Phe Leu Gly Lys
```

```
   290                 295                 300
Asn Val Ser Lys Val Arg His Arg Asp His Lys Ser Val Glu Pro Pro
305                 310                 315                 320
Gln Arg Met Leu His Gly Phe Gly Cys Asn Leu Phe Ser Val Asp Ile
                325                 330                 335
Gly Ala Asp Arg Ile Val Cys Gly Gly Glu Lys Gly Val Val Arg Ile
            340                 345                 350
Trp Asn Phe Ser Gln Ala Leu Glu Ala Glu Gln Arg Ala Arg Ala Leu
        355                 360                 365
Arg Gly Ile Arg Leu Glu Thr Arg Met Arg Arg Arg Arg Leu Gln Ile
    370                 375                 380
Glu Leu Thr Ser Lys Gly Ser Arg Thr Asp Gln Cys Ser Val Ala Ala
385                 390                 395                 400
Ser Lys Asn Pro Ile Asn Gly Asp Arg Ser Gly Val Trp Pro Asn Lys
                405                 410                 415
Arg Gly Met Ser Gly Arg Met Lys Ala
            420                 425
<210> 73
<211> 660
<212> DNA
<213> Senecio cambrensis
<400> 73
atggcatttg agttaaaccc gagcacatcg aattctgaaa aagataaacc tcaggataat     60
tcatctgaaa ataatttact gattgattcg gaaagcggga aacattttgc tgctaagtta    120
ctggttgacg agtgttgttt gttgaaaggt tataaaacaa ttaccgacct tcctccagcg    180
atagtttctg aaatatttaa cagccttgat ccaaaggaac tcgggatagt ttcctgtgtg    240
aatacatctt tatatcgaat tgcatttgat caccacgttt ggaaggagtt ttattgtgag    300
agatggggac ttcctattgg agtccccaat acgtctttag agaaatcatg gagagacctt    360
tttgtggaac gtgagtttcg tagtaagacg tttatgggac cttactctac tgaaactctt    420
tatggtcata ctgaagctgt tcgtacagtt tttcttttac tttcaagaaa gattataatt    480
acttcgggat atgattcagt tattcgaatg tgggacatgg aaggtggtta ttcaattgct    540
tcgtcccgtc ctcttggttg taccatccga gccgttacag cagattcaaa actgttaatt    600
gctggcggtt ccgatggttt tattcatggt tggagagccc aagaagaaga ggacatcctt    660
<210> 74
<211> 220
<212> PRT
<213> Senecio cambrensis
<400> 74
Met Ala Phe Glu Leu Asn Pro Ser Thr Ser Asn Ser Glu Lys Asp Lys
1               5                  10                  15
Pro Gln Asp Asn Ser Ser Glu Asn Asn Leu Leu Ile Asp Ser Glu Ser
            20                  25                  30
Gly Lys His Phe Ala Ala Lys Leu Leu Val Asp Glu Cys Cys Leu Leu
        35                  40                  45
Lys Gly Tyr Lys Thr Ile Thr Asp Leu Pro Pro Ala Ile Val Ser Glu
    50                  55                  60
Ile Phe Asn Ser Leu Asp Pro Lys Glu Leu Gly Ile Val Ser Cys Val
65                  70                  75                  80
Asn Thr Ser Leu Tyr Arg Ile Ala Phe Asp His His Val Trp Lys Glu
                85                  90                  95
Phe Tyr Cys Glu Arg Trp Gly Leu Pro Ile Gly Val Pro Asn Thr Ser
            100                 105                 110
Leu Glu Lys Ser Trp Arg Asp Leu Phe Val Glu Arg Glu Phe Arg Ser
        115                 120                 125
Lys Thr Phe Met Gly Pro Tyr Ser Thr Glu Thr Leu Tyr Gly His Thr
    130                 135                 140
Glu Ala Val Arg Thr Val Phe Leu Leu Leu Ser Arg Lys Ile Ile Ile
145                 150                 155                 160
Thr Ser Gly Tyr Asp Ser Val Ile Arg Met Trp Asp Met Glu Gly Gly
                165                 170                 175
Tyr Ser Ile Ala Ser Ser Arg Pro Leu Gly Cys Thr Ile Arg Ala Val
            180                 185                 190
Thr Ala Asp Ser Lys Leu Leu Ile Ala Gly Gly Ser Asp Gly Phe Ile
```

```
       195                 200                 205
His Gly Trp Arg Ala Gln Glu Glu Glu Asp Ile Leu
   210                 215                 220
<210> 75
<211> 847
<212> DNA
<213> Helianthus annuus
<220>
<221> misc_feature
<222> (789)..(789)
<223> n é a, c, g, ou t
<400> 75
gcgggcgacg gagtccccgt ctctttaaat gctgagtgtt ctgatgagag atcttggaag     60
gcattatttg tggaacggga attccggagc aaaacatttc tgggccggta tacaatcgat    120
acgctttacg gtcatacaga acccgttcgc actctttttg ttttgctctc gaaaaagctt    180
ataataactt ccggttatga ctcgattatt cgaatgtggg acgttgaaga tggttattca    240
atcgcttcat ctcggcctct gggttgcacc atccgagccg ttgcagctga ttctaagctg    300
ttaattgctg gcgggtctga tgggtttata catggctgga gagctgaaga aggacaccct    360
cacttgtttg acataaaagg accccaaaac cagaacaccg agtttcgact ttgggaacat    420
caaggcccga taacttgtat cgggctagat tcatctagga tttacagcgg gtcatgggac    480
atgaccatac gcgtctggga tcgtgtctcg ctcaagtgct tgactgtctt gatgcataac    540
gactgggtgt ggagcctggt cccacatgat ggtactgtag taagtacttc tggttcagat    600
gtatatatct gggagactaa tacctttca cttattgaca ttatcaaaga tgcccatgtg    660
ggtaatgctt attctctagc tagaagtcac accggtaatt tcatcttcac tggtggtgaa    720
gatggtgcta tacatatgtt tgagattact agtaatggat gtggttcagt tcgccggcta    780
gcaacgtgng gtccacatac gggtcctgta tattctcttt catttgagtt tccatggcta    840
gtttctg                                                              847
<210> 76
<211> 282
<212> PRT
<213> Helianthus annuus
<220>
<221> INCERTO
<222> (263)..(263)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 76
Ala Gly Asp Gly Val Pro Val Ser Leu Asn Ala Glu Cys Ser Asp Glu
1               5                  10                  15
Arg Ser Trp Lys Ala Leu Phe Val Glu Arg Glu Phe Arg Ser Lys Thr
           20                  25                  30
Phe Leu Gly Arg Tyr Thr Ile Asp Thr Leu Tyr Gly His Thr Glu Pro
       35                  40                  45
Val Arg Thr Leu Phe Val Leu Leu Ser Lys Lys Leu Ile Ile Thr Ser
   50                  55                  60
Gly Tyr Asp Ser Ile Ile Arg Met Trp Asp Val Glu Asp Gly Tyr Ser
65                  70                  75                  80
Ile Ala Ser Ser Arg Pro Leu Gly Cys Thr Ile Arg Ala Val Ala Ala
               85                  90                  95
Asp Ser Lys Leu Leu Ile Ala Gly Gly Ser Asp Gly Phe Ile His Gly
           100                 105                 110
Trp Arg Ala Glu Glu Gly His Pro His Leu Phe Asp Ile Lys Gly Pro
       115                 120                 125
Gln Asn Gln Asn Thr Glu Phe Arg Leu Trp Glu His Gln Gly Pro Ile
   130                 135                 140
Thr Cys Ile Gly Leu Asp Ser Ser Arg Ile Tyr Ser Gly Ser Trp Asp
145                 150                 155                 160
Met Thr Ile Arg Val Trp Asp Arg Val Ser Leu Lys Cys Leu Thr Val
               165                 170                 175
Leu Met His Asn Asp Trp Val Trp Ser Leu Val Pro His Asp Gly Thr
           180                 185                 190
Val Val Ser Thr Ser Gly Ser Asp Val Tyr Ile Trp Glu Thr Asn Thr
       195                 200                 205
Phe Ser Leu Ile Asp Ile Ile Lys Asp Ala His Val Gly Asn Ala Tyr
```

```
   210                 215                 220
Ser Leu Ala Arg Ser His Thr Gly Asn Phe Ile Phe Thr Gly Gly Glu
225                 230                 235                 240
Asp Gly Ala Ile His Met Phe Glu Ile Thr Ser Asn Gly Cys Gly Ser
                245                 250                 255
Val Arg Arg Leu Ala Thr Xaa Gly Pro His Thr Gly Pro Val Tyr Ser
            260                 265                 270
Leu Ser Phe Glu Phe Pro Trp Leu Val Ser
        275                 280
<210> 77
<211> 728
<212> DNA
<213> Euphorbia esula
<400> 77
ggatggaaag ccgtggaggg tcttccacat ttgttcaacc ttaagggtag ttccgagaat     60
cttccaaacg atgaatttcg actttgggaa cacgagggac ctataacctc tctcgccttg    120
gatcgcacga gaatttacag tggttcttgg gacatgactg ttcgtgtatg ggatcgttcc    180
acattgaagt gccttaaaat cttgaggcat agcgattggg tatggggcct tgttccgcac    240
gataccaccg ttgccacctc atcgggttcc gatgtgtata tttgggatac tcaaaccgga    300
actcttatag ccgatattaa taacgcccat gtcgggaaca tttattctct agcccgaagc    360
cacacgggag attttctctt caccggagga gaagatgggg ctatacatat gtttgagatc    420
attggtaata atggatccaa gcctaaggtc tacaagatcg caacttggat tccacactcg    480
ggtcccgtct actccctctc gtttgaattt ccgtggcttg tttcggcttc tagtgatggg    540
aaactctcgc ttatagatgt tcgaaagcta cttcgaacaa gtaggcgctc tatgggaaag    600
aagaatcttg gtagggttag caaaatggat agtaacaatg tggagccacc tcaacggatg    660
ctccacgggt ttggacccaa tttgttttcg gtagacattg gggctgaccg gattgtttgt    720
ggagggga                                                             728
<210> 78
<211> 241
<212> PRT
<213> Euphorbia esula
<400> 78
Gly Trp Lys Ala Val Glu Gly Leu Pro His Leu Phe Asn Leu Lys Gly
1               5                   10                  15
Ser Ser Glu Asn Leu Pro Asn Asp Glu Phe Arg Leu Trp Glu His Glu
            20                  25                  30
Gly Pro Ile Thr Ser Leu Ala Leu Asp Arg Thr Arg Ile Tyr Ser Gly
        35                  40                  45
Ser Trp Asp Met Thr Val Arg Val Trp Asp Arg Ser Thr Leu Lys Cys
    50                  55                  60
Leu Lys Ile Leu Arg His Ser Asp Trp Val Trp Gly Leu Val Pro His
65                  70                  75                  80
Asp Thr Thr Val Ala Thr Ser Ser Gly Ser Asp Val Tyr Ile Trp Asp
                85                  90                  95
Thr Gln Thr Gly Thr Leu Ile Ala Asp Ile Asn Asn Ala His Val Gly
            100                 105                 110
Asn Ile Tyr Ser Leu Ala Arg Ser His Thr Gly Asp Phe Leu Phe Thr
        115                 120                 125
Gly Gly Glu Asp Gly Ala Ile His Met Phe Glu Ile Ile Gly Asn Asn
    130                 135                 140
Gly Ser Lys Pro Lys Val Tyr Lys Ile Ala Thr Trp Ile Pro His Ser
145                 150                 155                 160
Gly Pro Val Tyr Ser Leu Ser Phe Glu Phe Pro Trp Leu Val Ser Ala
                165                 170                 175
Ser Ser Asp Gly Lys Leu Ser Leu Ile Asp Val Arg Lys Leu Leu Arg
            180                 185                 190
Thr Ser Arg Arg Ser Met Gly Lys Lys Asn Leu Gly Arg Val Ser Lys
        195                 200                 205
Met Asp Ser Asn Asn Val Glu Pro Pro Gln Arg Met Leu His Gly Phe
    210                 215                 220
Gly Pro Asn Leu Phe Ser Val Asp Ile Gly Ala Asp Arg Ile Val Cys
225                 230                 235                 240
Gly
```

<210> 79
<211> 713
<212> DNA
<213> Lycopersicon esculentum
<400> 79
atcgtcttgc ttcggagcac catgtgtgga aggagttcta ttgtgaaagg tgggggcagc 60
caatattgtt ggcacctttg ggcgcagagc atgcagatga gaagtcgtgg aaggagcttt 120
tcgtggagag ggagttccgt agtaaaacat tcatgggacg ctatagtatt gatacattgt 180
atggtcatac cgaggctgtt aggactgttt ctgttttatc ttcaaagaaa cttctgttta 240
cttcagggta tgatcagata gtgcgaatgt gggacttgga agaaggttta tctattgcat 300
catctcgccc tcttggctgc actattcgtg cagttgcggc tgattcgagg ctcttagtag 360
caggggtac agatggattc atacatggtt ggagagcaga ggatggaaat ccacatctct 420
ttgatcttag atcacctcag aaccagaaca tggaattcag agtttgggaa catgaaggac 480
caataacatg tctggcattg gatttgaata agatgtacag tggttcttgg gacatgagta 540
ttcgtgtttg ggatcgttct tctttgaaat gtctgaaagt tctaatgcac aatgactggg 600
tttggagcct tgctccccat gatacaacag tagcgagcgc tgcaggctca gatctttatg 660
tctgggaaca ctatattggg tcagaacttg ccaccatcac caatggtcat gct 713
<210> 80
<211> 236
<212> PRT
<213> Lycopersicon esculentum
<400> 80
Arg Leu Ala Ser Glu His His Val Trp Lys Glu Phe Tyr Cys Glu Arg
1 5 10 15
Trp Gly Gln Pro Ile Leu Leu Ala Pro Leu Gly Ala Glu His Ala Asp
20 25 30
Glu Lys Ser Trp Lys Glu Leu Phe Val Glu Arg Glu Phe Arg Ser Lys
35 40 45
Thr Phe Met Gly Arg Tyr Ser Ile Asp Thr Leu Tyr Gly His Thr Glu
50 55 60
Ala Val Arg Thr Val Ser Val Leu Ser Ser Lys Lys Leu Leu Phe Thr
65 70 75 80
Ser Gly Tyr Asp Gln Ile Val Arg Met Trp Asp Leu Glu Glu Gly Leu
85 90 95
Ser Ile Ala Ser Ser Arg Pro Leu Gly Cys Thr Ile Arg Ala Val Ala
100 105 110
Ala Asp Ser Arg Leu Leu Val Ala Gly Gly Thr Asp Gly Phe Ile His
115 120 125
Gly Trp Arg Ala Glu Asp Gly Asn Pro His Leu Phe Asp Leu Arg Ser
130 135 140
Pro Gln Asn Gln Asn Met Glu Phe Arg Val Trp Glu His Glu Gly Pro
145 150 155 160
Ile Thr Cys Leu Ala Leu Asp Leu Asn Lys Met Tyr Ser Gly Ser Trp
165 170 175
Asp Met Ser Ile Arg Val Trp Asp Arg Ser Ser Leu Lys Cys Leu Lys
180 185 190
Val Leu Met His Asn Asp Trp Val Trp Ser Leu Ala Pro His Asp Thr
195 200 205
Thr Val Ala Ser Ala Ala Gly Ser Asp Leu Tyr Val Trp Glu His Tyr
210 215 220
Ile Gly Ser Glu Leu Ala Thr Ile Thr Asn Gly His
225 230 235
<210> 81
<211> 717
<212> DNA
<213> Aquilegia formosa x Aquilegia pubescens
<400> 81
gtccctcgag actctactat tgctagtact gccggggtgg atatgtatgt ttgggacatt 60
gatagtgggg aattgcttgc tgtaattcaa aaggctcatg ttggcaacac tctctctttg 120
gcccgaagtc acacagggaa tttgattttc actggtgggg atgatgggac tataagtatg 180
tttgagcttt tggacgattc tacgctcttg catgtggcaa cttggagccc acatactggt 240
gctgtgcact ctcttgcatt tgagttccct tggcttgtgt cagcctctag cgacggaaag 300
ctctcattga ttgacattag acggttgctg aaatctagcc agcggtctgt gtcaagagac 360

ctcaaaaggg taaattgtcc agtttcgttg actgtggagg ctccacagag gatgttacat 420
ggttttggta gtaatttatt ctcagtgggc attggttctg atcggattgt atgtggtggt 480
gaggaaggtg tagttagaat ttggaacttt tcacaagcta tggagattga gcgtagggtt 540
caggccttaa ggggcatgag attagaaaac cgaatgaggc ggcggaaggc ccaaatagag 600
atgaacagca agggtggtcg ctctgatcag tgttcagttg cagctaagaa gaaccagatt 660
aatggagaca gggcctggca tggtaggcga ggagcgagtg gaaagctgaa ggcctag 717

<210> 82
<211> 238
<212> PRT
<213> Aquilegia formosa x Aquilegia pubescens
<400> 82

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

<210> 83
<211> 621
<212> DNA
<213> Gossypium hirsutum
<400> 83

ggccgcttcg accggagttg cactggagac ttccttttta ctggtggaga agatggagcg 60
atacacatgt ttgagattat aagtgggtct gacgaatcta gtgtcgtgca ggttgcaaca 120
tggattcctc actcaggtgc tgtttactca cttgcatttg agtttccctg gcttgtttca 180
gcttccagtg atggtaaact tgcactaatt gatgttcgga agttgttgag ggccggtaag 240
cgcaccttgg ggaaaagggt ttcaagggat aatgacattg accaaaggaa catagagcca 300
ccccagagaa tgctccatgg atttgggtgc aatttgttct cagtcggtat tggtgcagat 360
cgaattgtct gtgggggtga agaaggtgtt gttcgtatct ggaacttctc agaagctcta 420
gaaattgagc agagggcacg ggcactaaga ggaatacgtt tggagaatag gatgaggcga 480
cgcagacttc aaattgagat gaataataag ggtggtcgaa ctgatcaatg ttctgttgca 540
gccaagaaga aaacagtcaa tggtgatagg aatagtgtct ggcacagtaa acgtagcata 600
agctccaagg tgaagacata a 621

<210> 84
<211> 206
<212> PRT
<213> Gossypium hirsutum
<400> 84

Gly Arg Phe Asp Arg Ser Cys Thr Gly Asp Phe Leu Phe Thr Gly Gly

```
1               5                   10                  15
Glu Asp Gly Ala Ile His Met Phe Glu Ile Ile Ser Gly Ser Asp Glu
            20                  25                  30
Ser Ser Val Val Gln Val Ala Thr Trp Ile Pro His Ser Gly Ala Val
        35                  40                  45
Tyr Ser Leu Ala Phe Glu Phe Pro Trp Leu Val Ser Ala Ser Ser Asp
    50                  55                  60
Gly Lys Leu Ala Leu Ile Asp Val Arg Lys Leu Leu Arg Ala Gly Lys
65                  70                  75                  80
Arg Thr Leu Gly Lys Arg Val Ser Arg Asp Asn Asp Ile Asp Gln Arg
                85                  90                  95
Asn Ile Glu Pro Pro Gln Arg Met Leu His Gly Phe Gly Cys Asn Leu
            100                 105                 110
Phe Ser Val Gly Ile Gly Ala Asp Arg Ile Val Cys Gly Gly Glu Glu
        115                 120                 125
Gly Val Val Arg Ile Trp Asn Phe Ser Glu Ala Leu Glu Ile Glu Gln
    130                 135                 140
Arg Ala Arg Ala Leu Arg Gly Ile Arg Leu Glu Asn Arg Met Arg Arg
145                 150                 155                 160
Arg Arg Leu Gln Ile Glu Met Asn Asn Lys Gly Gly Arg Thr Asp Gln
                165                 170                 175
Cys Ser Val Ala Ala Lys Lys Lys Thr Val Asn Gly Asp Arg Asn Ser
            180                 185                 190
Val Trp His Ser Lys Arg Ser Ile Ser Ser Lys Val Lys Thr
        195                 200                 205
<210> 85
<211> 631
<212> DNA
<213> Sorghum bicolor
<400> 85
gatacatccc gtgcacttgg tggcaccatc cgggcgattg cagctgactc taggctttta     60
gtgagtggag gaactaatgc ctatattcag tgttggaggg ctgttgaagg caatgatcac    120
ttatttcaca tctctggaag tggtactgac cagaatgcgg agtttcgcct ctgggggcat    180
gaaggtcctg tgactagtct tgccttggat tcactgagga tttatagtgg ttcttgggac    240
atgactgttc gtgtttggga cagagcccag atggattgcg tgcaaaagtt catgcatgca    300
gactgggtgt ggtctcttgc tctgcgtgga aatactgttg ccagtactgc tggtagagat    360
gcatatgtat gggatatcag ggacggcgag ttaacaagct tagtttccaa tgcccatgtt    420
ggtaatgcat attcattggc ttggacacac ctggcggatg tgctgtttac tggtggagag    480
gatggcgcta ttcgcttgtt caatgtttct gatgtctctg atgatgagga ggacattaaa    540
ccagtggcaa cttgggtgcc acattcaggt cctgttcatt ctcttgcttt tgagtatcca    600
tggcttgtgt cagcctcgag tgatggcagg a                                   631
<210> 86
<211> 210
<212> PRT
<213> Sorghum bicolor
<400> 86
Asp Thr Ser Arg Ala Leu Gly Gly Thr Ile Arg Ala Ile Ala Ala Asp
1               5                   10                  15
Ser Arg Leu Leu Val Ser Gly Gly Thr Asn Ala Tyr Ile Gln Cys Trp
            20                  25                  30
Arg Ala Val Glu Gly Asn Asp His Leu Phe His Ile Ser Gly Ser Gly
        35                  40                  45
Thr Asp Gln Asn Ala Glu Phe Arg Leu Trp Gly His Glu Gly Pro Val
    50                  55                  60
Thr Ser Leu Ala Leu Asp Ser Leu Arg Ile Tyr Ser Gly Ser Trp Asp
65                  70                  75                  80
Met Thr Val Arg Val Trp Asp Arg Ala Gln Met Asp Cys Val Gln Lys
                85                  90                  95
Phe Met His Ala Asp Trp Val Trp Ser Leu Ala Leu Arg Gly Asn Thr
            100                 105                 110
Val Ala Ser Thr Ala Gly Arg Asp Ala Tyr Val Trp Asp Ile Arg Asp
        115                 120                 125
Gly Glu Leu Thr Ser Leu Val Ser Asn Ala His Val Gly Asn Ala Tyr
```

```
   130                 135                 140
Ser Leu Ala Trp Thr His Leu Ala Asp Val Leu Phe Thr Gly Gly Glu
145                 150                 155                 160
Asp Gly Ala Ile Arg Leu Phe Asn Val Ser Asp Val Ser Asp Asp Glu
                165                 170                 175
Glu Asp Ile Lys Pro Val Ala Thr Trp Val Pro His Ser Gly Pro Val
            180                 185                 190
His Ser Leu Ala Phe Glu Tyr Pro Trp Leu Val Ser Ala Ser Ser Asp
        195                 200                 205
Gly Arg
    210
<210> 87
<211> 573
<212> DNA
<213> Ipomea nil
<400> 87
ggagaagacg gagccataca catgtttgag gttgtaaaaa acgcaaggtt tcatgtaaag     60
aaggttgcaa catggattcc tcactcgggc cctgtttatt ctcttgcatt tgaattcccc    120
tggcttgttt ccgcttcaag cgatggaaaa ctggcactta tcgatgtgag gaagttattg    180
aagacaagta ggtattcagc aacaggaagc cgtccttgta aggttgacaa tctggaccat    240
acaagtgttg agcctccgca acgaatgctt catggatttg ggtccaattt atttgcagtg    300
gatgttggtc tggatcgtat tgtgtgtgga ggtgaagaag gcgctgttag gatctggaac    360
ttctcacaag cgttggagat agagcagagg gtccgtgctt tacgaggttt aaggttagaa    420
aacaggatga ggaggcgtaa gcttcagtcc aacatgaata gcaaaggcac ccggagtgat    480
cagtgctcgg ttgcagcaaa gaagaaccaa atcaatggcg ataggaatag ctggcagaac    540
aagcgtaggg taagcgggaa gctcaaggct tag                                 573
<210> 88
<211> 190
<212> PRT
<213> Ipomea nil
<400> 88
Gly Glu Asp Gly Ala Ile His Met Phe Glu Val Val Lys Asn Ala Arg
1               5                   10                  15
Phe His Val Lys Lys Val Ala Thr Trp Ile Pro His Ser Gly Pro Val
            20                  25                  30
Tyr Ser Leu Ala Phe Glu Phe Pro Trp Leu Val Ser Ala Ser Ser Asp
        35                  40                  45
Gly Lys Leu Ala Leu Ile Asp Val Arg Lys Leu Leu Lys Thr Ser Arg
    50                  55                  60
Tyr Ser Ala Thr Gly Ser Arg Pro Cys Lys Val Asp Asn Leu Asp His
65                  70                  75                  80
Thr Ser Val Glu Pro Pro Gln Arg Met Leu His Gly Phe Gly Ser Asn
                85                  90                  95
Leu Phe Ala Val Asp Val Gly Leu Asp Arg Ile Val Cys Gly Gly Glu
            100                 105                 110
Glu Gly Ala Val Arg Ile Trp Asn Phe Ser Gln Ala Leu Glu Ile Glu
        115                 120                 125
Gln Arg Val Arg Ala Leu Arg Gly Leu Arg Leu Glu Asn Arg Met Arg
    130                 135                 140
Arg Arg Lys Leu Gln Ser Asn Met Asn Ser Lys Gly Thr Arg Ser Asp
145                 150                 155                 160
Gln Cys Ser Val Ala Ala Lys Lys Asn Gln Ile Asn Gly Asp Arg Asn
                165                 170                 175
Ser Trp Gln Asn Lys Arg Arg Val Ser Gly Lys Leu Lys Ala
            180                 185                 190
<210> 89
<211> 615
<212> DNA
<213> Solanum tuberosum
<400> 89
ttttccttgg cccgaattca cacagggaag ctccttattc cactgaaggg gaagatggag     60
ctatacgcat gttcgagatc attagaaatg ataaagcttc atctcaggcg tgtggcgaca    120
tggattcctc actcgggtgc tgtttattct cttgcatttg agttcccttg gcttgtttca    180
```

```
gcttccagtg atggaaaact ctcacttatt gatgtgagaa agctgttgag gacgaatcgg    240
agttatgtga tggagaagcc ttcgaaggtt gacaataggg ctaaaaatgt agagccacct    300
caaagaatgt tacatggatt tgggagcaat ctgtttgctg tggatattgg tcttgatcgt    360
atcgtatgtg ctggagaaga aggcattgtt aggatctgga acttctcaga agctttggag    420
gttgagcaga gagttcgagc attaagagga ttaaggttag agaacagaat gaggaggcgt    480
aaacttcagt ctgaaatgac tggtaaaggc agtcgtgctg atcagtgctc agttgctgct    540
aagaagaatc aattgaatgg tgataggaac agctggcgca acaagcgtag ggtgactggg    600
aagctgaagg cctag                                                     615
<210> 90
<211> 204
<212> PRT
<213> Solanum tuberosum
<400> 90
Phe Ser Leu Ala Arg Ile His Thr Gly Lys Leu Leu Ile Pro Leu Lys
1               5                  10                  15
Gly Lys Met Glu Leu Tyr Ala Cys Ser Arg Ser Leu Glu Met Ile Lys
            20                  25                  30
Leu His Leu Arg Arg Val Ala Thr Trp Ile Pro His Ser Gly Ala Val
        35                  40                  45
Tyr Ser Leu Ala Phe Glu Phe Pro Trp Leu Val Ser Ala Ser Ser Asp
    50                  55                  60
Gly Lys Leu Ser Leu Ile Asp Val Arg Lys Leu Leu Arg Thr Asn Arg
65                  70                  75                  80
Ser Tyr Val Met Glu Lys Pro Ser Lys Val Asp Asn Arg Ala Lys Asn
                85                  90                  95
Val Glu Pro Pro Gln Arg Met Leu His Gly Phe Gly Ser Asn Leu Phe
            100                 105                 110
Ala Val Asp Ile Gly Leu Asp Arg Ile Val Cys Ala Gly Glu Glu Gly
        115                 120                 125
Ile Val Arg Ile Trp Asn Phe Ser Glu Ala Leu Glu Val Glu Gln Arg
    130                 135                 140
Val Arg Ala Leu Arg Gly Leu Arg Leu Glu Asn Arg Met Arg Arg Arg
145                 150                 155                 160
Lys Leu Gln Ser Glu Met Thr Gly Lys Gly Ser Arg Ala Asp Gln Cys
                165                 170                 175
Ser Val Ala Ala Lys Lys Asn Gln Leu Asn Gly Asp Arg Asn Ser Trp
            180                 185                 190
Arg Asn Lys Arg Arg Val Thr Gly Lys Leu Lys Ala
        195                 200
<210> 91
<211> 632
<212> DNA
<213> Zamia fischeri
<400> 91
ggcagacatt ctcaatcgcc tcaatgcaag agaactcagc atagtttcat gtgtatgtac     60
cctctttcga aggatttctt cagatagcca tggatggaag gagttctact gtgagagatg    120
ggggcttcct gcgcctagca aaatttccgc gtcttcggcc ccgggaccag agaagtcgtg    180
gagagagttg tatttggaga gagatgcccg aagcaaggcc ttcctgggcc gatttaatgt    240
agacatgctg catggccaca ctgcagctct tcgatctgtt tgccttctgc catctgcaaa    300
tctcatattt acagcaggat atgacacagt actcagaatg tggaacatgg aggaaggtct    360
cttgatcgct tgttcccggc ctcttggttg cacgctccgt gccatagcgg cagacatgga    420
tatgttggtg gtggcaggaa ctgatccttt cgtgcgctgt tggcaagcaa tttctgagct    480
tcctaactta tttgacattt cgggggtttc agggcagaac actgaatctt gccttcgtgg    540
acatgttgga cccataactt gccttggcct agattcgttg aaaatttaca gtggctcttg    600
ggacatgagc attcgggtat gggatagggt ta                                  632
<210> 92
<211> 210
<212> PRT
<213> Zamia fischeri
<400> 92
Ala Asp Ile Leu Asn Arg Leu Asn Ala Arg Glu Leu Ser Ile Val Ser
1               5                  10                  15
Cys Val Cys Thr Leu Phe Arg Arg Ile Ser Ser Asp Ser His Gly Trp
```

```
          20                  25                  30
Lys Glu Phe Tyr Cys Glu Arg Trp Gly Leu Pro Ala Pro Ser Lys Ile
       35                  40                  45
Ser Ala Ser Ser Ala Pro Gly Pro Glu Lys Ser Trp Arg Glu Leu Tyr
   50                  55                  60
Leu Glu Arg Asp Ala Arg Ser Lys Ala Phe Leu Gly Arg Phe Asn Val
65                  70                  75                  80
Asp Met Leu His Gly His Thr Ala Ala Leu Arg Ser Val Cys Leu Leu
               85                  90                  95
Pro Ser Ala Asn Leu Ile Phe Thr Ala Gly Tyr Asp Thr Val Leu Arg
           100                 105                 110
Met Trp Asn Met Glu Glu Gly Leu Leu Ile Ala Cys Ser Arg Pro Leu
       115                 120                 125
Gly Cys Thr Leu Arg Ala Ile Ala Ala Asp Met Asp Met Leu Val Val
   130                 135                 140
Ala Gly Thr Asp Pro Phe Val Arg Cys Trp Gln Ala Ile Ser Glu Leu
145                 150                 155                 160
Pro Asn Leu Phe Asp Ile Ser Gly Val Ser Gly Gln Asn Thr Glu Ser
               165                 170                 175
Cys Leu Arg Gly His Val Gly Pro Ile Thr Cys Leu Gly Leu Asp Ser
           180                 185                 190
Leu Lys Ile Tyr Ser Gly Ser Trp Asp Met Ser Ile Arg Val Trp Asp
       195                 200                 205
Arg Val
   210
<210> 93
<211> 507
<212> DNA
<213> Persea americana
<400> 93
gcgacatggc ttcctcacac gggctctgtt aattctctag catttgagtt cccatggctt     60
gtgtcatctt ccagtgatgg acgacttgct ctgattgatg tgagaaagct gttaaggtca    120
agccgacgta cgtcatcaag acaccatgtt aaagctaaac ggtctgcccc cgtaaatgtg    180
gagccgcctc aaaggatgtt gcatgggtat gggtgcaatt tgttttctgt ggatattggc    240
gcggatagga ttgtttgtgg tggagaggag ggtgttgtgc gtatctggaa cttctctaag    300
gcacttgaga ttgagcagag gattcgggct ctgaaaggga tacggttgga gaaccgaatg    360
aggcggcgga agatacaatt agagatgagc agtaagaatc gacctggaga tcaatgctct    420
gttgcagcca aaaggaatca gatgcctggt gatagaaaca gggtttggcg tggaaggcgc    480
aatatgagtg aaaagccgaa ggcctaa                                        507
<210> 94
<211> 168
<212> PRT
<213> Persea americana
<400> 94
Ala Thr Trp Leu Pro His Thr Gly Ser Val Asn Ser Leu Ala Phe Glu
1               5                   10                  15
Phe Pro Trp Leu Val Ser Ser Ser Ser Asp Gly Arg Leu Ala Leu Ile
           20                  25                  30
Asp Val Arg Lys Leu Leu Arg Ser Ser Arg Arg Thr Ser Ser Arg His
       35                  40                  45
His Val Lys Ala Lys Arg Ser Ala Pro Val Asn Val Glu Pro Pro Gln
   50                  55                  60
Arg Met Leu His Gly Tyr Gly Cys Asn Leu Phe Ser Val Asp Ile Gly
65                  70                  75                  80
Ala Asp Arg Ile Val Cys Gly Gly Glu Glu Gly Val Val Arg Ile Trp
               85                  90                  95
Asn Phe Ser Lys Ala Leu Glu Ile Glu Gln Arg Ile Arg Ala Leu Lys
           100                 105                 110
Gly Ile Arg Leu Glu Asn Arg Met Arg Arg Arg Lys Ile Gln Leu Glu
       115                 120                 125
Met Ser Ser Lys Asn Arg Pro Gly Asp Gln Cys Ser Val Ala Ala Lys
   130                 135                 140
Arg Asn Gln Met Pro Gly Asp Arg Asn Arg Val Trp Arg Gly Arg Arg
```

145 150 155 160
Asn Met Ser Glu Lys Pro Lys Ala
165

<210> 95
<211> 498
<212> DNA
<213> Glycine max
<220>
<221> misc_feature
<222> (127)..(127)
<223> n é a, c, g, ou t
<400> 95
gttgctgctt ggattcctca ctctgcccct gtgtattccc tggcctttga gtttccatgg 60
cttgtttctg catcaagtga tggccagcta gcactaattg atgtgagaaa gctgttgagg 120
attagcnagc gagctctagg gaaaagagtt tcaaaggtaa agcatttggg tggagacata 180
gtggagcctc cacagaggat gttgcatgga tttaagagcc atcttttctc tgtggatata 240
ggagctgaac gaattgtcag tggaggtgaa gaaggtgttg tcaggatctg gaatttcacg 300
gaagctttgg aaattgaacg gagagcccga gcattaagag gaatacgatt agagcacaga 360
atgaggcgac ggaagcttca aacagagctg agccataaaa gtggtcggag tgatcagtgt 420
tcagttgcag cccagaaaaa ttctgtcact tgtatttggc ccactaaacg tggcatgagc 480
ggaaagctga aagcatag 498

<210> 96
<211> 165
<212> PRT
<213> Glycine max
<220>
<221> INCERTO
<222> (43)..(43)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 96
Val Ala Ala Trp Ile Pro His Ser Ala Pro Val Tyr Ser Leu Ala Phe
1 5 10 15
Glu Phe Pro Trp Leu Val Ser Ala Ser Ser Asp Gly Gln Leu Ala Leu
20 25 30
Ile Asp Val Arg Lys Leu Leu Arg Ile Ser Xaa Arg Ala Leu Gly Lys
35 40 45
Arg Val Ser Lys Val Lys His Leu Gly Gly Asp Ile Val Glu Pro Pro
50 55 60
Gln Arg Met Leu His Gly Phe Lys Ser His Leu Phe Ser Val Asp Ile
65 70 75 80
Gly Ala Glu Arg Ile Val Ser Gly Gly Glu Glu Gly Val Val Arg Ile
85 90 95
Trp Asn Phe Thr Glu Ala Leu Glu Ile Glu Arg Arg Ala Arg Ala Leu
100 105 110
Arg Gly Ile Arg Leu Glu His Arg Met Arg Arg Arg Lys Leu Gln Thr
115 120 125
Glu Leu Ser His Lys Ser Gly Arg Ser Asp Gln Cys Ser Val Ala Ala
130 135 140
Gln Lys Asn Ser Val Thr Cys Ile Trp Pro Thr Lys Arg Gly Met Ser
145 150 155 160
Gly Lys Leu Lys Ala
165

<210> 97
<211> 12
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 1
<220>
<221> INCERTO
<222> (3)..(3)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>

<221> VARIANTE
<222> (4)..(4)
<223> /substituir = "Val" /substituir = "Leu"
<220>
<221> VARIANTE
<222> (6)..(6)
<223> /substituir = "Arg" /substituir = "Gly"
<220>
<221> INCERTO
<222> (11)..(11)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 97
Trp Lys Xaa Phe Tyr Cys Glu Arg Trp Gly Xaa Pro
1 5 10
<210> 98
<211> 16
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 2
<220>
<221> VARIANTE
<222> (3)..(3)
<223> /substituir = "Ser"
<220>
<221> VARIANTE
<222> (6)..(6)
<223> /substituir = "Tyr"
<220>
<221> VARIANTE
<222> (12)..(12)
<223> /substituir = "Ser"
<220>
<221> VARIANTE
<222> (14)..(14)
<223> /substituir = "Ala" /substituir = "Gly"
<400> 98
Ser Leu Ala Phe Glu Phe Pro Trp Leu Val Ser Ala Ser Ser Asp Gly
1 5 10 15
<210> 99
<211> 15
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 3
<220>
<221> VARIANTE
<222> (2)..(2)
<223> /substituir = "Phe"
<220>
<221> VARIANTE
<222> (8)..(8)
<223> substituir = "Thr"
<220>
<221> VARIANTE
<222> (9)..(9)
<223> substituir = "Ser"
<220>
<221> VARIANTE
<222> (10)..(10)
<223> substituir = "Ile"
<220>
<221> VARIANTE

<222> (12)..(12)
<223> substituir = "Ile"
<400> 99
Ile Tyr Ser Gly Ser Trp Asp Met Thr Val Arg Val Trp Asp Arg
1 5 10 15
<210> 100
<211> 7
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 4
<220>
<221> VARIANTE
<222> (5)..(5)
<223> /substituir = "Asp"
<400> 100
Phe Thr Gly Gly Glu Asp Gly
1 5
<210> 101
<211> 9
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 5
<220>
<221> VARIANTE
<222> (2)..(2)
<223> /substituir = "Ala"
<400> 101
Glu Pro Pro Gln Arg Met Leu His Gly
1 5
<210> 102
<211> 38
<212> PRT
<213> Seqüência artificial
<220>
<223> região conservada 1
<400> 102
Ser Gln Trp Met Pro His Thr Ser Pro Val Tyr Ser Leu Ser Phe Glu
1 5 10 15
Phe Pro Trp Leu Val Ser Ala Ser Gly Asp Gly Lys Leu Ala Leu Ile
20 25 30
Asp Val Arg Lys Leu Leu
35
<210> 103
<211> 65
<212> PRT
<213> Seqüência artificial
<220>
<223> região conservada 2
<400> 103
Val Glu Pro Pro Gln Arg Met Leu His Gly Phe Gly Ser Asn Leu Phe
1 5 10 15
Ser Val Asp Val Gly Tyr Asp Arg Ile Val Cys Gly Gly Glu Glu Gly
20 25 30
Thr Val Arg Ile Trp Asn Phe Thr Gln Ala Leu Glu Ile Glu Arg Arg
35 40 45
Thr Arg Ala Leu Lys Gly Met Arg His Glu Asn Arg Met Arg Arg Arg
50 55 60
Arg
65
<210> 104
<211> 1167

<212> DNA
<213> Oryza sativa
<400> 104
atccaagcca agaagaggga gagcaccaag gacacgcgac tagcagaagc cgagcgaccg 60
ccttcttcga tccatatctt ccggtcgagt tcttggtcga tctcttccct cctccacctc 120
ctcctcacag ggtatgtgcc cttcggttgt tcttggattt attgttctag gttgtgtagt 180
acgggcgttg atgttaggaa aggggatctg tatctgtgat gattcctgtt cttggatttg 240
ggatagaggg gttcttgatg ttgcatgtta tcggttcggt ttgattagta gtatggtttt 300
caatcgtctg gagagctcta tggaaatgaa atggtttagg gtacggaatc ttgcgatttt 360
gtgagtacct tttgtttgag gtaaaatcag agcaccggtg attttgcttg gtgtaataaa 420
agtacggttg tttggtcctc gattctggta gtgatgcttc tcgatttgac gaagctatcc 480
tttgtttatt ccctattgaa caaaaataat ccaactttga agacggtccc gttgatgaga 540
ttgaatgatt gattcttaag cctgtccaaa atttcgcagc tggcttgttt agatacagta 600
gtccccatca cgaaattcat ggaaacagtt ataatcctca ggaacagggg attccctgtt 660
cttccgattt gctttagtcc cagaattttt tttcccaaat atcttaaaaa gtcactttct 720
ggttcagttc aatgaattga ttgctacaaa taatgctttt atagcgttat cctagctgta 780
gttcagttaa taggtaatac ccctatagtt tagtcaggag aagaacttat ccgatttctg 840
atctccattt ttaattatat gaaatgaact gtagcataag cagtattcat ttggattatt 900
ttttttatta gctctcaccc cttcattatt ctgagctgaa agtctggcat gaactgtcct 960
caattttgtt ttcaaattca catcgattat ctatgcatta tcctcttgta tctacctgta 1020
gaagtttctt tttggttatt ccttgactgc ttgattacag aaagaaattt atgaagctgt 1080
aatcgggata gttatactgc ttgttcttat gattcatttc ctttgtgcag ttcttggtgt 1140
agcttgccac tttcaccagc aaagttc 1167
<210> 105
<211> 55
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm6999
<400> 105
ggggacaagt ttgtacaaaa aagcaggctt aaacaatgaa tcgtttttct cgttt 55
<210> 106
<211> 49
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm7000
<400> 106
ggggaccact ttgtacaaga aagctgggta tccaatctta tcgcttagg 49
<210> 107
<211> 1602
<212> DNA
<213> Brassica rapa
<400> 107
atggaattag agtgccaaga gagagctgaa gttgcgattg ttggcgaagg tttatccaat 60
caaggagctg agttgagaat tggagatggg tctgtcgagg ttccatgtgt gaagctttgt 120
tatcagcaaa agaagtcaac tttggtggtg cccagggaca aggatttgtc tacaaatact 180
catcgataca ccattatcga tttgcctcag gctttgatat ccgagattct taactgcctt 240
gacccaaaag agttgggcct tgtgtcctgt gtttcgactt gtctgcatag gttagcttcg 300
gagcatcacg cgtggaaagg cttctactgc gagagatggg ggctccccgt cgtcttaggg 360
ggtaaaactt ctgaagagag gtcatggaaa gagctgttta ttgagaggga gtttaggagc 420
aagacttttt taggacgtca tagtattgat acccctttatg gtcacactga agcagttcgc 480
actgtgtttc ttttagcttc tgctaagctc attttcactt ctggatatga ctgcgttgtt 540
cggatgtggg gaatggaaga aggcttgtct attgcagcgt ctaaaccgct tggttgtact 600
attagagcag ttgcggctga tacgaagctc ttggtagctg gcggtactga tggttttata 660
cattgctgga aagcagtgga tggtctgaga gacttgtttg acctcacggg ttttcagaaa 720
gagaagacgg agtttcgcct ttgggagcat gagggtccta ttacatctct tgccctggat 780
atgacgagta tctttagcgg ttcgtgggac atgtctgttc gtatatggga tagatcttca 840
ttcaagtgta tcaaaacgtt gaggcatagc gattgggttt ggggactagc gcctcatgag 900
acaagcatcg cgtgtgctgc tggttcggat gtgtacattt gggacattag cagtgagact 960
ccagaggcga ttatccatga tgctcatgag ggtaacactt actctcttgc aagaagccac 1020
agtggggatt tcttgtttac tggtggggaa gatggaggga tcaaaatgtt tgagatcaga 1080
aaacacggta atgaaacaag cgtcctgctc atatctcaat ggatgcctca cactggtccc 1140

gtctactctc tttctttcga gttcccctgg cttgtatcag catctggtga cggtaaactc 1200
gctcttatcg acgttaggaa gttgttgaag actaaccgtc gtggcttacc caagagggtt 1260
tcttcgagaa ttgtggaacc cccacagaga atgctgcacg ggttcggttc aaacctgttc 1320
tctgtggacg tgggctgtga ccgtatagtt tgtggaggtg aagaaggcgt agttcggata 1380
tggaacttca cacaagcgtt ggagattgag aggagagctc gagctctgaa aggaatgagg 1440
cttgagaaca ggatgaggcg aaggaggatg cagatggaga tgaatgcgaa gagtggaagg 1500
cctgaccaat gctcgattgc tgctcataag aaccctgtta atggagatag gaatcgagcg 1560
tggcatacaa aacgaagagc aagtggcaag gcgaaggcct aa 1602

<210> 108
<211> 533
<212> PRT
<213> Brassica rapa
<400> 108

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

```
Ala Leu Ile Asp Val Arg Lys Leu Leu Lys Thr Asn Arg Arg Gly Leu
                405                 410                 415
Pro Lys Arg Val Ser Ser Arg Ile Val Glu Pro Pro Gln Arg Met Leu
            420                 425                 430
His Gly Phe Gly Ser Asn Leu Phe Ser Val Asp Val Gly Cys Asp Arg
        435                 440                 445
Ile Val Cys Gly Gly Glu Glu Gly Val Val Arg Ile Trp Asn Phe Thr
    450                 455                 460
Gln Ala Leu Glu Ile Glu Arg Arg Ala Arg Ala Leu Lys Gly Met Arg
465                 470                 475                 480
Leu Glu Asn Arg Met Arg Arg Arg Arg Met Gln Met Glu Met Asn Ala
                485                 490                 495
Lys Ser Gly Arg Pro Asp Gln Cys Ser Ile Ala Ala His Lys Asn Pro
            500                 505                 510
Val Asn Gly Asp Arg Asn Arg Ala Trp His Thr Lys Arg Arg Ala Ser
        515                 520                 525
Gly Lys Ala Lys Ala
    530
<210> 109
<211> 1716
<212> DNA
<213> Sorghum bicolor
<400> 109
atggcctttg actgcaacaa ggaaagagga gattcttcgc ctgataattg ctcccgcatt     60
tgcactgagg gtactctcgt ccaggcaaat cccttatctc actactggaa gcctaagggt    120
ttgaagagcc tcaacaaatt ggagaaccag aagtccagtc atggatctgt tccgagagac    180
gataatccag aacaagaagc tgaagcgagc gatgaggcat ctgcctcaac ctgtggcatc    240
aggtgcttca ccgatctgcc ggctgctttg gtctgcgagg tccttgcacg cctcgatgca    300
aaggaccttg gaattgtatc ctgtgtctcc accctcctgc atgccttagc cacagatcat    360
cagggatgga agaagttata ctgcgaaagg tgggggcttc cagaccttcc cgacgccctg    420
aatgggccct tgttgcctgg tgccccccta gatgggaagt cttggaaaac atttttcgtg    480
gagcgggagt tcaggagtaa atcattcatg ggtagattca atgtggatat tctccgtggc    540
cacaatgagg atgttcgagc cgtcttcctt ttggtatcag caaatctgat attcactggc    600
gtcactggcg gccgcgattc tgtggttaga atgtggaaaa tggaggaagg gttattgatt    660
gatacatccc gtgcacttgg tggcaccatc cgggcgattg cagctgactc taggctttta    720
gtgagtggag gaactaatgc ctatattcag tgttggaggg ctgttgaagg caatgatcac    780
ttatttcaca tctctggaag tggtactgac cagaatgcgg agtttcgcct ctgggggcat    840
gaaggtcctg tgactagtct tgccttggat tcactgagga tttatagtgg ttcttgggac    900
atgactgttc gtgtttggga cagagcccag atggattgcg tgcaaaagtt catgcatgca    960
gactgggtgt ggtctcttgc tctgcgtgga aatactgttg ccagtactgc tggtagagat   1020
gcatatgtat gggatatcag ggacggcgag ttaacaagct tagtttccaa tgcccatgtt   1080
ggtaatgcat attcattggc ttggacacac ctggcggatg tgctgtttac tggtggagag   1140
gatggcgcta ttcgcttgtt caatgtttct gatgtctctg atgatgagga ggacattaaa   1200
ccagtggcaa cttgggtgcc acattcaggt cctgttcatt ctcttgcttt tgagtatcca   1260
tggcttgtgt cagcctcgag tgatggcagg agtgatggca ggattgcact tattgatttg   1320
aggaagcttc tgactcccaa aagatcatca aaaggtctgc gtgttaagag ctttgatgca   1380
agtgccatag agcctccaca gaggatgctt catggcgttg gatgtgatct cttctccatc   1440
gccattggtg cagataggat tgtatgtgca ggtgaggatg gttctgttag agtctggaac   1500
ttctcagaag cactggagat tgagaggagg gcacaggctc taaggagttt gaggcaggag   1560
aaccgcatga ggcggaggaa ggcacaagta gagatgaatg caaatggtag aaggcctgac   1620
cagtgctcaa tagccatgaa aaagaaccaa ctgaaggctg ataagagtgc cacttggcac   1680
aacaagcgtg ccgttaatga gaaggccaag tcttag                             1716
<210> 110
<211> 571
<212> PRT
<213> Sorghum bicolor
<400> 110
Met Ala Phe Asp Cys Asn Lys Glu Arg Gly Asp Ser Ser Pro Asp Asn
1               5                   10                  15
Cys Ser Arg Ile Cys Thr Glu Gly Thr Leu Val Gln Ala Asn Pro Leu
            20                  25                  30
Ser His Tyr Trp Lys Pro Lys Gly Leu Lys Ser Leu Asn Lys Leu Glu
        35                  40                  45
```

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

545 550 555 560
Asn Lys Arg Ala Val Asn Glu Lys Ala Lys Ser
565 570

<210> 111
<211> 1779
<212> DNA
<213> Vitis vinifera
<400> 111
atggcatttg aatgccaaga gagtatagag gtttgtttgg tgaaaaattc aatagattct 60
gatgaacaac aggstggatt gagtgatttt aattccaatt ttaatcatat cacttcaatt 120
tttgatgcca aatctcaatt ggaaaccgaa actgcacacc gagaaagtgg agtggtttgt 180
ttgttgaaga attcaattga agaacgggct ggattgactc agtttagtcc cgattctgat 240
cacaatactt taatacrtga ctcgggaaat tcccaatcgg aaattgaaat cggaatccaa 300
aaaccttcaa cttcaaaccc caaagttaac gacacttcag gacattatcg taggttgata 360
actgatcttc cgtccgcgtt gatatcggaa attcttaatt gccttgaccc gaaagagctt 420
ggtgtggttt cgtgtgtctc aactgttctc aataggctag cgtcygagca ctccgtttgg 480
aaagatttct attgtgagag gtggggagct ccggttgttt cggggttttc agatgagaaa 540
tcatggaggg aattgtttgt ggagagggag tttaggagca aaacctttag gggacgattt 600
agcattgatg ttctgtatgg tcacactgag gcggttcgag ctgttttcct tttggcctct 660
gcaaagctca ttttcacttc tgggtatgat tccattgttc gaatgtggga tatggaagaa 720
gggttgtcaa ttgcgtgttc acggcctctc ggttgcacaa taagagcagt ggctgcggat 780
aagaaactgt tgcttgcggg gggtagtgat gggtttattc attgttggag agctgtagag 840
gggctctctt gcttgtttga ccttgtgggt tctcaaaacc tgagtactga gttccgaatt 900
tgggaacatg aaggtcctgt gacttgtctt gctttggata ttaagagaat ttatagtggc 960
tcatgggaca tgactgtgcg catatgggac cgttcttctt ttaaggttgt aaaggtcttg 1020
aggcacacag actgggtttg gggccttgtt cctcgtgata ctacagttgc tagcacttct 1080
ggctcagatg tgtatgtttg ggacgctgat agtgggactc tgctgacgat aatctctaat 1140
gctcatgtgg gtaatgctta tgctttggca cggagccaca caggggattt tctattcact 1200
gggggagaag atggggcaat acatatgttt gaggtcgtga gtgattgcat ggaaaggaat 1260
gtattggagg tttcaacgtg gattcctcat tctggtcctg ttcattccct ggcatttgag 1320
tttccctggc ttgtttcagc ttcagctgat gggaggatgt cactgattga tgtgagaaag 1380
cttttgcaaa cttgcaagcc ttccttaggg aagaatgttt ccaaggttag gcacagggac 1440
cacaaaagtg tggagccccc tcagaggatg ttacatgggt ttggctgcaa tttattctca 1500
gtggacattg gtgcagatcg tattgtctgt ggaggtgagg aaggtgtggt tagaatttgg 1560
aacttctcac aagctttaga agcagagcag agggcccgtg ctttaagagg aatacgtcta 1620
gaaaacagga tgaggcggcg taggcttcaa atagagctga ctagtaaggg tagtcgaact 1680
gatcaatgtt cagttgcagc cagtaagaat cctattaatg gtgataggag tggtgtgtgg 1740
cacaataagc ggggaatgag tggcaggatg aaagcatag 1779

<210> 112
<211> 592
<212> PRT
<213> Vitis vinifera
<220>
<221> INCERTO
<222> (25)..(25)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> INCERTO
<222> (86)..(86)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 112
Met Ala Phe Glu Cys Gln Glu Ser Ile Glu Val Cys Leu Val Lys Asn
1 5 10 15
Ser Ile Asp Ser Asp Glu Gln Gln Xaa Gly Leu Ser Asp Phe Asn Ser
20 25 30
Asn Phe Asn His Ile Thr Ser Ile Phe Asp Ala Lys Ser Gln Leu Glu
35 40 45
Thr Glu Thr Ala His Arg Glu Ser Gly Val Val Cys Leu Leu Lys Asn
50 55 60
Ser Ile Glu Glu Arg Ala Gly Leu Thr Gln Phe Ser Pro Asp Ser Asp
65 70 75 80
His Asn Thr Leu Ile Xaa Asp Ser Gly Asn Ser Gln Ser Glu Ile Glu
85 90 95

```
Ile Gly Ile Gln Lys Pro Ser Thr Ser Asn Pro Lys Val Asn Asp Thr
            100                 105                 110
Ser Gly His Tyr Arg Arg Leu Ile Thr Asp Leu Pro Ser Ala Leu Ile
        115                 120                 125
Ser Glu Ile Leu Asn Cys Leu Asp Pro Lys Glu Leu Gly Val Val Ser
    130                 135                 140
Cys Val Ser Thr Val Leu Asn Arg Leu Ala Ser Glu His Ser Val Trp
145                 150                 155                 160
Lys Asp Phe Tyr Cys Glu Arg Trp Gly Ala Pro Val Val Ser Gly Phe
                165                 170                 175
Ser Asp Glu Lys Ser Trp Arg Glu Leu Phe Val Glu Arg Glu Phe Arg
            180                 185                 190
Ser Lys Thr Phe Arg Gly Arg Phe Ser Ile Asp Val Leu Tyr Gly His
        195                 200                 205
Thr Glu Ala Val Arg Ala Val Phe Leu Leu Ala Ser Ala Lys Leu Ile
    210                 215                 220
Phe Thr Ser Gly Tyr Asp Ser Ile Val Arg Met Trp Asp Met Glu Glu
225                 230                 235                 240
Gly Leu Ser Ile Ala Cys Ser Arg Pro Leu Gly Cys Thr Ile Arg Ala
                245                 250                 255
Val Ala Ala Asp Lys Lys Leu Leu Leu Ala Gly Gly Ser Asp Gly Phe
            260                 265                 270
Ile His Cys Trp Arg Ala Val Glu Gly Leu Ser Cys Leu Phe Asp Leu
        275                 280                 285
Val Gly Ser Gln Asn Leu Ser Thr Glu Phe Arg Ile Trp Glu His Glu
    290                 295                 300
Gly Pro Val Thr Cys Leu Ala Leu Asp Ile Lys Arg Ile Tyr Ser Gly
305                 310                 315                 320
Ser Trp Asp Met Thr Val Arg Ile Trp Asp Arg Ser Ser Phe Lys Val
                325                 330                 335
Val Lys Val Leu Arg His Thr Asp Trp Val Trp Gly Leu Val Pro Arg
            340                 345                 350
Asp Thr Thr Val Ala Ser Thr Ser Gly Ser Asp Val Tyr Val Trp Asp
        355                 360                 365
Ala Asp Ser Gly Thr Leu Leu Thr Ile Ile Ser Asn Ala His Val Gly
    370                 375                 380
Asn Ala Tyr Ala Leu Ala Arg Ser His Thr Gly Asp Phe Leu Phe Thr
385                 390                 395                 400
Gly Gly Glu Asp Gly Ala Ile His Met Phe Glu Val Val Ser Asp Cys
                405                 410                 415
Met Glu Arg Asn Val Leu Glu Val Ser Thr Trp Ile Pro His Ser Gly
            420                 425                 430
Pro Val His Ser Leu Ala Phe Glu Phe Pro Trp Leu Val Ser Ala Ser
        435                 440                 445
Ala Asp Gly Arg Met Ser Leu Ile Asp Val Arg Lys Leu Leu Gln Thr
    450                 455                 460
Cys Lys Pro Ser Leu Gly Lys Asn Val Ser Lys Val Arg His Arg Asp
465                 470                 475                 480
His Lys Ser Val Glu Pro Pro Gln Arg Met Leu His Gly Phe Gly Cys
                485                 490                 495
Asn Leu Phe Ser Val Asp Ile Gly Ala Asp Arg Ile Val Cys Gly Gly
            500                 505                 510
Glu Glu Gly Val Val Arg Ile Trp Asn Phe Ser Gln Ala Leu Glu Ala
        515                 520                 525
Glu Gln Arg Ala Arg Ala Leu Arg Gly Ile Arg Leu Glu Asn Arg Met
    530                 535                 540
Arg Arg Arg Arg Leu Gln Ile Glu Leu Thr Ser Lys Gly Ser Arg Thr
545                 550                 555                 560
Asp Gln Cys Ser Val Ala Ala Ser Lys Asn Pro Ile Asn Gly Asp Arg
                565                 570                 575
Ser Gly Val Trp His Asn Lys Arg Gly Met Ser Gly Arg Met Lys Ala
            580                 585                 590
<210> 113
```

<211> 745
<212> DNA
<213> Zea mays
<400> 113
atggcggaca aggagcctgt cgtggagcga tccgaggcgt ctgaggagga ggacgcctcc 60
gccgcggccg ccgcggggga ggaagaggac acgggtgccc aggtggcccc catcgtgcgg 120
cttgaggagg tactcgtcac caccggtgag gaggacgagg acgtcctgct cgacatgaag 180
gcgaagctgt accggtttga taaagacggg aatcaatgga aggaacgggg cacgggcacc 240
gtcaagcttc tcaagaacaa ggagaccggc aaggtccgac tggtcatgcg ccaggccaag 300
acgcttaaga tctgcgcgaa ccacctcgtg gcccccacca cgagaatgca ggaacatgcc 360
ggcagcgaca aatcgtgcgt ctggcacgct tctgactttg cggatggcga actcaaggag 420
gagatgtttg ccatcaggtt tggttcggta gaaaactgca agaagttcaa ggagttggtt 480
gaggagattg ctgagtcact tacagagaac gaggacaagg aaggtcaaga tggctcttcc 540
acggccggat tgctggagaa gctcactgtg agcgagggca aatctgagga aagtggcaag 600
tcggagtcca ctgattctgg caaagtgacc gaaaccaaag ccgatgcagc cccagccgag 660
tagttggtgt ggttgcacta cccagctttc ttgtacaaag ttggcattat aagaaagcat 720
tgcttatcaa tttgttgcaa cgaac 745
<210> 114
<211> 220
<212> PRT
<213> Zea mays
<400> 114
Met Ala Asp Lys Glu Pro Val Val Glu Arg Ser Glu Ala Ser Glu Glu
1 5 10 15
Glu Asp Ala Ser Ala Ala Ala Ala Ala Gly Glu Glu Glu Asp Thr Gly
20 25 30
Ala Gln Val Ala Pro Ile Val Arg Leu Glu Glu Val Leu Val Thr Thr
35 40 45
Gly Glu Glu Asp Glu Asp Val Leu Leu Asp Met Lys Ala Lys Leu Tyr
50 55 60
Arg Phe Asp Lys Asp Gly Asn Gln Trp Lys Glu Arg Gly Thr Gly Thr
65 70 75 80
Val Lys Leu Leu Lys Asn Lys Glu Thr Gly Lys Val Arg Leu Val Met
85 90 95
Arg Gln Ala Lys Thr Leu Lys Ile Cys Ala Asn His Leu Val Ala Pro
100 105 110
Thr Thr Arg Met Gln Glu His Ala Gly Ser Asp Lys Ser Cys Val Trp
115 120 125
His Ala Ser Asp Phe Ala Asp Gly Glu Leu Lys Glu Glu Met Phe Ala
130 135 140
Ile Arg Phe Gly Ser Val Glu Asn Cys Lys Lys Phe Lys Glu Leu Val
145 150 155 160
Glu Glu Ile Ala Glu Ser Leu Thr Glu Asn Glu Asp Lys Glu Gly Gln
165 170 175
Asp Gly Ser Ser Thr Ala Gly Leu Leu Glu Lys Leu Thr Val Ser Glu
180 185 190
Gly Lys Ser Glu Glu Ser Gly Lys Ser Glu Ser Thr Asp Ser Gly Lys
195 200 205
Val Thr Glu Thr Lys Ala Asp Ala Ala Pro Ala Glu
210 215 220
<210> 115
<211> 196
<212> PRT
<213> Zea mays
<400> 115
Met Ala Asp Lys Glu Pro Val Val Glu Arg Pro Glu Ala Gly Glu Glu
1 5 10 15
Glu Glu Asp Ala Ser Ala Ala Ala Ala Ala Gly Glu Glu Glu Asp Thr
20 25 30
Gly Ala Gln Val Ala Pro Ile Val Arg Leu Glu Glu Val Ala Val Thr
35 40 45
Thr Gly Glu Glu Asp Glu Asp Val Leu Leu Asp Met Lys Ala Lys Leu
50 55 60

```
Leu Pro Ile Arg Gln Arg Arg Met Arg Gln Ala Lys Thr Leu Lys Ile
65                  70                  75                  80
Cys Ala Asn His Leu Val Ala Ser Thr Thr Lys Met Gln Glu His Ala
                85                  90                  95
Gly Ser Asp Lys Ser Cys Val Trp His Ala Ala Asp Phe Ala Asp Gly
            100                 105                 110
Glu Leu Lys Glu Glu Met Phe Ala Ile Arg Phe Gly Ser Val Glu Asn
        115                 120                 125
Cys Lys Lys Phe Lys Glu Leu Val Glu Glu Ile Ala Glu Ser Leu Thr
    130                 135                 140
Lys Asn Glu Gly Lys Glu Gly Glu Asp Gly Gly Ser Thr Ala Gly Leu
145                 150                 155                 160
Leu Ala Lys Leu Thr Val Ser Glu Gly Lys Ser Glu Gly Ser Gly Lys
                165                 170                 175
Ser Glu Ser Thr Asp Ser Gly Lys Val Thr Glu Thr Lys Ala Asp Thr
            180                 185                 190
Ala Pro Ala Glu
        195
<210> 116
<211> 725
<212> DNA
<213> Arabidopsis thaliana
<400> 116
atgccaactt tgtacaaaaa agcaggcttc acaatggcga gcattagcaa cgagcccgag     60
cgtgagaaca gagacgaaga agagactgga gccaacgaag atgaagacac cggtgctcag    120
gttgctccta tcgtcaggct tgaagaagtc gccgtcacca ccggtgaaga agacgaagac    180
accatcctcg atctgaaatc gaagttgtat cgatttgata aagatggaag tcagtggaag    240
gagagaggtg ctggtactgt taagtttttg aaacatagag tttctgggaa gattcgtctc    300
gttatgaggc aatcgaaaac tttgaagatc tgtgctaatc atcttgttgg atcgggtatg    360
agtgttcagg aacacgctgg gaatgataag tcttgtgtat ggcacgctcg tgatttctcc    420
gatggtgaat tgaaggatga gctcttctgt atccggtttg cttcagttga gaattgcaaa    480
gcatttatgc aaaagttcaa ggaagtagct gaatctgaag aagagaaaga agagagcaaa    540
gatgcctctg ataccgctgg tcttcttgag aagttaacag tggaagagaa ggaaagtgag    600
aagaaaccag tggagaaggc agaggaaaac aaaaagagtg aagctgttga agaaaagaaa    660
acagaggagt ctgttccctc agcttaagat gcacccagct ttcttgtaca aagttggcat    720
tataa                                                                725
<210> 117
<211> 228
<212> PRT
<213> Arabidopsis thaliana
<400> 117
Met Pro Thr Leu Tyr Lys Lys Ala Gly Phe Thr Met Ala Ser Ile Ser
1               5                  10                  15
Asn Glu Pro Glu Arg Glu Asn Arg Asp Glu Glu Glu Thr Gly Ala Asn
            20                  25                  30
Glu Asp Glu Asp Thr Gly Ala Gln Val Ala Pro Ile Val Arg Leu Glu
        35                  40                  45
Glu Val Ala Val Thr Thr Gly Glu Glu Asp Glu Asp Thr Ile Leu Asp
    50                  55                  60
Leu Lys Ser Lys Leu Tyr Arg Phe Asp Lys Asp Gly Ser Gln Trp Lys
65                  70                  75                  80
Glu Arg Gly Ala Gly Thr Val Lys Phe Leu Lys His Arg Val Ser Gly
                85                  90                  95
Lys Ile Arg Leu Val Met Arg Gln Ser Lys Thr Leu Lys Ile Cys Ala
            100                 105                 110
Asn His Leu Val Gly Ser Gly Met Ser Val Gln Glu His Ala Gly Asn
        115                 120                 125
Asp Lys Ser Cys Val Trp His Ala Arg Asp Phe Ser Asp Gly Glu Leu
    130                 135                 140
Lys Asp Glu Leu Phe Cys Ile Arg Phe Ala Ser Val Glu Asn Cys Lys
145                 150                 155                 160
Ala Phe Met Gln Lys Phe Lys Glu Val Ala Glu Ser Glu Glu Glu Lys
                165                 170                 175
```

Glu Glu Ser Lys Asp Ala Ser Asp Thr Ala Gly Leu Leu Glu Lys Leu
180 185 190
Thr Val Glu Glu Lys Glu Ser Glu Lys Lys Pro Val Glu Lys Ala Glu
195 200 205
Glu Asn Lys Lys Ser Glu Ala Val Glu Glu Lys Lys Thr Glu Glu Ser
210 215 220
Val Pro Ser Ala
225

<210> 118
<211> 217
<212> PRT
<213> Arabidopsis thaliana
<400> 118

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

<210> 119
<211> 1416
<212> DNA
<213> Arabidopsis thaliana
<400> 119

gcgaaacccc caaattctct tctctctatc tctctcgcgt acgccgcacc gtagcgacca 60
attgaatcca cgaggaaatc tcagtaaata ctatggcgac caacgaaccc gagcacgagc 120
acagagacga ggaagaggcc ggagctaacg aggatgagga caccggagct caggtcgctc 180
cgatcgttag gcttgaggag gttgccgtca ctaccggcga ggaagacgaa gatgccgtcc 240
ttgatctgaa atcgaagctt tatcgattcg ataaggatgc gaatcagtgg aaggagagag 300
gagctggtac tgtgaagttc ttaaagcata agaacactgg gaagattcgt ctcgttatga 360
ggcaatctaa aactttgaag atctgtgcta atcacttcgt taaatcgggc atgagtgttc 420
aggaacacgt tgggaatgaa aagtcatgtg tgtggcacgc tcgtgacttt gctgatggtg 480
aactcaagga tgagctttc tgtatccgat ttgcttctat tgagaattgc aaaacattta 540
tgcaaaagtt caaggaagtt gctgagtctg aagaagagaa agaagagagc aaagatgccg 600
ctgacactgc tggccttctt gagaaattga ctgtggaaga gacaaaaacg gaggagaaaa 660
ccgaagcgaa agctgtggag acggcaaaga ctgaagtgaa agcagaagaa aagaaagaga 720
gcgaggcaga gaaatctggt gaagcaaaga aaacagaaga aagtggtccc tcaacataag 780
aagcgtcatc atttaagttg ccaaatcctg gcgaggtaat taagcctcga aatgttttga 840
tgcatcatga gtctaccatt gtcttggcac tatctatcta tacatgtttt gtcgagtcat 900
atcaagtggt tgggggatcg cttggggtcg ggttttactg aatgctgagt cgttatgggt 960
ctaatagttt ttgagtctaa agtgtcgggt gatatgagag atttgggtga aatatttttt 1020

catgttggtt atctgaaaga gtacattggt ttcattgttt taagttttct catggatcct 1080
ttttggatgg tcttattttg aggatacaaa tgtgtttgtc catggacaag aatcagaagt 1140
ctccattttt tttttttcaa aatttaatgc atgtgatttt ttaatcttaa ggtaattgcc 1200
aattgtttga caaaaaaaag gtaattgcca atctactttg ctctcttgtg ttagtcgatg 1260
ctagtcttga ccctgataaa gatccaaaat gtatattaac agtattcata aaattcttca 1320
gttataacga actgttcacg gaaaaaaaat gacattgtac tatactgtgc taaaattacc 1380
aaagcatgat ttatataaat catatccagt aagacc 1416

<210> 120
<211> 228
<212> PRT
<213> Arabidopsis thaliana
<400> 120

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

<210> 121
<211> 660
<212> DNA
<213> Arabidopsis thaliana
<400> 121

atggcgagca ctgagccaga gcgtgagaac cgtgaagatg aaaccgaagt caacgaagat 60
gaagacaccg gagctcaggt agctccgatc gttaggcttg aagaggttgc agtcaccacc 120
ggcgaagaag atgaagacgc cgtcctcgat ctgaaatcga agatgtatcg attcgataaa 180
gaagggaacc aatggaagga gagaggagct ggaactgtga agttattgaa gcataaggaa 240
actggaaaag ttcgtcttgt tatgagacaa tctaaaaccc taaagatctg tgccaatcac 300
ctcatttcgt ctgggatgag tgttcaggaa catagtggga atgaaaagtc ttgtttatgg 360
cacgctactg atttctccga cggcgagttg aaagatgagc ttttctgcat tcgatttgct 420
tccattgaga attgcaaaac atttatggag aagttcactg aaatagctga aagccagcaa 480
gtagggaaag agagcaccca gggcgacgag gctgctggtt taatagagaa tctttcggtt 540
gaggaaaata taagtgagga gaaagccaag gaagcagaag agaaagagcc tgcaaaggaa 600
gataaagaaa caaaaaagga gaaggtagaa gaagagaaga aaacagaggc aagcacttaa 660

<210> 122
<211> 260
<212> PRT
<213> Arabidopsis thaliana
<400> 122

```
Met Ala Ser Thr Glu Pro Glu Arg Glu Asn Arg Glu Asp Glu Thr Glu
1               5                  10                  15
Val Asn Glu Asp Glu Asp Thr Gly Ala Gln Val Ala Pro Ile Val Arg
            20                  25                  30
Leu Glu Glu Val Ala Val Thr Thr Gly Glu Glu Asp Glu Asp Ala Val
        35                  40                  45
Leu Asp Leu Tyr Val Thr Arg Ser Val Asn Ile Leu Val Ser Leu Pro
    50                  55                  60
Leu Val Lys Met Tyr Arg Phe Asp Lys Glu Gly Asn Gln Trp Lys Glu
65                  70                  75                  80
Arg Gly Ala Gly Thr Val Lys Leu Leu Lys His Lys Glu Thr Gly Lys
                85                  90                  95
Val Arg Leu Val Met Arg Gln Ser Lys Thr Leu Lys Ile Cys Ala Asn
            100                 105                 110
His Leu Ile Ser Ser Gly Met Ser Val Gln Glu His Ser Gly Asn Glu
        115                 120                 125
Lys Ser Cys Leu Trp His Ala Thr Asp Phe Ser Asp Gly Glu Leu Lys
    130                 135                 140
Asp Glu Leu Phe Cys Ile Arg Phe Ala Ser Ile Glu Arg Lys Met Val
145                 150                 155                 160
Trp Lys Ile Leu Glu Trp Leu Thr Asp Ser Val Ala Ser Thr Asp Cys
                165                 170                 175
Lys Thr Phe Met Glu Lys Phe Thr Glu Ile Ala Glu Ser Gln Gln Val
            180                 185                 190
Gly Lys Glu Ser Thr Gln Gly Asp Glu Ala Ala Gly Leu Ile Glu Asn
        195                 200                 205
Leu Ser Val Glu Glu Asn Ile Ser Glu Glu Lys Ala Lys Glu Ala Glu
    210                 215                 220
Glu Lys Glu Pro Ala Lys Glu Asp Lys Glu Thr Lys Lys Glu Lys Val
225                 230                 235                 240
Thr Gln Ser Val Ile Asp Met Phe Gly Val Ala Ala Lys Gly Thr Ile
                245                 250                 255
Met Trp Cys Phe
            260
<210> 123
<211> 1104
<212> DNA
<213> Lycopersicon esculentum
<400> 123
aaaaaaagca aaaagaaaag aggcaaagtt atcaaatcaa aaaccctact tcctcccttg     60
ttgttcttct tcttcttcaa atccggggta aatctttcta aaaccaaaaa tctagagtga    120
tggcaagcac tactgttgaa cccaaattgg agaagaaaga agaggaagta gaagcagaag    180
tagaagaaaa cccaactggc gatgatgaag acactggtgc acaagttgct cccattgtta    240
ggctacaaga agttgctgtc tccactggtg aagaaaatga acatgttctt cttgatctga    300
aatcaaagtt ataccgattt gacaaggaag ggagtcaatg gaaagagaga ggtgttggga    360
ctgtgaagct tctcaagcac aaggaaactg gaaaggttag gcttgtgatg aggcaatcca    420
agacgctgaa aatctgtgcc aaccacttgg tccttcctac tatgtcgatc caagaacatg    480
ctgggaatga gaagtcttgt gtgtggcatg ctgctgactt tgctgatgga gaactgaaag    540
atgagacttt ttgcatccgc tttgcttcag ttgagaattg caaggctttc aaagagaagg    600
ttgaagaaat tgctgaatct caacagacaa agtctggaca aagtgaagaa gctggtgctg    660
ctgctactga acttattgaa aagcttagtg ttgaaagcaa agataaaaat gacaaacctg    720
aagacaaaga ggcccctgca gctacagagg aaaaggaaga taagaaggaa gagaaagctg    780
aagaaaagaa ttaaaatgtt tgtattgtcg gtcagttttt ttttttgaaa ggttaatagc    840
attcgttgtt tctgtctgat ccaatagata tgatagatat aagcaatgtg tctcttgtgg    900
tcttatttgt tgttgttggg atggtttgga ttttcatttg ggtcttgagt cgagtgcagc    960
ttcatgtttt attgtggtct ttggtgtttc cttgtactta tttattgtgt gtttgcaaat   1020
tttggttcat ggggtacacc acccagtata agggattacg tttgttaaaa cctttagcac   1080
tgttagtgag atcatttttg agtt                                          1104
<210> 124
<211> 224
<212> PRT
<213> Lycopersicon esculentum
<400> 124
```

```
Met Ala Ser Thr Thr Val Glu Pro Lys Leu Glu Lys Lys Glu Glu Glu
1               5                  10                  15
Val Glu Ala Glu Val Glu Glu Asn Pro Thr Gly Asp Asp Glu Asp Thr
            20                  25                  30
Gly Ala Gln Val Ala Pro Ile Val Arg Leu Gln Glu Val Ala Val Ser
        35                  40                  45
Thr Gly Glu Glu Asn Glu His Val Leu Leu Asp Leu Lys Ser Lys Leu
    50                  55                  60
Tyr Arg Phe Asp Lys Glu Gly Ser Gln Trp Lys Glu Arg Gly Val Gly
65                  70                  75                  80
Thr Val Lys Leu Leu Lys His Lys Glu Thr Gly Lys Val Arg Leu Val
                85                  90                  95
Met Arg Gln Ser Lys Thr Leu Lys Ile Cys Ala Asn His Leu Val Leu
            100                 105                 110
Pro Thr Met Ser Ile Gln Glu His Ala Gly Asn Glu Lys Ser Cys Val
        115                 120                 125
Trp His Ala Ala Asp Phe Ala Asp Gly Glu Leu Lys Asp Glu Thr Phe
    130                 135                 140
Cys Ile Arg Phe Ala Ser Val Glu Asn Cys Lys Ala Phe Lys Glu Lys
145                 150                 155                 160
Val Glu Glu Ile Ala Glu Ser Gln Gln Thr Lys Ser Gly Gln Ser Glu
                165                 170                 175
Glu Ala Gly Ala Ala Ala Thr Glu Leu Ile Glu Lys Leu Ser Val Glu
            180                 185                 190
Ser Lys Asp Lys Asn Asp Lys Pro Glu Asp Lys Glu Ala Pro Ala Ala
        195                 200                 205
Thr Glu Glu Lys Glu Asp Lys Lys Glu Glu Lys Ala Glu Glu Lys Asn
    210                 215                 220
<210> 125
<211> 442
<212> DNA
<213> Oryza sativa
<400> 125
caacctctcc ccccacgatc gccctcctcc ccgaaacttc tggaaccgag agcagcagac     60
gttagctctt ctcctccgat ggcggacaag gagcccgtcg tggaccgccc cgaggacgag    120
gaggaggcct ccgccgccgc cgccgccgcg ggcggcgagg aggaggacac gggcgcccag    180
gtcgccccca tcgtgcggct cgaggaggtc gccgtcacca ccggcgagga ggacgaggac    240
gtgctcctcg acatgaaggc gaagctttac cggtttgaca aggaggggaa ccagtggaag    300
gagcgggggа cgggcaccgt caagctcctc aagcacaagg agaacggcaa ggtccgcctc    360
gtcatgcgcc aggccaagac gctcaagatc tgcgcgaacc acctagttgc ttcgaccacg    420
aagatgcagg agcatgcggg ca                                             442
<210> 126
<211> 121
<212> PRT
<213> Oryza sativa
<400> 126
Met Ala Asp Lys Glu Pro Val Val Asp Arg Pro Glu Asp Glu Glu Glu
1               5                  10                  15
Ala Ser Ala Ala Ala Ala Ala Ala Gly Gly Glu Glu Glu Asp Thr Gly
            20                  25                  30
Ala Gln Val Ala Pro Ile Val Arg Leu Glu Glu Val Ala Val Thr Thr
        35                  40                  45
Gly Glu Glu Asp Glu Asp Val Leu Leu Asp Met Lys Ala Lys Leu Tyr
    50                  55                  60
Arg Phe Asp Lys Glu Gly Asn Gln Trp Lys Glu Arg Gly Thr Gly Thr
65                  70                  75                  80
Val Lys Leu Leu Lys His Lys Glu Asn Gly Lys Val Arg Leu Val Met
                85                  90                  95
Arg Gln Ala Lys Thr Leu Lys Ile Cys Ala Asn His Leu Val Ala Ser
            100                 105                 110
Thr Thr Lys Met Gln Glu His Ala Gly
        115                 120
<210> 127
```

<211> 1063
<212> DNA
<213> Zea mays
<400> 127
gcacgcggta aacaccccac gtctcaagct gccccttttct caccgccgcg gccgccacgc 60
gcagcagcaa ccgagcagga gcgcagccag caagctcagg cccgagccct actgcaaccg 120
ccgctgctac cgcactgaac accatggccg acccggcgga gcaccgtcca gcggaggagg 180
acgaggaagc cgcggcggcc ggcgaggatg aggacaccgg cgctcaggtc gcgcccatcg 240
tgaagctgga ggaggtcgcc gttaccaccg gagaggagga cgaggacgta cttgtcgaca 300
tgaaggcgaa gctctaccgg ttcgacaagg aggcgaacca gtggaaagaa cgtggcacag 360
gcactgtgaa gctgcttaag cacaaggaga ccgccaaggt ccgcctcgtc atgcgtcagg 420
cgaagacgct taagatctgt gctaaccatc ttgtggtggc gacgactaag atgcaggagc 480
acgcagggag cgacaagtcg tgcgtgtggc acgcgctgga cttcgccgac ggtgagctca 540
aggaggagat gtttgctatc cgttttggat ctgtcgagaa ttgcaagaag ttcaaggaca 600
ctgttgaaga gatagctgaa gagcaaggaa agaccgaggt gaaagaaaac gaggaagtct 660
ccattgccgc cgcagatttg gtacagaagt taacagtgac agaggctagc aacgagggag 720
atgcagagga agaggaggct cctgcatctg accataagaa ggatgctaaa gggtaaagat 780
tgaaggcaaa caaggccaaa tgattatgat tccattcatt tcgttgtcaa ggttcatctt 840
ccccacaaat tttcattaca aagtttcatt gcattttctc ctgagatgat ttgtgtgtgt 900
gtgtgtgtgt gtgtgtgtgt gttgggtgaa atctgagttg tcagtcatct acatgtcttt 960
cttggttgtt agacttattc tcagtcgcct gtttatctgg gatggctagg tacatcatgt 1020
ttgtacaatt atttggggtt gatttaaaaa aaagaaaaaa aaa 1063
<210> 128
<211> 210
<212> PRT
<213> Zea mays
<400> 128
Met Ala Asp Pro Ala Glu His Arg Pro Ala Glu Glu Asp Glu Glu Ala
1 5 10 15
Ala Ala Ala Gly Glu Asp Glu Asp Thr Gly Ala Gln Val Ala Pro Ile
20 25 30
Val Lys Leu Glu Glu Val Ala Val Thr Thr Gly Glu Glu Asp Glu Asp
35 40 45
Val Leu Val Asp Met Lys Ala Lys Leu Tyr Arg Phe Asp Lys Glu Ala
50 55 60
Asn Gln Trp Lys Glu Arg Gly Thr Gly Thr Val Lys Leu Leu Lys His
65 70 75 80
Lys Glu Thr Ala Lys Val Arg Leu Val Met Arg Gln Ala Lys Thr Leu
85 90 95
Lys Ile Cys Ala Asn His Leu Val Val Ala Thr Thr Lys Met Gln Glu
100 105 110
His Ala Gly Ser Asp Lys Ser Cys Val Trp His Ala Leu Asp Phe Ala
115 120 125
Asp Gly Glu Leu Lys Glu Glu Met Phe Ala Ile Arg Phe Gly Ser Val
130 135 140
Glu Asn Cys Lys Lys Phe Lys Asp Thr Val Glu Glu Ile Ala Glu Glu
145 150 155 160
Gln Gly Lys Thr Glu Val Lys Glu Asn Glu Glu Val Ser Ile Ala Ala
165 170 175
Ala Asp Leu Val Gln Lys Leu Thr Val Thr Glu Ala Ser Asn Glu Gly
180 185 190
Asp Ala Glu Glu Glu Glu Ala Pro Ala Ser Asp His Lys Lys Asp Ala
195 200 205
Lys Gly
210
<210> 129
<211> 1111
<212> DNA
<213> Oryza sativa
<400> 129
ggacgcgcag caaccggcag cagcagcaga ggaggaggcc agggcaaccc gcaaccccaa 60
accctactcc cgccgcgcca ccgccatggc cgacccagcg gagcaccgtg aggaggagga 120
ggaggccgcg gcggcggcgg cgggcgagga cgaggacacc ggcgcccagg tcgcgcccat 180

cgtcaagctc gaggaggtcg ccgtcaccac cggcgaggag gacgaggagg tcctcctcga 240
catgaagtcg aagctgtacc gcttcgacaa ggaggggaac caatggaagg agagaggcac 300
cggcacggtg aagctgctga agcacaagga gaccggcaag gtccgtctcg tcatgcgcca 360
ggctaagacg ctcaagatct gcgccaatca cctcgtggcg acgaccacga agatgcagga 420
gcacgccggc agcgacaagt cgtgcgtgtg gcacgcgctg gacttcgccg acggcgagct 480
caaggaggaa atgttcgcca tacgctttgg atccgtcgag aactgcaaga agttcaggga 540
gatggtagaa gagattgctg agcagcaagg aaagaacgag gagaaagaaa atgaggaagt 600
ctcctctact gcagggttgg ttgagaagct ttcagtgacc gagacaaaaa aggaggaaaa 660
tgcagagaaa gaggagactc ctgcagaaga ggataagaag gacgctaaag agtgaaagca 720
cccgcaacct ctaggcagct ttgatatgcc tgaagcgagt gaagtgttaa atgagcgtca 780
tttcattcat tttgtggtca aagttcttcc tttcacaaat tttcattgtg ttttcttttg 840
gatgaatgtt tgtgctaggc aaaatttgag ttgtatcagt cgggttcttg gttactacac 900
tgttacttaa accttgagtt gtttatccga gatggggtgg cggagtgggg gcacagcatg 960
tttgtacaat tatttgaagt tgcttttgga atgggcacgg tttgggttgt ccaaaatttg 1020
gacggtgatg ccctgtcact cacaattctt atctctgtct tgcaattata tgcgatgtga 1080
agctcttcgc ctcttcggtg acgagttgtc g 1111

<210> 130
<211> 209
<212> PRT
<213> Oryza sativa
<400> 130

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

<210> 131
<211> 767
<212> DNA
<213> Populus balsamifera subsp. trichocarpa
<400> 131

tctcctctcc aagcaagaac atcaaatcta atggcaagca cagaacccga acacgagcac 60
agagaagatg aggaagctcc ggctggcgaa gacgaagaca ccggtgctca ggttgctccg 120
atcgtcaagc tcgaagaagt tgctgtctct accggtgaag aagatgaaga tgcgatcctc 180
gatctgaaat cgaagcttta tagatttgat aaggacggga atcagtggaa agagagaggt 240
gctggcactg tcaagttatt gaagcataaa gagtctggaa aagttcgtct tgttatgaga 300
caatctaaga ctctcaagat ctgcgctaat catctcgtgc tgccgactat gtccgtgcag 360
gagcacgcag ggaatgataa gtcgtgtgtg tggcacgcta cagatttcgc tgatggtgaa 420
ttgaaggatg agttgttctg cattcgtttc gcatctgttg aaaattgcaa aacctttatg 480
gaaatgtttc aagaagttgc tgaatcccaa gagagcaaag aggaaaatga ggatgcaact 540
gttgctgctg atgcattgga gaagttgagt gttgaaggga agaaaactga ggagaatgct 600

ggcgaagagg cacctgctgc aaccaagaat gaggaaacta agactgatac agataagaaa 660
ggggagaagc ctgctccctc aacttgaggg ttgatttgtt tgttttgcca ttcccttggc 720
agtatgatga gttcagttgt caaccatatt tcttgtccat tttgtgg 767

<210> 132
<211> 218
<212> PRT
<213> Populus balsamifera subsp. trichocarpa
<400> 132

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

<210> 133
<211> 1105
<212> DNA
<213> Saccharum officinarum
<400> 133

ggagaacacc tacccgtctc cccaccctag ccccgccgtc gcgtcctctt cgaatcgaat 60
cgcgccgcga tcacctcatc tcatggcgga caaggagcct gtcgtggagc gacccgaggc 120
ggctgaggag gaggacgcct cagccgccgc cgctgccgcg ggggaggagg aggacacggg 180
cgcccaggtg gcccccatcg tgcggcttga ggaggtagac gtcaccaccg gcgaggagga 240
cgaggacgtc ctgctcgaca tgaaggcgaa gttgtaccga tttgacaaag acgggaacca 300
atggaaggaa cggggcaccg gcaccgtcaa gcttctcaag cacaaggaga ccggcaaggt 360
ccgactcgtc atgcgccagg ccaagacgct taagatctgc gcgaaccacc tcgtggcctc 420
gaccacgaag atgcaggagc atgccggcag cgacaagtca tgcgtctggc acgctgctga 480
ctttgccgat ggcgaactca aggaggagat gtttgccatc aggtttggtt ccgtagaaaa 540
ttgtaagaag ttcaaggagt tggttgagga gatttctgag tcgcttacaa agaacgaggg 600
caaggaaagt gaagatggtt cttccacagc tggattgctg gagaagctta ctgtgagtga 660
gcacaaatct gaggaaagtg acaagtcgga gtccactgat tctggcaaag tgaccgaaac 720
caaagccgac acagccccag ctgagtagtt ggtggtggca gatcctcccg gtgccaaatc 780
agtttgtgtc cttccagtgg aagtttggtg cccatttgcc atgaaatatg actgctaaca 840
tttgggggcg agttggagca gtctcagatg tttggatttg gtctagtctg gtttggtccg 900
ggcttgattt tgttaaaaac ttagtacatt ggggttggaa cgtttggtat gcgagtcgct 960
agtatttcac tgcatggatt gttatggatt gcggtataag gtgcatctta tatttttttt 1020
atttcgttca atacacatac atgtgttgta ttaatacttt atgccaatgc ccatggggag 1080
agttgaattt gaatgtcgag agtgg 1105

<210> 134
<211> 221
<212> PRT

<213> Saccharum officinarum
<400> 134
Met Ala Asp Lys Glu Pro Val Val Glu Arg Pro Glu Ala Ala Glu Glu
1 5 10 15
Glu Asp Ala Ser Ala Ala Ala Ala Ala Ala Gly Glu Glu Glu Asp Thr
20 25 30
Gly Ala Gln Val Ala Pro Ile Val Arg Leu Glu Glu Val Asp Val Thr
35 40 45
Thr Gly Glu Glu Asp Glu Asp Val Leu Leu Asp Met Lys Ala Lys Leu
50 55 60
Tyr Arg Phe Asp Lys Asp Gly Asn Gln Trp Lys Glu Arg Gly Thr Gly
65 70 75 80
Thr Val Lys Leu Leu Lys His Lys Glu Thr Gly Lys Val Arg Leu Val
85 90 95
Met Arg Gln Ala Lys Thr Leu Lys Ile Cys Ala Asn His Leu Val Ala
100 105 110
Ser Thr Thr Lys Met Gln Glu His Ala Gly Ser Asp Lys Ser Cys Val
115 120 125
Trp His Ala Ala Asp Phe Ala Asp Gly Glu Leu Lys Glu Glu Met Phe
130 135 140
Ala Ile Arg Phe Gly Ser Val Glu Asn Cys Lys Lys Phe Lys Glu Leu
145 150 155 160
Val Glu Glu Ile Ser Glu Ser Leu Thr Lys Asn Glu Gly Lys Glu Ser
165 170 175
Glu Asp Gly Ser Ser Thr Ala Gly Leu Leu Glu Lys Leu Thr Val Ser
180 185 190
Glu His Lys Ser Glu Glu Ser Asp Lys Ser Glu Ser Thr Asp Ser Gly
195 200 205
Lys Val Thr Glu Thr Lys Ala Asp Thr Ala Pro Ala Glu
210 215 220
<210> 135
<211> 1034
<212> DNA
<213> Saccharum officinarum
<400> 135
acagcagcaa ccggtagcag gagcgcagcc agcaagcgca ggccccgagc cgtactgcaa 60
ccgccgctgc caccgcgccg aacgccatgg ccgacccggc ggagcaccgc cctgcggagg 120
aggaggagga ggccgcggcg gccggcgagg atgaggacac cggcgcccag gtcgcgccca 180
tcgtaaagct ggaggaggtc gccgttacca ccggagagga ggacgaggac gtgctcctcg 240
acatgaaggc gaagctttac cggttcgaca aggaggggaa ccagtggaaa gagcgcggca 300
ctggcactgt gaagctgctc aagcacaagg agaccgccaa ggtccgcctc gttatgcgtc 360
aggcgaagac gcttaagatc tgcgctaacc atctcgtggt ggcgacgact aagatgcagg 420
agcacgcggg gagcgacaag tcgtgcgtgt ggcacgcgct ggacttcgcc gacggcgagc 480
tcaaggagga gatgttcgct atccgttttg gatctgtcga gaattgtaag aagtttaagg 540
atattgttga agagatagct gaacagcaag gaaagactga ggagaaagaa aacgaggaag 600
cctccactgc cgctgatttg gtacagaagt taacagtgac ggaggccagc aaggaggaaa 660
ctgcggagaa agaggaggct cctgcatctg gcgataagaa ggatgctaaa gattgaaggc 720
gtttatatac gcaaacaatg gtcaaatgag tgtccttcca ttcattttgt tgtcaaggtt 780
catctaccct ccacaaattt tcatcgtgtt ttctctggga tgaatgcttg tgttaggtga 840
aatcagagtc atcagtcatc tacatggttt tcttggtcat agactgttat ttttaagtcg 900
cctgtttatc tgggatggct ggggtcagca tgtctgtaca attatttgga gttgcttttg 960
gaatggacgc ggtttgatga gtagcctaaa gcttgagatg gtggtgatgt cttttcgctc 1020
cacttttctt attc 1034
<210> 136
<211> 209
<212> PRT
<213> Saccharum officinarum
<400> 136
Met Ala Asp Pro Ala Glu His Arg Pro Ala Glu Glu Glu Glu Glu Ala
1 5 10 15
Ala Ala Ala Gly Glu Asp Glu Asp Thr Gly Ala Gln Val Ala Pro Ile
20 25 30
Val Lys Leu Glu Glu Val Ala Val Thr Thr Gly Glu Glu Asp Glu Asp

```
       35                  40                  45
Val Leu Leu Asp Met Lys Ala Lys Leu Tyr Arg Phe Asp Lys Glu Gly
   50                  55                  60
Asn Gln Trp Lys Glu Arg Gly Thr Gly Thr Val Lys Leu Leu Lys His
65                  70                  75                  80
Lys Glu Thr Ala Lys Val Arg Leu Val Met Arg Gln Ala Lys Thr Leu
                85                  90                  95
Lys Ile Cys Ala Asn His Leu Val Val Ala Thr Thr Lys Met Gln Glu
            100                 105                 110
His Ala Gly Ser Asp Lys Ser Cys Val Trp His Ala Leu Asp Phe Ala
       115                 120                 125
Asp Gly Glu Leu Lys Glu Glu Met Phe Ala Ile Arg Phe Gly Ser Val
   130                 135                 140
Glu Asn Cys Lys Lys Phe Lys Asp Ile Val Glu Glu Ile Ala Glu Gln
145                 150                 155                 160
Gln Gly Lys Thr Glu Glu Lys Glu Asn Glu Glu Ala Ser Thr Ala Ala
                165                 170                 175
Asp Leu Val Gln Lys Leu Thr Val Thr Glu Ala Ser Lys Glu Glu Thr
            180                 185                 190
Ala Glu Lys Glu Glu Ala Pro Ala Ser Gly Asp Lys Lys Asp Ala Lys
       195                 200                 205
Asp
<210> 137
<211> 1039
<212> DNA
<213> Medicago sativa
<400> 137
atgtctacaa gcaccgatca tcacgaagag gaagatgttc ccgccggcga tgaagaagac     60
accggcgctc aaatcgctcc gatcatccaa cttcatgaag tcgctgtttc tactggtgaa    120
gaagatgaag aagctattct cgacctaaaa gcgaaactgt accgatttga caaggtaggg    180
aatcaatgga aggaacgagg ggcaggaact gttaagtttt tgaagcataa ggttactgga    240
aaagttaggc ttgttatgag acaatcaaag actcttaaga tctgtgctaa tcatctcatt    300
ttgcctaaaa tgactgtgca agagcatgct gggaatgaga aatcttgtgt ttggcacgcg    360
aaggactttg ctgatggtga attgaaagac gagtttttct gcattcgatt tgcgtcaatt    420
gaaaattgca gaaagttcat agagacattt caagaaattg ctgaatccct gaacaaagag    480
gaaagcaagg atgcatccac tgctgctgat ctccttgaga atttgagtgt tgaagggaaa    540
gcagatgcag aaaagaaaga taaagagaaa tctgaggaga aaactaagga gaaagagaca    600
ccagaaaaag aaagcaagga agatacagaa aagaaagttg aagagcctgc ttcatctgct    660
tgattatgat atgttcatat tttcgacaag gcaatatgga aaagtttggt ggttgacctt    720
cgtgccactc ttatcaatac tctctcatag tactagtctt caggttgttt tgttgctacg    780
ttattgagtg gttgatatcc ttcgttgaat tattgtcagc aaggttggtt ttttgagtcg    840
ggtttgaatc attgagattg gcgcttttgg tctatgggtc tatgggagtt gttattttcg    900
ttctattcat ttatttagtc gaaatttgaa tgagtcatgt tggtttggtg tcggggatgg    960
ggagggtgca gtcgggaaaa attagtagta agtacaaaca aaggttttga atgcatatta   1020
ttgagtataa cattttata                                               1039
<210> 138
<211> 220
<212> PRT
<213> Medicago sativa
<400> 138
Met Ser Thr Ser Thr Asp His His Glu Glu Glu Asp Val Pro Ala Gly
1               5                   10                  15
Asp Glu Glu Asp Thr Gly Ala Gln Ile Ala Pro Ile Ile Gln Leu His
            20                  25                  30
Glu Val Ala Val Ser Thr Gly Glu Glu Asp Glu Glu Ala Ile Leu Asp
       35                  40                  45
Leu Lys Ala Lys Leu Tyr Arg Phe Asp Lys Val Gly Asn Gln Trp Lys
   50                  55                  60
Glu Arg Gly Ala Gly Thr Val Lys Phe Leu Lys His Lys Val Thr Gly
65                  70                  75                  80
Lys Val Arg Leu Val Met Arg Gln Ser Lys Thr Leu Lys Ile Cys Ala
                85                  90                  95
Asn His Leu Ile Leu Pro Lys Met Thr Val Gln Glu His Ala Gly Asn
```

```
           100                 105                 110
Glu Lys Ser Cys Val Trp His Ala Lys Asp Phe Ala Asp Gly Glu Leu
        115                 120                 125
Lys Asp Glu Phe Phe Cys Ile Arg Phe Ala Ser Ile Glu Asn Cys Arg
    130                 135                 140
Lys Phe Ile Glu Thr Phe Gln Glu Ile Ala Glu Ser Leu Asn Lys Glu
145                 150                 155                 160
Glu Ser Lys Asp Ala Ser Thr Ala Ala Asp Leu Leu Glu Asn Leu Ser
                165                 170                 175
Val Glu Gly Lys Ala Asp Ala Glu Lys Lys Asp Lys Glu Lys Ser Glu
            180                 185                 190
Glu Lys Thr Lys Glu Lys Glu Thr Pro Glu Lys Glu Ser Lys Glu Asp
        195                 200                 205
Thr Glu Lys Lys Val Glu Glu Pro Ala Ser Ser Ala
    210                 215                 220
<210> 139
<211> 30
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 2 de Zea mays
<400> 139
Glu Glu Glu Asp Thr Gly Ala Gln Val Ala Pro Ile Val Arg Leu Glu
1               5                   10                  15
Glu Val Ala Val Thr Thr Gly Glu Glu Asp Glu Asp Val Leu
            20                  25                  30
<210> 140
<211> 10
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 3 de Zea mays
<400> 140
Arg Gln Ala Lys Thr Leu Lys Ile Cys Ala
1               5                   10
<210> 141
<211> 5
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 4 de Zea mays
<400> 141
Asn His Leu Val Ala
1               5
<210> 142
<211> 17
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 5 de Zea mays
<400> 142
Asp Lys Ser Cys Val Trp His Ala Ala Asp Phe Ala Asp Gly Glu Leu
1               5                   10                  15
Lys
<210> 143
<211> 5
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 6 de Zea mays
<400> 143
Glu Ile Ala Glu Ser
1               5
```

<210> 144
<211> 9
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 7 de Zea mays
<400> 144
Leu Ala Lys Leu Thr Val Ser Glu Gly
1 5
<210> 145
<211> 30
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 2 de Arabidopsis thaliana
<400> 145
Glu Asp Glu Asp Thr Gly Ala Gln Val Ala Pro Ile Val Arg Leu Glu
1 5 10 15
Glu Val Ala Val Thr Thr Gly Glu Glu Asp Glu Asp Thr Ile
20 25 30
<210> 146
<211> 10
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 3 de Arabidopsis thaliana
<400> 146
Arg Gln Ser Lys Thr Leu Lys Ile Cys Ala
1 5 10
<210> 147
<211> 5
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 4 de Arabidopsis thaliana
<400> 147
Asn His Leu Val Gly
1 5
<210> 148
<211> 17
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 5 de Arabidopsis thaliana
<400> 148
Asp Lys Ser Cys Val Trp His Ala Arg Asp Phe Ser Asp Gly Glu Leu
1 5 10 15
Lys
<210> 149
<211> 5
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 6 de Arabidopsis thaliana
<400> 149
Glu Val Ala Glu Ser
1 5
<210> 150
<211> 9
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 7 de Arabidopsis thaliana

<400> 150
Leu Glu Lys Leu Thr Val Glu Glu Lys
1 5
<210> 151
<211> 51
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm06703
<400> 151
ggggacaagt ttgtacaaaa aagcaggctt aaacaatggc ggacaaggag c 51
<210> 152
<211> 50
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm06704
<400> 152
ggggaccact ttgtacaaga aagctgggta gtgcaaccac accaactact 50
<210> 153
<211> 52
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm06491
<400> 153
ggggacaagt ttgtacaaaa aagcaggctt cacaatggcg agcattagca ac 52
<210> 154
<211> 49
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm06492
<400> 154
ggggaccact ttgtacaaga aagctgggtg catcttaagc tgagggaac 49
<210> 155
<211> 654
<212> DNA
<213> Oryza sativa
<400> 155
cttctacatc ggcttaggtg tagcaacacg actttattat tattattatt attattatta 60
ttattttaca aaaatataaa atagatcagt ccctcaccac aagtagagca agttggtgag 120
ttattgtaaa gttctacaaa gctaatttaa aagttattgc attaacttat ttcatattac 180
aaacaagagt gtcaatggaa caatgaaaac catatgacat actataattt tgtttttatt 240
attgaaatta tataattcaa agagaataaa tccacatagc cgtaaagttc tacatgtggt 300
gcattaccaa aatatatata gcttacaaaa catgacaagc ttagtttgaa aaattgcaat 360
ccttatcaca ttgacacata aagtgagtga tgagtcataa tattattttc tttgctaccc 420
atcatgtata tatgatagcc acaaagttac tttgatgatg atatcaaaga acatttttag 480
gtgcacctaa cagaatatcc aaataatatg actcacttag atcataatag agcatcaagt 540
aaaactaaca ctctaaagca accgatggga aagcatctat aaatagacaa gcacaatgaa 600
aatcctcatc atccttcacc acaattcaaa tattatagtt gaagcatagt agta 654
<210> 156
<211> 1394
<212> DNA
<213> Oryza sativa
<400> 156
atgcttgccg tgtcgccggc gatgtgcccc gacattgagg accgcgccgc ggtggccggc 60
gatgctggca tggaggtcgt cgggatgtcg tcggacgaca tggatcagtt cgacttctcc 120
gtcgatgaca tagacttcgg ggacttcttc ctgaggctgg aggacggtga tgtgctcccg 180
gacctcgagg tcgacccggc cgagatcttc accgacttcg aggcaattgc gacgagtgga 240
ggcgaaggtg tgcaggacca ggaggtgccc accgtcgagc tcttggcgcc tgcggacgac 300
gtcggtgtgc tggatccgtg cggcgatgtc gtcgtcgggg aggagaacgc ggcgtttgcc 360
ggggctggag aggagaaggg agggtgtaac caggacgatg atgcgggggа agcgaatgtc 420

```
gacgatggag ccgcggcggt tgaggccaag tcttcgtcgc cgtcatcgac gacgtcgtcg    480
tcgcaggagg ctgagagccg gcacaagtca tccagcaaga gctcccatgg gaagaagaaa    540
gcgaaggtgg actggacgcc tgagcttcac cggaggttcg tgcaggcggt ggagcagctc    600
ggcatcgaca aggccgtgcc gtcgaggata cttgagatca tggggatcga ctctctcacc    660
cggcacaaca tagccagcca tcttcagaag taccggtcac acagaaaaca catgattgcg    720
agagaggcgg aggcagcgag ttggacccaa cggcggcaga tttacgccgc cggtggaggt    780
gctgtttgcg aagaggccgg agtccaacgc gtggaccgtg ccaaccattg gcttccctcc    840
tcctccgcca ccaccaccat caccggctcc gattcaacat tttgctcgcc cgttgcatgt    900
tggggccacc cgacgatgga cccgtcccga gttccagtgt ggccaccgcg gcacctcgtt    960
ccccgtggcc cggcgccacc atgggttcca ccgccgccgc cgtcggaccc tgctttctgg   1020
caccaccctt acatgagggg gccagcacat gtgccaactc aagggacacc ttgcatggcg   1080
atgcccatgc cagctgcgag atttcctgct ccaccggtgc caggagttgt cccgtgtcca   1140
atgtataggc cattgactcc accagcactg gcgagcaaga atcagcagga cgcacagctt   1200
caactccagg ttcaaccatc aagcgagagc atcgacgcag ctatcggtga tgttttatcg   1260
aaaccgtggt tgcctttgcc tcttggactg aagccacctt cagtggacag tgtgatgggc   1320
gagctgcaga ggcaaggcgt agcaaacgtt cctccagcgt gtggatgata ttacccagct   1380
ttcttgtaca aagt                                                     1394
<210> 157
<211> 455
<212> PRT
<213> Oryza sativa
<400> 157
Met Leu Ala Val Ser Pro Ala Met Cys Pro Asp Ile Glu Asp Arg Ala
1               5                  10                  15
Ala Val Ala Gly Asp Ala Gly Met Glu Val Val Gly Met Ser Ser Asp
            20                  25                  30
Asp Met Asp Gln Phe Asp Phe Ser Val Asp Asp Ile Asp Phe Gly Asp
        35                  40                  45
Phe Phe Leu Arg Leu Glu Asp Gly Asp Val Leu Pro Asp Leu Glu Val
    50                  55                  60
Asp Pro Ala Glu Ile Phe Thr Asp Phe Glu Ala Ile Ala Thr Ser Gly
65                  70                  75                  80
Gly Glu Gly Val Gln Asp Gln Glu Val Pro Thr Val Glu Leu Leu Ala
                85                  90                  95
Pro Ala Asp Asp Val Gly Val Leu Asp Pro Cys Gly Asp Val Val Val
            100                 105                 110
Gly Lys Glu Asn Ala Ala Phe Ala Gly Ala Gly Glu Glu Lys Gly Gly
        115                 120                 125
Cys Asn Gln Asp Asp Asp Ala Gly Glu Ala Asn Ala Asp Asp Gly Ala
    130                 135                 140
Ala Ala Val Glu Ala Lys Ser Ser Ser Pro Ser Ser Ser Thr Ser Ser
145                 150                 155                 160
Ser Gln Glu Ala Glu Ser Arg His Lys Ser Ser Ser Lys Ser Ser His
                165                 170                 175
Gly Lys Lys Lys Ala Lys Val Asp Trp Thr Pro Glu Leu His Arg Arg
            180                 185                 190
Phe Val Gln Ala Val Glu Gln Leu Gly Ile Asp Lys Ala Val Pro Ser
        195                 200                 205
Arg Ile Leu Glu Ile Met Gly Ile Asp Ser Leu Thr Arg His Asn Ile
    210                 215                 220
Ala Ser His Leu Gln Lys Tyr Arg Ser His Arg Lys His Met Ile Ala
225                 230                 235                 240
Arg Glu Ala Glu Ala Ala Ser Trp Thr Gln Arg Arg Gln Ile Tyr Ala
                245                 250                 255
Ala Gly Gly Gly Ala Val Ala Lys Arg Pro Glu Ser Asn Ala Trp Thr
            260                 265                 270
Val Pro Thr Ile Gly Phe Pro Pro Pro Pro Pro Pro Pro Pro Ser Pro
        275                 280                 285
Ala Pro Met Gln His Phe Ala Arg Pro Leu His Val Trp Gly His Pro
    290                 295                 300
Thr Met Asp Pro Ser Arg Val Pro Val Trp Pro Pro Arg His Leu Val
305                 310                 315                 320
Pro Arg Gly Pro Ala Pro Pro Trp Val Pro Pro Pro Pro Pro Ser Asp
```

```
                325                 330                 335
Pro Ala Phe Trp His His Pro Tyr Met Arg Gly Pro Ala His Val Pro
            340                 345                 350
Thr Gln Gly Thr Pro Cys Met Ala Met Pro Met Pro Ala Ala Arg Phe
        355                 360                 365
Pro Ala Pro Pro Val Pro Gly Val Val Pro Cys Pro Met Tyr Arg Pro
    370                 375                 380
Leu Thr Pro Pro Ala Leu Thr Ser Lys Asn Gln Gln Asp Ala Gln Leu
385                 390                 395                 400
Gln Leu Gln Val Gln Pro Ser Ser Glu Ser Ile Asp Ala Ala Ile Gly
                405                 410                 415
Asp Val Leu Ser Lys Pro Trp Leu Pro Leu Pro Leu Gly Leu Lys Pro
            420                 425                 430
Pro Ser Val Asp Ser Val Met Gly Glu Leu Gln Arg Gln Gly Val Ala
        435                 440                 445
Asn Val Pro Pro Ala Cys Gly
    450                 455
<210> 158
<211> 50
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm02251
<400> 158
ggggacaagt ttgtacaaaa aagcaggctt cacaatgctt gccgtgtcgc                50
<210> 159
<211> 49
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm02252
<400> 159
ggggaccact ttgtacaaga aagctgggta atatcatcca cacgctgga                49
<210> 160
<211> 2193
<212> DNA
<213> Oryza sativa
<400> 160
aatccgaaaa gtttctgcac cgttttcacc ccctaactaa caatataggg aacgtgtgct     60
aaatataaaa tgagacctta tatatgtagc gctgataact agaactatgc aagaaaaact    120
catccaccta ctttagtggc aatcgggcta ataaaaaaag agtcgctaca ctagtttcgt    180
tttccttagt aattaagtgg gaaaatgaaa tcattattgc ttagaatata cgttcacatc    240
tctgtcatga agttaaatta ttcgaggtag ccataattgt catcaaactc ttcttgaata    300
aaaaaatctt tctagctgaa ctcaatgggt aaagagagag attttttta aaaaaataga     360
atgaagatat tctgaacgta ttggcaaaga tttaaacata taattatata attttatagt    420
ttgtgcattc gtcatatcgc acatcattaa ggacatgtct tactccatcc caatttttat    480
ttagtaatta aagacaattg acttatttt attatttatc ttttttcgat tagatgcaag     540
gtacttacgc acacactttg tgctcatgtg catgtgtgag tgcacctcct caatacacgt    600
tcaactagca acacatctct aatatcactc gcctatttaa tacatttagg tagcaatatc    660
tgaattcaag cactccacca tcaccagacc acttttaata atatctaaaa tacaaaaaat    720
aattttacag aatagcatga aaagtatgaa acgaactatt taggtttttc acatacaaaa    780
aaaaaaagaa ttttgctcgt gcgcgagcgc caatctccca tattgggcac acaggcaaca    840
acagagtggc tgcccacaga acaacccaca aaaaacgatg atctaacgga ggacagcaag    900
tccgcaacaa ccttttaaca gcaggctttg cggccaggag agaggaggag aggcaaagaa    960
aaccaagcat cctcctcctc ccatctataa attcctcccc ccttttcccc tctctatata   1020
ggaggcatcc aagccaagaa gagggagagc accaaggaca cgcgactagc agaagccgag   1080
cgaccgcctt cttcgatcca tatcttccgg tcgagttctt ggtcgatctc ttccctcctc   1140
cacctcctcc tcacagggta tgtgcccttc ggttgttctt ggatttattg ttctaggttg   1200
tgtagtacgg gcgttgatgt taggaaaggg gatctgtatc tgtgatgatt cctgttcttg   1260
gatttgggat agaggggttc ttgatgttgc atgttatcgg ttcggtttga ttagtagtat   1320
ggttttcaat cgtctggaga gctctatgga aatgaaatgg tttagggtac ggaatcttgc   1380
gattttgtga gtacctttg tttgaggtaa aatcagagca ccggtgattt tgcttggtgt    1440
aataaaagta cggttgtttg gtcctcgatt ctggtagtga tgcttctcga tttgacgaag   1500
```

ctatcctttg tttattccct attgaacaaa aataatccaa ctttgaagac ggtcccgttg 1560
atgagattga atgattgatt cttaagcctg tccaaaattt cgcagctggc ttgtttagat 1620
acagtagtcc ccatcacgaa attcatggaa acagttataa tcctcaggaa caggggattc 1680
cctgttcttc cgatttgctt tagtcccaga attttttttc ccaaatatct taaaaagtca 1740
ctttctggtt cagttcaatg aattgattgc tacaaataat gcttttatag cgttatccta 1800
gctgtagttc agttaatagg taataccccc atagtttagt caggagaaga acttatccga 1860
tttctgatct ccatttttaa ttatatgaaa tgaactgtag cataagcagt attcatttgg 1920
attatttttt ttattagctc tcacccccttc attattctga gctgaaagtc tggcatgaac 1980
tgtcctcaat tttgttttca aattcacatc gattatctat gcattatcct cttgtatcta 2040
cctgtagaag tttctttttg gttattcctt gactgcttga ttacagaaag aaatttatga 2100
agctgtaatc gggatagtta tactgcttgt tcttatgatt catttccttt gtgcagttct 2160
tggtgtagct tgccactttc accagcaaag ttc 2193

<210> 161
<211> 21
<212> PRT
<213> Seqüência artificial
<220>
<223> Seqüência de consenso GARP 1
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir = "Arg"
<220>
<221> VARIANTE
<222> (2)..(2)
<223> /substituir = "Met" /substituir = "Val" /substituir = "Ala"
<220>
<221> VARIANTE
<222> (3)..(3)
<223> /substituir = "Lys" /substituir = "Met"
<220>
<221> VARIANTE
<222> (4)..(4)
<223> /substituir = "Leu"
<220>
<221> VARIANTE
<222> (5)..(5)
<223> /substituir = "Asp"
<220>
<221> VARIANTE
<222> (7)..(7)
<223> /substituir = "Thr" /substituir = "Ile" /substituir = "Asn"
<220>
<221> VARIANTE
<222> (8)..(8)
<223> /substituir = "Ala" /substituir = "Pro" /substituir = "Ser" /substituir = "Thr" /substituir = "Cys" /substituir = "His" /substituir = "Gln" /substituir = "Asp"
<220>
<221> VARIANTE
<222> (9)..(9)
<223> /substituir = "Gln" /substituir = "Thr" /substituir = "Asp" /substituir = "Ser"
<220>
<221> VARIANTE
<222> (11)..(11)
<223> /substituir = "Asp"
<220>
<221> VARIANTE
<222> (12)..(12)
<223> /substituir = "Lys" /substituir = "Gln" /substituir = "Ala" /substituir = "His" /substituir = "Asp" /substituir = "Glu" /substituir = "Leu" /substituir = "Ile"

<220>
<221> VARIANTE
<222> (13)..(13)
<223> /substituir = "Arg" /substituir = "Gln" /substituir = "Ser" /substituir = "Cys" /substituir = "Val" /substituir = "Ala" /substituir = "His"
<220>
<221> VARIANTE
<222> (15)..(15)
<223> /substituir = "Leu" /substituir = "Ile"
<220>
<221> VARIANTE
<222> (16)..(16)
<223> /substituir = "Lys" /substituir = "Gln" /substituir = "Asp" /substituir = "His" /substituir = "Asp" /substituir = "Asn" /substituir = "Arg" /substituir = "Ser"
<220>
<221> VARIANTE
<222> (17)..(17)
<223> /substituir = "Val" /substituir = "Cys"
<220>
<221> VARIANTE
<222> (18)..(18)
<223> /substituir = "Gly" /substituir = "Leu" /substituir = "Ile"
<220>
<221> VARIANTE
<222> (19)..(19)
<223> /substituir = "Glu" /substituir = "Ala" /substituir = "Gln" /substituir = "Asp" /substituir = "Gly" /substituir = "Thr" /substituir = "Ile" /substituir = "Lys" /substituir = "His"
<220>
<221> VARIANTE
<222> (20)..(20)
<223> /substituir = "Glu" /substituir = "His" /substituir = "Leu" /substituir = "Ile" /substituir = "Met" /substituir = "Lys" /substituir = "Arg" /substituir = "Ser"
<400> 161
Lys Pro Arg Val Val Trp Ser Val Glu Leu His Arg Lys Phe Val Ala
1 5 10 15
Ala Val Asn Gln Leu
20
<210> 162
<211> 3
<212> PRT
<213> Seqüência artificial
<220>
<223> Seqüência de consenso GARP 2
<220>
<221> VARIANTE
<222> (2)..(2)
<223> /substituir = "Val" /substituir = "Leu" /substituir = "Pro" /substituir = "Ser" /substituir = "His" /substituir = "Gln" /substituir = "Ala" /substituir = "Gly"
<220>
<221> VARIANTE
<222> (3)..(3)
<223> /substituir = "Glu" /substituir = "Lys" /substituir = "His" /substituir = "Gln" /substituir = "Asn" /substituir = "Ala"
<400> 162
Gly Ile Asp
1
<210> 163
<211> 11

<212> PRT
<213> Seqüência artificial
<220>
<223> Seqüência de consenso GARP 3
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir = "Thr"
<220>
<221> VARIANTE
<222> (2)..(2)
<223> /substituir = "Ile" /substituir = "Tyr" /substituir = "Phe" /substituir = "Thr"
<220>
<221> VARIANTE
<222> (4)..(4)
<223> /substituir = "Ser"
<220>
<221> VARIANTE
<222> (5)..(5)
<223> /substituir = "Arg" /substituir = "Thr" /substituir = "Gln" /substituir = "Leu" /substituir = "Ser" /substituir = "Gly" /substituir = "Ala"
<220>
<221> VARIANTE
<222> (6)..(6)
<223> /substituir = "Val" /substituir = "Leu"
<220>
<221> VARIANTE
<222> (7)..(7)
<223> /substituir = "Met" /substituir = "Arg" /substituir = "Lys"
<220>
<221> VARIANTE
<222> (8)..(8)
<223> /substituir = "Glu" /substituir = " Gln" /substituir = "Lys" /substituir = "Arg" /substituir = "Ser"
<220>
<221> VARIANTE
<222> (9)..(9)
<223> /substituir = "Ile" /substituir = "Phe" /substituir = "His" /substituir = "Val" /substituir = "Met" /substituir = "Thr" /substituir = "Arg" /substituir = "Ala"
<220>
<221> VARIANTE
<222> (10)..(10)
<223> /substituir = "Ile" /substituir = "Leu"
<220>
<221> VARIANTE
<222> (11)..(11)
<223> /substituir = "Lys" /substituir = "Gly" /substituir = "Ser" /substituir = "Asp" /substituir = "Gln" /substituir = "Glu"
<400> 163
Ala Val Pro Lys Lys Ile Leu Asp Leu Met Asn
1 5 10
<210> 164
<211> 8
<212> PRT
<213> Seqüência artificial
<220>
<223> Seqüência de consenso GARP 4
<220>
<221> VARIANTE
<222> (1)..(1)

<223> /substituir = "Ile" /substituir = "Met" /substituir = "Thr" /substituir = "Glu" /substituir = "Leu" /substituir = "Ser" /substituir = "Asn"
<220>
<221> VARIANTE
<222> (2)..(2)
<223> /substituir = "Asp" /substituir = "Gly" /substituir = "Asn" /substituir = "Tyr" /substituir = "Lys" /substituir = "His" /substituir = "Gln" /substituir = "Pro"
<220>
<221> VARIANTE
<222> (3)..(3)
<223> /substituir = "Gly" /substituir = "Lys" /substituir = "Thr" /substituir = "Ser" /substituir = "Cys" /substituir = "Arg" /substituir = "Asp"
<220>
<221> VARIANTE
<222> (4)..(4)
<223> /substituir = "Leu"
<220>
<221> VARIANTE
<222> (5)..(5)
<223> /substituir = "Asp" /substituir = "Ala" /substituir = "Ser"
<220>
<221> VARIANTE
<222> (6)..(6)
<223> /substituir = "Asn" /substituir = "Ile" /substituir = "Leu" /substituir = "Val"
<220>
<221> VARIANTE
<222> (7)..(7)
<223> /substituir = "His" /substituir = "Asp" /substituir = "Ser" /substituir = "Ala" /substituir = "Tyr" /substituir = "Phe"
<220>
<221> VARIANTE
<222> (8)..(8)
<223> /substituir = "Glu" /substituir = "His"
<400> 164
Val Glu Asn Ile Thr Arg Glu Asn
1 5
<210> 165
<211> 9
<212> PRT
<213> Seqüência artificial
<220>
<223> Seqüência de consenso GARP 5
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir = "Ile" /substituir = "Leu"
<220>
<221> VARIANTE
<222> (2)..(2)
<223> /substituir = "Lys"
<220>
<221> VARIANTE
<222> (7)..(7)
<223> /substituir = "Met" /substituir = "Ile"
<220>
<221> VARIANTE
<222> (8)..(8)
<223> /substituir = "Phe"
<220>

<221> VARIANTE
<222> (9)..(9)
<223> /substituir = "Val"
<400> 165
Val Ala Ser His Leu Gln Lys Tyr Arg
1 5
<210> 166
<211> 16
<212> PRT
<213> Seqüência artificial
<220>
<223> Seqüência de consenso 6
<220>
<221> VARIANTE
<222> (4)..(4)
<223> /substituir = "Arg"
<220>
<221> VARIANTE
<222> (6)..(6)
<223> /substituir = "Met"
<220>
<221> VARIANTE
<222> (7)..(7)
<223> /substituir = "Ala" /substituir = "Met" /substituir = "Ile"
<220>
<221> VARIANTE
<222> (11)..(11)
<223> /substituir = "Gly" /substituir = "Val"
<220>
<221> VARIANTE
<222> (14)..(14)
<223> /substituir = "Gly"
<220>
<221> VARIANTE
<222> (15)..(15)
<223> /substituir = "Asn" /substituir = "Thr"
<400> 166
Ser His Arg Lys His Leu Leu Ala Arg Glu Ala Glu Ala Ala Ser Trp
1 5 10 15
<210> 167
<211> 48
<212> PRT
<213> Seqüência artificial
<220>
<223> Seqüência de consenso do domínio GCT
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir = "Gln"
<220>
<221> VARIANTE
<222> (2)..(2)
<223> /substituir = "Leu"
<220>
<221> VARIANTE
<222> (4)..(4)
<223> /substituir = "Lys" /substituir = "Ser"
<220>
<221> VARIANTE
<222> (6)..(6)
<223> /substituir = "Val"
<220>
<221> VARIANTE

<222> (7)..(7)
<223> /substituir = "Val" /substituir = "Leu"
<220>
<221> VARIANTE
<222> (13)..(13)
<223> /substituir = "Glu"
<220>
<221> VARIANTE
<222> (14)..(14)
<223> /substituir = "Ala"
<220>
<221> VARIANTE
<222> (15)..(15)
<223> /substituir = "Leu"
<220>
<221> VARIANTE
<222> (16)..(16)
<223> /substituir = "Thr" /substituir = "Ala" /substituir = "Val"
<220>
<221> VARIANTE
<222> (17)..(17)
<223> /substituir = "Lys" /substituir = "Arg"
<220>
<221> VARIANTE
<222> (20)..(20)
<223> /substituir = "Thr"
<220>
<221> VARIANTE
<222> (22)..(22)
<223> /substituir = "Pro"
<220>
<221> VARIANTE
<222> (27)..(27)
<223> /substituir = "Asn"
<220>
<221> VARIANTE
<222> (30)..(30)
<223> /substituir = "Ala"
<220>
<221> VARIANTE
<222> (31)..(31)
<223> /substituir = "Met" /substituir = "Leu"
<220>
<221> VARIANTE
<222> (32)..(32)
<223> /substituir = "Glu" /substituir = "Gly"
<220>
<221> VARIANTE
<222> (33)..(33)
<223> /substituir = "Ser"
<220>
<221> VARIANTE
<222> (35)..(35)
<223> /substituir = "Ile"
<220>
<221> VARIANTE
<222> (36)..(36)
<223> /substituir = "Ala" /substituir = "Ser" /substituir = "Gly"
<220>
<221> VARIANTE
<222> (39)..(39)
<223> /substituir = "His" /substituir = "Glu"
<220>

<221> VARIANTE
<222> (40)..(40)
<223> /substituir = "Lys"
<220>
<221> VARIANTE
<222> (41)..(41)
<223> /substituir = "His"
<220>
<221> VARIANTE
<222> (43)..(43)
<223> /substituir = "Ile"
<220>
<221> VARIANTE
<222> (44)..(44)
<223> /substituir = "Asn" /substituir = "Pro" /substituir = "Ala"
<220>
<221> VARIANTE
<222> (45)..(45)
<223> /substituir = "Glu" /substituir = "Thr" /substituir = "Lys"
<220>
<221> VARIANTE
<222> (46)..(46)
<223> /substituir = "Ile"
<220>
<221> VARIANTE
<222> (48)..(48)
<223> /substituir = "Gln"
<400> 167

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

<210> 168
<211> 1542
<212> DNA
<213> Oryza sativa
<400> 168

```
atgcttgagg tgtccacgct gcgaagccct aaggcggatc agcgggcggg cgtcggcggc     60
caccatgtcg tcggcttcgt cccggcgccg ccgtcgccgg ccgacgtcgc cgacgaggtc    120
gacgcgttca tcgtcgacga cagctgcctg ctcgagtaca tcgacttcag ctgctgcgac    180
gtgccgttct tccacgccga cgacggcgac atcctcccgg acctcgaggt cgaccccacg    240
gagctcctcg ccgagttcgc cagctccccg gacgacgagc cgccgccgac gacgtcggct    300
ccgggccccg gcgagccagc tgctgctgca ggagccaagg aagacgtgaa ggaagatgga    360
gccgccgccg ccgccgccgc cgccgccgct gactacgacg ggtcgccgcc gccaccgcgg    420
gggaagaaga agaaggacga cgaggaaagg tcgtcgtcgt tgccggagga gaaagacgcg    480
aagaacggcg gcggcgacga ggtcctgagc gcggtgacga cggaggattc ctcggccggt    540
gccgccaagt cgtgctcgcc gtcggcagag ggccacagca agaggaagcc gtcgtcgtcg    600
tcgtcatcgg cggcggccgg caagaactct cacggcaagc gcaaggtgaa ggtggactgg    660
acgccggagt tgcaccggcg gttcgtgcag gcggtggagc agctcgggat agacaaggcc    720
gtgccgtcca ggatcctgga gctcatgggc atcgagtgcc tcactcgcca caacatcgcc    780
agccatctcc agaaatatcg gtcgcacagg aaacatctga tggcgaggga ggcggaggcg    840
gcgagctgga cgcagaagcg gcagatgtac accgccgccg ccgccgccgc cgcggtggca    900
gccggcggcg ggccaaggaa ggacgccgcc gccgccactg cggcggtggc cccgtgggtc    960
atgccgacca tcggtttccc tccgccgcac gcggcggcga tggtgcctcc cccgccgcac   1020
cctccaccgt tctgccggcc gccgctgcac gtgtggggcc acccgaccgc cggcgtcgag   1080
ccgaccaccg cggcggcgcc accaccaccc tcgccgcacg cgcagccgcc gttgctgccc   1140
gtctggccgc gccacctggc gccgccgccg ccgccgctgc cggcggcgtg ggcgcacggc   1200
caccagccgg cgccggtgga cccggcggcg tactggcagc aacagtataa cttgcagagg   1260
tttcctgtac cgccggtgcc tggaatggtg ccgcacccca tgtacagacc gataccgccg   1320
ccgtcaccgc cgcaggggaa taaactcgct gccttgcagc ttcagcttga tgcccacccg   1380
tctaaggaga gcatagacgc agccatcgga gatgttttag tgaagccatg gctgccgctt   1440
```

```
cccctcggcc tcaagccacc gtcgctggac agcgtcatgt ctgagctgca caagcagggc   1500
atccccaagg tgccaccggc ggcgagcggt gccgccggct ga                      1542
<210> 169
<211> 513
<212> PRT
<213> Oryza sativa
<400> 169
Met Leu Glu Val Ser Thr Leu Arg Ser Pro Lys Ala Asp Gln Arg Ala
1               5                  10                  15
Gly Val Gly Gly His His Val Val Gly Phe Val Pro Ala Pro Pro Ser
            20                  25                  30
Pro Ala Asp Val Ala Asp Glu Val Asp Ala Phe Ile Val Asp Asp Ser
        35                  40                  45
Cys Leu Leu Glu Tyr Ile Asp Phe Ser Cys Cys Asp Val Pro Phe Phe
    50                  55                  60
His Ala Asp Asp Gly Asp Ile Leu Pro Asp Leu Glu Val Asp Pro Thr
65                  70                  75                  80
Glu Leu Leu Ala Glu Phe Ala Ser Ser Pro Asp Asp Glu Pro Pro Pro
                85                  90                  95
Thr Thr Ser Ala Pro Gly Pro Gly Glu Pro Ala Ala Ala Ala Gly Ala
            100                 105                 110
Lys Glu Asp Val Lys Glu Asp Gly Ala Ala Ala Ala Ala Ala Ala Ala
        115                 120                 125
Ala Ala Asp Tyr Asp Gly Ser Pro Pro Pro Pro Arg Gly Lys Lys Lys
    130                 135                 140
Lys Asp Asp Glu Glu Arg Ser Ser Ser Leu Pro Glu Glu Lys Asp Ala
145                 150                 155                 160
Lys Asn Gly Gly Gly Asp Glu Val Leu Ser Ala Val Thr Thr Glu Asp
                165                 170                 175
Ser Ser Ala Gly Ala Ala Lys Ser Cys Ser Pro Ser Ala Glu Gly His
            180                 185                 190
Ser Lys Arg Lys Pro Ser Ser Ser Ser Ser Ser Ala Ala Ala Gly Lys
        195                 200                 205
Asn Ser His Gly Lys Arg Lys Val Lys Val Asp Trp Thr Pro Glu Leu
    210                 215                 220
His Arg Arg Phe Val Gln Ala Val Glu Gln Leu Gly Ile Asp Lys Ala
225                 230                 235                 240
Val Pro Ser Arg Ile Leu Glu Leu Met Gly Ile Glu Cys Leu Thr Arg
                245                 250                 255
His Asn Ile Ala Ser His Leu Gln Lys Tyr Arg Ser His Arg Lys His
            260                 265                 270
Leu Met Ala Arg Glu Ala Glu Ala Ala Ser Trp Thr Gln Lys Arg Gln
        275                 280                 285
Met Tyr Thr Ala Ala Ala Ala Ala Ala Ala Val Ala Ala Gly Gly Gly
    290                 295                 300
Pro Arg Lys Asp Ala Ala Ala Ala Thr Ala Ala Val Ala Pro Trp Val
305                 310                 315                 320
Met Pro Thr Ile Gly Phe Pro Pro Pro His Ala Ala Ala Met Val Pro
                325                 330                 335
Pro Pro Pro His Pro Pro Pro Phe Cys Arg Pro Pro Leu His Val Trp
            340                 345                 350
Gly His Pro Thr Ala Gly Val Glu Pro Thr Thr Ala Ala Ala Pro Pro
        355                 360                 365
Pro Pro Ser Pro His Ala Gln Pro Pro Leu Leu Pro Val Trp Pro Arg
    370                 375                 380
His Leu Ala Pro Pro Pro Pro Pro Leu Pro Ala Ala Trp Ala His Gly
385                 390                 395                 400
His Gln Pro Ala Pro Val Asp Pro Ala Ala Tyr Trp Gln Gln Gln Tyr
                405                 410                 415
Asn Leu Gln Arg Phe Pro Val Pro Pro Val Pro Gly Met Val Pro His
            420                 425                 430
Pro Met Tyr Arg Pro Ile Pro Pro Pro Ser Pro Pro Gln Gly Asn Lys
        435                 440                 445
```

Leu Ala Ala Leu Gln Leu Gln Leu Asp Ala His Pro Ser Lys Glu Ser
450 455 460
Ile Asp Ala Ala Ile Gly Asp Val Leu Val Lys Pro Trp Leu Pro Leu
465 470 475 480
Pro Leu Gly Leu Lys Pro Pro Ser Leu Asp Ser Val Met Ser Glu Leu
485 490 495
His Lys Gln Gly Ile Pro Lys Val Pro Pro Ala Ala Ser Gly Ala Ala
500 505 510
Gly

<210> 170
<211> 1263
<212> DNA
<213> Arabidopsis thaliana
<400> 170

atgttagctc tgtctccggc gacaagagat ggttgcgacg gagcgtcaga gtttcttgat 60
acgtcgtgtg gattcacgat tataaacccg gaggaggagg aggagtttcc ggatttcgct 120
gaccacggtg atcttcttga catcattgac ttcgacgata tattcggtgt ggccggagat 180
gtgcttcctg acttggagat cgaccctgag atcttatccg gggatttctc caatcacatg 240
aacgcttctt caacgattac tacgacgtcg gataagactg atagtcaagg ggagactact 300
aagggtagtt cggggaaagg tgaagaagtc gtaagcaaaa gagacgatgt tgcggcggag 360
acggtgactt atgacggtga cagtgaccgg aaaaggaagt attcctcttc agcttcttcc 420
aagaacaatc ggatcagtaa caacgaaggg aagagaaagg tgaaggtgga ttggacacca 480
gagctacaca ggagattcgt ggaggcagtg gaacagttag gagtggacaa agctgttcct 540
tctcgaattc tggagcttat gggagtccat tgtctcactc gtcacaacgt tgctagtcac 600
ctccaaaaat ataggtctca tcggaaacat ttgctagctc gtgaggccga agcggctaat 660
tggacacgca aaaggcatat ctatggagta gacaccggtg ctaatcttaa tggtcggact 720
aaaaatggat ggcttgcacc ggcacccact ctcgggtttc caccaccacc acccgtggct 780
gttgcaccgc cacctgtcca ccaccatcat tttaggcccc tgcatgtgtg gggacatccc 840
acggttgatc agtccattat gccgcatgtg tggcccaaac acttacctcc gccttctacc 900
gccatgccta atccgccgtt ttgggtctcc gattctccct attggcatcc aatgcataac 960
gggacgactc cgtatttacc gaccgtagct acgagattta gagcaccgcc agttgccgga 1020
atcccgcatg ctctgccgcc gcatcacacg atgtacaaac caaatcttgg atttggtggt 1080
gctcgtcctc cggtagactt acatccgtca aaagagagcg tggatgcagc cataggagat 1140
gtattgacga ggccatggct gccacttccg ttgggattaa atccgccggc tgttgacggt 1200
gttatgacag agcttcaccg tcacggtgtc tctgaggttc ctccgaccgc gtcttgtgcc 1260
tga 1263

<210> 171
<211> 420
<212> PRT
<213> Arabidopsis thaliana
<400> 171

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

```
Lys Ala Val Pro Ser Arg Ile Leu Glu Leu Met Gly Val His Cys Leu
            180                 185                 190
Thr Arg His Asn Val Ala Ser His Leu Gln Lys Tyr Arg Ser His Arg
        195                 200                 205
Lys His Leu Leu Ala Arg Glu Ala Glu Ala Ala Asn Trp Thr Arg Lys
    210                 215                 220
Arg His Ile Tyr Gly Val Asp Thr Gly Ala Asn Leu Asn Gly Arg Thr
225                 230                 235                 240
Lys Asn Gly Trp Leu Ala Pro Ala Pro Thr Leu Gly Phe Pro Pro Pro
                245                 250                 255
Pro Pro Val Ala Val Ala Pro Pro Pro Val His His His His Phe Arg
            260                 265                 270
Pro Leu His Val Trp Gly His Pro Thr Val Asp Gln Ser Ile Met Pro
        275                 280                 285
His Val Trp Pro Lys His Leu Pro Pro Pro Ser Thr Ala Met Pro Asn
    290                 295                 300
Pro Pro Phe Trp Val Ser Asp Ser Pro Tyr Trp His Pro Met His Asn
305                 310                 315                 320
Gly Thr Thr Pro Tyr Leu Pro Thr Val Ala Thr Arg Phe Arg Ala Pro
                325                 330                 335
Pro Val Ala Gly Ile Pro His Ala Leu Pro Pro His His Thr Met Tyr
            340                 345                 350
Lys Pro Asn Leu Gly Phe Gly Gly Ala Arg Pro Pro Val Asp Leu His
        355                 360                 365
Pro Ser Lys Glu Ser Val Asp Ala Ala Ile Gly Asp Val Leu Thr Arg
    370                 375                 380
Pro Trp Leu Pro Leu Pro Leu Gly Leu Asn Pro Pro Ala Val Asp Gly
385                 390                 395                 400
Val Met Thr Glu Leu His Arg His Gly Val Ser Glu Val Pro Pro Thr
                405                 410                 415
Ala Ser Cys Ala
            420
<210> 172
<211> 1403
<212> DNA
<213> Arabidopsis thaliana
<400> 172
gtatcattct tgttttctct aaatttttca aaaaaccttt agaattttca ttttttttacg     60
attccgatgt taactgtttc tccggctcca gtactcatcg gaaacaactc aaaggatact    120
tacatggcgg cagatttcgc agattttacg acggaagact tgccggactt tacgacggtc    180
ggggattttt ccgatgatct tcttgatgga atcgattact acgacgatct tttcattggt    240
ttcgatggag acgatgtttt gccggatttg gagatagatt cggagattct tggggaatat    300
tccggtagcg gaagagatga ggaacaagaa atggagggta acacttcgac ggcatcggag    360
acatcggaga gagacgttgg tgtgtgtaag caagagggtg gtggtggtgg tgacggtggt    420
tttagggaca aaacggtgcg tcgaggcaaa cgtaaaggga agaaaagtaa agattgttta    480
tccgatgaga acgatattaa gaaaaaacct aaggtggatt ggacgccgga gttacaccgg    540
aaatttgtac aagcggtgga gcaattaggg gtagacaagg cggtgccgtc tcgaatcttg    600
gaaattatga acgttaaatc tctcactcgt cacaacgttg ctagccatct tcagaaatat    660
aggtcacatc ggaaacatct actagcgcgt gaagcagaag ctgccagctg gaatctccgg    720
agacatgcca cggtggcagt gcccggagta ggaggaggag ggaagaagcc gtggacagct    780
cctgccttag gctatcctcc acacgtggca ccaatgcatc atggtcactt caggcctttg    840
cacgtatggg gtcatcctac gtggccaaaa cacaagccta atactccggc gtctgctcat    900
cggacgtatc caatgccggc cattgcggcg gctccggcat cttggccagg tcatccaccg    960
tactggcatc agcaaccact ctatccacag ggatatggta tggcatcatc gaatcattca   1020
agcatcggtg ttcccacaag acaattagga cccactaatc ctcccatcga cattcatccc   1080
tcgaatgaga gcatagatgc agctattggg gacgttatat caaagccgtg gctgccgctt   1140
cctttgggac tgaaaccgcc gtcggttgac ggtgttatga cggagttaca acgtcaagga   1200
gtttctaatg ttcctcctct tccttgagaa agatctccaa aatttgtcga aatctcaaac   1260
ttttaacttc attttttttgg tatcttctat gtatttttgc aagggaaaga cgataaatcc   1320
ttgttgcttg atcatatgta tttctctata tgagtgcatg tatcgaagtt aagcaacttt   1380
aatatatcgt tataatcttc gat                                           1403
<210> 173
<211> 386
```

<212> PRT
<213> Arabidopsis thaliana
<400> 173

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

<210> 174
<211> 2833
<212> DNA
<213> Physcomitrella patens
<400> 174

```
ttctaggagg tcagcttggt ccaagtccga atcgcatcca cgcgatagga tatgtcgatc     60
caaatcacac atttttacac tccccagaag gaggaaaaga agttttcgaa accagtagtg    120
aggaatcata gtttcgtttt ccgaaaccta caaattcatc atactcattt cggtaacaaa    180
tgctacttaa aagacatgtc aggaaaacat ataaagtgac agtgattctg catcaacgtg    240
acagggatcc atcagtggag aatgaataaa aaaaacaatt agagaatggt gagtagaatc    300
```

ctcagcaaaa ttacgtttta tttattaata ttatacttaa aagtaaaaca ggatgattca 360
tgatgtcatc accgaggttg tggagagatt caaaaagagg atggaggcta cttttatggg 420
gagggtgggt tcactttatg cagtaaagta gtagatctct atttaaaggg gggaaagagg 480
atgtctgctg aaggccagtg cagtggatga acgcaggcgc tcgtctatgt ttcatccatc 540
catccattgt gtcgatttag aggccaccat ggcttctgag ggtgagggga gaggatagat 600
agcatagagg gggggggtg gagcaggcaa gggcaggagt gagattgtgg tggtggtgtg 660
attaggcgat ggagatatgg cgatggacat ggcgaggata gaagtcgagc ctcttgtcga 720
agtccacaac aacaccaaca acttgcttgt gcattgcgtg ctcgatgctt tgccggattc 780
ttcaccttgc ttgaaatcca gtgagacttc gtttgaagct gtggttgtaa aggaggacga 840
ggagggaggg aggccgctgt ttggcaagcc tgagctccac cctgcctcac cctcgagtga 900
tacagcggct gctgcaaatg gtgaattcgc tagatgcttg aattttgtga cggaggctga 960
ttgtggagat gtaggcgtac aatgttttga ggattttggt acgttgccgg actgtggtga 1020
ggcggggata agcagggaag agggtggagg tagagacggg gagcaggtgg agttgttaga 1080
gtctatgagc ctcgatggta gtaggaattc ggagaattta gagctagggg agctcggcga 1140
attgttgcag gggtcggaga cgctggattc ggttcctggg aacgaggttg gggaggagga 1200
ggcgctgctg ttggcgaaag cggcaaaggc gacgggcgtt gttgtttcgg cctccgatag 1260
tggtgaatgt agcagtgtcg ataggaagga aaatcaacaa agtccgaaat catgtaaaag 1320
tgccgcaccg gggaagaaga aggcgaaggt cgattggacg ccggagcttc atcggcgctt 1380
tgtccacgcg gtggaacagc ttggagtgga gaaagctttt ccctcgcgca tactagaact 1440
gatgggagta caatgtctca cccggcacaa tatcgccagt catttgcaga agtatcgctc 1500
gcatcgtcgc catcttgcag ccagggaagc cgaggcagca tcctggactc atcgtcgagc 1560
gtacacccag atgccctggt ctcgaagttc acggcgcgat ggccttcctt atcttgtacc 1620
cttacacacc cctcacatac aacctcgccc atccatggtc atggcaatgc aaccacagct 1680
tcagacgcag cacaccccgg tgtcgacgcc tcttaaggtg tgggggtatc ctacagtaga 1740
tcattcaagt gtacacatgt ggcagcaacc tgcagtggcg accccatcgt attggcaagc 1800
ccccgatggc tcttactggc agcatcctgc caccaattat gatgcgtatt cagctcgcgc 1860
ttgttatccc catcccatgc gagtttcgtt aggcactacg catgctggct ctccaatgat 1920
ggctccagga tttcctgacg agagctacta cggtgaagat gttcttgcag ctaccatgta 1980
tctttgtaac caatcatatg acagtgaatt aggacgagct gcgggcgtcg ctgcgtgcag 2040
taaaccaccg gagacgcatt tatcgaagga ggttcttgat gcagccatcg gagaagcgct 2100
tgctaaccct tggactcccc cgcccctggg actgaagccg ccgtctatgg agggagtaat 2160
cgcagagctt cagcggcagg gaatcaacac tgtgcctccc tctacctgtt agctttagtt 2220
gtttgctgta gagaatattt cattctacga ttcgcattcc aatgaccgct gaccgctggt 2280
gtcgcttggc tggtgttcat ctcgaggaat caatttggtg aaattttggt tatgtcacgg 2340
taggggggtc tattttctcc agagttcctc cggagaggtg tatgaaagga tcgattttct 2400
ttcttctttt ttttttttct ttttttttct ttttttttct ttttttttc cttcgaataa 2460
ggctgccttg gtggtgtttc ttcttagttt tttaacgagg gataccttat catatctgtc 2520
aagatatgtc actgaacgtt gctgcatgta gacttacgtt tatggtaaca ctttctcaat 2580
gccattaaaa accgaaacca gttcttctga ttgtttcatt ggaagtatcc tgacaacctg 2640
tcagcatctg ttggcatgtg gtttcctcac ccacaccctt ctttcagttg gttttgaatc 2700
tgcatctact gactgttaag gttttacgtg tatcaagtgt acgatttttc cacaagataa 2760
tttctcaaca actctgcttt cgaagcacct ttcttctcca ccaatcaaga gtggtgggaa 2820
gatggaccaa ggt 2833

<210> 175
<211> 511
<212> PRT
<213> Physcomitrella patens
<400> 175

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

```
       115                 120                 125
Glu Gln Val Glu Leu Leu Glu Ser Met Ser Leu Asp Gly Ser Arg Asn
   130                 135                 140
Ser Glu Asn Leu Glu Leu Gly Glu Leu Gly Glu Leu Leu Gln Gly Ser
145                 150                 155                 160
Glu Thr Leu Asp Ser Val Pro Gly Asn Glu Val Gly Glu Glu Glu Ala
                165                 170                 175
Leu Leu Leu Ala Lys Ala Ala Lys Ala Thr Gly Val Val Val Ser Ala
            180                 185                 190
Ser Asp Ser Gly Glu Cys Ser Ser Val Asp Arg Lys Glu Asn Gln Gln
        195                 200                 205
Ser Pro Lys Ser Cys Lys Ser Ala Ala Pro Gly Lys Lys Lys Ala Lys
   210                 215                 220
Val Asp Trp Thr Pro Glu Leu His Arg Arg Phe Val His Ala Val Glu
225                 230                 235                 240
Gln Leu Gly Val Glu Lys Ala Phe Pro Ser Arg Ile Leu Glu Leu Met
                245                 250                 255
Gly Val Gln Cys Leu Thr Arg His Asn Ile Ala Ser His Leu Gln Lys
            260                 265                 270
Tyr Arg Ser His Arg Arg His Leu Ala Ala Arg Glu Ala Glu Ala Ala
        275                 280                 285
Ser Trp Thr His Arg Arg Ala Tyr Thr Gln Met Pro Trp Ser Arg Ser
   290                 295                 300
Ser Arg Arg Asp Gly Leu Pro Tyr Leu Val Pro Leu His Thr Pro His
305                 310                 315                 320
Ile Gln Pro Arg Pro Ser Met Val Met Ala Met Gln Pro Gln Leu Gln
                325                 330                 335
Thr Gln His Thr Pro Val Ser Thr Pro Leu Lys Val Trp Gly Tyr Pro
            340                 345                 350
Thr Val Asp His Ser Ser Val His Met Trp Gln Gln Pro Ala Val Ala
        355                 360                 365
Thr Pro Ser Tyr Trp Gln Ala Pro Asp Gly Ser Tyr Trp Gln His Pro
   370                 375                 380
Ala Thr Asn Tyr Asp Ala Tyr Ser Ala Arg Ala Cys Tyr Pro His Pro
385                 390                 395                 400
Met Arg Val Ser Leu Gly Thr Thr His Ala Gly Ser Pro Met Met Ala
                405                 410                 415
Pro Gly Phe Pro Asp Glu Ser Tyr Tyr Gly Glu Asp Val Leu Ala Ala
            420                 425                 430
Thr Met Tyr Leu Cys Asn Gln Ser Tyr Asp Ser Glu Leu Gly Arg Ala
        435                 440                 445
Ala Gly Val Ala Ala Cys Ser Lys Pro Pro Glu Thr His Leu Ser Lys
   450                 455                 460
Glu Val Leu Asp Ala Ala Ile Gly Glu Ala Leu Ala Asn Pro Trp Thr
465                 470                 475                 480
Pro Pro Pro Leu Gly Leu Lys Pro Pro Ser Met Glu Gly Val Ile Ala
                485                 490                 495
Glu Leu Gln Arg Gln Gly Ile Asn Thr Val Pro Pro Ser Thr Cys
            500                 505                 510
<210> 176
<211> 2270
<212> DNA
<213> Physcomitrella patens
<400> 176
ggtggtggtg gtagaaggcg atggagatat ggcgatggac atggcgagga tagatgaatc     60
aaccgccgtc gaggtcaact cgctttcgct tgtgcattgc gtgttggatg ggttgcccga    120
ttcgccttgc ttgaaatcca gcccgacgtc attcgaggag gctgtggcgg aagggaggtc    180
ggtgttcggg gacgaggagg acatcatcaa caacagcaac gaccaggaca actcgtcctc    240
gtgcggtgca gtggtcacca cccacgaaga tttcgccgag tgcttgaatt ttgtgacgga    300
ggcggagtgt ggagatgtgg gtgtgcggtg tttcgaggat tttgacaagc tgccggactg    360
cggcgacgag ggggagacta gcaaagcgga ggaggagggg tgtgtacgaa taggcggagg    420
aggggagcag ggcgagctgt tagagtctgt gagccttgac tgtagtagga attcggagaa    480
tttagagctt cgggatctcg gggaattgtg ggaagggtcg gaacggcctg actcagtgcc    540
```

```
agggaacgag gtgggtgagg aggaggcgtt gctgttggcg gaggcggcca aggcgacggg    600
cgatgtcgtg tcggcctcgg atagtggaga atgtagcagt gtcgatagga aggacaatca    660
agccagtccg aaatccagta aaaatgcagc gccggggaag aagaaggcca aggtggactg    720
gacgcctgag cttcaccggc gcttcgtcca tgcagtggag cagctgggtg tggagaaagc    780
ctatccatcg cgtatcctag aactgatggg cgtgcaatgc ttgactcggc acaacatcgc    840
cagtcacttg cagaagtacc ggtcccaccg tcgccacctc gcagctcgag aagcagaggc    900
cgcgtcctgg acgcatcgtc gcacatacac tcaggctccc tggcctcgta gctcacggcg    960
cgatggcctc ccttatcttg tacctataca cacccctcac atacaacctc ggccttccat   1020
ggccatggca atgcaaccgc agcttcagac gccgcatcat ccgatatcca ctcctctcaa   1080
ggtctggggc taccccacag tagatcattc aaatgtacat atgtggcagc aacctgctgt   1140
ggcgacccca tcttactggc aagctgccga tggctcatac tggcaacatc ccgcgaccgg   1200
ttacgacgct ttctcagctc gtgcctgcta ctcgcatccc atgcagcgag ttcctgtaac   1260
gaccacgcat gcgggtttac caattgtggc gccaggattt cctgacgaga gctgctacta   1320
cggcgacgac atgcttgcag gctccatgta tctatgtaac caatcatatg atagtgaaat   1380
aggacgagct gcgggtgttg ctgcgtgcag caagccgata gagacgcatt tgtccaaaga   1440
ggtgttggat gcggccattg gcgaagctct cgccaatccc tggactcccc cacctctggg   1500
tctgaagcca ccatccatgg agggcgtcat tgcagagctt cagcggcagg ggatcaacac   1560
tgtgcctcct tcaacttgtt agctctcgac gatgtttttt tcttcctctt cctcttcctc   1620
ttcttctaga gagcaatttt tctttctacg actcatattc caatgaccgc tgaccgctgg   1680
tgtcgctggt gtcgctggtg ttcatcccga ggaactaatt tggtgaaaat tcggttaagt   1740
tgcgatagag gtctgatctc cggaaggttt atttttgtc ttctggagag gttgtataag   1800
aggttcccct gttttgcaat aaggcttcct tgatgagatt ttcctttatt tttcccctga   1860
ttctttctcc ttccatttta gtttcacaag aaggcatctt ctggccatgg cggtcacgat   1920
gtgtcacttt gatgctgcat gtggaccttt tttcttatat ctgtagtagc actcctccat   1980
ttcatgagga ataaagatca aacatcacac atcatcactc gaattcttca tctcctcctc   2040
ctccctttc cactaggatg ccttctattg cattggttgt catgtgactc agtctttctc   2100
ttacacgcac tttaactcgc tggatgtctc actttctgac tgttcaaggt gaggtctgat   2160
aggttcgaga cgggattgct tgcgatgact caccatctcc taatttggaa ccaacattcc   2220
tcctcaacct atcaactctc actcaaactt gcattgtgtg agcattggtc              2270
<210> 177
<211> 517
<212> PRT
<213> Physcomitrella patens
<400> 177
Met Ala Met Asp Met Ala Arg Ile Asp Glu Ser Thr Ala Val Glu Val
1               5                  10                  15
Asn Ser Leu Ser Leu Val His Cys Val Leu Asp Gly Leu Pro Asp Ser
            20                  25                  30
Pro Cys Leu Lys Ser Ser Pro Thr Ser Phe Glu Glu Ala Val Ala Glu
        35                  40                  45
Gly Arg Ser Val Phe Gly Asp Glu Glu Asp Ile Ile Asn Asn Ser Asn
    50                  55                  60
Asp Gln Asp Asn Ser Ser Ser Cys Gly Ala Val Val Thr Thr His Glu
65                  70                  75                  80
Asp Phe Ala Glu Cys Leu Asn Phe Val Thr Glu Ala Glu Cys Gly Asp
                85                  90                  95
Val Gly Val Arg Cys Phe Glu Asp Phe Asp Lys Leu Pro Asp Cys Gly
            100                 105                 110
Asp Glu Gly Glu Thr Ser Lys Ala Glu Glu Glu Gly Cys Val Arg Ile
        115                 120                 125
Gly Gly Gly Gly Glu Gln Gly Glu Leu Leu Glu Ser Val Ser Leu Asp
    130                 135                 140
Cys Ser Arg Asn Ser Glu Asn Leu Glu Leu Arg Asp Leu Gly Glu Leu
145                 150                 155                 160
Trp Glu Gly Ser Glu Arg Pro Asp Ser Val Pro Gly Asn Glu Val Gly
                165                 170                 175
Glu Glu Glu Ala Leu Leu Leu Ala Glu Ala Ala Lys Ala Thr Gly Asp
            180                 185                 190
Val Val Ser Ala Ser Asp Ser Gly Glu Cys Ser Ser Val Asp Arg Lys
        195                 200                 205
Asp Asn Gln Ala Ser Pro Lys Ser Ser Lys Asn Ala Ala Pro Gly Lys
    210                 215                 220
Lys Lys Ala Lys Val Asp Trp Thr Pro Glu Leu His Arg Arg Phe Val
```

```
225                 230                 235                 240
His Ala Val Glu Gln Leu Gly Val Glu Lys Ala Tyr Pro Ser Arg Ile
                245                 250                 255
Leu Glu Leu Met Gly Val Gln Cys Leu Thr Arg His Asn Ile Ala Ser
            260                 265                 270
His Leu Gln Lys Tyr Arg Ser His Arg Arg His Leu Ala Ala Arg Glu
        275                 280                 285
Ala Glu Ala Ala Ser Trp Thr His Arg Arg Thr Tyr Thr Gln Ala Pro
    290                 295                 300
Trp Pro Arg Ser Ser Arg Arg Asp Gly Leu Pro Tyr Leu Val Pro Ile
305                 310                 315                 320
His Thr Pro His Ile Gln Pro Arg Pro Ser Met Ala Met Ala Met Gln
                325                 330                 335
Pro Gln Leu Gln Thr Pro His His Pro Ile Ser Thr Pro Leu Lys Val
            340                 345                 350
Trp Gly Tyr Pro Thr Val Asp His Ser Asn Val His Met Trp Gln Gln
        355                 360                 365
Pro Ala Val Ala Thr Pro Ser Tyr Trp Gln Ala Ala Asp Gly Ser Tyr
    370                 375                 380
Trp Gln His Pro Ala Thr Gly Tyr Asp Ala Phe Ser Ala Arg Ala Cys
385                 390                 395                 400
Tyr Ser His Pro Met Gln Arg Val Pro Val Thr Thr Thr His Ala Gly
                405                 410                 415
Leu Pro Ile Val Ala Pro Gly Phe Pro Asp Glu Ser Cys Tyr Tyr Gly
            420                 425                 430
Asp Asp Met Leu Ala Gly Ser Met Tyr Leu Cys Asn Gln Ser Tyr Asp
        435                 440                 445
Ser Glu Ile Gly Arg Ala Ala Gly Val Ala Ala Cys Ser Lys Pro Ile
    450                 455                 460
Glu Thr His Leu Ser Lys Glu Val Leu Asp Ala Ala Ile Gly Glu Ala
465                 470                 475                 480
Leu Ala Asn Pro Trp Thr Pro Pro Pro Leu Gly Leu Lys Pro Pro Ser
                485                 490                 495
Met Glu Gly Val Ile Ala Glu Leu Gln Arg Gln Gly Ile Asn Thr Val
            500                 505                 510
Pro Pro Ser Thr Cys
        515
<210> 178
<211> 1843
<212> DNA
<213> Zea mays
<400> 178
aaacagcgaa acagtgcaga gacaagctac aacaacctac cctaacaggc cattccttcc     60
actgcacttc attcgatcct gagctatagc tggctgcctg agctcgctcc aagaagtgta    120
tctatagtat agacactagg cttcgtgtgg cgtgctcgag atatgcttgc agtgtcgccg    180
tcgccggtgc ggtgtgccga tgcggaggag tgcggcggag gaggcgccag caaggaaatg    240
gaggagaccg ccgtcgggcc tgtgtccgac tcggacctgg atttcgactt cacggtcgac    300
gacatagact tcggggactt cttcctcagg ctagacgacg gggatgacgc gctgccgggc    360
ctcgaggtcg accctgccga gatcgtcttc gctgacttcg aggcaatcgc caccgccggc    420
ggcgatggcg gcgtcacgga ccaggaggtg cccagtgtcc tgccctttgc ggacgcggcg    480
cacataggcg ccgtggatcc gtgttgtggt gtccttggcg aggacaacga cgcagcgtgc    540
gcagacgtgg aagaagggaa aggggagtgc gaccatgccg acgaggtagc agccgccggt    600
aataataata gcgactccgg tgaggccggc tgtggaggag cctttgccgg cgaaaaatca    660
ccgtcgtcga cggcatcgtc gtcgcaggag gctgagagcc ggcgcaaggt gtccaagaag    720
cactcccaag ggaagaagaa agcaaaggtg gattggacgc cggagcttca ccggagattc    780
gttcaggcgg tggaggagct gggcatcgac aaggcggtgc cgtccaggat cctcgagatc    840
atggggatcg actccctcac gcggcataac atagccagcc atctgcagaa gtaccgttcc    900
cacaggaagc acatgcttgc gagggaggtg gaggcagcga cgtggacgac gcaccggcgg    960
ccgatgtacg ctgcccccag cggcgccgtg aagaggcccg actctaacgc gtggaccgtg   1020
ccgaccatcg gtttccctcc gccggcgggg acccctcctc gtccggtgca gcacttcggg   1080
aggccactgc acgtctgggg ccatccgagt ccgacgccag cggtggagtc accccgggtg   1140
ccaatgtggc ctcggcatct cgcccccccgc gccccgccgc cgccgccgtg ggctccgcca   1200
ccgccagctg acccggcgtc gttctggcac catgcttaca tgaggggccc tgctgcccat   1260
```

atgccagacc aggtggcggt gactccatgc gtggcagtgc caatggcagc agcgcgtttc 1320
cctgctccac acgtgagggg ttctttgcca tggccacctc cgatgtacag acctctcgtt 1380
cctccagcac tcgcaggcaa gagccagcaa gacgcgctgt ttcagctaca gatacagcca 1440
tcaagcgaga gcatagatgc agcaataggt gatgtcttaa cgaagccgtg gctgccgctg 1500
cccctcggac tgaagccccc ttcggtagac agtgtcatgg gcgagctgca gaggcaaggc 1560
gtagcgaatg tgccgcaagc ttgtggatga tccccggggc tgcatagctg ctagctcgtc 1620
cctgcacgca gtgcaacaaa aggaatttgt cgacatgcct tttatgtttt gaattcccgt 1680
gaaccgttta tggagcactc tcaactcgtt cagggtcaga tgaacagaac attataggag 1740
tacattcttt gaggtttcag gtttgaggga aatcaacctg agtgcatcca agtacagtgt 1800
actgccacag tgccacgtcg gaaaatgaag tcttgacccg aat 1843

<210> 179
<211> 475
<212> PRT
<213> Zea mays
<400> 179

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

```
Arg Phe Pro Ala Pro His Val Arg Gly Ser Leu Pro Trp Pro Pro Pro
385                 390                 395                 400
Met Tyr Arg Pro Leu Val Pro Pro Ala Leu Ala Gly Lys Ser Gln Gln
                405                 410                 415
Asp Ala Leu Phe Gln Leu Gln Ile Gln Pro Ser Ser Glu Ser Ile Asp
            420                 425                 430
Ala Ala Ile Gly Asp Val Leu Thr Lys Pro Trp Leu Pro Leu Pro Leu
        435                 440                 445
Gly Leu Lys Pro Pro Ser Val Asp Ser Val Met Gly Glu Leu Gln Arg
    450                 455                 460
Gln Gly Val Ala Asn Val Pro Gln Ala Cys Gly
465                 470                 475
<210> 180
<211> 2192
<212> DNA
<213> Zea mays
<400> 180
ctcatcttct cttcttcttc ccccagtcct ccccccctttt cccctcttcg gcctctctcg     60
ctcagctcac tcttcattaa gcgagctcac gtcgttcctc ccttcttctt cttcttatcc    120
gttgttcaat tcgttcagct agccggccag agatcgagca tcatctccat catcgattca    180
tccatctcat ctttctcttc ttctattcta tgtcgtcgtc gtcccagatt agatcgaagc    240
tgctagcagt ctatccagtc tctagctagc tagctagatc aagcccgcag actatacaat    300
ataatacaag ctagctacgt gcttattatt gctttattag tctagaataa tcttgatata    360
tacgatcgaa catatatctc gatctccacc gatacacacc cggccctatc catgcttgag    420
gtgtcgacgc tgcgcggccc tactagcagc ggcagcaagg cggagcagca ctgcggcggc    480
ggcggcggct tcgtcggcga ccaccatgtg gtgttcccga cgtccggcga ctgcttcgcc    540
atggtggacg acaacctcct ggactacatc gacttcagct gcgacgtgcc cttcttcgac    600
gctgacgggg acatcctccc cgacctggag gtagacacca cggagctcct cgccgagttc    660
tcgtccaccc ctcctgcgga cgacctgctg gcagtggcag tattcggcgc cgacgaccag    720
ccggcggcgg cagtagcaca agagaagccg tcgtcgtcgt tggagcaaac atgtggtgac    780
gacaaaggtg tagcagtagc cgccgccaga agaaagctgc agacgacgac gacgacgacg    840
acgacggagg aggaggattc ttctcctgcc gggtccgggg ccaacaagtc gtcggcgtcg    900
gcagagggcc acagcagcaa gaagaagtcg gcgggcaaga actccaacgg cggcaagcgc    960
aaggtgaagg tggactggac gccggagctg caccggcggt tcgtgcaggc ggtggagcag   1020
ctgggcatcg acaaggccgt gccgtccagg atcctggaga tcatgggcac ggactgcctc   1080
acaaggcaca acattgccag ccacctccag aagtaccggt cgcacagaaa gcacctgatg   1140
gcgcgggagg cggaggccgc cacctgggcg cagaagcgcc acatgtacgc gccgccagct   1200
ccaaggacga cgacgacgac ggacgccgcc aggccgccgt gggtggtgcc gacgaccatc   1260
gggttcccgc cgccgcgctt ctgccgcccg ctgcacgtgt ggggccaccc gccgccgcac   1320
gccgccgcgg ctgaagcagc agcggcgact cccatgctgc ccgtgtggcc gcgtcacctg   1380
gcgccgcccc ggcacctggc gccgtgggcg cacccgacgc cggtggaccc ggcgttctgg   1440
caccagcagt acagcgctgc caggaaatgg ggcccacagg cagccgccgt gacgcaaggg   1500
acgccatgcg tgccgctgcc gaggtttccg gtgcctcacc ccatctacag cagaccggcg   1560
atggtacctc cgccgccaag caccaccaag ctagctcaac tgcatctgga gctccaagcg   1620
cacccgtcca aggagagcat cgacgcagcc atcggagatg ttttagtgaa gccatggctg   1680
ccgcttccac tggggctcaa gccgccgtcg ctcgacagcg tcatgtcgga gctgcacaag   1740
caaggcgtac caaaaatccc accggcggct gccaccacca ccggcgccac cggatgacat   1800
actatctcgt cgacaataca tgcatgtaca atagacggat gtctagtagt atagtagatt   1860
gctgctagct agctagccgc tagagtgtag tagcatatgc atgccttttt tttcttcttc   1920
ttttttttgcc cttattatta agctggctaa tagcgattga gatggagctt gacacagatc   1980
tgctagctag ctatttgagg gtttctcttg tatgctacct attgctgctg cttgctgctg   2040
agacaagtaa tgtacgtacg cctgtgcaaa cgacacatcg gattgtattg tattactagt   2100
tatatgaaca acaacaataa taataggcac atgcatgcct gcctagtatg tgagctgcta   2160
gctagctaca tgcatgttgc gcattccttt cc                                 2192
<210> 181
<211> 461
<212> PRT
<213> Zea mays
<400> 181
Met Leu Glu Val Ser Thr Leu Arg Gly Pro Thr Ser Ser Gly Ser Lys
1               5                   10                  15
Ala Glu Gln His Cys Gly Gly Gly Gly Gly Phe Val Gly Asp His His
            20                  25                  30
```

```
Val Val Phe Pro Thr Ser Gly Asp Cys Phe Ala Met Val Asp Asp Asn
        35                  40                  45
Leu Leu Asp Tyr Ile Asp Phe Ser Cys Asp Val Pro Phe Phe Asp Ala
    50                  55                  60
Asp Gly Asp Ile Leu Pro Asp Leu Glu Val Asp Thr Thr Glu Leu Leu
65                  70                  75                  80
Ala Glu Phe Ser Ser Thr Pro Pro Ala Asp Asp Leu Leu Ala Val Ala
                85                  90                  95
Val Phe Gly Ala Asp Asp Gln Pro Ala Ala Ala Val Ala Gln Glu Lys
            100                 105                 110
Pro Ser Ser Ser Leu Glu Gln Thr Cys Gly Asp Asp Lys Gly Val Ala
        115                 120                 125
Val Ala Ala Ala Arg Arg Lys Leu Gln Thr Thr Thr Thr Thr Thr Thr
    130                 135                 140
Thr Glu Glu Glu Asp Ser Ser Pro Ala Gly Ser Gly Ala Asn Lys Ser
145                 150                 155                 160
Ser Ala Ser Ala Glu Gly His Ser Ser Lys Lys Lys Ser Ala Gly Lys
                165                 170                 175
Asn Ser Asn Gly Gly Lys Arg Lys Val Lys Val Asp Trp Thr Pro Glu
            180                 185                 190
Leu His Arg Arg Phe Val Gln Ala Val Glu Gln Leu Gly Ile Asp Lys
        195                 200                 205
Ala Val Pro Ser Arg Ile Leu Glu Ile Met Gly Thr Asp Cys Leu Thr
    210                 215                 220
Arg His Asn Ile Ala Ser His Leu Gln Lys Tyr Arg Ser His Arg Lys
225                 230                 235                 240
His Leu Met Ala Arg Glu Ala Glu Ala Ala Thr Trp Ala Gln Lys Arg
                245                 250                 255
His Met Tyr Ala Pro Pro Ala Pro Arg Thr Thr Thr Thr Thr Asp Ala
            260                 265                 270
Ala Arg Pro Pro Trp Val Val Pro Thr Thr Ile Gly Phe Pro Pro Pro
        275                 280                 285
Arg Phe Cys Arg Pro Leu His Val Trp Gly His Pro Pro Pro His Ala
    290                 295                 300
Ala Ala Ala Glu Ala Ala Ala Ala Thr Pro Met Leu Pro Val Trp Pro
305                 310                 315                 320
Arg His Leu Ala Pro Pro Arg His Leu Ala Pro Trp Ala His Pro Thr
                325                 330                 335
Pro Val Asp Pro Ala Phe Trp His Gln Gln Tyr Ser Ala Ala Arg Lys
            340                 345                 350
Trp Gly Pro Gln Ala Ala Ala Val Thr Gln Gly Thr Pro Cys Val Pro
        355                 360                 365
Leu Pro Arg Phe Pro Val Pro His Pro Ile Tyr Ser Arg Pro Ala Met
    370                 375                 380
Val Pro Pro Pro Pro Ser Thr Thr Lys Leu Ala Gln Leu His Leu Glu
385                 390                 395                 400
Leu Gln Ala His Pro Ser Lys Glu Ser Ile Asp Ala Ala Ile Gly Asp
                405                 410                 415
Val Leu Val Lys Pro Trp Leu Pro Leu Pro Leu Gly Leu Lys Pro Pro
            420                 425                 430
Ser Leu Asp Ser Val Met Ser Glu Leu His Lys Gln Gly Val Pro Lys
        435                 440                 445
Ile Pro Pro Ala Ala Ala Thr Thr Thr Gly Ala Thr Gly
    450                 455                 460
<210> 182
<211> 1221
<212> DNA
<213> Triticum aestivum
<400> 182
gaattgggca cgaggaccgc cccattgtgc cggacgcata atcgccgtcg tcgacaacgt     60
cgtcgtccac ggaggccgag agccgccaca agtcatcaag caagaactcc cacggaaaga    120
agaaagccaa ggtggactgg acgccggagc tgcaccggag gttcgtgcag gccgtagagc    180
agctcggcat cgacaaggca gtgccgtcta ggattttgga gatcatgggg atcaactcac    240
```

```
tcactcggca caacatagca agccatttgc agaagtaccg gtctcaccgg aagcacatga    300
ttgcgcggga ggcggaggcg gcgagctgga cccaacgacg acagatgtac gccgccggcg    360
ggccggccgc ggccgtgaag aggcaggact cgaacatgtg gaccgtgcca accatcggct    420
tcgcgccggc gcaccctcct cctcctcctc cccctccttc accggcagcc atgcagcact    480
acgctcggcc gttgcacgtc tggggtcacc ctacgatgga ctcgccccgg atgccgatgt    540
ggccgaggca cacaatatcc cgcgccccga tgccggcgtg ggctcccccg ccgccgccgc    600
cgtctgaccc ggctttctgg caccaccctt acatgagggg gcccgcagcg tatatgccga    660
cccatgggac tccttgcatg gcaatgccga tgacaccgaa atttcctgct gcacccgtgc    720
ccgtggccat gccgtgccca gtctatgcgt ccccctcccc ggcaccagcg ctggcgagca    780
agagccaaca agattcgcag ctccagctcc aagcacaacc atcaaatgag agcatagacg    840
cggccattgg tgacgtttta tccaaaccgt ggctgccgct gccgctgggg ctgaagcccc    900
cttcgttggg cagcgtcatg ggcgagcttg aaaggcaagg cgtggccaat gtgccccaag    960
cctgcgggtg agcgttggag ggtgcatccg cgtgcactcc atgcatgact gctgctccgt   1020
tcgatggagc agctagcagc aacaacaaca tcgtcgacat gtctttgttc ctttgaacgt   1080
tttgtggatc tctctctctc tctcttggtc aacttgtgca taatttcgat caagagaaac   1140
atcgttattt tgagttcact cccgagacac atttttgtta cacctaaact ttaagtttgc   1200
ggtaacagca gcatcaacaa c                                             1221
<210> 183
<211> 248
<212> PRT
<213> Triticum aestivum
<400> 183
Met Gly Ile Asn Ser Leu Thr Arg His Asn Ile Ala Ser His Leu Gln
1               5                  10                  15
Lys Tyr Arg Ser His Arg Lys His Met Ile Ala Arg Glu Ala Glu Ala
            20                  25                  30
Ala Ser Trp Thr Gln Arg Arg Gln Met Tyr Ala Ala Gly Gly Pro Ala
        35                  40                  45
Ala Ala Val Lys Arg Gln Asp Ser Asn Met Trp Thr Val Pro Thr Ile
    50                  55                  60
Gly Phe Ala Pro Ala His Pro Pro Pro Pro Pro Pro Pro Pro Ser Pro
65                  70                  75                  80
Ala Ala Met Gln His Tyr Ala Arg Pro Leu His Val Trp Gly His Pro
                85                  90                  95
Thr Met Asp Ser Pro Arg Met Pro Met Trp Pro Arg His Thr Ile Ser
            100                 105                 110
Arg Ala Pro Met Pro Ala Trp Ala Pro Pro Pro Pro Pro Pro Ser Asp
        115                 120                 125
Pro Ala Phe Trp His His Pro Tyr Met Arg Gly Pro Ala Ala Tyr Met
    130                 135                 140
Pro Thr His Gly Thr Pro Cys Met Ala Met Pro Met Thr Pro Lys Phe
145                 150                 155                 160
Pro Ala Ala Pro Val Pro Val Ala Met Pro Cys Pro Val Tyr Ala Ser
                165                 170                 175
Pro Ser Pro Ala Pro Ala Leu Ala Ser Lys Ser Gln Gln Asp Ser Gln
            180                 185                 190
Leu Gln Leu Gln Ala Gln Pro Ser Asn Glu Ser Ile Asp Ala Ala Ile
        195                 200                 205
Gly Asp Val Leu Ser Lys Pro Trp Leu Pro Leu Pro Leu Gly Leu Lys
    210                 215                 220
Pro Pro Ser Leu Gly Ser Val Met Gly Glu Leu Glu Arg Gln Gly Val
225                 230                 235                 240
Ala Asn Val Pro Gln Ala Cys Gly
                245
<210> 184
<211> 839
<212> DNA
<213> Allium cepa
<400> 184
tgcaaatatt gttcaacgta tataaagaca atttagtatg ctaacaatct ctccccttga     60
aagttctaat ccagatgcaa aagacgaggg agattttgta ctagaagaca tcagctttga    120
tgacttgttc gtaggattcg atgaacttcc cgaccttgaa attgaccctt gtgatatatt    180
tgctgatttt tctttcaata gtagcgagga gggagatgga aataacaagg gttcaacatc    240
```

ggaagaaaat gatgtacagc ataaaagaga cggatactat aaggagcact caggtaaaga 300
agagacggag gttgtgagtg caagaacaag ggaggatgaa aagaagccgc cttcatctaa 360
atcagagaat gttaaaagcc taaagtcatc aggcaaaaag tcatctcagt ctaaaaagaa 420
agctaaggtg gactggactc cggagcttca tcggagattt gttcaggcag tggagcaact 480
tggggtagac aaagcagtac catccaggat tttagagctc atgggaattg attgtctcac 540
aagacataat attgcaagcc atttacagaa atatcggtcg cacagaaagc atttgctagc 600
aagagaagca gaagcagcaa gctggagtca cagacgtcag ttgtacgtga gcagcactaa 660
caatggaaag aggaggagga aaaacaatga acatgaaacc tagttgggcc ccacagtcac 720
agcccattgg tggatttcct cccaacaatg catcatccct caacatcaga acttcagaac 780
ctttgcacgt ctggggccca tccacaggca gtggagcatc ctcagctagt tccagtctg 839

<210> 185
<211> 221
<212> PRT
<213> Allium cepa
<400> 185

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

<210> 186
<211> 454
<212> DNA
<213> Hordeum vulgare
<400> 186

catcgtgggc gatgaggttt gcagcgcggt gacgacggac gattcttccg ctgcggtcgg 60
gtccgagaac agcaagtcgt cggcgtcggc agagggccac agcaagagga cgtcggcagc 120
cgcagcgacc aagagctccc acggcaggcg gaaggtgaag gtggactgga cgccggagtt 180
gcaccggcgg ttcgtgcagg cgggggagca gctggggctg gacaaggcgg tgccgtcgcg 240
gatcctggag ctcatgggca acgagtaccg tctcacgcgc cacaacatcg ccagccatct 300
ccagaagtac cggtcacaca ggaagcacct gatggcgagg gagggcgagg cgggtagctg 360
gacgccgcag ccgcaaatgt gctggggtgc gaaggtggtg agggtggggc gggctctgac 420
gttgtagggg ttatatagcg gggtggcggc gggg 454

<210> 187
<211> 144
<212> PRT
<213> Hordeum vulgare
<400> 187

Ser Leu Ser Ile Val Gly Asp Glu Val Cys Ser Ala Val Thr Thr Asp
1 5 10 15

```
Asp Ser Ser Ala Ala Val Gly Ser Glu Asn Ser Lys Ser Ser Ala Ser
            20                  25                  30
Ala Glu Gly His Ser Lys Arg Thr Ser Ala Ala Ala Ala Thr Lys Ser
        35                  40                  45
Ser His Gly Arg Arg Lys Val Lys Val Asp Trp Thr Pro Glu Leu His
    50                  55                  60
Arg Arg Phe Val Gln Ala Gly Glu Gln Leu Gly Leu Asp Lys Ala Val
65                  70                  75                  80
Pro Ser Arg Ile Leu Glu Leu Met Gly Asn Glu Tyr Arg Leu Thr Arg
                85                  90                  95
His Asn Ile Ala Ser His Leu Gln Lys Tyr Arg Ser His Arg Lys His
            100                 105                 110
Leu Met Ala Arg Glu Gly Glu Ala Gly Ser Trp Thr Pro Gln Pro Gln
        115                 120                 125
Met Cys Trp Gly Ala Lys Val Val Arg Val Gly Arg Ala Leu Thr Leu
    130                 135                 140
<210> 188
<211> 811
<212> DNA
<213> Sorghum bicolor
<400> 188
gcacgaggct tgaaattccc tccgccgccg ccgccgccgc cgcctcctcc tcctccttct     60
cctcatccga tgcagcactt cggaaggccg ctgcatgcct ggggccatcc gacgccaacg    120
gtggagtcac cccgggtgcc aatgtggcct cggcatctcg tcccccgcac cccgccgccg    180
ccgtgggctc ctccgccgcc atctgacccg gcgttctggc accatgctta catgaggggg    240
cctgcacaca tgccgggcca ggtgactccc tgcgtggcag tgccaatgcc agctgcgcgt    300
ttccctgctc cacccgtgag gggtgctttg ccatgtccac ctccaatgta cagacctctc    360
gttcctccaa cactcgatac gcagtttcag ctacagacac agccatcaag tgagagcata    420
gatgcagcaa taggtgatgt tttaacgaaa ccgtggctgc cactgcccct gggactgaag    480
ccccccttcag tagacagtgt catgggcgag ctgcagaggc aaggtgtggc aaatgtgcca    540
ccagcttgtg gatgatcccc ggggctccat agctagctgc ttgtccctgc agtccaacaa    600
aaagaatttg tcgacatgcc ttttctgttt caaattttca tgaaccattt atggaccact    660
ctctcttagt taacttgttc agggtcagag agcaaaacag tacatccttt caggtttgag    720
ggaaatcaac ctgagtacat ccaagtacaa tgtgccatga cggaataatg aagtcttggc    780
cttggcccga ataatcagta gctgccatct c                                   811
<210> 189
<211> 184
<212> PRT
<213> Sorghum bicolor
<400> 189
Ala Arg Gly Leu Lys Phe Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro
1               5                  10                  15
Pro Pro Pro Ser Pro His Pro Met Gln His Phe Gly Arg Pro Leu His
            20                  25                  30
Ala Trp Gly His Pro Thr Pro Thr Val Glu Ser Pro Arg Val Pro Met
        35                  40                  45
Trp Pro Arg His Leu Val Pro Arg Thr Pro Pro Pro Pro Trp Ala Pro
    50                  55                  60
Pro Pro Pro Ser Asp Pro Ala Phe Trp His His Ala Tyr Met Arg Gly
65                  70                  75                  80
Pro Ala His Met Pro Gly Gln Val Thr Pro Cys Val Ala Val Pro Met
                85                  90                  95
Pro Ala Ala Arg Phe Pro Ala Pro Pro Val Arg Gly Ala Leu Pro Cys
            100                 105                 110
Pro Pro Pro Met Tyr Arg Pro Leu Val Pro Pro Thr Leu Asp Thr Gln
        115                 120                 125
Phe Gln Leu Gln Thr Gln Pro Ser Ser Glu Ser Ile Asp Ala Ala Ile
    130                 135                 140
Gly Asp Val Leu Thr Lys Pro Trp Leu Pro Leu Pro Leu Gly Leu Lys
145                 150                 155                 160
Pro Pro Ser Val Asp Ser Val Met Gly Glu Leu Gln Arg Gln Gly Val
                165                 170                 175
Ala Asn Val Pro Pro Ala Cys Gly
```

180
<210> 190
<211> 593
<212> DNA
<213> Saccharum officinarum
<220>
<221> misc_feature
<222> (521)..(521)
<223> n é a, c, g, ou t
<220>
<221> misc_feature
<222> (524)..(524)
<223> n é a, c, g, ou t
<220>
<221> misc_feature
<222> (534)..(534)
<223> n é a, c, g, ou t
<400> 190
catcgacaag gccgtgccgt cccggatcct cgagataatg ggcatggatt gcctcacaag 60
acacaacatt gccagccacc tccagaagta ccggtcgcac agaaagcacc tgatggcgcg 120
ggaggcggag gccgccacct gggcgcagaa gcggcacatg tacgcggcgg ccggcggagt 180
agccccgagg acggacgcac cacacggcag ccggccgtgg gtggtgccga ccatcgggtt 240
cccgccgccg gcgccccccgc ccttctgccg gccgctgcac gtgtggggcc acccgcccca 300
cgccgccgcc gctgaagccg ctgctgcggt ggcggcgact ccgactccga tgctgcccgt 360
gtggccgcgg cacctggcgc cgccccggcc gctgccgccg tgggcgcacc cgcacccgcc 420
ggccgtggac ccggcgttct ggcaccagca gtacaacgcg gccaggaaat ggggtccaca 480
ggcaaccgca gtgacgcaag ggacgccggc gtgcgtgccg ncgnccgccg ccgncatgct 540
gccgaggttt tcccgggttg gccccggggg ggggggggga aaaaaaattt ttt 593
<210> 191
<211> 197
<212> PRT
<213> Saccharum officinarum
<220>
<221> INCERTO
<222> (174)..(175)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> INCERTO
<222> (178)..(178)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 191
Ile Asp Lys Ala Val Pro Ser Arg Ile Leu Glu Ile Met Gly Met Asp
1 5 10 15
Cys Leu Thr Arg His Asn Ile Ala Ser His Leu Gln Lys Tyr Arg Ser
20 25 30
His Arg Lys His Leu Met Ala Arg Glu Ala Glu Ala Ala Thr Trp Ala
35 40 45
Gln Lys Arg His Met Tyr Ala Ala Ala Gly Gly Val Ala Pro Arg Thr
50 55 60
Asp Ala Pro His Gly Ser Arg Pro Trp Val Val Pro Thr Ile Gly Phe
65 70 75 80
Pro Pro Pro Ala Pro Pro Pro Phe Cys Arg Pro Leu His Val Trp Gly
85 90 95
His Pro Pro His Ala Ala Ala Ala Glu Ala Ala Ala Ala Val Ala Ala
100 105 110
Thr Pro Thr Pro Met Leu Pro Val Trp Pro Arg His Leu Ala Pro Pro
115 120 125
Arg Pro Leu Pro Pro Trp Ala His Pro His Pro Pro Ala Val Asp Pro
130 135 140
Ala Phe Trp His Gln Gln Tyr Asn Ala Ala Arg Lys Trp Gly Pro Gln
145 150 155 160
Ala Thr Ala Val Thr Gln Gly Thr Pro Ala Cys Val Pro Xaa Xaa Ala
165 170 175

Ala Xaa Met Leu Pro Arg Phe Ser Arg Val Gly Pro Gly Gly Gly Gly
180 185 190
Gly Lys Lys Ile Phe
195

<210> 192
<211> 1816
<212> DNA
<213> Oryza sativa
<400> 192

ggcaactaac aatacctcgc acctagtcac caccgcacca accgcgcgcg accggccccc 60
cgtcagtcga aagctgagcc tgagcgagca actgatgccg ctagctatag ctgctgcctg 120
cagcacgcag cgcccgtaac ctaacgtatc atttacttcc attgcactgc acatcctgtg 180
agttcgttgt ctgagtgaga gctggacgtg ctgcaccgag ctcctggccg agatgcttgc 240
cgtgtcgccg gcgatgtgcc ccgacattga ggaccgcgcc gcggtggccg gcgatgctgg 300
catggaggtc gtcgggatgt cgtcggacga catggatcag ttcgacttct ccgtcgatga 360
catagacttc ggggacttct tcctgaggct ggaggacggt gatgtgctcc cggacctcga 420
ggtcgacccg gccgagatct tcaccgactt cgaggcaatc gcgacgagtg gaggcgaagg 480
tgtgcaggac caggaggtgc ccaccgtcga gctcttggcg cctgcggacg acgtcggtgt 540
gctggatccg tgcggcgatg tcgtcgtcgg ggaggagaac gcggcgtttg ccggggctgg 600
agaggagaag gtagggtgta accaggacga tgatgcgggg gaagcgaatg tcgacgatgg 660
agccgcggcg gttgaggcca agtcttcgtc gccgtcatcg acgacgtcgt cgtcgcagga 720
ggctgagagc cggcacaagt catccagcaa gagctcccat gggaagaaga aagcgaaggt 780
ggactggacg cctgagcttc accggaggtt cgtgcaggcg gtggagcagc tcggcatcga 840
caaggccgtg ccgtcgagga tacttgagat catggggatc gactctctca cccggcacaa 900
catagccagc catcttcaga agtaccggtc acacagaaaa cacatgattg cgagagaggc 960
ggaggcagcg agttggaccc aacggcggca gatttacgcc gccggtggag gtgctgttgc 1020
gaagaggccg gagtccaacg cgtggaccgt gccaaccatt ggcttccctc ctcctccgcc 1080
accaccacca tcaccggctc cgattcaaca ttttgctcgc ccgttgcatg tttggggcca 1140
cccgacgatg gacccgtccc gagttccagt gtggccaccg cggcacctcg ttccccgtgg 1200
cccggcgcca ccatgggttc caccgccgcc gccgtcggac cctgctttct ggcaccaccc 1260
ttacatgagg gggccagcac atgtgccaac tcaagggaca ccttgcatgg cgatgcccat 1320
gccagctgcg agatttcctg ctccaccggt gccaggagtt gtcccgtgtc caatgtatag 1380
gccattgact ccaccagcac tggcgagcaa gaatcagcag gacgcacagc ttcaactcca 1440
ggttcaacca tcaagcgaga gcatcgacgc agctatcggt gatgttttat cgaaaccgtg 1500
gttgcctttg cctcttggac tgaagccacc ttcagtggac agtgtgatgg gcgagctgca 1560
gaggcaaggc gtagcaaacg ttcctccagc gtgtggatga tatttgcatg atagattgat 1620
ggttccgtta ctgactgatt gcttatctgg tcagttcacc aaaaaatggg aaaattcgcc 1680
gacatgtcct tttatttttc ttttggccta gcgtttcttg gacatctcgt ttaattttta 1740
acttgtgcag agttcgaaag acatctttgt ttcgactccg ggagaaatta acctgcgtat 1800
tgtaaatgta aaatgt 1816

<210> 193
<211> 455
<212> PRT
<213> Oryza sativa
<400> 193

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

```
Ala Ala Val Glu Ala Lys Ser Ser Ser Pro Ser Ser Thr Thr Ser Ser
145                 150                 155                 160
Ser Gln Glu Ala Glu Ser Arg His Lys Ser Ser Ser Lys Ser Ser His
                165                 170                 175
Gly Lys Lys Lys Ala Lys Val Asp Trp Thr Pro Glu Leu His Arg Arg
            180                 185                 190
Phe Val Gln Ala Val Glu Gln Leu Gly Ile Asp Lys Ala Val Pro Ser
        195                 200                 205
Arg Ile Leu Glu Ile Met Gly Ile Asp Ser Leu Thr Arg His Asn Ile
    210                 215                 220
Ala Ser His Leu Gln Lys Tyr Arg Ser His Arg Lys His Met Ile Ala
225                 230                 235                 240
Arg Glu Ala Glu Ala Ala Ser Trp Thr Gln Arg Arg Gln Ile Tyr Ala
                245                 250                 255
Ala Gly Gly Gly Ala Val Ala Lys Arg Pro Glu Ser Asn Ala Trp Thr
            260                 265                 270
Val Pro Thr Ile Gly Phe Pro Pro Pro Pro Pro Pro Pro Pro Ser Pro
        275                 280                 285
Ala Pro Ile Gln His Phe Ala Arg Pro Leu His Val Trp Gly His Pro
    290                 295                 300
Thr Met Asp Pro Ser Arg Val Pro Val Trp Pro Pro Arg His Leu Val
305                 310                 315                 320
Pro Arg Gly Pro Ala Pro Pro Trp Val Pro Pro Pro Pro Pro Ser Asp
                325                 330                 335
Pro Ala Phe Trp His His Pro Tyr Met Arg Gly Pro Ala His Val Pro
            340                 345                 350
Thr Gln Gly Thr Pro Cys Met Ala Met Pro Met Pro Ala Ala Arg Phe
        355                 360                 365
Pro Ala Pro Pro Val Pro Gly Val Val Pro Cys Pro Met Tyr Arg Pro
    370                 375                 380
Leu Thr Pro Pro Ala Leu Ala Ser Lys Asn Gln Gln Asp Ala Gln Leu
385                 390                 395                 400
Gln Leu Gln Val Gln Pro Ser Ser Glu Ser Ile Asp Ala Ala Ile Gly
                405                 410                 415
Asp Val Leu Ser Lys Pro Trp Leu Pro Leu Pro Leu Gly Leu Lys Pro
            420                 425                 430
Pro Ser Val Asp Ser Val Met Gly Glu Leu Gln Arg Gln Gly Val Ala
        435                 440                 445
Asn Val Pro Pro Ala Cys Gly
    450                 455
<210> 194
<211> 1096
<212> DNA
<213> Oryza sativa
<400> 194
tgtgctaagg catccatgct actgcagagt gtcccacctg cagttttggt tcgatttttg     60
agggaacatc gttctgaatg ggcggattat aacttcgatg catattcagc ttcatctctg    120
aagacaagct catgttcact tcctgggttg cggcctatga gattttctgg gagccagatc    180
attatgccac ttgctcacac ggtggagaat gaagagattt tagaagttgt ccgtcttgaa    240
ggacaagcac ttacacatga tgatggtctt atgtctagag atattcacct gcttcagctt    300
tgcactggaa tagatgagaa atcaatggga tcctgcttcc agcttgtctt tgcaccaatc    360
gatgagcttt tccctgatga tgctccgtta atatcttcag gctttcgtgt tataccgctg    420
gacatgaaaa cagatggtac acctgctggt agaacattag atttggcatc tagccttgag    480
gttggttcaa ctgcacagcc cacaggggat gcatctatgg atgactgtaa tctacgatca    540
gtgctgacaa ttgcctttca gttcccttat gaaatgcatc tccaagacag cgttgcaact    600
atggcccggc aatatgtccg cagtattgtt ttctctgttc agagagtatc aatggctgtt    660
tctccttctc ggtctggctt gaatgctggg cagaagataa tttcaggctt ccctgaagcc    720
ccaacgctag ctcgttggat ttgccaaagc taccagttcc atttgggggt ggagttactt    780
aggcaggcag atgatgctgg ggaagcacta ttgaaaatgc tatgggatta cgaagacgct    840
attttgtgct gttctttcaa ggaaaagcct gtatttactt ttgccaacga gatgggacta    900
aacatgctag aaacatctct cgtcgctctc caagatctct cactggacaa gatatttgat    960
gaagccggta ggaaggccct atacaacgag atcccgaaat tgatggaaca gggttacgtg   1020
tacctgcctg gtggagtgtg cttgtccggg atggggcgcc atgtttcttt cgagcaagct   1080
```

gtagcatgga aggtgc 1096

<210> 195
<211> 365
<212> PRT
<213> Oryza sativa
<400> 195

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

<210> 196
<211> 1395
<212> DNA
<213> Oryza sativa
<400> 196

```
gctcaagaga caagcgggga ggttgtatac gctttgggga ggcaacctgc tgttttgcgg     60
acatttagtc agaggttgag tagaggcttc aatgatgcta taagtggttt caatgatgat    120
ggttggtctg tcatgggtgg ggatggcatt gaagatgtga tcattgcttg caatgcaaag    180
aaggttagga atactagcac ttcggccaat gcttttgtaa ctccaggagg tgttatatgt    240
gctaaggcat ccatgctact gcagagtgtc ccacctgcag ttttggttcg atttttgagg    300
gaacatcgtt ctgaatgggc ggattataac ttcgatgcat attcagcttc atctctgaag    360
```

```
acaagctcat gttcacttcc tgggttgcgg cctatgagat tttctgggag ccagatcatt    420
atgccacttg ctcacacggt ggagaatgag gagattttag aagttgtccg tcttgaagga    480
caagcactta cacatgatga tggtcttatg tctagagata ttcacctgct tcagctttgc    540
actggaatag atgagaaatc aatgggatcc tgcttccagc ttgtctctgc accaatcgat    600
gagcttttcc ctgatgatgc tccgttaata tcttcaggct ttcgtgttat accgctggac    660
atgaaaacag atggtacacc tgctggtaga acattagatt tggcatctag ccttgaagtt    720
ggttcaactg cacagcccac aggggatgca tctatggatg actgtaatct acgatcagtg    780
ctgacaattg cctttcagtt cccttatgaa atgcatctcc aagacagcgt tgcaactatg    840
gcccggcaat atgtccgcag tattgtttcc tctgttcaga gagtatcaat ggctatttct    900
ccttctcggt ctggcttgaa tgctgggcag aagataattt caggcttccc tgaagcccca    960
acgctagctc gttggatttg ccaaagctac cagttccatt tgggggtgga gttacttagg   1020
caggcagatg atgctgggga agcactattg aaaatgctat gggattacga agacgctatt   1080
ttgtgctgtt ctttcaagga aaagcctgtg tttacttttg ccaacgagat gggactaaac   1140
atgctagaaa catctctcgt cgctctccaa gatctctcac tggacaagat atttgatgaa   1200
gccggtagga aggccctata caacgagatc ccgaaattga tggaacaggg ttacgtgtac   1260
ctgcctggtg gagtgtgctt gtccgggatg gggcgccatg tttctttcga gcaagctgta   1320
gcatggaagg tgctcggaga agacaacaat gtgcactgcc tcgccttctg cttcgtcaac   1380
tggtccttcg tgtga                                                    1395
<210> 197
<211> 464
<212> PRT
<213> Oryza sativa
<400> 197
Ala Gln Glu Thr Ser Gly Glu Val Val Tyr Ala Leu Gly Arg Gln Pro
1               5                  10                  15
Ala Val Leu Arg Thr Phe Ser Gln Arg Leu Ser Arg Gly Phe Asn Asp
            20                  25                  30
Ala Ile Ser Gly Phe Asn Asp Asp Gly Trp Ser Val Met Gly Gly Asp
        35                  40                  45
Gly Ile Glu Asp Val Ile Ile Ala Cys Asn Ala Lys Lys Val Arg Asn
    50                  55                  60
Thr Ser Thr Ser Ala Asn Ala Phe Val Thr Pro Gly Gly Val Ile Cys
65                  70                  75                  80
Ala Lys Ala Ser Met Leu Leu Gln Ser Val Pro Pro Ala Val Leu Val
                85                  90                  95
Arg Phe Leu Arg Glu His Arg Ser Glu Trp Ala Asp Tyr Asn Phe Asp
            100                 105                 110
Ala Tyr Ser Ala Ser Ser Leu Lys Thr Ser Ser Cys Ser Leu Pro Gly
        115                 120                 125
Leu Arg Pro Met Arg Phe Ser Gly Ser Gln Ile Ile Met Pro Leu Ala
    130                 135                 140
His Thr Val Glu Asn Glu Glu Ile Leu Glu Val Val Arg Leu Glu Gly
145                 150                 155                 160
Gln Ala Leu Thr His Asp Asp Gly Leu Met Ser Arg Asp Ile His Leu
                165                 170                 175
Leu Gln Leu Cys Thr Gly Ile Asp Glu Lys Ser Met Gly Ser Cys Phe
            180                 185                 190
Gln Leu Val Ser Ala Pro Ile Asp Glu Leu Phe Pro Asp Asp Ala Pro
        195                 200                 205
Leu Ile Ser Ser Gly Phe Arg Val Ile Pro Leu Asp Met Lys Thr Asp
    210                 215                 220
Gly Thr Pro Ala Gly Arg Thr Leu Asp Leu Ala Ser Ser Leu Glu Val
225                 230                 235                 240
Gly Ser Thr Ala Gln Pro Thr Gly Asp Ala Ser Met Asp Asp Cys Asn
                245                 250                 255
Leu Arg Ser Val Leu Thr Ile Ala Phe Gln Phe Pro Tyr Glu Met His
            260                 265                 270
Leu Gln Asp Ser Val Ala Thr Met Ala Arg Gln Tyr Val Arg Ser Ile
        275                 280                 285
Val Ser Ser Val Gln Arg Val Ser Met Ala Ile Ser Pro Ser Arg Ser
    290                 295                 300
Gly Leu Asn Ala Gly Gln Lys Ile Ile Ser Gly Phe Pro Glu Ala Pro
305                 310                 315                 320
```

```
Thr Leu Ala Arg Trp Ile Cys Gln Ser Tyr Gln Phe His Leu Gly Val
                325                 330                 335
Glu Leu Leu Arg Gln Ala Asp Asp Ala Gly Glu Ala Leu Leu Lys Met
            340                 345                 350
Leu Trp Asp Tyr Glu Asp Ala Ile Leu Cys Cys Ser Phe Lys Glu Lys
        355                 360                 365
Pro Val Phe Thr Phe Ala Asn Glu Met Gly Leu Asn Met Leu Glu Thr
    370                 375                 380
Ser Leu Val Ala Leu Gln Asp Leu Ser Leu Asp Lys Ile Phe Asp Glu
385                 390                 395                 400
Ala Gly Arg Lys Ala Leu Tyr Asn Glu Ile Pro Lys Leu Met Glu Gln
                405                 410                 415
Gly Tyr Val Tyr Leu Pro Gly Gly Val Cys Leu Ser Gly Met Gly Arg
            420                 425                 430
His Val Ser Phe Glu Gln Ala Val Ala Trp Lys Val Leu Gly Glu Asp
        435                 440                 445
Asn Asn Val His Cys Leu Ala Phe Cys Phe Val Asn Trp Ser Phe Val
    450                 455                 460
<210> 198
<211> 3000
<212> DNA
<213> Oryza sativa
<400> 198
gacgtcaggt tatttttttt tctcctcgtc tcctgacgat ttcgcttttg ctcgaaatct     60
cccaggtttg ttgttgttgt gttgaaggcg gagaagaggg gaggctttgg tggtggtgga    120
ggaggaggat ggcggcggcg gtggcgatgc ggagcggcag cggcagcgac ggcggcggcg    180
gcgggtacga caaggccggg atggactccg gcaagtacgt gcggtacacg ccggagcagg    240
tggaggcgct ggagagggtg tacgccgagt gccccaagcc cagctcctcc cgccgccagc    300
agctgctccg cgactgtccc atcctcgcca acatcgagcc caagcagatc aaggtctggt    360
tccagaacag aaggtgccga gataagcagc ggaaggaggc atcaaggctt caggccgtga    420
accgaaaatt gacggcgatg aataagcttc tcatggagga gaatgagcgt cttcagaagc    480
aggtctccca gctggtccat gagaacgcgt acatgaagca gcaacttcag aatccgtcat    540
tgggcaatga tacaagctgt gaatcaaatg tgaccactcc tcagaaccct ctgagagatg    600
caagtaaccc gtctggactc cttacaattg cggaggagac cctgacagag ttcctctcca    660
aggctacagg gactgctgtt gattgggtgc caatgcctgg gatgaagcct ggtccggatt    720
cgtttggtat tgtggccgtt tcacatggtt gccgtggtgt tgctgcccgt gcctgtggtt    780
tggtgaatct agaaccaaca aagatcgtgg agatcttaaa agaccgccca tcttggttcc    840
gtgattgtcg aagtcttgaa gtcttcacaa tgtttccagc tggaaatggt ggcacgatcg    900
aacttgttta catgcagatg tatgctccta ctactttggt tcctgcacga gatttttgga    960
cacttagata cacaactaca atggaggatg gcagccttgt ggtctgtgag agatcattga   1020
gtggttctgg aggtggtcca agtacagcct ccgcacagca atttgtaaga gctgagatgc   1080
ttcctagcgg ctatctagtg cgcccatgcg agggtggtgg ctccatcgtg catattgtgg   1140
accatctgga tcttgaggct tggagtgttc cagaagtgct tcggccactc tacgagtcat   1200
ctagggtagt tgctcagaaa atgactactg cagcactacg gcacatcaga caaattgctc   1260
aagagacaag cggggaggtt gtatacgctt tggggaggca acctgctgtt ttgcggacat   1320
ttagtcagag gttgagtaga ggcttcaatg atgctataag tggtttcagt gatgatggtt   1380
ggtctgtcat gggtggggat ggcattgaag atgtgatcat tgcttgcaat gcaaagaagg   1440
ttaggaatac tagcacttcg gccaatgctt ttgtaactcc aggaggtgtt atatgtgcta   1500
aggcatccat gctactgcag agtgtcccac ctgcagtttt ggttcgattt ttgagggaac   1560
atcgttctga atgggcggat tataacttcg atgcatattc agcttcatct ctgaagacaa   1620
gctcatgttc acttcctggg ttgcggccta tgagattttc tgggagccag atcattatgc   1680
cacttgctca cacggtggag aatgaggaga ttttagaagt tgtccgtctt gaaggacaag   1740
cacttacaca tgatgatggt cttatgtcta gagatattca cctgcttcag ctttgcactg   1800
gaatagatga gaaatcaatg ggatcctgct tccagcttgt ctctgcacca atcgatgagc   1860
ttttccctga tgatgctccg ttaatatctt caggctttcg tgttataccg ctggacatga   1920
aaacagatgg tacacctgct ggtagaacat tagatttggc atctagcctt gaagttggtt   1980
caactgcaca gcccacaggg gatgcatcta tggatgactg taatctacga tcagtgctga   2040
caattgcctt tcagttccct tatgaaatgc atctccaaga cagcgttgca actatggccc   2100
ggcaatatgt ccgcagtatt gtttcctctg ttcagagagt atcaatggct atttctcctt   2160
ctcggtctgg cttgaatgct gggcagaaga taatttcagg cttccctgaa gccccaacgc   2220
tagctcgttg gatttgccaa agctaccagt tccatttggg ggtggagtta cttaggcagg   2280
cagatgatgc tggggaagca ctattgaaaa tgctatggga ttacgaagac gctattttgt   2340
gctgttcttt caaggaaaag cctgtgttta cttttgccaa cgagatggga ctaaacatgc   2400
```

tagaaacatc tctcgtcgct ctccaagatc tctcactgga caagatattt gatgaagccg 2460
gtaggaaggc cctatacaac gagatcccga aattgatgga acagggttac gtgtacctgc 2520
ctggtggagt gtgcttgtcc gggatggggc gccatgtttc tttcgagcaa gctgtagcat 2580
ggaaggtgct cggagaagac aacaatgtgc actgcctcgc cttctgcttc gtcaactggt 2640
ccttcgtgtg acccaaaatc caatccacca tgctcgcagt catcgtcgtt gtcgtcagat 2700
ctccattttt tgcaatctct gtagaagaga ggacaccaaa ccaaccacaa acctgtcagc 2760
tgttagctaa tgattctact agcttttcta gatctttgaa ggcaaaatgc tgccggtgcc 2820
tgtaaagagt gtgtgcgtgc gtgtgtaaat ttgggcgcgt tttgcactaa tttgatctgg 2880
ggtaaggagg cgctggctgc tgctgtgctg tagtttggta aaagttgagt acgctaacct 2940
tggatggtct cgaactgtat gatgaaattc ctagcagtaa aatttcaact tggatccgcg 3000

<210> 199
<211> 840
<212> PRT
<213> Oryza sativa
<400> 199

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

```
Val Tyr Ala Leu Gly Arg Gln Pro Ala Val Leu Arg Thr Phe Ser Gln
385                 390                 395                 400
Arg Leu Ser Arg Gly Phe Asn Asp Ala Ile Ser Gly Phe Asn Asp Asp
                405                 410                 415
Gly Trp Ser Val Met Gly Gly Asp Gly Ile Glu Asp Val Ile Ile Ala
            420                 425                 430
Cys Asn Ala Lys Lys Val Arg Asn Thr Ser Thr Ser Ala Asn Ala Phe
        435                 440                 445
Val Thr Pro Gly Gly Val Ile Cys Ala Lys Ala Ser Met Leu Leu Gln
    450                 455                 460
Ser Val Pro Pro Ala Val Leu Val Arg Phe Leu Arg Glu His Arg Ser
465                 470                 475                 480
Glu Trp Ala Asp Tyr Asn Phe Asp Ala Tyr Ser Ala Ser Ser Leu Lys
                485                 490                 495
Thr Ser Ser Cys Ser Leu Pro Gly Leu Arg Pro Met Arg Phe Ser Gly
            500                 505                 510
Ser Gln Ile Ile Met Pro Leu Ala His Thr Val Glu Asn Glu Glu Ile
        515                 520                 525
Leu Glu Val Val Arg Leu Glu Gly Gln Ala Leu Thr His Asp Asp Gly
    530                 535                 540
Leu Met Ser Arg Asp Ile His Leu Leu Gln Leu Cys Thr Gly Ile Asp
545                 550                 555                 560
Glu Lys Ser Met Gly Ser Cys Phe Gln Leu Val Ser Ala Pro Ile Asp
                565                 570                 575
Glu Leu Phe Pro Asp Asp Ala Pro Leu Ile Ser Ser Gly Phe Arg Val
            580                 585                 590
Ile Pro Leu Asp Met Lys Thr Asp Gly Thr Pro Ala Gly Arg Thr Leu
        595                 600                 605
Asp Leu Ala Ser Ser Leu Glu Val Gly Ser Thr Ala Gln Pro Thr Gly
    610                 615                 620
Asp Ala Ser Met Asp Asp Cys Asn Leu Arg Ser Val Leu Thr Ile Ala
625                 630                 635                 640
Phe Gln Phe Pro Tyr Glu Met His Leu Gln Asp Ser Val Ala Thr Met
                645                 650                 655
Ala Arg Gln Tyr Val Arg Ser Ile Val Ser Ser Val Gln Arg Val Ser
            660                 665                 670
Met Ala Ile Ser Pro Ser Arg Ser Gly Leu Asn Ala Gly Gln Lys Ile
        675                 680                 685
Ile Ser Gly Phe Pro Glu Ala Pro Thr Leu Ala Arg Trp Ile Cys Gln
    690                 695                 700
Ser Tyr Gln Phe His Leu Gly Val Glu Leu Leu Arg Gln Ala Asp Asp
705                 710                 715                 720
Ala Gly Glu Ala Leu Leu Lys Met Leu Trp Asp Tyr Glu Asp Ala Ile
                725                 730                 735
Leu Cys Cys Ser Phe Lys Glu Lys Pro Val Phe Thr Phe Ala Asn Glu
            740                 745                 750
Met Gly Leu Asn Met Leu Glu Thr Ser Leu Val Ala Leu Gln Asp Leu
        755                 760                 765
Ser Leu Asp Lys Ile Phe Asp Glu Ala Gly Arg Lys Ala Leu Tyr Asn
    770                 775                 780
Glu Ile Pro Lys Leu Met Glu Gln Gly Tyr Val Tyr Leu Pro Gly Gly
785                 790                 795                 800
Val Cys Leu Ser Gly Met Gly Arg His Val Ser Phe Glu Gln Ala Val
                805                 810                 815
Ala Trp Lys Val Leu Gly Glu Asp Asn Asn Val His Cys Leu Ala Phe
            820                 825                 830
Cys Phe Val Asn Trp Ser Phe Val
        835                 840
<210> 200
<211> 3252
<212> DNA
<213> Oryza sativa
<400> 200
```

ctctctcccg tgctgctctc ccttctcctc tcctcaatta ccacgccgca gccgcagccc 60
cccaacccta gctagtacgg cgggcgacct gaggggggtag ctgcctacct ctatacctgc 120
tatctccgtt cgatccgggt cgggctcgtg ctgctggggg cgggggcaat caatcggcga 180
gctcggtgat ccatggccgt ggctgtgggg tggctgtgag ggtgcccccg gcggcgctcc 240
cctccgcgcc tgccggcgag ggggctcgga ctgaagggat ctaggcgagc tgaaaattga 300
agtgcaggca aggagataag agcagcgtcc aaattgtgag tacttcatta gcaggaggta 360
gtggttgtgc ttgcttggct cctttgcaat ttggctttgg cgaggtagca atggctgcgg 420
cagtggcaat gcgagggagt agcagtgatg gaggtggcta tgataaggtt tccgggatgg 480
actccggtaa atatgtgcgc tacacgcctg agcaggtgga ggcgcttgag cgggtgtacg 540
ccgattgccc caagccaacc tcctcccgca ggcagcaact gctgcgtgag tgccccatac 600
ttgctaacat tgagcccaag cagatcaagg tctggttcca gaacagaagg tgccgggata 660
agcagcggaa ggagtcttca cggcttcagg ctgtcaacag gaaattgacg gcaatgaaca 720
agctacttat ggaagagaat gagcgactcc agaagcaggt ctcccaattg gttcatgaga 780
atgcccacat gcgacagcag ctgcagaata ctccgctggc aaatgataca agctgtgaat 840
caaatgtgac tacccctcaa aacccttaa gggatgcaag taacccctct gggctccttt 900
caattgcaga ggagaccttg acagagttcc tctcaaaggc tactggtaca gctattgatt 960
gggtccagat gcctgggatg aagcctggtc cggattcggt tggtattgtg gccatttcac 1020
atggttgccg tggtgttgct gcccgtgcct gtggtttggt gaacctagaa ccaacaaaag 1080
tggtagagat attgaaagat cgtccatctt ggttccgtga ttgtcgaaac ctggaagtct 1140
ttacaatgat tccagcagga aatggaggaa cggttgaact tgtctacaca cagttgtatg 1200
ctccaacaac tttagttcct gcacgagatt tttggacgtt acggtacaca accacaatgg 1260
aagatggcag tcttgtggtc tgtgagagat ctttaagtgg ttcaggggc ggtccaagtg 1320
ctgcctctgc tcagcaatat gtgagagcgg aaatgcttcc aagtggatac ctggttcgcc 1380
catgtgaagg tgggggatca attgtgcaca tagtggacca tctggatctt gaggcatgga 1440
gtgttcctga ggtgcttcgg ccactctatg aatcttcaag ggtagtcgct cagaaaatga 1500
ctactgcggc actccggcac atcagacaaa ttgctcaaga aacaagtggg gaagtggtgt 1560
atgccttggg gaggcaacca gcagtgctac ggacttttag tcaaaggctg agcagaggct 1620
ttaacgatgc cattagtggt ttcaatgatg atgggtggtc tataatgggc ggagacggtg 1680
ttgaagatgt agttattgct tgcaactcaa ctaagaaaat taggagtaac agcaatgctg 1740
gcatcgcctt tggagccccc ggaggtatta tatgtgctaa ggcatcaatg ttactgcaga 1800
gtgttcctcc agcagtactg gttcgatttc tgagggagca tagatctgaa tgggccgatt 1860
acaatattga tgcatatttg gcttcaactc tgaaaacaag tgcatgttca cttactgggt 1920
tgcgacccat gagattttct gggagccaaa tcatcattcc acttgctcac acagttgaga 1980
atgaggagat tcttgaagtt gttcgccttg agggtcaacc tcttactcat gatgaagctc 2040
ttctttcaag ggatatccac ctgcttcagc tctgcactgg aatagacgag aagtctgtgg 2100
gatcctcctt tcagcttgtg tttgcaccga ttgatgattt cccagatgaa actccattga 2160
tttcttctgg cttccgtgtt ataccacttg atatgaaaac agatggtgca tcctctggta 2220
ggacattaga tttggcatct agtcttgaag taggttcagc aacagctcaa gcctccggag 2280
atgcatctgc agatgattgt aacttgcgat ctgttctgac gatcgctttt caattccctt 2340
acgagttgca tctccaagac agtgttgcag ctatggctcg ccaatatgtc cgtagcattg 2400
tttctgctgt gcaaagagtg tcaatggcta tctctccctc tcaaactggt ctaaatgccg 2460
gacagaggat aatctctggt ttccctgaag cagcaaccct tgctcgatgg gtttgccaga 2520
gctaccatta ccatctaggg gtagagttac ttagtcaatc agatggagat gcagaacaat 2580
tgttgaagat gctatggcat taccaagatg ctattttgtg ctgctcattc aaggaaaaac 2640
cggtgtttac atttgccaac aaagcaggac tggacatgct agaaacttcc cttgtcgcct 2700
tacaggacct cacattggac aggatctttg atgagcctgg aaaagaagca ttgttctcaa 2760
acattcccaa attgatggag cagggccatg tctacctgcc atcaggcgtg tgcatgtcag 2820
gaatgggtcg gcatgtttct ttcgatcagg ccgtggcttg gaaagtgctt gccgaggata 2880
gcaatgttca ctgcctggcc ttctgtttcg tcaactggtc ctttgtgtga ccatcccaca 2940
tccactgcga agtgaagatt cattttgct tgttcaagac tgttttgcac tagatctaga 3000
cctgagaatc tagtagtatt tgcgaaattt ctcgcttatg taaatgtaat ctcgctccca 3060
ttccatcaag tacccaagta caaggcagtg gccaagtggt tttagttgta gggggatcgg 3120
agcaatcctg acggttgggt acttgggttc cttccagcaa tgtaaaatcg gatcctgtag 3180
ggcgtcgaaa actgcattgg caagattctt cgaaatgcag accaaattta gctagtacat 3240
cgtaactttg gc 3252

<210> 201
<211> 839
<212> PRT
<213> Oryza sativa
<400> 201

Met Ala Ala Ala Val Ala Met Arg Gly Ser Ser Ser Asp Gly Gly Gly
1 5 10 15
Tyr Asp Lys Val Ser Gly Met Asp Ser Gly Lys Tyr Val Arg Tyr Thr

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

Glu Val Val Arg Leu Glu Gly Gln Pro Leu Thr His Asp Glu Ala Leu
530 535 540
Leu Ser Arg Asp Ile His Leu Leu Gln Leu Cys Thr Gly Ile Asp Glu
545 550 555 560
Lys Ser Val Gly Ser Ser Phe Gln Leu Val Phe Ala Pro Ile Asp Asp
565 570 575
Phe Pro Asp Glu Thr Pro Leu Ile Ser Ser Gly Phe Arg Val Ile Pro
580 585 590
Leu Asp Met Lys Thr Asp Gly Ala Ser Ser Gly Arg Thr Leu Asp Leu
595 600 605
Ala Ser Ser Leu Glu Val Gly Ser Ala Thr Ala Gln Ala Ser Gly Asp
610 615 620
Ala Ser Ala Asp Asp Cys Asn Leu Arg Ser Val Leu Thr Ile Ala Phe
625 630 635 640
Gln Phe Pro Tyr Glu Leu His Leu Gln Asp Ser Val Ala Ala Met Ala
645 650 655
Arg Gln Tyr Val Arg Ser Ile Val Ser Ala Val Gln Arg Val Ser Met
660 665 670
Ala Ile Ser Pro Ser Gln Thr Gly Leu Asn Ala Gly Gln Arg Ile Ile
675 680 685
Ser Gly Phe Pro Glu Ala Ala Thr Leu Ala Arg Trp Val Cys Gln Ser
690 695 700
Tyr His Tyr His Leu Gly Val Glu Leu Leu Ser Gln Ser Asp Gly Asp
705 710 715 720
Ala Glu Gln Leu Leu Lys Met Leu Trp His Tyr Gln Asp Ala Ile Leu
725 730 735
Cys Cys Ser Phe Lys Glu Lys Pro Val Phe Thr Phe Ala Asn Lys Ala
740 745 750
Gly Leu Asp Met Leu Glu Thr Ser Leu Val Ala Leu Gln Asp Leu Thr
755 760 765
Leu Asp Arg Ile Phe Asp Glu Pro Gly Lys Glu Ala Leu Phe Ser Asn
770 775 780
Ile Pro Lys Leu Met Glu Gln Gly His Val Tyr Leu Pro Ser Gly Val
785 790 795 800
Cys Met Ser Gly Met Gly Arg His Val Ser Phe Asp Gln Ala Val Ala
805 810 815
Trp Lys Val Leu Ala Glu Asp Ser Asn Val His Cys Leu Ala Phe Cys
820 825 830
Phe Val Asn Trp Ser Phe Val
835

<210> 202
<211> 3169
<212> DNA
<213> Arabidopsis thaliana
<400> 202

agagtcttca aaacttttgc agcttcaatt gtacctgggt ttcttcttca ttgttcctaa 60
ggtttctgtg tccttcaatt cttctgatat aatgcttctt taagagagtt gacatcatca 120
ctttcttggg gtactcttct ctgtttctcc ccagaaaatc caactctgta attttgggtc 180
tttattctgt ttttctcttt gaagaatctt taaaattctc agatcttctg aatctctctt 240
ctttaaaact ttttttaact ttatttttg tactcgcttc tttgccttca tttttctcgt 300
atccacatgt cgttggtctt tcgctacaag ccacgaccgt agaatcttct tttgtctgaa 360
aagaattaca atttacgttt ctcttacgat acgacggact ttccgaagaa attaatttaa 420
agagaaaaga agaagaagcc aaagaagaag aagaagctag aagaaacagt aaagtttgag 480
actttttttg agggtcgagc taaaatggag atggcggtgg ctaaccaccg tgagagaagc 540
agtgacagta tgaatagaca tttagatagt agcggtaagt acgttaggta cacagctgag 600
caagtcgagg ctcttgagcg tgtctacgct gagtgtccta agcctagctc tctccgtcga 660
caacaattga tccgtgaatg ttccattttg gccaatattg agcctaagca gatcaaagtc 720
tggtttcaga accgcaggtg tcgagataag cagaggaaag aggcgtcgag gctccagagc 780
gtaaaccgga agctctctgc gatgaataaa ctgttgatgg aggagaatga taggttgcag 840
aagcaggttt ctcagcttgt ctgcgaaaat ggatatatga aacagcagct aactactgtt 900
gttaacgatc caagctgtga atctgtggtc acaactcctc agcattcgct tagagatgcg 960
aatagtcctg ctggattgct ctcaatcgca gaggagactt tggcagagtt cctatccaag 1020
gctacaggaa ctgctgttga ttgggttcag atgcctggga tgaagcctgg tccggattcg 1080

```
gttggcatct ttgccatttc gcaaagatgc aatggagtgg cagctcgagc ctgtggtctt   1140
gttagcttag aacctatgaa gattgcagag atcctcaaag atcggccatc ttggttccgt   1200
gactgtagga gccttgaagt tttcactatg ttcccggctg gtaatggtgg cacaatcgag   1260
cttgtttata tgcagacgta tgcaccaacg actctggctc ctgcccgcga tttctggacc   1320
ctgagataca caacgagcct cgacaatggg agttttgtgg tttgtgagag gtcgctatct   1380
ggctctggag ctgggcctaa tgctgcttca gcttctcagt ttgtgagagc agaaatgctt   1440
tctagtgggt atttaataag gccttgtgat ggtggtggtt ctattattca cattgtcgat   1500
caccttaatc ttgaggcttg gagtgttccg gatgtgcttc gaccccttta tgagtcatcc   1560
aaagtcgttg cacaaaaaat gaccatttcc gcgttgcggt atatcaggca attagcccaa   1620
gagtctaatg gtgaagtagt gtatggatta ggaaggcagc ctgctgttct tagaaccttt   1680
agccaaagat taagcagggg cttcaatgat gcggttaatg ggtttggtga cgacgggtgg   1740
tctacgatgc attgtgatgg agcggaagat attatcgttg ctattaactc tacaaagcat   1800
ttgaataata tttctaattc tctttcgttc cttggaggcg tgctctgtgc caaggcttca   1860
atgcttctcc aaaatgttcc tcctgcggtt ttgatccggt tccttagaga gcatcgatct   1920
gagtgggctg atttcaatgt tgatgcatat tccgctgcta cacttaaagc tggtagcttt   1980
gcttatccgg gaatgagacc aacaagattc actgggagtc agatcataat gccactagga   2040
catacaattg aacacgaaga aatgctagaa gttgttagac tggaaggtca ttctcttgct   2100
caagaagatg catttatgtc acgggatgtc catctccttc agatttgtac cgggattgac   2160
gagaatgccg ttggagcttg ttctgaactg atatttgctc cgattaatga gatgttcccg   2220
gatgatgctc cacttgttcc ctctggattc cgagtcatac ccgttgatgc taaaacggga   2280
gatgtacaag atctgttaac cgctaatcac cgtacactag acttaacttc tagccttgaa   2340
gtcggtccat cacctgagaa tgcttctgga aactcttttt ctagctcaag ctcgagatgt   2400
attctcacta tcgcgtttca attccctttt gaaaacaact tgcaagaaaa tgttgctggt   2460
atggcttgtc agtatgtgag gagcgtgatc tcatcagttc aacgtgttgc aatggcgatc   2520
tcaccgtctg ggataagccc gagtctgggc tccaaattgt ccccaggatc tcctgaagct   2580
gttactcttg ctcagtggat ctctcaaagt tacagtcatc acttaggctc ggagttgctg   2640
acgattgatt cacttggaag cgacgactcg gtactaaaac ttctatggga tcaccaagat   2700
gccatcctgt gttgctcatt aaagccacag ccagtgttca tgtttgcgaa ccaagctggt   2760
ctagacatgc tagagacaac acttgtagcc ttacaagata taacactcga aaagatattc   2820
gatgaatcgg gtcgtaaggc tatctgttcg gacttcgcca agctaatgca acagggattt   2880
gcttgcttgc cttcaggaat ctgtgtgtca acgatgggaa gacatgtgag ttatgaacaa   2940
gctgttgctt ggaaagtgtt tgctgcatct gaagaaaaca acaacaatct gcattgtctt   3000
gccttctcct ttgtaaactg gtcttttgtg tgattcgatt gacagaaaaa gactaattta   3060
aatttacgtt agagaactca aatttttggt tgttgtttag gtgtctctgt tttgtttttt   3120
aaaattattt tgatcaaatg ttactcactt tcttctttca aaaaaaaaa               3169
<210> 203
<211> 842
<212> PRT
<213> Arabidopsis thaliana
<400> 203
Met Glu Met Ala Val Ala Asn His Arg Glu Arg Ser Ser Asp Ser Met
1               5                  10                  15
Asn Arg His Leu Asp Ser Ser Gly Lys Tyr Val Arg Tyr Thr Ala Glu
           20                  25                  30
Gln Val Glu Ala Leu Glu Arg Val Tyr Ala Glu Cys Pro Lys Pro Ser
       35                  40                  45
Ser Leu Arg Arg Gln Gln Leu Ile Arg Glu Cys Ser Ile Leu Ala Asn
   50                  55                  60
Ile Glu Pro Lys Gln Ile Lys Val Trp Phe Gln Asn Arg Arg Cys Arg
65                  70                  75                  80
Asp Lys Gln Arg Lys Glu Ala Ser Arg Leu Gln Ser Val Asn Arg Lys
                85                  90                  95
Leu Ser Ala Met Asn Lys Leu Leu Met Glu Glu Asn Asp Arg Leu Gln
           100                 105                 110
Lys Gln Val Ser Gln Leu Val Cys Glu Asn Gly Tyr Met Lys Gln Gln
       115                 120                 125
Leu Thr Thr Val Val Asn Asp Pro Ser Cys Glu Ser Val Val Thr Thr
   130                 135                 140
Pro Gln His Ser Leu Arg Asp Ala Asn Ser Pro Ala Gly Leu Leu Ser
145                 150                 155                 160
Ile Ala Glu Glu Thr Leu Ala Glu Phe Leu Ser Lys Ala Thr Gly Thr
               165                 170                 175
Ala Val Asp Trp Val Gln Met Pro Gly Met Lys Pro Gly Pro Asp Ser
```

```
          180                 185                 190
Val Gly Ile Phe Ala Ile Ser Gln Arg Cys Asn Gly Val Ala Ala Arg
       195                 200                 205
Ala Cys Gly Leu Val Ser Leu Glu Pro Met Lys Ile Ala Glu Ile Leu
   210                 215                 220
Lys Asp Arg Pro Ser Trp Phe Arg Asp Cys Arg Ser Leu Glu Val Phe
225                 230                 235                 240
Thr Met Phe Pro Ala Gly Asn Gly Gly Thr Ile Glu Leu Val Tyr Met
                245                 250                 255
Gln Thr Tyr Ala Pro Thr Thr Leu Ala Pro Ala Arg Asp Phe Trp Thr
            260                 265                 270
Leu Arg Tyr Thr Thr Ser Leu Asp Asn Gly Ser Phe Val Val Cys Glu
        275                 280                 285
Arg Ser Leu Ser Gly Ser Gly Ala Gly Pro Asn Ala Ala Ser Ala Ser
    290                 295                 300
Gln Phe Val Arg Ala Glu Met Leu Ser Ser Gly Tyr Leu Ile Arg Pro
305                 310                 315                 320
Cys Asp Gly Gly Gly Ser Ile Ile His Ile Val Asp His Leu Asn Leu
                325                 330                 335
Glu Ala Trp Ser Val Pro Asp Val Leu Arg Pro Leu Tyr Glu Ser Ser
            340                 345                 350
Lys Val Val Ala Gln Lys Met Thr Ile Ser Ala Leu Arg Tyr Ile Arg
        355                 360                 365
Gln Leu Ala Gln Glu Ser Asn Gly Glu Val Val Tyr Gly Leu Gly Arg
    370                 375                 380
Gln Pro Ala Val Leu Arg Thr Phe Ser Gln Arg Leu Ser Arg Gly Phe
385                 390                 395                 400
Asn Asp Ala Val Asn Gly Phe Gly Asp Asp Gly Trp Ser Thr Met His
                405                 410                 415
Cys Asp Gly Ala Glu Asp Ile Ile Val Ala Ile Asn Ser Thr Lys His
            420                 425                 430
Leu Asn Asn Ile Ser Asn Ser Leu Ser Phe Leu Gly Gly Val Leu Cys
        435                 440                 445
Ala Lys Ala Ser Met Leu Leu Gln Asn Val Pro Pro Ala Val Leu Ile
    450                 455                 460
Arg Phe Leu Arg Glu His Arg Ser Glu Trp Ala Asp Phe Asn Val Asp
465                 470                 475                 480
Ala Tyr Ser Ala Ala Thr Leu Lys Ala Gly Ser Phe Ala Tyr Pro Gly
                485                 490                 495
Met Arg Pro Thr Arg Phe Thr Gly Ser Gln Ile Ile Met Pro Leu Gly
            500                 505                 510
His Thr Ile Glu His Glu Glu Met Leu Glu Val Val Arg Leu Glu Gly
        515                 520                 525
His Ser Leu Ala Gln Glu Asp Ala Phe Met Ser Arg Asp Val His Leu
    530                 535                 540
Leu Gln Ile Cys Thr Gly Ile Asp Glu Asn Ala Val Gly Ala Cys Ser
545                 550                 555                 560
Glu Leu Ile Phe Ala Pro Ile Asn Glu Met Phe Pro Asp Asp Ala Pro
                565                 570                 575
Leu Val Pro Ser Gly Phe Arg Val Ile Pro Val Asp Ala Lys Thr Gly
            580                 585                 590
Asp Val Gln Asp Leu Leu Thr Ala Asn His Arg Thr Leu Asp Leu Thr
        595                 600                 605
Ser Ser Leu Glu Val Gly Pro Ser Pro Glu Asn Ala Ser Gly Asn Ser
    610                 615                 620
Phe Ser Ser Ser Ser Ser Arg Cys Ile Leu Thr Ile Ala Phe Gln Phe
625                 630                 635                 640
Pro Phe Glu Asn Asn Leu Gln Glu Asn Val Ala Gly Met Ala Cys Gln
                645                 650                 655
Tyr Val Arg Ser Val Ile Ser Ser Val Gln Arg Val Ala Met Ala Ile
            660                 665                 670
Ser Pro Ser Gly Ile Ser Pro Ser Leu Gly Ser Lys Leu Ser Pro Gly
        675                 680                 685
```

```
Ser Pro Glu Ala Val Thr Leu Ala Gln Trp Ile Ser Gln Ser Tyr Ser
   690                 695                 700
His His Leu Gly Ser Glu Leu Leu Thr Ile Asp Ser Leu Gly Ser Asp
705                 710                 715                 720
Asp Ser Val Leu Lys Leu Leu Trp Asp His Gln Asp Ala Ile Leu Cys
                725                 730                 735
Cys Ser Leu Lys Pro Gln Pro Val Phe Met Phe Ala Asn Gln Ala Gly
            740                 745                 750
Leu Asp Met Leu Glu Thr Thr Leu Val Ala Leu Gln Asp Ile Thr Leu
        755                 760                 765
Glu Lys Ile Phe Asp Glu Ser Gly Arg Lys Ala Ile Cys Ser Asp Phe
   770                 775                 780
Ala Lys Leu Met Gln Gln Gly Phe Ala Cys Leu Pro Ser Gly Ile Cys
785                 790                 795                 800
Val Ser Thr Met Gly Arg His Val Ser Tyr Glu Gln Ala Val Ala Trp
                805                 810                 815
Lys Val Phe Ala Ala Ser Glu Glu Asn Asn Asn Asn Leu His Cys Leu
            820                 825                 830
Ala Phe Ser Phe Val Asn Trp Ser Phe Val
        835                 840
<210> 204
<211> 2700
<212> DNA
<213> Zea mays
<400> 204
gagagctaag agcaacaagg cgacgagatt tgtgtggcta cgattttgcg ttgccgacgg     60
aacatgggaa caatggttgc agcggtggcg ttgcgagggg gaagcagcga tagcggagga    120
tttgataagg ttcccgggat ggactcggga aaatacgtgc gctacacccc ggagcaagtt    180
gaggtgctcg agcggctcta catagattgc cccaagccaa gctcctcacg gcggcagcaa    240
ctgctgcgcg agtgtcctat actctccaac attgagccaa agcagatcaa ggtctggttc    300
cagaaccgga ggtgccgcga taagcagcgg aaggagtctt cgcggcttca ggctgtgaac    360
agaaagttga cggcaatgaa caagcttcta atggaagaga atgagcggct tcagaagcaa    420
gtctctcagt tggttcatga aaacgcgcac atgcggcagc agctgcagaa tacttcattg    480
gccaatgaca caagctgtga atcaaatgtc actacccctc caaaccctat aagggacgca    540
agtaaccctt ctggactcct tgcgattgca gaggagacct tcacagagtt cctctcaaag    600
gctactggga cagctattga ttgggtccag atgcctggga tgaagcctgg tccggattca    660
gttggtatcg tggccatttc gcatggttgc cgtggcgttg ctgcccgcgc ctgtggtttg    720
gtgaatctag aaccaacaaa aggcatagag atcttgaaag atcgtccctc ttggttccgt    780
gattgccgaa gtcttgaagt gtttacaagg tttccagctg gaaatggggg aacaattgaa    840
cttatttaca tgcagatgta tgccccaaca actttagtcc ctgcacgtga tttttggaca    900
ctacgataca cgacaacaat ggaagatggc agccttgtgg tctgtgagag atccttgagt    960
ggttctggtg gtggtccaaa tgcagcctct acacaacaat ttgttagggc tgagatgctt   1020
ccaagtgggt atttagttcg cccatgcgaa ggtggaggat caattgtgca tatagtggac   1080
catctagatc tcgaagcatg gagtgttcct gaagtgcttc gaccactgta tgagtcttct   1140
agagttgttg ctcagaaaat gactactgtg gcactgcgcc accttagaca aattgctcaa   1200
gaaacaagtg gagaagtagt gtacgccttg ggaaggcaac ctgcagtcct acggaccttt   1260
agtcaaagac taagcagggg gtttaatgat gccattagcg gtttcaatga tgatggctgg   1320
tctgtaatgg ggggagatgg cattgaagac gttgttattg cttgcaactc aaccaagaaa   1380
attaggaata ccagcaatgc tgggattaca tttggagccc caggaggcat tatttgtgcg   1440
aaggcatcca tgttactgca gagcgtccca ccggcagtac tggtacgatt tctgagggag   1500
catagatctg aatgggctga ttacaatatt gatgcgtatt tggcttcatc attgaagacc   1560
agtgcgtgct cacttcctgg gttgcgaccc atgagatttt ctgaggggca gatgatcatg   1620
ccacttgctc acacagttga gaacgaggag attcttgaag ttgtccgcct tgaaggtcag   1680
cctcttactc atgatgaagc tcttctttca agggacatcc accttctcca gctttgcact   1740
ggaatagatg agaaatctgt gggttcctcc ttccagcttg tgtttgcacc aattgatgag   1800
catttcccag atgatgctcc attgatttct tctggctttc gtgtcatacc acttgatgtg   1860
aaaacagatg gtgtatcctc tggtaggacg ctagatttgg catctagtct tgatgtgggc   1920
tctgctgcac cccaagcctc aggggatgca tctccagatg actgcagttt gagatctgtg   1980
ctgacaatcg cctttcaatt cccgtatgag atgcaccttc aggacagcgt tgcagcaatg   2040
gcccgtcaat atgttcgtag tgtcatttct gctgtgcaaa gagtgtcgat ggctatatct   2100
ccctcccaat ctggtctaaa tgctgggcat aggatgcttt ctggcttccc tgaagctgcc   2160
acacttgcta gatgggtttg ccagagttat cactaccatc taggtatgga attacttaat   2220
caatcagatg gagctggtga agcattgttg aaaatgctct ggcatcatcc agatgctgtt   2280
```

ctgtgctgct cctttaagga gaaacctatg tttacgtttg caaacaaggc agggctggac 2340
atgttagaaa catctcttgt tgccctgcaa gacctgacgc tagacaagat cttcgacgag 2400
tcaggaagga aagcactatt ctcagacatc tcgaaactaa tggaacaggg ctacgcgtac 2460
ctgccgtcag gcgtgtgcat gtcaggaatg ggccgccatg tttctttcga ccaagctgta 2520
gcgtggaagg tgctcggcga ggatagcaac atccactgcc tggcgttctg cttcgtcaac 2580
tggtccttcg tgtgactgac gcccaacccg tcgggtggtt atggcgagct gacccttttt 2640
tgtatggaaa atcatcagct gtgacaaaag gcatatgttg catgcactag ttgaagaaac 2700

<210> 205
<211> 843
<212> PRT
<213> Zea mays
<400> 205

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

```
                405                 410                 415
Asp Gly Trp Ser Val Met Gly Gly Asp Gly Ile Glu Asp Val Val Ile
            420                 425                 430
Ala Cys Asn Ser Thr Lys Lys Ile Arg Asn Thr Ser Asn Ala Gly Ile
        435                 440                 445
Thr Phe Gly Ala Pro Gly Gly Ile Ile Cys Ala Lys Ala Ser Met Leu
    450                 455                 460
Leu Gln Ser Val Pro Pro Ala Val Leu Val Arg Phe Leu Arg Glu His
465                 470                 475                 480
Arg Ser Glu Trp Ala Asp Tyr Asn Ile Asp Ala Tyr Leu Ala Ser Ser
                485                 490                 495
Leu Lys Thr Ser Ala Cys Ser Leu Pro Gly Leu Arg Pro Met Arg Phe
            500                 505                 510
Ser Glu Gly Gln Met Ile Met Pro Leu Ala His Thr Val Glu Asn Glu
        515                 520                 525
Glu Ile Leu Glu Val Val Arg Leu Glu Gly Gln Pro Leu Thr His Asp
    530                 535                 540
Glu Ala Leu Leu Ser Arg Asp Ile His Leu Leu Gln Leu Cys Thr Gly
545                 550                 555                 560
Ile Asp Glu Lys Ser Val Gly Ser Ser Phe Gln Leu Val Phe Ala Pro
                565                 570                 575
Ile Asp Glu His Phe Pro Asp Asp Ala Pro Leu Ile Ser Ser Gly Phe
            580                 585                 590
Arg Val Ile Pro Leu Asp Val Lys Thr Asp Gly Val Ser Ser Gly Arg
        595                 600                 605
Thr Leu Asp Leu Ala Ser Ser Leu Asp Val Gly Ser Ala Ala Pro Gln
    610                 615                 620
Ala Ser Gly Asp Ala Ser Pro Asp Asp Cys Ser Leu Arg Ser Val Leu
625                 630                 635                 640
Thr Ile Ala Phe Gln Phe Pro Tyr Glu Met His Leu Gln Asp Ser Val
                645                 650                 655
Ala Ala Met Ala Arg Gln Tyr Val Arg Ser Val Ile Ser Ala Val Gln
            660                 665                 670
Arg Val Ser Met Ala Ile Ser Pro Ser Gln Ser Gly Leu Asn Ala Gly
        675                 680                 685
His Arg Met Leu Ser Gly Phe Pro Glu Ala Ala Thr Leu Ala Arg Trp
    690                 695                 700
Val Cys Gln Ser Tyr His Tyr His Leu Gly Met Glu Leu Leu Asn Gln
705                 710                 715                 720
Ser Asp Gly Ala Gly Glu Ala Leu Leu Lys Met Leu Trp His His Pro
                725                 730                 735
Asp Ala Val Leu Cys Cys Ser Phe Lys Glu Lys Pro Met Phe Thr Phe
            740                 745                 750
Ala Asn Lys Ala Gly Leu Asp Met Leu Glu Thr Ser Leu Val Ala Leu
        755                 760                 765
Gln Asp Leu Thr Leu Asp Lys Ile Phe Asp Glu Ser Gly Arg Lys Ala
    770                 775                 780
Leu Phe Ser Asp Ile Ser Lys Leu Met Glu Gln Gly Tyr Ala Tyr Leu
785                 790                 795                 800
Pro Ser Gly Val Cys Met Ser Gly Met Gly Arg His Val Ser Phe Asp
                805                 810                 815
Gln Ala Val Ala Trp Lys Val Leu Gly Glu Asp Ser Asn Ile His Cys
            820                 825                 830
Leu Ala Phe Cys Phe Val Asn Trp Ser Phe Val
        835                 840
<210> 206
<211> 2669
<212> DNA
<213> Populus trichocarpa
<400> 206
ctgtgtttga ttatttattt tgtggtggaa atggccatgg aagtggcccc tctgcaccga     60
gagagcagca gtagtgggag cataaacaag caccttactg atgatgggaa gtatgttcgg    120
tacacagcag agcaagtgga ggctcttgaa agagtttatg cggagtgccc taagcccagc    180
```

```
tctttgcgta ggcaacaatt gattagagag tgccccattt tagctaatat tgagccgaag    240
cagatcaaag tctggtttca aaatcgcagg tgtagagaga agcagagaaa agagtcctca    300
agactccaga ctgtgaacag gaaattgaca gcaatgaata agctgttgat ggaggagaat    360
gatcgccttc aaaaacaggt gtcccagctg gtgtgcgaaa atggtttcat gcggcaacaa    420
ctgcaaactg caccaacagc aactgatgca agctgtgact ctgtggttgc cactccacaa    480
cattcactaa gagatgccaa taaccctgct ggactcctct caatagctga ggagaccttg    540
tcagagttcc ttgcaaaggc cacaggaact gctcttgagt gggtccagat gcctggaatg    600
aagcctggtc cggattcgat tgggattttt tccatttcac aaaggtgcgg tggagtggca    660
gcaagagcct gcggtcttgt aagtttagag cctaaaaaga ttgcagagat ccttaaagat    720
cgttcatctt ggttccgtga ttgtcggaac cttgaagttt tcactatgtt tcctgctgga    780
aatggtggaa caattgaact agtgtacagt cagatatatg ctccaactac tcttgctcca    840
gcacgggata tgtggactct gagatacact acaagtttag agaatggcag ccttgtggtc    900
tgtgagagat cactttctgg ttatggtgct ggccccgatg cagctgccgc tgcccagttt    960
gtgagggctg aaatgcttcc tagtggctat ttgataaggc catgtgaagg agggtcaatt   1020
atccacatcg tggatcacct gaatcttcag gcatggagtg tgcctgaggt gcttcgacca   1080
ctttatgaat catcaaaagc agtggcacag aaaatgacca ttgcggcact gcgttatgtc   1140
aggcaagtag ctcatgaaac cagtggtgag gtggtgtatg gtttgggcag gcaaccagct   1200
gttctgagaa cctttaacca aagattaagc agaggtttca atgatgccat caatgggttt   1260
aatgatgatg gctggtcact gatgaatgcg gatggagctg aggatgtaat aattgcagtt   1320
aactcaacca agaatttgat tggtgctaat aattctgctc attctctttc gttcttggga   1380
gggattcttt gtgcgaaagc ttccatgcta ctgcaaaatg tgcatcctgc tgtgctggtc   1440
tgtttcctaa gggagcatca cgctgagtgg gctgatttca gtgttgatgc ctattctgct   1500
gcattgtgga aggctggttc atatgcatat ccaggaatga ggcccatgag gtttactggg   1560
agccaaatca ccatgccact gggccacaca attgaacaag aagacttgct tgaagttatt   1620
cgacttgaag gccattcttt tgcacaagaa gatgcttttg tttcacagga catccatctc   1680
ctacagatat gtagtggaat tgatgagaat gctgttggag cctgttctga actagttttt   1740
gctccaattg atgaaacgtt tccagatgat gctccactgc taccttctgg tttccgcatc   1800
atttctctgg aatcaaaagc aaaggataca caggaggtat tgactacaaa ttgtactctg   1860
gatctaacat caagtcttga agcggggctg gcaataaacc acactgcagt ggatggctca   1920
tcgtgtcata gtttgcgatc agtgttgact attgccttcc agttcccttt tgagagcaat   1980
ctacaggata atgttgccac catggcacgt cagtacgtac gtagtgtgat ttcatctgtc   2040
cagagggttg ccatggccat atctccatca ggattgagtc cagttttggg accaaagcta   2100
tctgcaggat ctcctgaagc tctaaccttg gctcactgga tctgccaaag tcacaggcaa   2160
gttctgctta agttttcatc atgttaccat ttaggagcag aattactgag atctgattct   2220
gtgggtggag attctgtcct gaaacatcta tggcatcatc cagatgcaat tttatgttgt   2280
tcattgaagt cgctgccagt tttcatcttt gccaatcagg cagggcttga catgctggag   2340
acaactctag tggctctaca agatattaca ctggataaga tattcaatga atctggacgc   2400
caggcattgt atacagaatt tgcaaagtta atgcaacagg gatttgcatg cttgcctgct   2460
ggaatctgca tgtcaacaat ggggcgtaat gtttcatatg agcaagctgt tgcttggaaa   2520
gtgctatctg cagaagagaa cgctgttcat tgcattgctt tctcttttgt aaactggtct   2580
tttttgtgaa ctccacagtt catttatcca actatgcagt cgaatacgag aacaacggta   2640
tggtagagat ttttgaaaat gaaaagaga                                     2669
<210> 207
<211> 852
<212> PRT
<213> Populus trichocarpa
<400> 207
Met Ala Met Glu Val Ala Pro Leu His Arg Glu Ser Ser Ser Ser Gly
1               5                  10                  15
Ser Ile Asn Lys His Leu Thr Asp Asp Gly Lys Tyr Val Arg Tyr Thr
            20                  25                  30
Ala Glu Gln Val Glu Ala Leu Glu Arg Val Tyr Ala Glu Cys Pro Lys
        35                  40                  45
Pro Ser Ser Leu Arg Arg Gln Gln Leu Ile Arg Glu Cys Pro Ile Leu
    50                  55                  60
Ala Asn Ile Glu Pro Lys Gln Ile Lys Val Trp Phe Gln Asn Arg Arg
65                  70                  75                  80
Cys Arg Glu Lys Gln Arg Lys Glu Ser Ser Arg Leu Gln Thr Val Asn
                85                  90                  95
Arg Lys Leu Thr Ala Met Asn Lys Leu Leu Met Glu Glu Asn Asp Arg
            100                 105                 110
Leu Gln Lys Gln Val Ser Gln Leu Val Cys Glu Asn Gly Phe Met Arg
        115                 120                 125
```

```
Gln Gln Leu Gln Thr Ala Pro Thr Ala Thr Asp Ala Ser Cys Asp Ser
   130                 135                 140
Val Val Ala Thr Pro Gln His Ser Leu Arg Asp Ala Asn Asn Pro Ala
145                 150                 155                 160
Gly Leu Leu Ser Ile Ala Glu Glu Thr Leu Ser Glu Phe Leu Ala Lys
                165                 170                 175
Ala Thr Gly Thr Ala Leu Glu Trp Val Gln Met Pro Gly Met Lys Pro
            180                 185                 190
Gly Pro Asp Ser Ile Gly Ile Phe Ser Ile Ser Gln Arg Cys Gly Gly
        195                 200                 205
Val Ala Ala Arg Ala Cys Gly Leu Val Ser Leu Glu Pro Lys Lys Ile
   210                 215                 220
Ala Glu Ile Leu Lys Asp Arg Ser Ser Trp Phe Arg Asp Cys Arg Asn
225                 230                 235                 240
Leu Glu Val Phe Thr Met Phe Pro Ala Gly Asn Gly Gly Thr Ile Glu
                245                 250                 255
Leu Val Tyr Ser Gln Ile Tyr Ala Pro Thr Thr Leu Ala Pro Ala Arg
            260                 265                 270
Asp Met Trp Thr Leu Arg Tyr Thr Thr Ser Leu Glu Asn Gly Ser Leu
        275                 280                 285
Val Val Cys Glu Arg Ser Leu Ser Gly Tyr Gly Ala Gly Pro Asp Ala
   290                 295                 300
Ala Ala Ala Ala Gln Phe Val Arg Ala Glu Met Leu Pro Ser Gly Tyr
305                 310                 315                 320
Leu Ile Arg Pro Cys Glu Gly Gly Ser Ile Ile His Ile Val Asp His
                325                 330                 335
Leu Asn Leu Gln Ala Trp Ser Val Pro Glu Val Leu Arg Pro Leu Tyr
            340                 345                 350
Glu Ser Ser Lys Ala Val Ala Gln Lys Met Thr Ile Ala Ala Leu Arg
        355                 360                 365
Tyr Val Arg Gln Val Ala His Glu Thr Ser Gly Glu Val Val Tyr Gly
   370                 375                 380
Leu Gly Arg Gln Pro Ala Val Leu Arg Thr Phe Asn Gln Arg Leu Ser
385                 390                 395                 400
Arg Gly Phe Asn Asp Ala Ile Asn Gly Phe Asn Asp Asp Gly Trp Ser
                405                 410                 415
Leu Met Asn Ala Asp Gly Ala Glu Asp Val Ile Ile Ala Val Asn Ser
            420                 425                 430
Thr Lys Asn Leu Ile Gly Ala Asn Asn Ser Ala His Ser Leu Ser Phe
        435                 440                 445
Leu Gly Gly Ile Leu Cys Ala Lys Ala Ser Met Leu Leu Gln Asn Val
   450                 455                 460
His Pro Ala Val Leu Val Cys Phe Leu Arg Glu His His Ala Glu Trp
465                 470                 475                 480
Ala Asp Phe Ser Val Asp Ala Tyr Ser Ala Ala Leu Trp Lys Ala Gly
                485                 490                 495
Ser Tyr Ala Tyr Pro Gly Met Arg Pro Met Arg Phe Thr Gly Ser Gln
            500                 505                 510
Ile Thr Met Pro Leu Gly His Thr Ile Glu Gln Glu Asp Leu Leu Glu
        515                 520                 525
Val Ile Arg Leu Glu Gly His Ser Phe Ala Gln Glu Asp Ala Phe Val
   530                 535                 540
Ser Gln Asp Ile His Leu Leu Gln Ile Cys Ser Gly Ile Asp Glu Asn
545                 550                 555                 560
Ala Val Gly Ala Cys Ser Glu Leu Val Phe Ala Pro Ile Asp Glu Thr
                565                 570                 575
Phe Pro Asp Asp Ala Pro Leu Leu Pro Ser Gly Phe Arg Ile Ile Ser
            580                 585                 590
Leu Glu Ser Lys Ala Lys Asp Thr Gln Glu Val Leu Thr Thr Asn Cys
        595                 600                 605
Thr Leu Asp Leu Thr Ser Ser Leu Glu Ala Gly Leu Ala Ile Asn His
   610                 615                 620
Thr Ala Val Asp Gly Ser Ser Cys His Ser Leu Arg Ser Val Leu Thr
```

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

<210> 208
<211> 2634
<212> DNA
<213> Medicago trunculata
<400> 208

agggtgtgat tgtttaagtt aatttgagat taggaagatg gctatggctg ttgcacaaca 60
acaaagagat aacagcattg agagacacct tgattcgtct ggcaaatatg tgaggtacac 120
tgctgaacag attgaagctt tggaaaaggt ttatgtggaa tgccctaagc ctagttcatt 180
gagaaggcaa cagctgattc gggagtgccc ggttctggcc aacattgagc ctaagcagat 240
caaggtttgg tttcagaata ggaggtgtag ggagaagcag agaaaagagg cttctcagct 300
tcagagtgtg aacaggaaac tttctgcgat gaataagctg ttgatggagg aaaatgagag 360
gctgcagaag caggtttcac agctggtgaa tgagaatgga tttatgcgcc agcaactaca 420
ccctacccca gcagctccaa atgctgacgg tagtggcgtt gattccgcgg ctgctgctcc 480
tatgaactca ttgagagatg ctaatagccc tgctggattc ctatcaattg cggaggagac 540
attgacagag ttcctttcaa aggctacagg aactgctgtc gattgggtcc agatgcctgg 600
gatgaagcct ggtccggatt cggttgggat atttgccatt tctcaaggtg gcaacggagt 660
ggcagctcga gcctgtggtc ttgttagttt agaacctact aagattgtgg agatattaaa 720
agatcgccca acttggtacc gtgattgtcg gagttcagaa gttttcacaa tgttcccagc 780
tggaaatgga ggaacaattg aacttgttta cacacagaca tatgctccaa tgacactggc 840
ttctgctcgc gacttctgga ctctaagata cactacaaat ttggaaaacg gaagtgttgt 900
ggtttgtgaa aggtcactgt ctggtactgg tgctggccct aatgctgcag ccgcctcaca 960
gtttgagagg gctgaaatgc tccctagtgg ctatttgatt cgaccatgtg aaggtggagg 1020
atcgatcatc cacattgtag accacctaaa cctgcaggca tggagtgtgc cagaagtgct 1080
gcggccgatc tatgaatcgt cgcaaatggt agctcagaga ctgacaattg cggcacttcg 1140
ctatatcagg caagtagctc aagaaacaag tggtgacgtg gtgtatagca tgggtcggca 1200
acctgcagtt cttagaactt ttagccaacg gttgagcaga ggtttcaatg acgctgtcaa 1260
tggattcaat gataatggtt ggtctgttct gaactgtgat ggtgctgagg gtgttactat 1320
ttcagtaaat tcaatcaaga atttgagtgg cacttctaat ccagcaagtt ccctttcact 1380
ccttggagga attgtctgtg caaaagcttc tatgttactc caaaacacca ctcctgctgt 1440
tttagttcgc tttctgaggg agcatcgctc ggagtgggct gattttagtg ttgatgcctt 1500
ttctgctgca tcacttaaag ctggctccta tggctatcct ggaatgaggt ctacaaagtt 1560
caccggcaat caagcaatca tgcctcttgg acatacaatt gaacatgaag agatgctaga 1620
aattatccgc cttgaaggtc ttgctcaaga tgattctttt gtttctaggg atgttcatct 1680
cttacaggtg cttcctctga cccttgttat gttatgtact ggaattgatg agaatgctgt 1740

```
gggggcttgt tccgagctca tatttgctcc aattgatgac atgttcccag aagatgctcc   1800
cttagtgcct tctggtttcc gcattgtcct gttgaattct caaccaggtg atacaaagaa   1860
cacaacaaca gcaaatcgaa ccttggattt gacatctggt cttgaagtaa gcccggcaac   1920
agctcatgct aacggagacg catcgtgtcc taacaatcga tgtgtgttga ctgttgcctt   1980
tcagtttcct tttgagagcg gtctgcagga taatgttgca gccatggcac gtcaatatgt   2040
ccggcgtgta gtttctgccg tgcaggcggt tgcaacggct atatctccat ccagtgttaa   2100
cacttctggt ggagcaaagc tctcccctgg cactccagaa gcacttacac tagctcaatg   2160
gatctgccag agttatagtc atcatctggg cgcgcaactg ctgagatctg attctcttat   2220
tggtgatatg ctactgaaac atttgtggca tcatccagat gctattttat gctgctcttt   2280
gaagcaagtg cccgtattca tctttgctaa ccaggctggc cttgacatgt tggaaacaac   2340
tctagtggct ctacaagata tcacactgga caaaatattt gatgagtctg cacgcaagaa   2400
tttgattgca tattttgcga agttaatgca gcaggggttt gcttgtatgc cagctgggat   2460
ctgcatgtca acaatggggc gacatgcttc atatgatcaa gccgtcgcgt ggaaagtgca   2520
tgctgaagac aacagtgttc attgcttggc tttctcattc attaattggt catttatatg   2580
acctattgct ggatttaaac ccccactttt ttttcccact tgttgaatac aaaa          2634
<210> 209
<211> 847
<212> PRT
<213> Medicago trunculata
<400> 209
Met Ala Met Ala Val Ala Gln Gln Gln Arg Asp Asn Ser Ile Glu Arg
1               5                  10                  15
His Leu Asp Ser Ser Gly Lys Tyr Val Arg Tyr Thr Ala Glu Gln Ile
            20                  25                  30
Glu Ala Leu Glu Lys Val Tyr Val Glu Cys Pro Lys Pro Ser Ser Leu
        35                  40                  45
Arg Arg Gln Gln Leu Ile Arg Glu Cys Pro Val Leu Ala Asn Ile Glu
    50                  55                  60
Pro Lys Gln Ile Lys Val Trp Phe Gln Asn Arg Arg Cys Arg Glu Lys
65                  70                  75                  80
Gln Arg Lys Glu Ala Ser Gln Leu Gln Ser Val Asn Arg Lys Leu Ser
                85                  90                  95
Ala Met Asn Lys Leu Leu Met Glu Glu Asn Glu Arg Leu Gln Lys Gln
            100                 105                 110
Val Ser Gln Leu Val Asn Glu Asn Gly Phe Met Arg Gln Gln Leu His
        115                 120                 125
Pro Thr Pro Ala Ala Pro Asn Ala Asp Gly Ser Gly Val Asp Ser Ala
    130                 135                 140
Ala Ala Ala Pro Met Asn Ser Leu Arg Asp Ala Asn Ser Pro Ala Gly
145                 150                 155                 160
Phe Leu Ser Ile Ala Glu Glu Thr Leu Thr Glu Phe Leu Ser Lys Ala
                165                 170                 175
Thr Gly Thr Ala Val Asp Trp Val Gln Met Pro Gly Met Lys Pro Gly
            180                 185                 190
Pro Asp Ser Val Gly Ile Phe Ala Ile Ser Gln Gly Gly Asn Gly Val
        195                 200                 205
Ala Ala Arg Ala Cys Gly Leu Val Ser Leu Glu Pro Thr Lys Ile Val
    210                 215                 220
Glu Ile Leu Lys Asp Arg Pro Thr Trp Tyr Arg Asp Cys Arg Ser Ser
225                 230                 235                 240
Glu Val Phe Thr Met Phe Pro Ala Gly Asn Gly Gly Thr Ile Glu Leu
                245                 250                 255
Val Tyr Thr Gln Thr Tyr Ala Pro Met Thr Leu Ala Ser Ala Arg Asp
            260                 265                 270
Phe Trp Thr Leu Arg Tyr Thr Thr Asn Leu Glu Asn Gly Ser Val Val
        275                 280                 285
Val Cys Glu Arg Ser Leu Ser Gly Thr Gly Ala Gly Pro Asn Ala Ala
    290                 295                 300
Ala Ala Ser Gln Phe Glu Arg Ala Glu Met Leu Pro Ser Gly Tyr Leu
305                 310                 315                 320
Ile Arg Pro Cys Glu Gly Gly Gly Ser Ile Ile His Ile Val Asp His
                325                 330                 335
Leu Asn Leu Gln Ala Trp Ser Val Pro Glu Val Leu Arg Pro Ile Tyr
```

```
          340                 345                 350
Glu Ser Ser Gln Met Val Ala Gln Arg Leu Thr Ile Ala Ala Leu Arg
       355                 360                 365
Tyr Ile Arg Gln Val Ala Gln Glu Thr Ser Gly Asp Val Val Tyr Ser
   370                 375                 380
Met Gly Arg Gln Pro Ala Val Leu Arg Thr Phe Ser Gln Arg Leu Ser
385                 390                 395                 400
Arg Gly Phe Asn Asp Ala Val Asn Gly Phe Asn Asp Asn Gly Trp Ser
                405                 410                 415
Val Leu Asn Cys Asp Gly Ala Glu Gly Val Thr Ile Ser Val Asn Ser
            420                 425                 430
Ile Lys Asn Leu Ser Gly Thr Ser Asn Pro Ala Ser Ser Leu Ser Leu
       435                 440                 445
Leu Gly Gly Ile Val Cys Ala Lys Ala Ser Met Leu Leu Gln Asn Thr
   450                 455                 460
Thr Pro Ala Val Leu Val Arg Phe Leu Arg Glu His Arg Ser Glu Trp
465                 470                 475                 480
Ala Asp Phe Ser Val Asp Ala Phe Ser Ala Ala Ser Leu Lys Ala Gly
                485                 490                 495
Ser Tyr Gly Tyr Pro Gly Met Arg Ser Thr Lys Phe Thr Gly Asn Gln
            500                 505                 510
Ala Ile Met Pro Leu Gly His Thr Ile Glu His Glu Glu Met Leu Glu
       515                 520                 525
Ile Ile Arg Leu Glu Gly Leu Ala Gln Asp Asp Ser Phe Val Ser Arg
   530                 535                 540
Asp Val His Leu Leu Gln Val Leu Pro Leu Thr Leu Val Met Leu Cys
545                 550                 555                 560
Thr Gly Ile Asp Glu Asn Ala Val Gly Ala Cys Ser Glu Leu Ile Phe
                565                 570                 575
Ala Pro Ile Asp Asp Met Phe Pro Glu Asp Ala Pro Leu Val Pro Ser
            580                 585                 590
Gly Phe Arg Ile Val Leu Leu Asn Ser Gln Pro Gly Asp Thr Lys Asn
       595                 600                 605
Thr Thr Thr Ala Asn Arg Thr Leu Asp Leu Thr Ser Gly Leu Glu Val
   610                 615                 620
Ser Pro Ala Thr Ala His Ala Asn Gly Asp Ala Ser Cys Pro Asn Asn
625                 630                 635                 640
Arg Cys Val Leu Thr Val Ala Phe Gln Phe Pro Phe Glu Ser Gly Leu
                645                 650                 655
Gln Asp Asn Val Ala Ala Met Ala Arg Gln Tyr Val Arg Arg Val Val
            660                 665                 670
Ser Ala Val Gln Ala Val Ala Thr Ala Ile Ser Pro Ser Ser Val Asn
       675                 680                 685
Thr Ser Gly Gly Ala Lys Leu Ser Pro Gly Thr Pro Glu Ala Leu Thr
   690                 695                 700
Leu Ala Gln Trp Ile Cys Gln Ser Tyr Ser His His Leu Gly Ala Gln
705                 710                 715                 720
Leu Leu Arg Ser Asp Ser Leu Ile Gly Asp Met Leu Leu Lys His Leu
                725                 730                 735
Trp His His Pro Asp Ala Ile Leu Cys Cys Ser Leu Lys Gln Val Pro
            740                 745                 750
Val Phe Ile Phe Ala Asn Gln Ala Gly Leu Asp Met Leu Glu Thr Thr
       755                 760                 765
Leu Val Ala Leu Gln Asp Ile Thr Leu Asp Lys Ile Phe Asp Glu Ser
   770                 775                 780
Ala Arg Lys Asn Leu Ile Ala Tyr Phe Ala Lys Leu Met Gln Gln Gly
785                 790                 795                 800
Phe Ala Cys Met Pro Ala Gly Ile Cys Met Ser Thr Met Gly Arg His
                805                 810                 815
Ala Ser Tyr Asp Gln Ala Val Ala Trp Lys Val His Ala Glu Asp Asn
            820                 825                 830
Ser Val His Cys Leu Ala Phe Ser Phe Ile Asn Trp Ser Phe Ile
       835                 840                 845
```

<210> 210
<211> 2211
<212> DNA
<213> Saccharum officinarum
<220>
<221> misc_feature
<222> (926)..(926)
<223> n é a, c, g, ou t
<220>
<221> misc_feature
<222> (944)..(944)
<223> n é a, c, g, ou t
<400> 210
gattcggttg gtatcgtggc catttcgcat ggttgccgtg gtgttgctgc ccgcgcctgt 60
ggttttggtga atctagaacc aacaagagtc atagagatct tgaaagatcg tccctcttgg 120
ttccgtgatt gtcgaagtct tgaagtgttt acaatgtttc cagctggaaa tgggggaaca 180
gttgaactta tctacatgca gatgtatgcc ccaacaactt tagtccctgc acgtgatttt 240
tggacgttac gatacacgac aacaatggaa gatggtagcc ttgtggtctg tgagagatcc 300
ttgagtggtt ctggaggcgg tccaaatgca gcctctgcac agcaatttgt tagggctgaa 360
atgcttccaa gtgggtattt agttcgccca tgcgaaggtg gaggttcgat tgtgcatata 420
gtggaccatc tagatctcga ggcatggagt gttcctgaag tgcttcgacc gctgtatgag 480
tcttccagag tagttgctca gaaaatgact acggtggcac tgcgccacct tagacaaatt 540
gctcaagaaa caagtggaga agtagtgtac gccttgggaa ggcaacctgc agtactacgg 600
acctttagtc aaagactaag caggggtttt aatgatgcca ttagtggttt caatgatgat 660
ggctggtctg taatgggggg agatggcatt gaagacgttg ctgttgcttg cacctcaacc 720
aagaaaatta ggaataacag caatgcgggg attacatttg gagcccctgg aggcattatt 780
tgtgcgaagg catccatgtt actgcagagt gtcccaccgg cagtactggt ccgatttctg 840
agggagcata gatctgaatg ggctgattac aatattgatg cgtatttggc ttcatcactg 900
aagaccagtg catgctcact tccggngttg caacctatga gatnttctgg ggggcagatg 960
atcatgccac ttgctcacac agtggagaat gaggagattc ttgaagttat ccgccttgaa 1020
ggacatcctc ttactcatga tgaagctctt ctttcaagag acatccacct tctccagctt 1080
tgcactggaa tagatgagaa atctgtgggt tcctccttcc agcttgtgtt tgcaccaatt 1140
gatgagcact ttccagatga tgctccattg atttcttctg gctttcgtgt cataccactt 1200
gatatgaaaa cagatggtgt atcctctggc aggacgctag atttggcatc tagtcttgat 1260
gtgggttctg ctgcacccca agcctcaggg gatgcatctc cagatgactg taatttgaga 1320
tctgtgctga caatcgcctt tcaattccct tacgagatgc accttcagga cagtgttgca 1380
actatggctc gtcaatatgt tcgtagtgtt gtttctgctg tgcaaagagt gtcgatggct 1440
atatctccct cccaatctgg tctaaatgct cagaggacac tttctggctt ccctgaaact 1500
gccacacttg ctagatgggt ttgccagagt tatcactacc atctaggcgt ggaattactt 1560
aatcaatcag atgaagctgg cgaagcattg ttgaaaatgc tctggcatca tccagacgct 1620
gttctgtgct gctcttttaa ggagaaacct atgtttacgt ttgcaaacaa ggcagggctt 1680
gacatgttag aaacatctct tattgccctg caagacctga cactagacaa gattttcgac 1740
gagtcaggaa ggaaagcaat attctcagat atctcaaaac taatggaaca gggctacgca 1800
tacctgccgt caggtgtgtg catgtcagga atgggtcgcc atgtttcttt cgaccaagct 1860
gtggcgtgga aggtgctcgg tgaggacagc aacgtgcact gcctggcttt ctgcttcgtc 1920
aactggtcct tcgtgtaacg cccacccatc cgatggggtg gcgagcctgt cggtctgtgt 1980
gatgagccag cccaattttg tatgatgatc ctgataagga aatcattgac ccgggcaaaa 2040
tgcttatggt gcatgcacta ttttgaggcc ctgaaatccc gggttttatt aaaaaaactg 2100
gagaattttc ccaggttttt tcccactttt cgttggcccc cccacccaca gggtgtttgt 2160
acaatttctt tggcgagggg caccccactc ctgcagtttt tttttccata c 2211
<210> 211
<211> 645
<212> PRT
<213> Saccharum officinarum
<220>
<221> INCERTO
<222> (309)..(309)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> INCERTO
<222> (315)..(315)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 211

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

```
          500                 505                 510
Tyr His Leu Gly Val Glu Leu Leu Asn Gln Ser Asp Glu Ala Gly Glu
       515                 520                 525
Ala Leu Leu Lys Met Leu Trp His His Pro Asp Ala Val Leu Cys Cys
   530                 535                 540
Ser Phe Lys Glu Lys Pro Met Phe Thr Phe Ala Asn Lys Ala Gly Leu
545                 550                 555                 560
Asp Met Leu Glu Thr Ser Leu Ile Ala Leu Gln Asp Leu Thr Leu Asp
                565                 570                 575
Lys Ile Phe Asp Glu Ser Gly Arg Lys Ala Ile Phe Ser Asp Ile Ser
            580                 585                 590
Lys Leu Met Glu Gln Gly Tyr Ala Tyr Leu Pro Ser Gly Val Cys Met
        595                 600                 605
Ser Gly Met Gly Arg His Val Ser Phe Asp Gln Ala Val Ala Trp Lys
    610                 615                 620
Val Leu Gly Glu Asp Ser Asn Val His Cys Leu Ala Phe Cys Phe Val
625                 630                 635                 640
Asn Trp Ser Phe Val
                645
<210> 212
<211> 2076
<212> DNA
<213> Triticum aestivum
<400> 212
agggtattgt tgccgtttca catggttgcc gaggtgtcgc tgcccgtgcc tgtggtctgg     60
tgaatctaca accaacaaag attgtggaga tcttgaaaga ccgcccatct tggttccgtg    120
attgtcggag tcttgaattt tttacgatgc ttccagctgg aaacggcggg accattgaac    180
tcgtttacat gcagatgtat gctcctacca ctttagttcc tgcacgtgat ttttggacgc    240
tgagatacac aactacaatg gaagatggaa gtctcgtggt ttgtgagaga tctttgagtg    300
gttcaggagg tggtccaagt actgcctcag cacagcaatt tgtaagggct gagatgcttc    360
ctagtggcta tttagttcgg ccatgcgacg gtgggggttc aattgtgcat atagtggatc    420
atctggacct tgaggcttgg agtgttccag aagtgcttcg cccgctctat gagtcttcta    480
gggtagttgc tcagaaaatg actactgcgg cattacgaca catcagacag attgctcaag    540
aaactagtgg ggaggttgta tatgctctag ggaggcaacc tgctgttctg cggacattta    600
gtcagaggct gagtagagga tttaatgatg ctataagtgg cttcaatgat gatggttggt    660
ctgtaatggc tggagacggc attgaagatg tgattattgc ttgcaactca aaaaagatta    720
gaagcaataa cactgctccc aatgcgttta tagctcctgg aggtgttata tgtgctaaag    780
catcaatgtt actgcagagt gttccaccag cagtactggt tcggtttctg agggaacatc    840
ggtctgaatg ggctgattat aactttgatg catattcggc ttcagcgctg aaatcaagct    900
catgctcact tcctgggttg cgtcccatga gattttctgg gagccagatc atcatgccac    960
ttgctcacac agtggagaat gaggagattc tagaagttgt tcgtcttgaa gggcaagcgc   1020
tcgatgaggg tcttttatca agagatatcc acctgcttca gttttgcact ggactagacg   1080
agaaatcaat gggatcctgc ttccaacttg tctttgcacc aattgatgag cttttccctg   1140
atgatgctcc attaatatct tcaggctttc gtgttatacc actggacatg aaaacagatg   1200
gtgcacccgc cggtagaaca ctagatttgg cctctagcct tgaagctggt tcaactacac   1260
tgcaagcctc aggcggtgcg gatgattgta atctacgatc agtgctgaca attgcctttc   1320
aattccctta cgagatgcat ctccaagata gtgttgcaac tatggcccgg cagtatgtcc   1380
gcagcattgt ctccgccgtt cagagagtgt cgatggctat ctctccctct cgatctggct   1440
tgaatgctga acagaagata atttctggct tccctgaagc tgcaacacta gctcgctgga   1500
tatgccaaag ttaccggttc catctggggg tcgagttatt tagacaggca gatgaagctg   1560
gggaatcttt attgagaatg ctctgggatc atgaagatgc tattttgtgc tgttctttca   1620
aggaaaagcc tgtatttacg tttgcaaacg agatgggaat taacatgttg gaaacatctt   1680
tcgttgccct tcaagatctc tcgttggaca agatatttga tgaagctggt agaaaggcac   1740
tatactccga gatcccaaag ttgatggagc agggctttgt ttacctgcca gggggtgtgt   1800
gcttgtctgg gatgggccgc catgtctcat ttgagagtgc tgtagcatgg aaagtagttg   1860
gtgaggacaa caatgtgcac tgcctcgcct tctgcttcgt caactggtct ttcgtgtgat   1920
catccatcac ccttgcctgt cccatgcagc tccatttttg ctctctctgt aggggagaac   1980
cgccgccact ggaaagtagt tttacgttat ctttaactag tttgtttcta gatttgaaga   2040
gccagcatcg ggaagaattc tgcctagtga aaattg                             2076
<210> 213
<211> 638
<212> PRT
<213> Triticum aestivum
```

<400> 213

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

```
Ala Arg Trp Ile Cys Gln Ser Tyr Arg Phe His Leu Gly Val Glu Leu
            500                 505                 510
Phe Arg Gln Ala Asp Glu Ala Gly Glu Ser Leu Leu Arg Met Leu Trp
        515                 520                 525
Asp His Glu Asp Ala Ile Leu Cys Cys Ser Phe Lys Glu Lys Pro Val
    530                 535                 540
Phe Thr Phe Ala Asn Glu Met Gly Ile Asn Met Leu Glu Thr Ser Phe
545                 550                 555                 560
Val Ala Leu Gln Asp Leu Ser Leu Asp Lys Ile Phe Asp Glu Ala Gly
                565                 570                 575
Arg Lys Ala Leu Tyr Ser Glu Ile Pro Lys Leu Met Glu Gln Gly Phe
            580                 585                 590
Val Tyr Leu Pro Gly Gly Val Cys Leu Ser Gly Met Gly Arg His Val
        595                 600                 605
Ser Phe Glu Ser Ala Val Ala Trp Lys Val Val Gly Glu Asp Asn Asn
    610                 615                 620
Val His Cys Leu Ala Phe Cys Phe Val Asn Trp Ser Phe Val
625                 630                 635
<210> 214
<211> 1554
<212> DNA
<213> Hordeum vulgare
<400> 214
gcgctgggga ggcagcctgc tgttctgcgg acatttagtc agaggctgag tagaggattt     60
aatgatgcta taagtggctt caatgatgat ggttggtctg taatggccgg agatggcatt    120
gaagatgtga ttatcgcttg caactcaaag aagattagga gcaataacac tgctcccaat    180
gcttttatag ctcctggagg tgttatatgt gctaaagcat caatgttact gcagagtgtt    240
ccaccagcag tactggttcg atttctaagg gaacatcggt ctgaatgggc tgattataac    300
tttgatgcat attcggcttc agcgctgaaa tcaagttcat gctcacttcc tgggttgcgc    360
cctatgagat tttctgggag ccagatcatc atgccacttg ctcacacggt ggagaatgag    420
gagattctag aagttgttcg tcttgaaggg caagcgctcg atgagggtct tttatcaaga    480
gatatccacc tgcttcagtt ttgcactgga atagacgaga aatcaatggg gtcctgcttc    540
caacttgtct ttgccccaat tgatgagctt ttccctgatg atgctccatt aatatcttca    600
ggcttccgtg ttataccact ggacatgaaa acagatggtg cacccaccgg tagaacacta    660
gatttggcct ctagccttga agctggttca actacactgc aagcctcagg caatgcggac    720
gattgtaatc tacgatcagt gctgacaatt gcctttcaat tcccttacga gatgcatctc    780
caagatagtg ttgcaaccat ggcccggcag tatgtccgca gcattgtctc tgccgttcag    840
agagtgtcga tggctatctc tccctctcga tctggtttga atgctgaaca gaagataatt    900
tctggcttcc ctgaagctgc aacacttgct cgctggatat gccaaagcta ccggttccat    960
ctgggggtcg agttatttag acaggcagat gaagctgggg aatctttatt gagaatgctc   1020
tgggatcatg aagatgctat tttgtgctgt tctttcaagg aaaagcctgt atttacgttt   1080
gcaaacgaga tggggattaa catgttggaa acatctttcg ttgcccttca agacctctcg   1140
ttggacaaga tatttgatga agctggtaga aaggcactat actctgagat cccaaagttg   1200
atggagcagg gctttgttta cctgccaggt ggtgtgtgct tgtctgggat gggccgccat   1260
gtctccttcg agaatgctat agcatggaaa gtagtcggtg aggacaacaa tgtgcactgc   1320
cttgccttct gcttcgtcaa ctggtctttc gtgtgatcat cctcccaagg caatccgtgc   1380
ccgtcgcacg cagctccatt tttgctctct ctctctgtag gagagaaccg ctgccgctgg   1440
aaagtagttt catgctatct ttaactagtt tgtttgtaga tttgaagagc cagcatccgg   1500
aagaattctg ccttgtgtaa atctgctcac atttccacta atttacccag gcaa         1554
<210> 215
<211> 451
<212> PRT
<213> Hordeum vulgare
<400> 215
Ala Leu Gly Arg Gln Pro Ala Val Leu Arg Thr Phe Ser Gln Arg Leu
1               5                   10                  15
Ser Arg Gly Phe Asn Asp Ala Ile Ser Gly Phe Asn Asp Asp Gly Trp
            20                  25                  30
Ser Val Met Ala Gly Asp Gly Ile Glu Asp Val Ile Ile Ala Cys Asn
        35                  40                  45
Ser Lys Lys Ile Arg Ser Asn Asn Thr Ala Pro Asn Ala Phe Ile Ala
    50                  55                  60
Pro Gly Gly Val Ile Cys Ala Lys Ala Ser Met Leu Leu Gln Ser Val
```

```
65                  70                  75                  80
Pro Pro Ala Val Leu Val Arg Phe Leu Arg Glu His Arg Ser Glu Trp
                85                  90                  95
Ala Asp Tyr Asn Phe Asp Ala Tyr Ser Ala Ser Ala Leu Lys Ser Ser
            100                 105                 110
Ser Cys Ser Leu Pro Gly Leu Arg Pro Met Arg Phe Ser Gly Ser Gln
        115                 120                 125
Ile Ile Met Pro Leu Ala His Thr Val Glu Asn Glu Glu Ile Leu Glu
    130                 135                 140
Val Val Arg Leu Glu Gly Gln Ala Leu Asp Glu Gly Leu Leu Ser Arg
145                 150                 155                 160
Asp Ile His Leu Leu Gln Phe Cys Thr Gly Ile Asp Glu Lys Ser Met
                165                 170                 175
Gly Ser Cys Phe Gln Leu Val Phe Ala Pro Ile Asp Glu Leu Phe Pro
            180                 185                 190
Asp Asp Ala Pro Leu Ile Ser Ser Gly Phe Arg Val Ile Pro Leu Asp
        195                 200                 205
Met Lys Thr Asp Gly Ala Pro Thr Gly Arg Thr Leu Asp Leu Ala Ser
    210                 215                 220
Ser Leu Glu Ala Gly Ser Thr Thr Leu Gln Ala Ser Gly Asn Ala Asp
225                 230                 235                 240
Asp Cys Asn Leu Arg Ser Val Leu Thr Ile Ala Phe Gln Phe Pro Tyr
                245                 250                 255
Glu Met His Leu Gln Asp Ser Val Ala Thr Met Ala Arg Gln Tyr Val
            260                 265                 270
Arg Ser Ile Val Ser Ala Val Gln Arg Val Ser Met Ala Ile Ser Pro
        275                 280                 285
Ser Arg Ser Gly Leu Asn Ala Glu Gln Lys Ile Ile Ser Gly Phe Pro
    290                 295                 300
Glu Ala Ala Thr Leu Ala Arg Trp Ile Cys Gln Ser Tyr Arg Phe His
305                 310                 315                 320
Leu Gly Val Glu Leu Phe Arg Gln Ala Asp Glu Ala Gly Glu Ser Leu
                325                 330                 335
Leu Arg Met Leu Trp Asp His Glu Asp Ala Ile Leu Cys Cys Ser Phe
            340                 345                 350
Lys Glu Lys Pro Val Phe Thr Phe Ala Asn Glu Met Gly Ile Asn Met
        355                 360                 365
Leu Glu Thr Ser Phe Val Ala Leu Gln Asp Leu Ser Leu Asp Lys Ile
    370                 375                 380
Phe Asp Glu Ala Gly Arg Lys Ala Leu Tyr Ser Glu Ile Pro Lys Leu
385                 390                 395                 400
Met Glu Gln Gly Phe Val Tyr Leu Pro Gly Gly Val Cys Leu Ser Gly
                405                 410                 415
Met Gly Arg His Val Ser Phe Glu Asn Ala Ile Ala Trp Lys Val Val
            420                 425                 430
Gly Glu Asp Asn Asn Val His Cys Leu Ala Phe Cys Phe Val Asn Trp
        435                 440                 445
Ser Phe Val
    450
<210> 216
<211> 2622
<212> DNA
<213> Phyllostachys praecox
<400> 216
cggcgggtac gacaaggccg ggatagactc gggcaagtac gtgcgataca cgccggagca     60
ggtggaggcg ctcgagcgga tgtatgctga gtgccccaag cccagctcca cgcgcaggca    120
gcagctgctg cgcgagtgcc cgattctggc caacatcgag cccaagcaga tcaaggtctg    180
gttccagaac cgaaggtgcc gtgataagca gcggaaggag gcttcccggc ttcaggccgt    240
gaacagaaaa ttgaccgcaa tgaacaagct tctcatggaa gagaatgatc gtctccagaa    300
gcaggtttcc cagctggttc atgagaatgc atacatgaag cagcagctgc agaatccatc    360
attggccaat gatacaagct gtgaatcaaa tgtgaccact caaaaccctc tgaaggatgc    420
aagtaaccct tctggactcc tttcaattgc ggaggagacc ttgacagagt tcctctccaa    480
ggctacaggg actgctgttg attgggttca gatgcctggg atgaagcctg gtccggattc    540
```

```
ggttggtatt gtggccattt cacatggttg ccgtggtgtt gctgcccgtg cctgtgattt    600
ggtgaatcta gaaccaacaa aagttgtgga gatcttgaaa gaccgtccat cttggttctg    660
tgatcgtcaa agcctggaag tcttcacaat gtttccagct ggaaatggag gaacaattga    720
acttgtctac acgcagttgt atgctccaac aactttagtc cctgcacgtg atttttggac    780
tctaaggtac acaaccacaa tggaagatgg cagccttgtg gtctgtgaga gatccttgag    840
tggttcaggg ggtggtccaa gtgcagcctc tgcacagcaa tttgtaagag ccgaaatgct    900
tccaagtgga tatttagttc gcccatgcga gggtgggggc tcaattgtgc acatagtgga    960
ccatctggac cttgaggctt ggagtgttcc tgaagtgctt cggccgctct acgagtcgtc   1020
tagggtagtt gcccagaaaa tgactacagc ggcactacgg cacatcagac aaattgctca   1080
agaaactagt ggggaggttg tatatgcttt ggggaggcag cctgccgttc tacggacatt   1140
tagtcagagg ttgagtagag gatttaacga tgctataagt ggtttcaatg atgatggttg   1200
gtctgtcatg ggtggagatg gcattgaaga tgtaatcatt gcttgcaact caaagaagat   1260
taggaacaat agcactgctg ccaatgcttt tggagcccct ggaggcgtta tatgtgctaa   1320
ggcatccatg ttactgcaga gtgttccacc agcagtactg gttcggtttc tgagggaaca   1380
tcggtctgaa tgggctgatt ataactttga tgcatattcg gctttagcac tgaagacgag   1440
ctcatgttca cttcctgggt tgcggcctac gagattttct gggagccaga tcatcatgcc   1500
acttgctcac acagtggaga atgaggagat tctagaagtc attcgtcttg aagggcaggc   1560
acttacacat gatgaaggtc ttttatcaag agatatccac ctgcttcagc tttgcactgg   1620
aatagatgag aaatcaatgg gatcctgctt ccagcttgtc tttgcaccca tcgatgagct   1680
tttccctgat gatgccccat tgatatcttc aggctttcgt gttataccac tggacatgaa   1740
aacagatggt gcacctactg gtagaacatt agatttggcc tctagccttg aagttggttc   1800
aactacacaa caagctacag gggatgcatc tttggatgat tgtaggaatc tgcgatcagt   1860
gctgacaatt gcctttcaat tcccttatga aattcatctc caggatagtg ttgcaactat   1920
ggcccggcaa tacgtccgta gcgttgtttc tgctgtgcag agagtgtcga tggctatttc   1980
tccccctcga tctggtgtga atgctgggca gaagatattt tctggcttcc ctgaagccgc   2040
aacacttgct cgctggattt gccaaagcta ccagttccat ttgggagtgg agttacttag   2100
gcaggcagat gaagccgggg aatcattatt gagaatgctc tgggattacg aagatgctat   2160
cttgtgctgt tctttcaagg aaaagcctgt atttactttt gccaacgaga tgggacttaa   2220
catgttagaa acatctctcg ttgctctaca agacctctca ttggacaaga tatttgatga   2280
aactggtaga aaggcactac attcggagat cccaaaattg atggaacagg gctatgttta   2340
cctgccggct ggtgtgtgct tgtccggaat gggccgccat gtctctttcg agcaagctgt   2400
agcatggaaa gtacttggtg aggacaacaa tgtgcactgc ctggccttct gcttcgtcaa   2460
ctggtctttc gtgtgatcca tccgacgcaa tccatgtccg ttgtgagcag caccattttc   2520
gcattctctg tagaagagaa ccatcgtcgc tgttagttaa tgctgtcttt aactagtttc   2580
tagatttgga aaagccagtc tgcaaaaaaa aaaaaaaaaa aa                      2622
<210> 217
<211> 824
<212> PRT
<213> Phyllostachys praecox
<400> 217
Gly Gly Tyr Asp Lys Ala Gly Ile Asp Ser Gly Lys Tyr Val Arg Tyr
1               5                  10                  15
Thr Pro Glu Gln Val Glu Ala Leu Glu Arg Met Tyr Ala Glu Cys Pro
           20                  25                  30
Lys Pro Ser Ser Thr Arg Arg Gln Gln Leu Leu Arg Glu Cys Pro Ile
       35                  40                  45
Leu Ala Asn Ile Glu Pro Lys Gln Ile Lys Val Trp Phe Gln Asn Arg
   50                  55                  60
Arg Cys Arg Asp Lys Gln Arg Lys Glu Ala Ser Arg Leu Gln Ala Val
65                  70                  75                  80
Asn Arg Lys Leu Thr Ala Met Asn Lys Leu Leu Met Glu Glu Asn Asp
                85                  90                  95
Arg Leu Gln Lys Gln Val Ser Gln Leu Val His Glu Asn Ala Tyr Met
           100                 105                 110
Lys Gln Gln Leu Gln Asn Pro Ser Leu Ala Asn Asp Thr Ser Cys Glu
       115                 120                 125
Ser Asn Val Thr Thr Gln Asn Pro Leu Lys Asp Ala Ser Asn Pro Ser
   130                 135                 140
Gly Leu Leu Ser Ile Ala Glu Glu Thr Leu Thr Glu Phe Leu Ser Lys
145                 150                 155                 160
Ala Thr Gly Thr Ala Val Asp Trp Val Gln Met Pro Gly Met Lys Pro
               165                 170                 175
Gly Pro Asp Ser Val Gly Ile Val Ala Ile Ser His Gly Cys Arg Gly
```

```
          180                 185                 190
Val Ala Ala Arg Ala Cys Asp Leu Val Asn Leu Glu Pro Thr Lys Val
       195                 200                 205
Val Glu Ile Leu Lys Asp Arg Pro Ser Trp Phe Cys Asp Arg Gln Ser
   210                 215                 220
Leu Glu Val Phe Thr Met Phe Pro Ala Gly Asn Gly Gly Thr Ile Glu
225                 230                 235                 240
Leu Val Tyr Thr Gln Leu Tyr Ala Pro Thr Thr Leu Val Pro Ala Arg
                245                 250                 255
Asp Phe Trp Thr Leu Arg Tyr Thr Thr Thr Met Glu Asp Gly Ser Leu
            260                 265                 270
Val Val Cys Glu Arg Ser Leu Ser Gly Ser Gly Gly Gly Pro Ser Ala
        275                 280                 285
Ala Ser Ala Gln Gln Phe Val Arg Ala Glu Met Leu Pro Ser Gly Tyr
    290                 295                 300
Leu Val Arg Pro Cys Glu Gly Gly Gly Ser Ile Val His Ile Val Asp
305                 310                 315                 320
His Leu Asp Leu Glu Ala Trp Ser Val Pro Glu Val Leu Arg Pro Leu
                325                 330                 335
Tyr Glu Ser Ser Arg Val Val Ala Gln Lys Met Thr Thr Ala Ala Leu
            340                 345                 350
Arg His Ile Arg Gln Ile Ala Gln Glu Thr Ser Gly Glu Val Val Tyr
        355                 360                 365
Ala Leu Gly Arg Gln Pro Ala Val Leu Arg Thr Phe Ser Gln Arg Leu
    370                 375                 380
Ser Arg Gly Phe Asn Asp Ala Ile Ser Gly Phe Asn Asp Asp Gly Trp
385                 390                 395                 400
Ser Val Met Gly Gly Asp Gly Ile Glu Asp Val Ile Ile Ala Cys Asn
                405                 410                 415
Ser Lys Lys Ile Arg Asn Asn Ser Thr Ala Ala Asn Ala Phe Gly Ala
            420                 425                 430
Pro Gly Gly Val Ile Cys Ala Lys Ala Ser Met Leu Leu Gln Ser Val
        435                 440                 445
Pro Pro Ala Val Leu Val Arg Phe Leu Arg Glu His Arg Ser Glu Trp
    450                 455                 460
Ala Asp Tyr Asn Phe Asp Ala Tyr Ser Ala Leu Ala Leu Lys Thr Ser
465                 470                 475                 480
Ser Cys Ser Leu Pro Gly Leu Arg Pro Thr Arg Phe Ser Gly Ser Gln
                485                 490                 495
Ile Ile Met Pro Leu Ala His Thr Val Glu Asn Glu Glu Ile Leu Glu
            500                 505                 510
Val Ile Arg Leu Glu Gly Gln Ala Leu Thr His Asp Glu Gly Leu Leu
        515                 520                 525
Ser Arg Asp Ile His Leu Leu Gln Leu Cys Thr Gly Ile Asp Glu Lys
    530                 535                 540
Ser Met Gly Ser Cys Phe Gln Leu Val Phe Ala Pro Ile Asp Glu Leu
545                 550                 555                 560
Phe Pro Asp Asp Ala Pro Leu Ile Ser Ser Gly Phe Arg Val Ile Pro
                565                 570                 575
Leu Asp Met Lys Thr Asp Gly Ala Pro Thr Gly Arg Thr Leu Asp Leu
            580                 585                 590
Ala Ser Ser Leu Glu Val Gly Ser Thr Thr Gln Gln Ala Thr Gly Asp
        595                 600                 605
Ala Ser Leu Asp Asp Cys Arg Asn Leu Arg Ser Val Leu Thr Ile Ala
    610                 615                 620
Phe Gln Phe Pro Tyr Glu Ile His Leu Gln Asp Ser Val Ala Thr Met
625                 630                 635                 640
Ala Arg Gln Tyr Val Arg Ser Val Val Ser Ala Val Gln Arg Val Ser
                645                 650                 655
Met Ala Ile Ser Pro Pro Arg Ser Gly Val Asn Ala Gly Gln Lys Ile
            660                 665                 670
Phe Ser Gly Phe Pro Glu Ala Ala Thr Leu Ala Arg Trp Ile Cys Gln
        675                 680                 685
```

Ser Tyr Gln Phe His Leu Gly Val Glu Leu Leu Arg Gln Ala Asp Glu
690 695 700
Ala Gly Glu Ser Leu Leu Arg Met Leu Trp Asp Tyr Glu Asp Ala Ile
705 710 715 720
Leu Cys Cys Ser Phe Lys Glu Lys Pro Val Phe Thr Phe Ala Asn Glu
725 730 735
Met Gly Leu Asn Met Leu Glu Thr Ser Leu Val Ala Leu Gln Asp Leu
740 745 750
Ser Leu Asp Lys Ile Phe Asp Glu Thr Gly Arg Lys Ala Leu His Ser
755 760 765
Glu Ile Pro Lys Leu Met Glu Gln Gly Tyr Val Tyr Leu Pro Ala Gly
770 775 780
Val Cys Leu Ser Gly Met Gly Arg His Val Ser Phe Glu Gln Ala Val
785 790 795 800
Ala Trp Lys Val Leu Gly Glu Asp Asn Asn Val His Cys Leu Ala Phe
805 810 815
Cys Phe Val Asn Trp Ser Phe Val
820

<210> 218
<211> 2193
<212> DNA
<213> Oryza sativa
<400> 218

aatccgaaaa gtttctgcac cgttttcacc ccctaactaa caatataggg aacgtgtgct 60
aaatataaaa tgagacctta tatatgtagc gctgataact agaactatgc aagaaaaact 120
catccaccta ctttagtggc aatcgggcta aataaaaaag agtcgctaca ctagtttcgt 180
tttccttagt aattaagtgg gaaaatgaaa tcattattgc ttagaatata cgttcacatc 240
tctgtcatga agttaaatta ttcgaggtag ccataattgt catcaaactc ttcttgaata 300
aaaaaatctt tctagctgaa ctcaatgggt aaagagagag atttttttta aaaaaataga 360
atgaagatat tctgaacgta ttggcaaaga tttaaacata taattatata attttatagt 420
ttgtgcattc gtcatatcgc acatcattaa ggacatgtct tactccatcc caatttttat 480
ttagtaatta aagacaattg acttatttt attatttatc ttttttcgat tagatgcaag 540
gtacttacgc acacactttg tgctcatgtg catgtgtgag tgcacctcct caatacacgt 600
tcaactagca acacatctct aatatcactc gcctatttaa tacatttagg tagcaatatc 660
tgaattcaag cactccacca tcaccagacc acttttaata atatctaaaa tacaaaaaat 720
aattttacag aatagcatga aaagtatgaa acgaactatt taggtttttc acatacaaaa 780
aaaaaaagaa ttttgctcgt gcgcgagcgc caatctccca tattgggcac acaggcaaca 840
acagagtggc tgcccacaga acaacccaca aaaaacgatg atctaacgga ggacagcaag 900
tccgcaacaa ccttttaaca gcaggctttg cggccaggag agaggaggag aggcaaagaa 960
aaccaagcat cctcctcctc ccatctataa attcctcccc ccttttcccc tctctatata 1020
ggaggcatcc aagccaagaa gagggagagc accaaggaca cgcgactagc agaagccgag 1080
cgaccgcctt cttcgatcca tatcttccgg tcgagttctt ggtcgatctc ttccctcctc 1140
cacctcctcc tcacagggta tgtgcccttc ggttgttctt ggatttattg ttctaggttg 1200
tgtagtacgg gcgttgatgt taggaaaggg gatctgtatc tgtgatgatt cctgttcttg 1260
gatttgggat agaggggttc ttgatgttgc atgttatcgg ttcggtttga ttagtagtat 1320
ggttttcaat cgtctggaga gctctatgga aatgaaatgg tttagggtac ggaatcttgc 1380
gattttgtga gtacctttg tttgaggtaa aatcagagca ccggtgattt tgcttggtgt 1440
aataaaagta cggttgtttg gtcctcgatt ctggtagtga tgcttctcga tttgacgaag 1500
ctatcctttg tttattccct attgaacaaa aataatccaa ctttgaagac ggtcccgttg 1560
atgagattga atgattgatt cttaagcctg tccaaaattt cgcagctggc ttgtttagat 1620
acagtagtcc ccatcacgaa attcatggaa acagttataa tcctcaggaa caggggattc 1680
cctgttcttc cgatttgctt tagtcccaga attttttttc ccaaatatct taaaaagtca 1740
ctttctggtt cagttcaatg aattgattgc tacaaataat gcttttatag cgttatccta 1800
gctgtagttc agttaatagg taataccccc atagtttagt caggagaaga acttatccga 1860
tttctgatct ccatttttaa ttatatgaaa tgaactgtag cataagcagt attcatttgg 1920
attatttttt ttattagctc tcacccccttc attattctga gctgaaagtc tggcatgaac 1980
tgtcctcaat tttgttttca aattcacatc gattatctat gcattatcct cttgtatcta 2040
cctgtagaag tttctttttg gttattcctt gactgcttga ttacagaaag aaatttatga 2100
agctgtaatc gggatagtta tactgcttgt tcttatgatt catttccttt gtgcagttct 2160
tggtgtagct tgccactttc accagcaaag ttc 2193

<210> 219
<211> 51
<212> DNA

<213> Seqüência artificial
<220>
<223> iniciador: prm03263
<400> 219
ggggacaagt ttgtacaaaa aagcaggctt gtgctaaggc atccatgcta c 51
<210> 220
<211> 51
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm03264
<400> 220
ggggaccact ttgtacaaga aagctgggtg caccttccat gctacagctt g 51
<210> 221
<211> 50
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm01983
<400> 221
ggggacaagt ttgtacaaaa aagcaggctt cacaatggcg gcggcggtgg 50
<210> 222
<211> 54
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm01984
<400> 222
ggggaccact ttgtacaaga aagctgggtg gattttgggt cacacgaagg acca 54
<210> 223
<211> 1096
<212> DNA
<213> Oryza sativa
<400> 223
tgtgctaagg catccatgct actgcagagt gtccctcctg cagttttggt tcgatttttg 60
agggaacatc gttctgaatg ggcggattat aacttcgatg catattcagc ttcatctctg 120
aagacaagct catgttcact tcctgggttg cggcctatga gattttctgg gagccagatc 180
attatgccac ttgctcacac ggtggagaat gaagagattt tagaagttgt ccgtcttgaa 240
ggacaagcac ttacacatga tgatggtctt atgtctagag atattcacct gcttcagctt 300
tgcactggaa tagatgagaa atcaatggga tcctgcttcc agcttgtctt tgcaccaatc 360
gatgagcttt tccctgatga tgctccgtta atatcttcag gctttcgtgt tataccgctg 420
gacatgaaaa cagatggtac acctgctggt agaacattag atttggcatc tagccttgag 480
gttggttcaa ctgcacagcc cacaggggat gcatctatgg atgactgtaa tctacgatca 540
gtgctgacaa ttgcctttca gttcccttat gaaatgcatc tccaagacag cgttgcaact 600
atggcccggc aatatgtccg cagtattgtt tcctctgttc agagagtatc aatggctgtt 660
tctccttctc ggtctggctt gaatgctggg cagaagataa tttcaggctt ccctgaagcc 720
ccaacgctag ctcgttggat ttgccaaagc taccagttcc atttgggggt ggagttactt 780
aggcaggcag atgatgctgg ggaagcacta ttgaaaatgc tatgggatta cgaagacgct 840
attttgtgct gttctttcaa ggaaaagcct gtatttactt ttgccaacga gatgggacta 900
aacatgctag aaacatctct cgtcgctctc caagatctct cactggacaa gatatttgat 960
gaagccggta ggaaggccct atacaacgag atcccgaaat tgatggaaca gggttacgtg 1020
tacctgcctg gtggagtgtg cttgtccggg atggggcgcc atgtttcttt cgagcaagct 1080
gtagcatgga aggtgc 1096
<210> 224
<211> 2553
<212> DNA
<213> Brassica rapa
<400> 224
atggagatgg cggtggcgaa tcaccgagag agaagcagtg atagcatgaa cagacattta 60
gacagcagcg gcaagtacgt caggtacacc gctgagcaag tcgaggccct cgagcgtgtc 120
tacgccgagt gtcccaagcc tagctcgctc cggagacagc agttgatccg tgaatgttct 180
atcttggcca acatcgagcc taagcagatc aaagtctggt tccaaaaccg gaggtgtcga 240
gataagcaga ggaaagaggc gtcgaggctc cagagcgtaa acaggaagct ttcggctatg 300

```
aacaagctgt tgatggagga gaacgatagg ttgcagaagc aagtttctca gctcgtctgt    360
gaaaatggat acatgaagca gcagcttact actactgttg tgaatgatcc aagctgtgat    420
tctgtggtga caactcctca gcattcgctt agagacgcta atagtcctgc tgggctgctc    480
tcgatcgcag aggagacatt ggcagagttc ctatccaagg ctacaggaac tgctgttgat    540
tgggttcaga tgcctgggat gaagcctggt ccggattcgg ttgggatctt tgctatatcg    600
caaagatgca gtggggtggc agctcgagcc tgtggtcttg ttagtttaga gcctgtcaag    660
attgcagaga tactcaaaga taggccatct tggttccgtg actgtaggag ccttgaagtc    720
ttcactatgt tcccggctgg taacggcggc accattgagc tcgtctatat gcagacatat    780
gcaccaacga ctctggctcc tgcccgcgat ttctggaccc tgagatacac aacgagccta    840
gacaatggca gttttgtggt ttgtgagagg tcactctctg gttctggtgc tggtcctaac    900
gctgcatcag cttctcagtt tgtaagagca gaaatgcttt ctagtgggta tctaataagg    960
ccttgtgatg gtggcggttc cattattcac attgtcgatc accttaatct tgaggcttgg   1020
agtgttcccg atgtgcttcg acccctttac gagtcatcta aagtcgtagc acagaaaatg   1080
accatttcag cattgaggta tatcaggcaa ctagctcaag agtctaatgg tgaattagtg   1140
tatggattag gaagacagcc tgctgtttta agaacattta gccaaagatt aagcaggggt   1200
tttaatgatg cggttaatgg gtttggtgac gacggatggt ctacaatgca ttgtgatggg   1260
gctgaagaca taatcgttgc tattaactct acaaagcatt tgaataatat gtctaattct   1320
ctttccttcc ttggaggtgt cctttgtgcc aaggcttcta tgcttctcca aaatgttcct   1380
cctgcggttt tgatccggtt tctaagagag catcggtccg agtgggctga cttcaatgtt   1440
gatgcatatt ctgctgcaac gcttaaagcc ggttcttttg cttatccggg tatgaggcct   1500
acaaggttca ccgggagcca aatcataatg cctctaggac ataccatcga acacgaagaa   1560
atgcttgaag ttgttagact agaaggtcat tctcttgcac aagaagatgc gttcatgtcc   1620
cgtgatgtcc atcttcttca ggtatgtact gggattgatg agaacgctgt tggagcatgt   1680
tcagaactaa tatttgctcc catcaatgag atgttcccgg atgatgctcc acttgttcct   1740
tctggattca gagtcatacc cgttgatgct aaaacgggag atgcgcaaga tctgttgacc   1800
gctaaccacc ggacgctaga cttgacttct agccttgagg ttggtccaac accagagaac   1860
gcttctggaa actcttcttc tagctcaagc tcaagatgta tcctcaccat cgcgtttcag   1920
ttccctttg aaaacaactt gcaagacaat gttgctggta tggcttgtca gtacgtgcgg   1980
agcgtgatct catcggttca acgtgttgca atggccatct caccctctgg gataagccca   2040
agtctgggat ccaaactgtc cccggggtct cctgaagctg ttacacttgc ccaatggatc   2100
tcacaaagtt acagtcacca cttaggctcg gagttgctga cggttgactc gcttggaagc   2160
aacgactcgg tattgaaact tctatgggat caccaagatg ccattttgtg ttgctcttta   2220
aagcctcagc cagtgtttat gttcgcgaac caagctggtt tagacatgtt agagacgacg   2280
cttgtagcct tacaagacat tacactcgaa aagatctttg atgaatctgg tcgaaaggct   2340
ctctgctccg atttcgccaa gctaatgcaa cagggatatg catgcttgcc ttcaggaata   2400
tgtttgtcta cgatgggaag acatgtgact tatgaacaag ctgttgcttg gaaagtgttt   2460
gctcctgctg ctgcatcatc cgaaaacaac gatggcaatg ataatacttt gcactgtctt   2520
gctttctcct ttgtaaactg gtcgtttgtg taa                                2553
<210> 225
<211> 850
<212> PRT
<213> Brassica rapa
<400> 225
Met Glu Met Ala Val Ala Asn His Arg Glu Arg Ser Ser Asp Ser Met
1               5                  10                  15
Asn Arg His Leu Asp Ser Ser Gly Lys Tyr Val Arg Tyr Thr Ala Glu
           20                  25                  30
Gln Val Glu Ala Leu Glu Arg Val Tyr Ala Glu Cys Pro Lys Pro Ser
       35                  40                  45
Ser Leu Arg Arg Gln Gln Leu Ile Arg Glu Cys Ser Ile Leu Ala Asn
   50                  55                  60
Ile Glu Pro Lys Gln Ile Lys Val Trp Phe Gln Asn Arg Arg Cys Arg
65                  70                  75                  80
Asp Lys Gln Arg Lys Glu Ala Ser Arg Leu Gln Ser Val Asn Arg Lys
                85                  90                  95
Leu Ser Ala Met Asn Lys Leu Leu Met Glu Glu Asn Asp Arg Leu Gln
           100                 105                 110
Lys Gln Val Ser Gln Leu Val Cys Glu Asn Gly Tyr Met Lys Gln Gln
       115                 120                 125
Leu Thr Thr Thr Val Val Asn Asp Pro Ser Cys Asp Ser Val Val Thr
   130                 135                 140
Thr Pro Gln His Ser Leu Arg Asp Ala Asn Ser Pro Ala Gly Leu Leu
145                 150                 155                 160
```

```
Ser Ile Ala Glu Glu Thr Leu Ala Glu Phe Leu Ser Lys Ala Thr Gly
                165                 170                 175
Thr Ala Val Asp Trp Val Gln Met Pro Gly Met Lys Pro Gly Pro Asp
            180                 185                 190
Ser Val Gly Ile Phe Ala Ile Ser Gln Arg Cys Ser Gly Val Ala Ala
        195                 200                 205
Arg Ala Cys Gly Leu Val Ser Leu Glu Pro Val Lys Ile Ala Glu Ile
    210                 215                 220
Leu Lys Asp Arg Pro Ser Trp Phe Arg Asp Cys Arg Ser Leu Glu Val
225                 230                 235                 240
Phe Thr Met Phe Pro Ala Gly Asn Gly Gly Thr Ile Glu Leu Val Tyr
                245                 250                 255
Met Gln Thr Tyr Ala Pro Thr Thr Leu Ala Pro Ala Arg Asp Phe Trp
            260                 265                 270
Thr Leu Arg Tyr Thr Thr Ser Leu Asp Asn Gly Ser Phe Val Val Cys
        275                 280                 285
Glu Arg Ser Leu Ser Gly Ser Gly Ala Gly Pro Asn Ala Ala Ser Ala
    290                 295                 300
Ser Gln Phe Val Arg Ala Glu Met Leu Ser Ser Gly Tyr Leu Ile Arg
305                 310                 315                 320
Pro Cys Asp Gly Gly Gly Ser Ile Ile His Ile Val Asp His Leu Asn
                325                 330                 335
Leu Glu Ala Trp Ser Val Pro Asp Val Leu Arg Pro Leu Tyr Glu Ser
            340                 345                 350
Ser Lys Val Val Ala Gln Lys Met Thr Ile Ser Ala Leu Arg Tyr Ile
        355                 360                 365
Arg Gln Leu Ala Gln Glu Ser Asn Gly Glu Leu Val Tyr Gly Leu Gly
    370                 375                 380
Arg Gln Pro Ala Val Leu Arg Thr Phe Ser Gln Arg Leu Ser Arg Gly
385                 390                 395                 400
Phe Asn Asp Ala Val Asn Gly Phe Gly Asp Asp Gly Trp Ser Thr Met
                405                 410                 415
His Cys Asp Gly Ala Glu Asp Ile Ile Val Ala Ile Asn Ser Thr Lys
            420                 425                 430
His Leu Asn Asn Met Ser Asn Ser Leu Ser Phe Leu Gly Gly Val Leu
        435                 440                 445
Cys Ala Lys Ala Ser Met Leu Leu Gln Asn Val Pro Pro Ala Val Leu
    450                 455                 460
Ile Arg Phe Leu Arg Glu His Arg Ser Glu Trp Ala Asp Phe Asn Val
465                 470                 475                 480
Asp Ala Tyr Ser Ala Ala Thr Leu Lys Ala Gly Ser Phe Ala Tyr Pro
                485                 490                 495
Gly Met Arg Pro Thr Arg Phe Thr Gly Ser Gln Ile Ile Met Pro Leu
            500                 505                 510
Gly His Thr Ile Glu His Glu Glu Met Leu Glu Val Val Arg Leu Glu
        515                 520                 525
Gly His Ser Leu Ala Gln Glu Asp Ala Phe Met Ser Arg Asp Val His
    530                 535                 540
Leu Leu Gln Val Cys Thr Gly Ile Asp Glu Asn Ala Val Gly Ala Cys
545                 550                 555                 560
Ser Glu Leu Ile Phe Ala Pro Ile Asn Glu Met Phe Pro Asp Asp Ala
                565                 570                 575
Pro Leu Val Pro Ser Gly Phe Arg Val Ile Pro Val Asp Ala Lys Thr
            580                 585                 590
Gly Asp Ala Gln Asp Leu Leu Thr Ala Asn His Arg Thr Leu Asp Leu
        595                 600                 605
Thr Ser Ser Leu Glu Val Gly Pro Thr Pro Glu Asn Ala Ser Gly Asn
    610                 615                 620
Ser Ser Ser Ser Ser Ser Ser Arg Cys Ile Leu Thr Ile Ala Phe Gln
625                 630                 635                 640
Phe Pro Phe Glu Asn Asn Leu Gln Asp Asn Val Ala Gly Met Ala Cys
                645                 650                 655
Gln Tyr Val Arg Ser Val Ile Ser Ser Val Gln Arg Val Ala Met Ala
```

660 665 670
Ile Ser Pro Ser Gly Ile Ser Pro Ser Leu Gly Ser Lys Leu Ser Pro
675 680 685
Gly Ser Pro Glu Ala Val Thr Leu Ala Gln Trp Ile Ser Gln Ser Tyr
690 695 700
Ser His His Leu Gly Ser Glu Leu Leu Thr Val Asp Ser Leu Gly Ser
705 710 715 720
Asn Asp Ser Val Leu Lys Leu Leu Trp Asp His Gln Asp Ala Ile Leu
725 730 735
Cys Cys Ser Leu Lys Pro Gln Pro Val Phe Met Phe Ala Asn Gln Ala
740 745 750
Gly Leu Asp Met Leu Glu Thr Thr Leu Val Ala Leu Gln Asp Ile Thr
755 760 765
Leu Glu Lys Ile Phe Asp Glu Ser Gly Arg Lys Ala Leu Cys Ser Asp
770 775 780
Phe Ala Lys Leu Met Gln Gln Gly Tyr Ala Cys Leu Pro Ser Gly Ile
785 790 795 800
Cys Leu Ser Thr Met Gly Arg His Val Thr Tyr Glu Gln Ala Val Ala
805 810 815
Trp Lys Val Phe Ala Pro Ala Ala Ala Ser Ser Glu Asn Asn Asp Gly
820 825 830
Asn Asp Asn Thr Leu His Cys Leu Ala Phe Ser Phe Val Asn Trp Ser
835 840 845
Phe Val
850

<210> 226
<211> 2529
<212> DNA
<213> Ginkgo biloba
<400> 226

atggcagtaa tagcacagaa agacgttaaa ggtgccatgg atacgagcaa gtatgttcgc 60
tatacttctg aacaagtcga agctcttgaa cgcgtttaca gcgagtgtcc gaagcctagt 120
tctcttcgtc ggcagcagct tattagagag tgtccaatac tttctaatat tgagcccaag 180
caaatcaaag tttggttcca aaatcgcagg tgtcgagaga aacagaggaa ggaggcatca 240
cgccttcaaa ctgttaacag gaagcttacg gcaatgaaca aattgctgat ggaggaaaat 300
gatcgccttc aaaagcaggt ttcacagttg gtgtacgaga acggttatat gagacagcag 360
ctacagaatg catctgtagc aacaacagac acaagttgtg agtctgttgt gactagtggt 420
cagcaccagc ataatccaac acctcagcat cctccaaggg atgctagccc cgctggactc 480
ctgtctatcg cagaggagac cttggcagag ttcctttcaa aggctacagg aactgctgtt 540
gactgggtcc agatgcctgg gatgaagcct ggtccggatt cgattggtat tgtggctatt 600
tcacacagtt gtagtggagt agctgcacga gcttgcggtc ttgtgggttt agaacctaca 660
aagattgcag aaattctcaa agatcgccca tcttggcttc gtgattgccg ttgccttgat 720
gtgttgactc cctttcctac tggaaatgga gggacaattg agcttttata catgcagaca 780
tatgctccca caactttagc ttctgctaga gatttttgga ctctgagata caccacagta 840
ttggaagatg gtagtcttgt ggtttgtgaa aggtccttaa gtggtgccca gggtggtcca 900
agcatagctc ctgcacagca ctttgtaaga gcagaaatgc ttcccagtgg gtatttgata 960
agaccttgtg aaggtggagg ttctattatc catattgttg accatatgga tttggagcca 1020
tggagtgtgc ctgaggtgtt acgtccacta tatgagtcat ctactgtact cgcccaaaag 1080
atgactattg cagccttgcg ccgcatacgc caaatagcac aggaggttac aggtgaagtg 1140
gtctttggtt ggggtaggca gccagctgtt ctgcggacat ttagccagag actgagcagg 1200
ggtttcaatg aggctgtgaa tggcttcaca gatgatggat ggtctttgat gggtaatgat 1260
ggaatggagg acgtgactat tgcaataaat tcatctccaa gcaaactcct aggttctcag 1320
gttaattctt ctaatgggct tacagctgtt ggtgggggta tactgtgtgc aaaggcatcc 1380
atgcttttac agaatgtgcc tccagcatta cttgttcgct tcttgcggga gcatcgatcg 1440
gagtgggcag attgtaacat tgacgcctat tctgcagctg cattaaaggc aagtccttat 1500
agtgtcccag gttcacgagc aggaggcttt tcaggaagtc aagttatcct ccccctggca 1560
cataccgtgg aacatgaaga gttcttggag gtcattaagc tggagggcca tggcctcaca 1620
caggaagaag ctgtgctatc tagggatatg tttctgttgc agctttgcag tggaattgat 1680
gaaaatgcag ctggtgcttg tgcccaactt gtgtttgcac caattgatga atcatttgct 1740
gatgatgctc ctttgttgcc gtctggtttc cgagttattc ctctggactc tagaacagat 1800
ggtactagtg gtcccaatcg gacattggat ctggcttcag ctcttgaggt tggatcagct 1860
ggaactagaa cttctggcga ttctggagcc aactcgttca atctaagatc tgtgttaact 1920
attgcattcc aattcactta tgagaatcat ttgcgagaaa atgtggcatc catggcccgt 1980

caatatgtac gcagtgttgt agcatctgtg cagagggttg ccatggcatt agccccctct 2040
cgactgagtt cacatgtggg gccaaggcta ccacctggca ctccagaagc acttactctt 2100
gctcggtgga tttgtcaaag ctacagattc cacctaggtg tagagctgct tcgcgctgat 2160
tgtgaggcca gtgaatccgt gctgaaacta ctctggcacc attcagatgc aattgtgtgc 2220
tgttctttga agtcgctacc agttttcaca tttgcaaatc aagcaggcct tgacatgctg 2280
gagacaaccc tggttgcctt gcaagatata tcattggaca aaatactcga tgaaaatgga 2340
cgcaaaagtt tatgctcaga ttttgcacaa attatgcaac agggctatgc ttatctgcct 2400
gcggggatat gtgtctctag tatgggcagg cctgtttcat atgaccgggc tgttgcttgg 2460
aaggtcctaa atgatgcaga cagcacccac tgcatggttt tcatgtttat gaattggtct 2520
ttcatgtga 2529

<210> 227
<211> 842
<212> PRT
<213> Ginkgo biloba
<400> 227

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

Gly Arg Gln Pro Ala Val Leu Arg Thr Phe Ser Gln Arg Leu Ser Arg
385 390 395 400
Gly Phe Asn Glu Ala Val Asn Gly Phe Thr Asp Asp Gly Trp Ser Leu
405 410 415
Met Gly Asn Asp Gly Met Glu Asp Val Thr Ile Ala Ile Asn Ser Ser
420 425 430
Pro Ser Lys Leu Leu Gly Ser Gln Val Asn Ser Ser Asn Gly Leu Thr
435 440 445
Ala Val Gly Gly Gly Ile Leu Cys Ala Lys Ala Ser Met Leu Leu Gln
450 455 460
Asn Val Pro Pro Ala Leu Leu Val Arg Phe Leu Arg Glu His Arg Ser
465 470 475 480
Glu Trp Ala Asp Cys Asn Ile Asp Ala Tyr Ser Ala Ala Ala Leu Lys
485 490 495
Ala Ser Pro Tyr Ser Val Pro Gly Ser Arg Ala Gly Gly Phe Ser Gly
500 505 510
Ser Gln Val Ile Leu Pro Leu Ala His Thr Val Glu His Glu Glu Phe
515 520 525
Leu Glu Val Ile Lys Leu Glu Gly His Gly Leu Thr Gln Glu Glu Ala
530 535 540
Val Leu Ser Arg Asp Met Phe Leu Leu Gln Leu Cys Ser Gly Ile Asp
545 550 555 560
Glu Asn Ala Ala Gly Ala Cys Ala Gln Leu Val Phe Ala Pro Ile Asp
565 570 575
Glu Ser Phe Ala Asp Asp Ala Pro Leu Leu Pro Ser Gly Phe Arg Val
580 585 590
Ile Pro Leu Asp Ser Arg Thr Asp Gly Thr Ser Gly Pro Asn Arg Thr
595 600 605
Leu Asp Leu Ala Ser Ala Leu Glu Val Gly Ser Ala Gly Thr Arg Thr
610 615 620
Ser Gly Asp Ser Gly Ala Asn Ser Phe Asn Leu Arg Ser Val Leu Thr
625 630 635 640
Ile Ala Phe Gln Phe Thr Tyr Glu Asn His Leu Arg Glu Asn Val Ala
645 650 655
Ser Met Ala Arg Gln Tyr Val Arg Ser Val Val Ala Ser Val Gln Arg
660 665 670
Val Ala Met Ala Leu Ala Pro Ser Arg Leu Ser Ser His Val Gly Pro
675 680 685
Arg Leu Pro Pro Gly Thr Pro Glu Ala Leu Thr Leu Ala Arg Trp Ile
690 695 700
Cys Gln Ser Tyr Arg Phe His Leu Gly Val Glu Leu Leu Arg Ala Asp
705 710 715 720
Cys Glu Ala Ser Glu Ser Val Leu Lys Leu Leu Trp His His Ser Asp
725 730 735
Ala Ile Val Cys Cys Ser Leu Lys Ser Leu Pro Val Phe Thr Phe Ala
740 745 750
Asn Gln Ala Gly Leu Asp Met Leu Glu Thr Thr Leu Val Ala Leu Gln
755 760 765
Asp Ile Ser Leu Asp Lys Ile Leu Asp Glu Asn Gly Arg Lys Ser Leu
770 775 780
Cys Ser Asp Phe Ala Gln Ile Met Gln Gln Gly Tyr Ala Tyr Leu Pro
785 790 795 800
Ala Gly Ile Cys Val Ser Ser Met Gly Arg Pro Val Ser Tyr Asp Arg
805 810 815
Ala Val Ala Trp Lys Val Leu Asn Asp Ala Asp Ser Thr His Cys Met
820 825 830
Val Phe Met Phe Met Asn Trp Ser Phe Met
835 840
<210> 228
<211> 2511
<212> DNA
<213> Gossypium barbadense
<400> 228

```
atggctatgg cgatgacaca tcatcacaac agggaaagca gcatagacaa gcatttagat     60
acaggcaagt atgttaggta cacagctgaa caagttgaag ctttggaacg tgtttatgct    120
gaatgtccaa aaccaagttc tttgcgtagg caacagttaa taagggaatg tcctattctt    180
tctaacattg agcctaaaca gttcaaagct ttgttccaaa atcgcaggtg tagagagaaa    240
caaaggaaag aagcttcaag gcttcaaaca gtaaatagaa aattaacagc aatgaataag    300
ctgttaatgg aagagaatga taggttgcaa aaacaggttt ctcagttggt ttgtgagaat    360
gggtacatga gacagcaatt gcatactgta aatgcatcgg cgactgatgc gagctgtgac    420
tctgcggtaa ccactcctca gcattcactg agaaatgcta ataaccctgc tggactcctc    480
tctatcgcgg aggagacctt ggcagagttc ctttctaagg ctacgggaac tgctgtcaat    540
tgggtccaga tgcctgggat gaagcctggt ccagattcag ttgggatatt tgccacttcc    600
caaagttgta gtggaatggc agctcgagct tgtggtcttg taagtttaga acctacaaag    660
attgcagaga ttcttaaaga tcgtccatct tggttccggg actgtcggaa gcttgaagtt    720
ttcaccatgt ttccagctgg caatggtggt acgattgagc ttgtatacac acagatgttt    780
gctcctacta cattggctcc tgcaagggac ttttggactc taaggtacac tacaacttta    840
gagaacggca gtctcgtggt gtgtgagagg tctctttcgg gttccggtgc tggcccgagt    900
gtggcttccg cagctcaatt tgtgagagct gaagtgctcc ctagtggcta tttgattagg    960
ccttgtgagg gtggagggtc gattatccat attgttgacc acttgaatct tgaggcatgg   1020
agtgtgccgg aggtcttgcg ccccctttat gaatcatcca gagtaattgc tcagaaaatg   1080
actattccgg cgctgcgcta tgttaggcag attgctcaag agacaagcgg cgaggtggtt   1140
tacagtttgg gcaggcagcc tgctgtgctt agaacattta gccaaagatt aagcaggggc   1200
ttcaatgagg caataaatgg attcaacgaa gatggctggt cgataatgaa ttgtgatggt   1260
acagaggatg tgataattgc tattaattcc ggcaagagct tgagcaacag ttcgaatttg   1320
accaccggtc tttcattcct cgggggcgtt ctatgtgcaa aggcatccat gctgcttcaa   1380
aatgttcctc ctgctgttct tgtccgattc ttgagggaac atcgtttgga atgggctgat   1440
ttcaatgtcg atgcttattc agctgcatca ttgaaggccg gaacatacac ttatcctgga   1500
atgagaccta caagttttac tgggagtcag atcatcatgc cactcggcca gacggtcgaa   1560
catgaagagc tgctcgaagt tataagactc gaaggccagt ctcttacgca agaagacgcg   1620
cttctatcaa gggacattca tctattacag atatgtagtg gaattgatga caatgcggtt   1680
ggggcctgtt cggagcttgt gtttgcacca atagatgaga tgtttccgga tgatgctgcc   1740
ttactacctt caggattccg cattatcccg ttggagtcaa aaccagattc gttggctaca   1800
aatcggactt tggatcttac ttcgagtctc gaagtggggc ctgcaacaag ccaagctgcc   1860
ggagattctc cgagtcagaa tgcacgatcg gtgttgacaa ttgccttcca gtttcccttc   1920
gatacaaatc ttcgggataa tgttgcgacc atggcacgcc agtatgtccg tagcgtcatt   1980
tcttctgtcc agaggrttgc tatggccata tctccatgyg gatcgagccc aactatagga   2040
ccaaagccat ctccaggttc tcccgaagcg cttacgttgg ctcattggat ctgccagagc   2100
tacagtttcc atttgggggа agagttgttg aaatccgaat cacttggtgg cgactcagta   2160
ttgaagaatc tttggcaaca tcaggatgca atattgtgtt gttcgttgaa gtctgtaccg   2220
gttttcatct tcgcaaatca ggccggtcta gacatgcttg agacaactct agtggatcta   2280
ccagacatca cactggacaa aatattcgat gagtcgggac ggaaggcatt gtgctctgat   2340
ttcaccaagt tgatgcagca gggattcact cacttgctgg ccggagtttg catgtcgaca   2400
atgggccgcc acgtctcgta tgaacaagct gttgcttgga aagtacttgc agctgatgca   2460
aacactgtcc actgcttggc attctctttt ataaactggt cttttgtgtg a            2511
<210> 229
<211> 836
<212> PRT
<213> Gossypium barbadense
<220>
<221> INCERTO
<222> (666)..(666)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 229
Met Ala Met Ala Met Thr His His His Asn Arg Glu Ser Ser Ile Asp
1               5                  10                  15
Lys His Leu Asp Thr Gly Lys Tyr Val Arg Tyr Thr Ala Glu Gln Val
            20                  25                  30
Glu Ala Leu Glu Arg Val Tyr Ala Glu Cys Pro Lys Pro Ser Ser Leu
        35                  40                  45
Arg Arg Gln Gln Leu Ile Arg Glu Cys Pro Ile Leu Ser Asn Ile Glu
    50                  55                  60
Pro Lys Gln Phe Lys Ala Leu Phe Gln Asn Arg Arg Cys Arg Glu Lys
65                  70                  75                  80
Gln Arg Lys Glu Ala Ser Arg Leu Gln Thr Val Asn Arg Lys Leu Thr
                85                  90                  95
```

```
Ala Met Asn Lys Leu Leu Met Glu Glu Asn Asp Arg Leu Gln Lys Gln
          100                 105                 110
Val Ser Gln Leu Val Cys Glu Asn Gly Tyr Met Arg Gln Gln Leu His
       115                 120                 125
Thr Val Asn Ala Ser Ala Thr Asp Ala Ser Cys Asp Ser Ala Val Thr
   130                 135                 140
Thr Pro Gln His Ser Leu Arg Asn Ala Asn Asn Pro Ala Gly Leu Leu
145                 150                 155                 160
Ser Ile Ala Glu Glu Thr Leu Ala Glu Phe Leu Ser Lys Ala Thr Gly
                165                 170                 175
Thr Ala Val Asn Trp Val Gln Met Pro Gly Met Lys Pro Gly Pro Asp
            180                 185                 190
Ser Val Gly Ile Phe Ala Thr Ser Gln Ser Cys Ser Gly Met Ala Ala
        195                 200                 205
Arg Ala Cys Gly Leu Val Ser Leu Glu Pro Thr Lys Ile Ala Glu Ile
    210                 215                 220
Leu Lys Asp Arg Pro Ser Trp Phe Arg Asp Cys Arg Lys Leu Glu Val
225                 230                 235                 240
Phe Thr Met Phe Pro Ala Gly Asn Gly Gly Thr Ile Glu Leu Val Tyr
                245                 250                 255
Thr Gln Met Phe Ala Pro Thr Thr Leu Ala Pro Ala Arg Asp Phe Trp
            260                 265                 270
Thr Leu Arg Tyr Thr Thr Thr Leu Glu Asn Gly Ser Leu Val Val Cys
        275                 280                 285
Glu Arg Ser Leu Ser Gly Ser Gly Ala Gly Pro Ser Val Ala Ser Ala
    290                 295                 300
Ala Gln Phe Val Arg Ala Glu Val Leu Pro Ser Gly Tyr Leu Ile Arg
305                 310                 315                 320
Pro Cys Glu Gly Gly Gly Ser Ile Ile His Ile Val Asp His Leu Asn
                325                 330                 335
Leu Glu Ala Trp Ser Val Pro Glu Val Leu Arg Pro Leu Tyr Glu Ser
            340                 345                 350
Ser Arg Val Ile Ala Gln Lys Met Thr Ile Pro Ala Leu Arg Tyr Val
        355                 360                 365
Arg Gln Ile Ala Gln Glu Thr Ser Gly Glu Val Val Tyr Ser Leu Gly
    370                 375                 380
Arg Gln Pro Ala Val Leu Arg Thr Phe Ser Gln Arg Leu Ser Arg Gly
385                 390                 395                 400
Phe Asn Glu Ala Ile Asn Gly Phe Asn Glu Asp Gly Trp Ser Ile Met
                405                 410                 415
Asn Cys Asp Gly Thr Glu Asp Val Ile Ile Ala Ile Asn Ser Gly Lys
            420                 425                 430
Ser Leu Ser Asn Ser Ser Asn Leu Thr Thr Gly Leu Ser Phe Leu Gly
        435                 440                 445
Gly Val Leu Cys Ala Lys Ala Ser Met Leu Leu Gln Asn Val Pro Pro
    450                 455                 460
Ala Val Leu Val Arg Phe Leu Arg Glu His Arg Leu Glu Trp Ala Asp
465                 470                 475                 480
Phe Asn Val Asp Ala Tyr Ser Ala Ala Ser Leu Lys Ala Gly Thr Tyr
                485                 490                 495
Thr Tyr Pro Gly Met Arg Pro Thr Ser Phe Thr Gly Ser Gln Ile Ile
            500                 505                 510
Met Pro Leu Gly Gln Thr Val Glu His Glu Glu Leu Leu Glu Val Ile
        515                 520                 525
Arg Leu Glu Gly Gln Ser Leu Thr Gln Glu Asp Ala Leu Leu Ser Arg
    530                 535                 540
Asp Ile His Leu Leu Gln Ile Cys Ser Gly Ile Asp Asp Asn Ala Val
545                 550                 555                 560
Gly Ala Cys Ser Glu Leu Val Phe Ala Pro Ile Asp Glu Met Phe Pro
                565                 570                 575
Asp Asp Ala Ala Leu Leu Pro Ser Gly Phe Arg Ile Ile Pro Leu Glu
            580                 585                 590
Ser Lys Pro Asp Ser Leu Ala Thr Asn Arg Thr Leu Asp Leu Thr Ser
```

595 600 605
Ser Leu Glu Val Gly Pro Ala Thr Ser Gln Ala Ala Gly Asp Ser Pro
610 615 620
Ser Gln Asn Ala Arg Ser Val Leu Thr Ile Ala Phe Gln Phe Pro Phe
625 630 635 640
Asp Thr Asn Leu Arg Asp Asn Val Ala Thr Met Ala Arg Gln Tyr Val
645 650 655
Arg Ser Val Ile Ser Ser Val Gln Arg Xaa Ala Met Ala Ile Ser Pro
660 665 670
Cys Gly Ser Ser Pro Thr Ile Gly Pro Lys Pro Ser Pro Gly Ser Pro
675 680 685
Glu Ala Leu Thr Leu Ala His Trp Ile Cys Gln Ser Tyr Ser Phe His
690 695 700
Leu Gly Glu Glu Leu Leu Lys Ser Glu Ser Leu Gly Gly Asp Ser Val
705 710 715 720
Leu Lys Asn Leu Trp Gln His Gln Asp Ala Ile Leu Cys Cys Ser Leu
725 730 735
Lys Ser Val Pro Val Phe Ile Phe Ala Asn Gln Ala Gly Leu Asp Met
740 745 750
Leu Glu Thr Thr Leu Val Asp Leu Pro Asp Ile Thr Leu Asp Lys Ile
755 760 765
Phe Asp Glu Ser Gly Arg Lys Ala Leu Cys Ser Asp Phe Thr Lys Leu
770 775 780
Met Gln Gln Gly Phe Thr His Leu Leu Ala Gly Val Cys Met Ser Thr
785 790 795 800
Met Gly Arg His Val Ser Tyr Glu Gln Ala Val Ala Trp Lys Val Leu
805 810 815
Ala Ala Asp Ala Asn Thr Val His Cys Leu Ala Phe Ser Phe Ile Asn
820 825 830
Trp Ser Phe Val
835

<210> 230
<211> 2526
<212> DNA
<213> Lycopersicon esculentum
<400> 230

atggctatgg tggctcaaca gcatagggag agtagtagtg gtagtattac taaacatctt 60
gatagtagtg gaaagtatgt tcgatatacg gctgagcaag ttgaggcttt ggagagggtt 120
tatgcagagt gtcctaagcc tagctcgttg cgtcgacagc aattgatccg tgaatgtcat 180
attctgtcga atatcgagcc taagcagatc aaagtttggt ttcagaacag aaggtgtcga 240
gagaagcaaa ggaaagagtc ttctcgattg cagactgtga acagaaagtt gtctgcgatg 300
aataaactgt tgatggagga gaatgaccgc ttgcagaaac aagtctcgca gcttgtatgt 360
gaaaatggct atatgcggca acaattgcaa agtgtatcgg cggccactac tgatgtaagt 420
tgtgaatcag tggtaaccac tcctcagcat tcccttagag atgctaacaa ccctgctgga 480
ctactaccaa ttgcagaaga aaccttggca gagttccttt ctaaggctac aggaactgct 540
gtcgattggg tcccgatgcc tgggatgaag cctggtccgg attcagttgg gatttttgcc 600
atctcacaca gttgcagtgg agtggcagcc cgagcatgtg gtcttgttag tttagagcca 660
acaaagattg ctgatatcct caaagatcga ccttcttggt tccgcgactg ccggaatgtt 720
gaagttatca caatgtttcc tgctggaaat ggtggtacag ttgagctttt gtatacccag 780
atatatgctc ccacaactct ggctcccgcg cgtgattttt ggacgctgag atacacaaca 840
accctagaca atggtagtct cgtggtttgt gaaagatccc tatctggtaa tgggcctggc 900
ccaaatccta ctgctgcttc ccagtttgta agagctcaaa tgcttccatc tggatatctg 960
atccgaccgt gtgatggtgg aggatcaatc atacatattg ttgatcacct gaatcttgag 1020
gcatggagtg cccctgagat tttgcgtcca ctctatgaat cgtcgaaagt tgtggcacag 1080
aaaatgacta ttgcagcact gcgatatgca aggcaactag ctcaagagac tagcgacgag 1140
gtagtatatg gtctaggaag gcaacctgct gttcttcgaa catttagcca gagattatgc 1200
agagggttca atgatgccat caatggattc ggtgacgatg gctggtcaat gttaagttca 1260
gatggtgctg aagatgtcat agttgctgtc aattcaagga agaacctcgc aaccacctcc 1320
attcctcttt ccccgcttgg tggcgtcctt tgtgccaaag catcaatgct actccagaat 1380
gtcccccctg ccgtactggt tcggtttctg agggagcacc gttcagagtg ggcagacttt 1440
aatgttgatg cttttgtagc ttctgcattg aagtcttgtc cgtatacata tcctgggatg 1500
aggcctacca gatttaccgg tagccagata ataatgccac ttggccacac aattgagcat 1560
gaagaaatgc ttgaagttat tagacttgaa gggcactcta ttggacagga agatgctttt 1620

```
atgccaagag atattcacct tttacagatg tgtagtggca ctgatgagaa tgcagttgga   1680
gcctgttctg aactagtttt tgcccctatt gatgagatgt ttccagatga tgcacctttg   1740
cttccctccg gatttcgcgt tattcctctc gaatcaaaat caggtgatgc ccaggatacg   1800
ttgaacgcac atagaacatt agatctagca tcaagtcttg aagtgggccc agcaacaaac   1860
tcgactactg gagatgcagc ttcttgttac agtgcacgat ctgtactgac aatcgctttc   1920
caatttccat ttgaggacaa tcttcaggac aatgtggcta ccatggcccg acagtatgtt   1980
cgcagtgtgg tttcgtctgt ccaacgagtt gccatggcaa tatctcctac gggaatgaat   2040
cctacactgg gggccaagct ttctccaggc tcaccagaag ctgtaacatt gtcgcattgg   2100
atctgccaga gctatagtta tcacatggga acagagttac ttcgagctga ttctagtggc   2160
gatgaatcag tactaaagaa tctctggcaa caccaggatg caatttttgtg ctgctcattg   2220
aagtcgctgc cagttttcat ttttgctaat aaggccggac tcgacatgct ggagacaaca   2280
ttagttgcat tacaggacat ttctctagat aagatatttg atgaatccgg aaggaaagtg   2340
ttgctctcag aatttgccaa gattatggaa cagggtttcg cgtgtttgcc tggtggtatc   2400
tgcatgtcaa caatgggacg acatatttca tatgaacaag ctatcgcgtg gaaagtgttt   2460
gcgtcgtctg aagaaaatgc tgtccactgt ttggcctttt cgtttatcaa ctggtcattc   2520
gtgtaa                                                              2526
<210> 231
<211> 841
<212> PRT
<213> Lycopersicon esculentum
<400> 231
Met Ala Met Val Ala Gln Gln His Arg Glu Ser Ser Ser Gly Ser Ile
1               5                  10                  15
Thr Lys His Leu Asp Ser Ser Gly Lys Tyr Val Arg Tyr Thr Ala Glu
            20                  25                  30
Gln Val Glu Ala Leu Glu Arg Val Tyr Ala Glu Cys Pro Lys Pro Ser
        35                  40                  45
Ser Leu Arg Arg Gln Gln Leu Ile Arg Glu Cys His Ile Leu Ser Asn
    50                  55                  60
Ile Glu Pro Lys Gln Ile Lys Val Trp Phe Gln Asn Arg Arg Cys Arg
65                  70                  75                  80
Glu Lys Gln Arg Lys Glu Ser Ser Arg Leu Gln Thr Val Asn Arg Lys
                85                  90                  95
Leu Ser Ala Met Asn Lys Leu Leu Met Glu Glu Asn Asp Arg Leu Gln
            100                 105                 110
Lys Gln Val Ser Gln Leu Val Cys Glu Asn Gly Tyr Met Arg Gln Gln
        115                 120                 125
Leu Gln Ser Val Ser Ala Ala Thr Thr Asp Val Ser Cys Glu Ser Val
    130                 135                 140
Val Thr Thr Pro Gln His Ser Leu Arg Asp Ala Asn Asn Pro Ala Gly
145                 150                 155                 160
Leu Leu Pro Ile Ala Glu Glu Thr Leu Ala Glu Phe Leu Ser Lys Ala
                165                 170                 175
Thr Gly Thr Ala Val Asp Trp Val Pro Met Pro Gly Met Lys Pro Gly
            180                 185                 190
Pro Asp Ser Val Gly Ile Phe Ala Ile Ser His Ser Cys Ser Gly Val
        195                 200                 205
Ala Ala Arg Ala Cys Gly Leu Val Ser Leu Glu Pro Thr Lys Ile Ala
    210                 215                 220
Asp Ile Leu Lys Asp Arg Pro Ser Trp Phe Arg Asp Cys Arg Asn Val
225                 230                 235                 240
Glu Val Ile Thr Met Phe Pro Ala Gly Asn Gly Gly Thr Val Glu Leu
                245                 250                 255
Leu Tyr Thr Gln Ile Tyr Ala Pro Thr Thr Leu Ala Pro Ala Arg Asp
            260                 265                 270
Phe Trp Thr Leu Arg Tyr Thr Thr Thr Leu Asp Asn Gly Ser Leu Val
        275                 280                 285
Val Cys Glu Arg Ser Leu Ser Gly Asn Gly Pro Gly Pro Asn Pro Thr
    290                 295                 300
Ala Ala Ser Gln Phe Val Arg Ala Gln Met Leu Pro Ser Gly Tyr Leu
305                 310                 315                 320
Ile Arg Pro Cys Asp Gly Gly Gly Ser Ile Ile His Ile Val Asp His
                325                 330                 335
```

```
Leu Asn Leu Glu Ala Trp Ser Ala Pro Glu Ile Leu Arg Pro Leu Tyr
            340                 345                 350
Glu Ser Ser Lys Val Val Ala Gln Lys Met Thr Ile Ala Ala Leu Arg
        355                 360                 365
Tyr Ala Arg Gln Leu Ala Gln Glu Thr Ser Asp Glu Val Val Tyr Gly
    370                 375                 380
Leu Gly Arg Gln Pro Ala Val Leu Arg Thr Phe Ser Gln Arg Leu Cys
385                 390                 395                 400
Arg Gly Phe Asn Asp Ala Ile Asn Gly Phe Gly Asp Asp Gly Trp Ser
                405                 410                 415
Met Leu Ser Ser Asp Gly Ala Glu Asp Val Ile Val Ala Val Asn Ser
            420                 425                 430
Arg Lys Asn Leu Ala Thr Thr Ser Ile Pro Leu Ser Pro Leu Gly Gly
        435                 440                 445
Val Leu Cys Ala Lys Ala Ser Met Leu Leu Gln Asn Val Pro Pro Ala
    450                 455                 460
Val Leu Val Arg Phe Leu Arg Glu His Arg Ser Glu Trp Ala Asp Phe
465                 470                 475                 480
Asn Val Asp Ala Phe Val Ala Ser Ala Leu Lys Ser Cys Pro Tyr Thr
                485                 490                 495
Tyr Pro Gly Met Arg Pro Thr Arg Phe Thr Gly Ser Gln Ile Ile Met
            500                 505                 510
Pro Leu Gly His Thr Ile Glu His Glu Glu Met Leu Glu Val Ile Arg
        515                 520                 525
Leu Glu Gly His Ser Ile Gly Gln Glu Asp Ala Phe Met Pro Arg Asp
    530                 535                 540
Ile His Leu Leu Gln Met Cys Ser Gly Thr Asp Glu Asn Ala Val Gly
545                 550                 555                 560
Ala Cys Ser Glu Leu Val Phe Ala Pro Ile Asp Glu Met Phe Pro Asp
                565                 570                 575
Asp Ala Pro Leu Leu Pro Ser Gly Phe Arg Val Ile Pro Leu Glu Ser
            580                 585                 590
Lys Ser Gly Asp Ala Gln Asp Thr Leu Asn Ala His Arg Thr Leu Asp
        595                 600                 605
Leu Ala Ser Ser Leu Glu Val Gly Pro Ala Thr Asn Ser Thr Thr Gly
    610                 615                 620
Asp Ala Ala Ser Cys Tyr Ser Ala Arg Ser Val Leu Thr Ile Ala Phe
625                 630                 635                 640
Gln Phe Pro Phe Glu Asp Asn Leu Gln Asp Asn Val Ala Thr Met Ala
                645                 650                 655
Arg Gln Tyr Val Arg Ser Val Val Ser Ser Val Gln Arg Val Ala Met
            660                 665                 670
Ala Ile Ser Pro Thr Gly Met Asn Pro Thr Leu Gly Ala Lys Leu Ser
        675                 680                 685
Pro Gly Ser Pro Glu Ala Val Thr Leu Ser His Trp Ile Cys Gln Ser
    690                 695                 700
Tyr Ser Tyr His Met Gly Thr Glu Leu Leu Arg Ala Asp Ser Ser Gly
705                 710                 715                 720
Asp Glu Ser Val Leu Lys Asn Leu Trp Gln His Gln Asp Ala Ile Leu
                725                 730                 735
Cys Cys Ser Leu Lys Ser Leu Pro Val Phe Ile Phe Ala Asn Lys Ala
            740                 745                 750
Gly Leu Asp Met Leu Glu Thr Thr Leu Val Ala Leu Gln Asp Ile Ser
        755                 760                 765
Leu Asp Lys Ile Phe Asp Glu Ser Gly Arg Lys Val Leu Leu Ser Glu
    770                 775                 780
Phe Ala Lys Ile Met Glu Gln Gly Phe Ala Cys Leu Pro Gly Gly Ile
785                 790                 795                 800
Cys Met Ser Thr Met Gly Arg His Ile Ser Tyr Glu Gln Ala Ile Ala
                805                 810                 815
Trp Lys Val Phe Ala Ser Ser Glu Glu Asn Ala Val His Cys Leu Ala
            820                 825                 830
Phe Ser Phe Ile Asn Trp Ser Phe Val
```

835 840

<210> 232
<211> 258
<212> DNA
<213> Saccharum officinarum
<400> 232
atgaggatgt tcttccggca cttgctatct tgcatcctcc tgctcttgct aatgtcacac 60
ttgccaagcc cgatcctcgg cttgagaaca ttgagaggag aggaggcgaa gagcgacttg 120
aggcgccatg agcatgagct tccgccagcg gtatctcctt caaaagaggt ggacaacgat 180
gacgctgccg cctccagtaa gttcacagtt tcaagaagaa tggttcctca aggtccaaac 240
cctctccaca acagatga 258
<210> 233
<211> 85
<212> PRT
<213> Saccharum officinarum
<400> 233
Met Arg Met Phe Phe Arg His Leu Leu Ser Cys Ile Leu Leu Leu Leu
1 5 10 15
Leu Met Ser His Leu Pro Ser Pro Ile Leu Gly Leu Arg Thr Leu Arg
20 25 30
Gly Glu Glu Ala Lys Ser Asp Leu Arg Arg His Glu His Glu Leu Pro
35 40 45
Pro Ala Val Ser Pro Ser Lys Glu Val Asp Asn Asp Asp Ala Ala Ala
50 55 60
Ser Ser Lys Phe Thr Val Ser Arg Arg Met Val Pro Gln Gly Pro Asn
65 70 75 80
Pro Leu His Asn Arg
85
<210> 234
<211> 53
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm05843
<400> 234
ggggacaagt ttgtacaaaa aagcaggctt aaacaatgag gatgttcttc cgg 53
<210> 235
<211> 49
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm05844
<400> 235
ggggaccact ttgtacaaga aagctgggtt cctctcatct gttgtggag 49
<210> 236
<211> 668
<212> DNA
<213> Oryza sativa
<400> 236
ccttctacat cggcttaggt gtagcaacac gactttatta ttattattat tattattatt 60
attattttac aaaaatataa aatagatcag tccctcacca caagtagagc aagttggtga 120
gttattgtaa agttctacaa agctaattta aaagttattg cattaactta tttcatatta 180
caaacaagag tgtcaatgga acaatgaaaa ccatatgaca tactataatt ttgtttttat 240
tattgaaatt atataattca aagagaataa atccacatag ccgtaaagtt ctacatgtgg 300
tgcattacca aaatatatat agcttacaaa acatgacaag cttagtttga aaaattgcaa 360
tccttatcac attgacacat aaagtgagtg atgagtcata atattatttt tcttgctacc 420
catcatgtat atatgatagc cacaaagtta ctttgatgat gatatcaaag aacattttta 480
ggtgcaccta acagaatatc caaataatat gactcactta gatcataata gagcatcaag 540
taaaactaac actctaaagc aaccgatggg aaagcatcta taaatagaca agcacaatga 600
aaatcctcat catccttcac cacaattcaa atattatagt tgaagcatag tagtagaatc 660
caacaaca 668
<210> 237
<211> 14

<212> PRT
<213> Seqüência artificial
<220>
<223> Seqüência de consenso, domínio CLE
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir="Glu" /substituir="Arg" /substituir="Met" /substituir="Pro" /substituir="Leu
<220>
<221> VARIANTE
<222> (2)..(2)
<223> /substituir="Glu" /substituir="Asp" /substituir="Arg" /substituir="Ser"
<220>
<221> VARIANTE
<222> (4)..(4)
<223> /substituir="Ile" /substituir="Leu" /substituir="Arg" /substituir="Phe" /substituir="Gln" /substituir="Val" /substituir="Met"
<220>
<221> VARIANTE
<222> (5)..(5)
<223> /substituir="Ile" /substituir="Ser" /substituir="Leu"
<220>
<221> VARIANTE
<222> (6)..(6)
<223> /substituir="Pro" /substituir="Leu" /substituir="Arg"
<220>
<221> VARIANTE
<222> (7)..(7)
<223> /substituir="Thr" /substituir="Gln" /substituir="Cys" /substituir="Asn" /substituir="Gly" /substituir="Lys" /substituir="Ser"
<220>
<221> VARIANTE
<222> (8)..(8)
<223> /substituir="Gly"
<220>
<221> VARIANTE
<222> (9)..(9)
<223> /substituir="Pro"
<220>
<221> VARIANTE
<222> (10)..(10)
<223> /substituir="Asn" /substituir="Tyr"
<220>
<221> VARIANTE
<222> (12)..(12)
<223> /substituir="Leu" /substituir="Gln" /substituir="Arg" /substituir="Tyr" /substituir="His"
<400> 237
Ser Lys Arg Lys Val Pro Arg Asn Ser Asp Pro Ile His His
1 5 10
<210> 238
<211> 14
<212> PRT
<213> Seqüência artificial
<220>
<223> Seqüência de consenso, domínio CLE
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir="Pro"
<220>

<221> VARIANTE
<222> (2)..(2)
<223> /substituir="Glu" /substituir="Lys"
<220>
<221> VARIANTE
<222> (4)..(4)
<223> /substituir="Leu" /substituir="Ile"
<220>
<221> VARIANTE
<222> (5)..(5)
<223> /substituir="Ser" /substituir="Ile"
<220>
<221> VARIANTE
<222> (7)..(7)
<223> /substituir="Gly" /substituir="Cys" /substituir="Thr" /substituir="Ser"
<220>
<221> VARIANTE
<222> (10)..(10)
<223> /substituir="Asp"
<220>
<221> VARIANTE
<222> (12)..(12)
<223> /substituir="Gln" /substituir="His"
<220>
<221> VARIANTE
<222> (14)..(14)
<223> /substituir="Asn"
<400> 238
Ser Arg Arg Met Val Pro Gln Gly Pro Asn Pro Leu His His
1 5 10
<210> 239
<211> 477
<212> DNA
<213> Populus trichocarpa x Populus deltoides
<400> 239
gcaaattccc cctgctctaa aatccatttt cctatctatc aacctctctg gtctctatat 60
ccccaccatt ttcatcatga ggaataaaaa tcctcagctg ttccttattt tcctgatagt 120
gttcctggta ctcgttcatg gcaccacttg ccgcgatgcc aaaagatcta ctagcaatgg 180
agaaacagta caagggtcga aaaccaaaca ttcttcaatg tttctacaag cgctttcttc 240
catttttaag gcttcagaat ccagcactaa caacatcaag gcacttcaca ccgtctcgcg 300
caggcttgtt ccttgtggac caaacccact ccacaactga tcatcgaatc aacaaaattc 360
caatctgttt ccatttgttg aaatatatat agtgttttta aaatattttt gttgaataat 420
ctatacatat tttccctata catccagttt tgaaaaaaga aaaaaaaaaa aaaaaaa 477
<210> 240
<211> 87
<212> PRT
<213> Populus trichocarpa x Populus deltoides
<400> 240
Met Arg Asn Lys Asn Pro Gln Leu Phe Leu Ile Phe Leu Ile Val Phe
1 5 10 15
Leu Val Leu Val His Gly Thr Thr Cys Arg Asp Ala Lys Arg Ser Thr
20 25 30
Ser Asn Gly Glu Thr Val Gln Gly Ser Lys Thr Lys His Ser Ser Met
35 40 45
Phe Leu Gln Ala Leu Ser Ser Ile Phe Lys Ala Ser Glu Ser Ser Thr
50 55 60
Asn Asn Ile Lys Ala Leu His Thr Val Ser Arg Arg Leu Val Pro Cys
65 70 75 80
Gly Pro Asn Pro Leu His Asn
85
<210> 241
<211> 443

<212> DNA
<213> Oryza sativa
<400> 241
tctctctctc tcacacacac acacacacaa actagagact tatgccatga ggaggttctc 60
caagcagcac ctggtccctt tcattctgct cctgctactt gtcatgtctc acttgccaat 120
ctcaagcctt ggttcaagga gagcattcag agaagaagct gtcagtggtt tcagaagcca 180
tgagcttgct ccaaccatgg ctccttctca agagaaagaa gcaggcgtcg tcgccggcgc 240
cggtagcatc tgtggccaga agtatgctgt ctcaagaaga atggttcctc agggaccaaa 300
ccctctccac aactgatgaa ctcatgcact gcaatctgcc gtgatcatca gcttctccaa 360
tacaggtgaa tggtgcagcc atcattggtt acctttaagt ccgtctgctt aattaactag 420
atatttcata gtattttatc aca 443
<210> 242
<211> 89
<212> PRT
<213> Oryza sativa
<400> 242
Met Arg Arg Phe Ser Lys Gln His Leu Val Pro Phe Ile Leu Leu Leu
1 5 10 15
Leu Leu Val Met Ser His Leu Pro Ile Ser Ser Leu Gly Ser Arg Arg
20 25 30
Ala Phe Arg Glu Glu Ala Val Ser Gly Phe Arg Ser His Glu Leu Ala
35 40 45
Pro Thr Met Ala Pro Ser Gln Glu Lys Glu Ala Gly Val Val Ala Gly
50 55 60
Ala Gly Ser Ile Cys Gly Gln Lys Tyr Ala Val Ser Arg Arg Met Val
65 70 75 80
Pro Gln Gly Pro Asn Pro Leu His Asn
85
<210> 243
<211> 677
<212> DNA
<213> Saccharum officinarum
<400> 243
tttgaaactg ttacaactgg ttcgcctgca cctgcctgcc cggaacgacc cacgcgtccg 60
cgctctctct ctccctctcc ctctacctct ctctaacata tgccatgagg atgttcttcc 120
ggcactagct atcttgcatc ctactgctct aggtaatgtc acacttgcca agctcgatcc 180
tcggcttgag aacatcgaga ggagaggagg cgaagagcga cttgaggcgc catgagcatg 240
agcttccgcc agccgtatct ccttcaaaag aggtggacaa cgacgacgct gccgccgcca 300
gtaagttaac agtttcaaga agaatggttc ctcaaggtcc aaaccctctc cacaacagat 360
gagaggatcg gagcagcatg ctgctgctgc ttctggtatg gtcttcagcc aaggtagcag 420
ctagctagct cttctcagtc tcagctaatg gtgtagtgag tggcacgcac tcaacaaagt 480
accttcctc tctttccttt ttttttaacc cgtcatatga atgatacaaa tggatgcata 540
tataagttga atactacttc ccatgtaatg tgtttttgtt gcattgattt catttatagg 600
cagctttgta catagatttt tatgatatta aggtgtaaaa gttgttgtgc tagatggatc 660
tagattttga atgtttg 677
<210> 244
<211> 68
<212> PRT
<213> Saccharum officinarum
<400> 244
Met Ser His Leu Pro Ser Ser Ile Leu Gly Leu Arg Thr Ser Arg Gly
1 5 10 15
Glu Glu Ala Lys Ser Asp Leu Arg Arg His Glu His Glu Leu Pro Pro
20 25 30
Ala Val Ser Pro Ser Lys Glu Val Asp Asn Asp Asp Ala Ala Ala Ala
35 40 45
Ser Lys Leu Thr Val Ser Arg Arg Met Val Pro Gln Gly Pro Asn Pro
50 55 60
Leu His Asn Arg
65
<210> 245
<211> 533
<212> DNA

<213> Arabidopsis thaliana
<400> 245
gcaacgagaa aaaccgtcct tggtgacaac tcatgcttgc actcaagcct taagctagct 60
aaacctatct cgcgcactac tagaattcaa ataaaactct ataaatagaa accctcatga 120
gatctcttct ttcctcatat acactcatac acaccacgtg aacaatctat ctctctttct 180
attgcttttc tatatataca gaaactaatt aattgtatct gtaatggcta agttaagctt 240
cactttctgc ttcttgttgt ttcttctgtt atcctcaatc gccgctggaa gccgccctct 300
tgaggggct cgggtcgggg tgaaggtgag aggcctaagc ccttctatcg aggctacgag 360
tccgactgta gaggatgatc aagctgcggg tagccatggg aaatctccag agcggttaag 420
cccaggagga cccgacccac aacatcacta gttatttgt gtttttcaat ttcttcgaca 480
tgtttattac ttatcaataa tttggttgca acgaagctgt ttttcttttt tgt 533
<210> 246
<211> 75
<212> PRT
<213> Arabidopsis thaliana
<400> 246
Met Ala Lys Leu Ser Phe Thr Phe Cys Phe Leu Leu Phe Leu Leu Leu
1 5 10 15
Ser Ser Ile Ala Ala Gly Ser Arg Pro Leu Glu Gly Ala Arg Val Gly
20 25 30
Val Lys Val Arg Gly Leu Ser Pro Ser Ile Glu Ala Thr Ser Pro Thr
35 40 45
Val Glu Asp Asp Gln Ala Ala Gly Ser His Gly Lys Ser Pro Glu Arg
50 55 60
Leu Ser Pro Gly Gly Pro Asp Pro Gln His His
65 70 75
<210> 247
<211> 417
<212> DNA
<213> Brassica napus
<400> 247
ctttcatatt gtatctcccg cagccaaaca aaaacttttt ttttgacaaa gatagaacat 60
gaagatcaag agtttgatat tggcttcctc ttttctgatt cttgccttca ttcatcactc 120
agaatcagct tcatttcgga gtttgctgat gaagaatgga ttgtacgaag aagaagaagc 180
aaaaattcta ttgggcgact ccaaagaaac gataactaat tctacagctt tggagtctaa 240
acggataatt ccgacgggtc caaatccact tcacaacagg taactttgat catttaagaa 300
caagatatgt tgtgagtatg tctcttcctc tgtttttgtt acaagtatct atcttactgt 360
gtaatgtaat ggtgaatgta taatacattt tctaattaaa ttaccttatt taaaaaa 417
<210> 248
<211> 74
<212> PRT
<213> Brassica napus
<400> 248
Met Lys Ile Lys Ser Leu Ile Leu Ala Ser Ser Phe Leu Ile Leu Ala
1 5 10 15
Phe Ile His His Ser Glu Ser Ala Ser Phe Arg Ser Leu Leu Met Lys
20 25 30
Asn Gly Leu Tyr Glu Glu Glu Glu Ala Lys Ile Leu Leu Gly Asp Ser
35 40 45
Lys Glu Thr Ile Thr Asn Ser Thr Ala Leu Glu Ser Lys Arg Ile Ile
50 55 60
Pro Thr Gly Pro Asn Pro Leu His Asn Arg
65 70
<210> 249
<211> 774
<212> DNA
<213> Arabidopsis thaliana
<400> 249
aacagtttct atatttctct ctgtatctct ctcactcagt cactttctct ctaaaaaatg 60
gattctaaaa gctttgtgct actactacta ctcttctgct tcttgttcct tcatgatgct 120
tctggtctca ctctctcttt cactcctcta tatgtctata acccatgtaa atggattctt 180
ataagtctca acataacttt tttttgttat ttactatttc accagatctc actcaagctc 240
atgctcacgt tcaaggactt tccaaccgca aggttatctt caacttgtac tcattaaggc 300

ctctcagcat tcatgtgtta tgttcatgta gatgtccggt ccagttcaac aactgtttca 360
ttgctttagt tgtcacgaga aatatttgta tatattatta tggtgtgcaa aacatagtaa 420
aatgttgttc aattggcaga tgatgatgat gaaaatggaa agtgaatggg ttggagcaaa 480
tggagaagca gagaaggcaa agacgaaggg tttaggacta catgaagagt taaggactgt 540
tccttcggga cctgacccgt tgcaccatca tgtgaaccca ccaagacagc caagaaacaa 600
ctttcagctc ccttgaccta atctcttgtt gctttaaatt atttcatatt gtaaattact 660
ttctgcttta tcggttttac catttcggga gtctttttg tgtgcaatct gtttcgtttg 720
gtgtagtttc atgaaagtga atgtaagata tgcattacgt ttgttgctga agtg 774

<210> 250
<211> 96
<212> PRT
<213> Arabidopsis thaliana
<400> 250

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

<210> 251
<211> 353
<212> DNA
<213> Oryza sativa
<400> 251

atggaaggtg taggtgctag gcagaggagg aaccctctga tacccagacc aaacggttca 60
aagaggcatc tgcagcatca gcatcagcca aatgctgccg agaagaagac cgccgcgaca 120
tcgaattact tcagtatcga ggcgttcctc gtgctcgtct tcctcaccat gtcattgctc 180
atacttccat tggtgcttcc cccattgcct ccgccgccat cgctgctgct gctgctgcca 240
gtctgcctgc tcatcctgct ggttgtgctg gccttcatgc caacggatgt gcggagcatg 300
gcttcctctt acttgtaaat acatctccta ggggaattta tttttgtttt tga 353

<210> 252
<211> 105
<212> PRT
<213> Oryza sativa
<400> 252

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

<210> 253
<211> 56
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm08170
<400> 253

ggggacaagt ttgtacaaaa aagcaggctt aaacaatgga aggtgtaggt gctagg 56

<210> 254
<211> 51
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm08171
<400> 254
ggggaccact ttgtacaaga aagctgggtc aaaaacaaaa ataaattccc c 51
<210> 255
<211> 2193
<212> DNA
<213> Oryza sativa
<400> 255
aatccgaaaa gtttctgcac cgttttcacc ccctaactaa caatataggg aacgtgtgct 60
aaatataaaa tgagacctta tatatgtagc gctgataact agaactatgc aagaaaaact 120
catccaccta ctttagtggc aatcgggcta aataaaaaag agtcgctaca ctagtttcgt 180
tttccttagt aattaagtgg gaaaatgaaa tcattattgc ttagaatata cgttcacatc 240
tctgtcatga agttaaatta ttcgaggtag ccataattgt catcaaactc ttcttgaata 300
aaaaaatctt tctagctgaa ctcaatgggt aaagagagag attttttta aaaaaataga 360
atgaagatat tctgaacgta ttggcaaaga tttaaacata taattatata attttatagt 420
ttgtgcattc gtcatatcgc acatcattaa ggacatgtct tactccatcc caatttttat 480
ttagtaatta aagacaattg acttatttt attatttatc ttttttcgat tagatgcaag 540
gtacttacgc acacactttg tgctcatgtg catgtgtgag tgcacctcct caatacacgt 600
tcaactagca acacatctct aatatcactc gcctatttaa tacatttagg tagcaatatc 660
tgaattcaag cactccacca tcaccagacc acttttaata atatctaaaa tacaaaaaat 720
aattttacag aatagcatga aaagtatgaa acgaactatt taggtttttc acatacaaaa 780
aaaaaaagaa ttttgctcgt gcgcgagcgc caatctccca tattgggcac acaggcaaca 840
acagagtggc tgcccacaga acaacccaca aaaaacgatg atctaacgga ggacagcaag 900
tccgcaacaa ccttttaaca gcaggctttg cggccaggag agaggaggag aggcaaagaa 960
aaccaagcat cctcctcctc ccatctataa attcctcccc ccttttcccc tctctatata 1020
ggaggcatcc aagccaagaa gagggagagc accaaggaca cgcgactagc agaagccgag 1080
cgaccgcctt cttcgatcca tatcttccgg tcgagttctt ggtcgatctc ttccctcctc 1140
cacctcctcc tcacagggta tgtgcccttc ggttgttctt ggatttattg ttctaggttg 1200
tgtagtacgg gcgttgatgt taggaaaggg gatctgtatc tgtgatgatt cctgttcttg 1260
gatttgggat agaggggttc ttgatgttgc atgttatcgg ttcggtttga ttagtagtat 1320
ggttttcaat cgtctggaga gctctatgga aatgaaatgg tttagggtac ggaatcttgc 1380
gattttgtga gtacctttg tttgaggtaa aatcagagca ccggtgattt tgcttggtgt 1440
aataaaagta cggttgtttg gtcctcgatt ctggtagtga tgcttctcga tttgacgaag 1500
ctatcctttg tttattccct attgaacaaa aataatccaa ctttgaagac ggtcccgttg 1560
atgagattga atgattgatt cttaagcctg tccaaaattt cgcagctggc ttgtttagat 1620
acagtagtcc ccatcacgaa attcatggaa acagttataa tcctcaggaa caggggattc 1680
cctgttcttc cgatttgctt tagtcccaga attttttttc ccaaatatct taaaaagtca 1740
ctttctggtt cagttcaatg aattgattgc tacaaataat gcttttatag cgttatccta 1800
gctgtagttc agttaatagg taatacccct atagtttagt caggagaaga acttatccga 1860
tttctgatct ccatttttaa ttatatgaaa tgaactgtag cataagcagt attcatttgg 1920
attatttttt ttattagctc tcacccctc attattctga gctgaaagtc tggcatgaac 1980
tgtcctcaat tttgttttca aattcacatc gattatctat gcattatcct cttgtatcta 2040
cctgtagaag tttctttttg gttattcctt gactgcttga ttacagaaag aaatttatga 2100
agctgtaatc gggatagtta tactgcttgt tcttatgatt catttccttt gtgcagttct 2160
tggtgtagct tgccactttc accagcaaag ttc 2193
<210> 256
<211> 3
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo conservado 1a
<400> 256
Tyr Phe Ser
1
<210> 257
<211> 3
<212> PRT
<213> Seqüência artificial

<220>
<223> motivo conservado 1b
<400> 257
Tyr Phe Thr
1
<210> 258
<211> 3
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo conservado 1c
<400> 258
Tyr Phe Gly
1
<210> 259
<211> 3
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo conservado 1d
<400> 259
Tyr Leu Gly
1
<210> 260
<211> 7
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo conservado 2
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir = "Ala" /substituir = "Ile"
<220>
<221> VARIANTE
<222> (7)..(7)
<223> /substituir = "Ser"
<400> 260
Val Leu Ala Phe Met Pro Thr
1 5
<210> 261
<211> 3
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo conservado 3
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir = "Pro"
<400> 261
Ser Tyr Leu
1
<210> 262
<211> 178
<212> PRT
<213> Oryza sativa
<400> 262
Met Tyr Leu Leu Ser Pro Arg Asn Gly Asp Glu Glu Asp Glu Gln Glu
1 5 10 15
Glu Ile Gln Glu Leu Ile Ser Asp Asp Glu Pro Pro Asn Leu Lys Leu
20 25 30
Ala Ser Cys Ala Thr Ala Ala Ser Ser Ser Ser Ser Ser Gly Ser Asp

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

<210> 263
<211> 126
<212> PRT
<213> Oryza sativa
<400> 263
Met Leu Leu Glu His Leu Met Ile Thr Met Glu Glu Gln Met Phe Arg
1               5                   10                  15
Glu Gln Gln Met Gln Arg Gly Gly Arg His His Gln His His Thr Thr
           20                  25                  30
Arg Glu Gln Glu Gln Gln Gln Lys Gln Gln Gln Arg Arg Arg Leu Met
       35                  40                  45
Asn Asn Ala Thr Asn Gly Gly Gly Gly Asp Gly Gly Ser Arg Cys Tyr
   50                  55                  60
Phe Ser Thr Glu Ala Ile Leu Val Leu Ala Cys Val Thr Val Ser Leu
65                  70                  75                  80
Leu Val Leu Pro Leu Ile Leu Pro Pro Leu Pro Pro Pro Pro Thr Leu
                85                  90                  95
Leu Leu Leu Leu Pro Val Cys Leu Leu Ala Leu Leu Val Val Leu Ala
           100                 105                 110
Phe Met Pro Thr Asp Met Arg Thr Met Ala Ser Ser Tyr Leu
       115                 120                 125
<210> 264
<211> 105
<212> PRT
<213> Zea mays
<220>
<221> INCERTO
<222> (65)..(75)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 264
Met Ala Ser Arg Ser Ser Ala Met Glu Gly Gly Ala Ala Ile Gln Arg
1               5                   10                  15
Arg Asn Ala Val Lys Arg His Leu Gln Gln Arg Gln Gln Glu Ala Asp
           20                  25                  30
Phe Leu Asp Lys Lys Val Ile Ala Ser Thr Tyr Phe Ser Ile Gly Ala
       35                  40                  45
Phe Leu Val Leu Ala Cys Leu Thr Val Ser Leu Leu Ile Leu Pro Leu
   50                  55                  60
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Leu Trp Leu Pro
65                  70                  75                  80
Val Cys Leu Leu Val Leu Leu Val Val Leu Ala Phe Met Pro Thr Asp
                85                  90                  95
Val Arg Ser Met Ala Ser Ser Tyr Leu
           100                 105
```

<210> 265
<211> 110
<212> PRT
<213> Triticum aestivum
<400> 265
Met Asp Ser Gln Phe Gly Ala Leu Glu Arg Gly Gly Ser Arg Gln Arg
1 5 10 15
Arg Ser Pro Val Leu Ala Arg Pro Asn Thr Thr Lys Arg His Ile Gln
20 25 30
Gln Gln Arg Ala Asn Ala Ala Asp Lys Lys Val Val Met Pro Asn Tyr
35 40 45
Phe Ser Ile Glu Ala Phe Phe Val Leu Ala Cys Leu Thr Val Ser Leu
50 55 60
Leu Ile Leu Pro Leu Val Leu Pro Pro Leu Pro Pro Pro Pro Ser Leu
65 70 75 80
Leu Leu Phe Val Pro Val Cys Leu Leu Ile Leu Leu Met Val Leu Ala
85 90 95
Phe Met Pro Thr Asp Met Arg Ser Met Ala Thr Ser Tyr Leu
100 105 110
<210> 266
<211> 110
<212> PRT
<213> Hordeum vulgare
<400> 266
Met Asp Ser Gln Phe Gly Ala Met Asp Arg Gly Gly Ser Arg Gln Arg
1 5 10 15
Ser Ser Pro Val Leu Ala Arg Pro Asn Thr Ala Lys Arg Gln Met Gln
20 25 30
Gln Gln Arg Ala Asn Ala Ala Asp Lys Lys Val Val Ile Pro Asn Tyr
35 40 45
Phe Gly Val Glu Ala Phe Phe Val Leu Ala Cys Leu Thr Val Ser Leu
50 55 60
Leu Ile Leu Pro Leu Val Leu Pro Pro Leu Pro Pro Pro Pro Ser Leu
65 70 75 80
Leu Leu Leu Leu Pro Val Cys Leu Leu Ile Leu Leu Met Val Leu Ala
85 90 95
Phe Met Pro Thr Asp Val Arg Ser Met Ala Thr Ser Tyr Leu
100 105 110
<210> 267
<211> 99
<212> PRT
<213> Saccharum officinarum
<220>
<221> INCERTO
<222> (62)..(62)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 267
Met Glu Gly Gly Gly Gln Ile Gln Arg Arg Asn Asn Ala Val Lys Arg
1 5 10 15
His Leu Gln Gln Arg Gln Gln Glu Ala Asp Phe Leu Asp Lys Lys Val
20 25 30
Ile Ala Ser Thr Tyr Phe Ser Ile Glu Ala Phe Leu Val Leu Ala Cys
35 40 45
Leu Thr Val Ser Leu Leu Ile Leu Pro Leu Val Leu Pro Xaa Leu Pro
50 55 60
Ala Pro Ala Ser Leu Leu Leu Trp Leu Pro Val Trp Leu Leu Glu Leu
65 70 75 80
Leu Ile Val Leu Ala Phe Met Pro Thr Asp Val Arg Ser Met Ala Ser
85 90 95
Ser Tyr Leu

<210> 268
<211> 99

<212> PRT
<213> Saccharum officinarum
<400> 268
Met Glu Gly Gly Gly Gln Ile Gln Arg Arg Asn Asn Ala Val Lys Arg
1 5 10 15
His Leu Gln Gln Arg Gln Gln Glu Ala Asp Phe Leu Asp Lys Lys Val
20 25 30
Ile Ala Ser Thr Tyr Phe Ser Ile Glu Ala Phe Leu Val Leu Ala Cys
35 40 45
Leu Thr Val Ser Leu Leu Ile Leu Pro Leu Val Leu Pro Pro Leu Pro
50 55 60
Pro Pro Pro Ser Leu Leu Leu Trp Leu Pro Val Cys Leu Leu Ile Leu
65 70 75 80
Leu Ile Val Leu Ala Phe Met Pro Thr Asp Val Arg Ser Met Ala Ser
85 90 95
Ser Tyr Leu

<210> 269
<211> 107
<212> PRT
<213> Sorghum bicolor
<400> 269
Met Ala Ser Arg Ser Ser Ala Leu Glu Gly Gly Gly Ala Ala Ile Gln
1 5 10 15
Arg Arg Asn Asn Ala Val Lys Arg His Leu Gln Gln Arg Gln Gln Glu
20 25 30
Ala Asp Phe His Asp Lys Lys Val Ile Ala Ser Thr Tyr Phe Ser Ile
35 40 45
Gly Ala Phe Leu Val Leu Ala Cys Leu Thr Phe Ser Leu Leu Ile Leu
50 55 60
Pro Leu Val Leu Pro Pro Leu Pro Pro Pro Pro Ser Leu Leu Leu Trp
65 70 75 80
Leu Pro Val Cys Leu Leu Val Leu Leu Val Val Leu Ala Phe Met Pro
85 90 95
Thr Asp Val Arg Ser Val Ala Ala Ser Tyr Leu
100 105
<210> 270
<211> 130
<212> PRT
<213> Arabidopsis thaliana
<400> 270
Met Ile Arg Glu Ile Ser Asn Leu Gln Lys Asp Ile Ile Asn Ile Gln
1 5 10 15
Asp Ser Tyr Ser Asn Asn Arg Val Met Asp Val Gly Arg Asn Asn Arg
20 25 30
Lys Asn Met Ser Phe Arg Ser Ser Pro Glu Lys Ser Lys Gln Glu Leu
35 40 45
Arg Arg Ser Phe Ser Ala Gln Lys Arg Met Met Ile Pro Ala Asn Tyr
50 55 60
Phe Ser Leu Glu Ser Leu Phe Leu Leu Val Gly Leu Thr Ala Ser Leu
65 70 75 80
Leu Ile Leu Pro Leu Val Leu Pro Pro Leu Pro Pro Pro Pro Phe Met
85 90 95
Leu Leu Leu Val Pro Ile Gly Ile Met Val Leu Leu Val Val Leu Ala
100 105 110
Phe Met Pro Ser Ser His Ser Asn Ala Asn Thr Asp Val Thr Cys Asn
115 120 125
Phe Met
130
<210> 271
<211> 135
<212> PRT
<213> Arabidopsis thaliana

<400> 271
Met Ile Arg Glu Phe Ser Ser Leu Gln Asn Asp Ile Ile Asn Ile Gln
1 5 10 15
Glu His Tyr Ser Leu Asn Asn Asn Met Asp Val Arg Gly Asp His Asn
20 25 30
Arg Lys Asn Thr Ser Phe Arg Gly Ser Ala Pro Ala Pro Ile Met Gly
35 40 45
Lys Gln Glu Leu Phe Arg Thr Leu Ser Ser Gln Asn Ser Pro Arg Arg
50 55 60
Leu Ile Ser Ala Ser Tyr Phe Ser Leu Glu Ser Met Val Val Leu Val
65 70 75 80
Gly Leu Thr Ala Ser Leu Leu Ile Leu Pro Leu Ile Leu Pro Pro Leu
85 90 95
Pro Pro Pro Pro Phe Met Leu Leu Leu Ile Pro Ile Gly Ile Met Val
100 105 110
Leu Leu Met Val Leu Ala Phe Met Pro Ser Ser Asn Ser Lys His Val
115 120 125
Ser Ser Ser Ser Thr Phe Met
130 135
<210> 272
<211> 139
<212> PRT
<213> Vitis vinifera
<400> 272
Met Ser Ile Glu Gln Pro Glu Ala Asp Ser Arg Leu Ser Glu Gly Pro
1 5 10 15
Leu Ile Asn Leu Gln Asp Arg Tyr Leu Ser Gly Ile Met Glu Ala Arg
20 25 30
Gly Arg Arg Asn Ser Ala Pro Leu Gln Val Glu Arg Lys Asn Pro Thr
35 40 45
Pro Pro Met Ala Glu Gly Lys Lys Met Glu Tyr Asn Arg Thr Pro Leu
50 55 60
Ser Arg Glu Asn Ser Arg Arg Leu Ile Pro Ala Ser Tyr Phe Ser Leu
65 70 75 80
Glu Ser Leu Leu Leu Leu Ile Cys Leu Thr Ala Ser Leu Leu Ile Leu
85 90 95
Pro Leu Ile Leu Pro Pro Leu Pro Pro Pro Pro Phe Met Leu Leu Leu
100 105 110
Leu Pro Ile Gly Ile Leu Ala Val Leu Met Ile Leu Ala Phe Met Pro
115 120 125
Ser Asn Val Arg Asp Leu Thr Tyr Thr Tyr Val
130 135
<210> 273
<211> 126
<212> PRT
<213> Citrus sp.
<220>
<221> INCERTO
<222> (103)..(103)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 273
Met Asn Ser Asp Asn Ser Glu Ser Arg Gln Arg Leu Ser Lys Gly Ile
1 5 10 15
Ile Asn Leu Gln Asp Arg Tyr Pro Thr Ser Ile Met Asp Arg Gly Val
20 25 30
Arg Lys Ile Ala Thr Pro Pro Val Glu Lys Arg Lys Val Glu Tyr His
35 40 45
Arg Ser Tyr Ser Gln Gly Ala Ser Arg Lys Leu Phe Ser Ala Ser Tyr
50 55 60
Phe Thr Leu Glu Ser Leu Leu Leu Leu Val Cys Leu Thr Ala Ser Leu
65 70 75 80
Leu Ile Leu Pro Leu Val Leu Pro Pro Leu Pro Pro Pro Pro Phe Leu
85 90 95

Leu Leu Leu Val Pro Ile Xaa Ile Leu Ala Val Leu Leu Val Leu Ala
100 105 110
Phe Met Pro Ser Asn Val Arg Asp Ile Thr Ser Thr Tyr Val
115 120 125
<210> 274
<211> 118
<212> PRT
<213> Lycorpersicon esculentum
<400> 274
Met Asn Met Asp Met Glu Ser Ser Glu Ala Lys Leu Arg Ser Ser Lys
1 5 10 15
Gly Phe Ile Asn Leu Glu Glu His Gln Gln Tyr Phe Asn Asn Ile Met
20 25 30
Glu Gly Asn Lys Met Glu His Lys Arg Ser Phe Thr Gln Gly His Gly
35 40 45
Lys Lys Met Leu Ser Met Asn Tyr Phe Ser Leu Glu Ser Ile Ile Leu
50 55 60
Leu Leu Gly Leu Thr Ala Ser Leu Leu Leu Leu Pro Leu Met Leu Pro
65 70 75 80
Pro Leu Pro Pro Pro Pro Phe Met Leu Leu Leu Val Pro Ile Phe Ile
85 90 95
Leu Val Val Leu Met Ile Leu Ala Phe Met Pro Ser Asn Val Arg Asn
100 105 110
Val Thr Cys Ser Tyr Leu
115
<210> 275
<211> 87
<212> PRT
<213> Lycorpersicon esculentum
<400> 275
Met Glu Gly Asn Lys Met Glu His Lys Arg Ser Phe Thr Gln Gly His
1 5 10 15
Gly Lys Lys Met Leu Ser Met Asn Tyr Phe Ser Leu Glu Ser Ile Ile
20 25 30
Leu Leu Leu Gly Leu Thr Ala Ser Leu Leu Leu Leu Pro Leu Met Leu
35 40 45
Pro Pro Leu Pro Pro Pro Pro Phe Met Leu Leu Leu Val Pro Ile Phe
50 55 60
Ile Leu Val Val Leu Met Ile Leu Ala Phe Met Pro Ser Asn Val Arg
65 70 75 80
Asn Val Thr Cys Ser Tyr Leu
85
<210> 276
<211> 106
<212> PRT
<213> Arabidopsis thaliana
<400> 276
Met Asp Val Gly Arg Asn Asn Arg Lys Asn Met Ser Phe Arg Ser Ser
1 5 10 15
Pro Glu Lys Ser Lys Gln Glu Leu Arg Arg Ser Phe Ser Ala Gln Lys
20 25 30
Arg Met Met Ile Pro Ala Asn Tyr Phe Ser Leu Glu Ser Leu Phe Leu
35 40 45
Leu Val Gly Leu Thr Ala Ser Leu Leu Ile Leu Pro Leu Val Leu Pro
50 55 60
Pro Leu Pro Pro Pro Pro Phe Met Leu Leu Leu Val Pro Ile Gly Ile
65 70 75 80
Met Val Leu Leu Val Val Leu Ala Phe Met Pro Ser Ser His Ser Asn
85 90 95
Ala Asn Thr Asp Val Thr Cys Asn Phe Met
100 105
<210> 277
<211> 537

<212> DNA
<213> Oryza sativa
<400> 277
atgtacttgt tgagcccaag aaatggcgac gaggaggacg aacaggagga aatccaggag 60
ctgatcagcg acgacgagcc gcccaatctc aagttggcat cctgcgccac tgcagccagc 120
agcagcagca gcagcggcag cgacatggag aagggaagag gtaaagcctg cggcggcggg 180
agtacggcgc cgccgccgcc gccgccgtcg tcgtcaggta aatccggcgg cggcggcggc 240
agcaatatca gggaggcggc ggctagcggc ggcggcggcg gcgtgtgggg caagtacttc 300
tcggtggagt cgctgctcct gctggtgtgc gtgacggcgt cgctggtgat cctcccgctc 360
gtgctgccgc cgctgccccc gccgccgtcg atgctgatgc tggtgccggt ggcgatgctg 420
gtgctgctgc tggcgctggc gttcatgccg acgacgacgt cgtcgtcgtc gtccgccggc 480
ggcggcggcg gcggcggccg caatggggcg acgacgggac atgctcccta cttgtag 537
<210> 278
<211> 538
<212> DNA
<213> Zea mays
<220>
<221> misc_feature
<222> (177)..(208)
<223> n é a, c, g, ou t
<220>
<221> misc_feature
<222> (493)..(493)
<223> n é a, c, g, ou t
<400> 278
ttcacattac actcatgact cgtcttagga cgaatcctgc agctgcaaaa cacagaaaat 60
ctggccaaga cctatttatc tatttacagg taagaggagg ccatgctgcg cacatctgtc 120
ggcatgaagg ccagtacaac cagcaagacg agcaggcaga ccggcagcca cagcagnnnn 180
nnnnnnnnnn nnnnnnnnnn nnnnnnnncc agcggcagta tcagcagcga gacggtgagg 240
caggcgagca cgaggaatgc cccgatgctg aagtaggtgg acgcgatgac cttcttgtcg 300
aggaaatccg cctcctgctg acgctgctgc agatgccgct tcacggcatt cctcctttgt 360
attgccgccc ctccttccat cgcgctagat cggcttgcca tgtgttcttt ggtgtaggcg 420
gaggtggaga ccagtcgcag tgagtggttg ccaaagtaag caagaagtga aaggcggtgt 480
agaaccgcct tgngttttct gaacgttttg taatcagatc tgagctcggg tggtcgaa 538
<210> 279
<211> 578
<212> DNA
<213> Vitis vinifera
<400> 279
ttttttttt ttttttaat caagaaataa aataactgat tttcatctga atatcaagac 60
aaataacaaa aattgtatga tgagaagagg tgcacttttg aacaccacca tttacacata 120
tgtataagtc aaatctctga cattagaagg catgaaagcc aagatcataa gcactgctag 180
aatgccaatg gggagcagaa gcagcatgaa aggaggggt ggcaatggtg gtagtatcag 240
gggaaggatc agcaatgaag ccgtgagaca gatgagcaaa agcaatgact ccaagctgaa 300
atagcttgct gggatcagtc tcctgctgtt ctctcgtgag aggggagttc tattatattc 360
catcttcttt ccttcggcca taggaggagt agggtttttc ctctcaactt gcagaggagc 420
agagttcctc cttcctctcg cttccatgat gccactcaaa tatcgatctt gcagattgat 480
caaaggtcct tcagacagtc ttgaatctgc ttcaggctgc tcaatactca tttaaatttc 540
tcctttaccg gtcaccaagt tccaaccggt tcaaagta 578
<210> 280
<211> 704
<212> DNA
<213> Citrus reticulata
<220>
<221> misc_feature
<222> (619)..(619)
<223> n é a, c, g, ou t
<400> 280
agggtttctt caaagatagg tagccatttg cacatttgaa tctgcttgtt ggatattgtc 60
aaggaggctg ttggaattag gccacatttc agaatctggt ttcatcctgg atcgctgggc 120
atttgaaggc attttgtgat catcgctgtt taaaatttgg ccgcatatta gaatctgggt 180
tctcatcggt tttccgtaca tttaccttga ccacattttg atatctgggt tgagctgcat 240
tttagccttc gtatttaaaa ggacttgatc taatctgggg tcttggtgag ccggggcaac 300

tgatcatagt aaatgaattc tgataattct gagtcgagac agagactatc aaagggcatt 360
ataaacttgc aagatcgata tccgaccagc attatggatc gtggtgtaag aaaaattgca 420
actcctccgg tcgagaagag gaaagttgag tatcaccgaa gttactcgca aggggcatcc 480
agaaaactgt tttcggcaag ctatttcacc ctggaatcat tgcttttgct cgtatgtctg 540
acggcctcat tgctgatcct gccattggtg cttccgccct tgccgccccc gccattcctg 600
ctgcttctgg ttcctatang tattctagcc gtgcttttgg tcttggcatt catgccttct 660
aatgtaagag atataacttc cacgtacgtg taaatggtgt tgct 704

<210> 281
<211> 382
<212> DNA
<213> Lycorpersicon esculentum
<400> 281
gaaaaaaatg tatttaatca ttatgtaaaa aacaagtgaa tctactttga tattttcttc 60
taaattaaac cacacaatta aagatatgag caagtcacat tcctaacatt agaaggcata 120
aaagctaaga tcataagaac aacaagaatg aaaattggga ctaacaacaa cataaaaggt 180
ggtggtggca atggtggaag catcaatggc aaaagtaaca aagatgctgt aagaccaagt 240
aacaaaataa ttgactctaa gctaaaataa ttcattgaca acattttctt gccatgtcct 300
tgtgtaaatg atctcttatg ctccatctta ttgccttcca taatgttgtt gaaatattgt 360
tgatgttcct ccaaattaat aa 382

<210> 282
<211> 624
<212> DNA
<213> Lycorpersicon esculentum
<400> 282
tttgttaaag attggcacat tttcaagttc agtattcatt cgatttttga tatctacata 60
aaaaaaaagt gtcctggtac tactcaatat tcctcagaac gacttcatat tcaggtctcg 120
aattcaaaac ctcacatcaa gattcttagg aaatttcaag attggttgaa aaactcatat 180
ccttctctaa gtttcaagat tggttccaaa ttaaaactcg agacttctga gtaagagcgt 240
acgactagta atgaacatgg acatggaatc atcagaggca aaattgagat catcaaaagg 300
gtttattaat ttggaggaac atcaacaata tttcaacaac attatggaag gcaataagat 360
ggagcataag agatcattta cacaaggaca tggcaagaaa atgttgtcaa tgaattattt 420
tagcttagag tcaattattt tgttacttgg tcttacagca tctttgttac ttttgccatt 480
gatgcttcca ccattgccac caccaccttt tatgttgttg ttagtcccaa ttttcattct 540
tgttgttctt atgatcttag cttttatgcc ttctaatgtt aggaatgtga cttgctcata 600
tctttaattg tgtggtttaa ttta 624

<210> 283
<211> 1130
<212> DNA
<213> Oryza sativa
<400> 283
catgcggcta atgtagatgc tcactgcgct agtagtaagg tactccagta cattatggaa 60
tatacaaagc tgtaatactc gtatcagcaa gagagaggca cacaagttgt agcagtagca 120
caggattaga aaaacgggac gacaaatagt aatggaaaaa caaaaaaaaa caaggaaaca 180
catggcaata taaatggaga aatcacaaga ggaacagaat ccgggcaata cgctgcgaaa 240
gtactcgtac gtaaaaaaaa gaggcgcatt catgtgtgga cagcgtgcag cagaagcagg 300
gatttgaaac cactcaaatc caccactgca aaccttcaaa cgaggccatg gtttgaagca 360
tagaaagcac aggtaagaag cacaacgccc tcgctctcca ccctcccacc caatcgcgac 420
gcacctcgcg gatcggtgac gtggcctcgc cccccaaaaa tatcccgcgg cgtgaagctg 480
acaccccggg cccacccacc tgtcacgttg gcacatgttg gttatggttc ccggccgcac 540
caaaatatca acgcggcgcg gcccaaaatt tccaaaatcc cgcccaagcc cctggcgcgt 600
gccgctcttc cacccaggtc cctctcgtaa tccataatgg cgtgtgtacc ctcggctggt 660
tgtacgtggg cgggttaccc tgggggtgtg ggtggatgac gggtgggccc ggaggaggtc 720
cggccccgcg cgtcatcgcg gggcggggtg tagcgggtgc gaaaaggagg cgatcggtac 780
gaaaattcaa attaggaggt gggggggcggg gcccttggag aataagcgga atcgcagata 840
tgcccctgac ttggcttggc tcctcttctt cttatccctt gtcctcgcaa ccccgcttcc 900
ttctctcctc tcctcttctc ttctcttctc tggtggtgtg ggtgtgtccc tgtctccccc 960
ctccttcctc ctctcctttc ccctcctctc ttccccccctc tcacaagaga gagagcgcca 1020
gactctcccc aggtgaggtg agaccagtct ttttgctcga ttcgacgcgc ctttcacgcc 1080
gcctcgcgcg gatctgaccg cttccctcgc ccttctcgca ggattcagcc 1130

<210> 284
<211> 77
<212> PRT
<213> Medicago sativa

<400> 284
Met Val Arg Cys Phe Ser Leu Gly Ser Val Leu Ile Leu Ile Ala Leu
1 5 10 15
Ala Ala Ser Met Val Val Leu Pro Leu Met Leu Pro Pro Leu Pro Pro
20 25 30
Pro Pro Leu Ala Leu Leu Phe Phe Pro Val Gly Ile Met Ala Ala Leu
35 40 45
Val Val Leu Ala Phe Ser Pro Ser Glu Asn Val Lys Asn Val Val Val
50 55 60
Tyr Ser Ser Ser Ser Ser Gly Ile Ala Asn Ser Lys Arg
65 70 75
<210> 285
<211> 78
<212> PRT
<213> Medicago sativa
<400> 285
Met Ile Met Val Ala Ser Lys Glu Lys Thr Asn Ser Gly Gly Cys Met
1 5 10 15
Phe Arg Tyr Ser Val Leu Ile Leu Ser Leu Leu Ala Leu Ser Ile Leu
20 25 30
Val Leu Pro Leu Val Met Pro Pro Leu Pro Pro Pro Pro Leu Leu Leu
35 40 45
Leu Leu Val Pro Val Phe Ile Met Leu Leu Leu Phe Phe Ile Ala Phe
50 55 60
Ser Pro Ser Lys Lys Val Pro Asn Lys Ala Ser Phe Val Ser
65 70 75
<210> 286
<211> 613
<212> DNA
<213> Hordeum vulgare
<400> 286
cttccgagaa agatgcttca tattgcactc atcttatgcc aacacacaat cagaaacaaa 60
aaccacgcag caagcctatt tacaagtaag aggtggccat gctccgcaca tcagtcggca 120
tgaaggccag caccatcaac aggatgagca agcagaccgg caagagcagc agtagcgatg 180
gcggtggggg caacggaggc agcaccaatg gcagtatgag caatgaaacg gtgaggcagg 240
cgagcacgaa gaacgcttcg acgccgaagt agttcggtat gacaaccttc ttgtcagcag 300
catttgccct ctgctgctgc atttgcctct ttgcagtatt tggtctggct aacacagggc 360
tgctcctctg cctactaccg cctctgtcca ttgcaccgaa ctggctatcc atctattctt 420
tggtgagtgc agatcagcgg ctatccagga actgagatga ataagaactt gtggaatctg 480
aaggttttta acagatctga agtctttgtg agtccgtgac tgaactgcag atcatctgct 540
gtggaagaac tgcaccgcaa gtggcaatac cttcgatcct gaaacttgca ttttggtgat 600
gcccgtacca cgc 613
<210> 287
<211> 617
<212> DNA
<213> Saccharum officinarum
<220>
<221> misc_feature
<222> (451)..(451)
<223> n é a, c, g, ou t
<400> 287
ggagaacgtg cttgttcagg tggttcaaca gtgcggacca gagaacaatg cgaggttcag 60
gattaaaggt attctggctt cagaggcagt tcttccaaag caagtgacag gcgattccat 120
caccggagct aagatcttat tacaacattc agaaacacag gagtcctccc accgcctttc 180
acttcttgct tacttctgca accactcgcc gtgactgatc tccacgcaca ccaaagaaca 240
caacacgtgg caagccgatc tagcgcgatg gaaggagggg ggcaaataca gaggaggaat 300
aatgccgtga agcggcacct gcagcagcgg cagcaggagg cggatttcct cgacaagaag 360
gtcatcgcgt ccacctattt cagcatcgag gcgttcctcg tgctcgcctg cctcaccgtc 420
tcgctgctga tactgccgct tgtgctgccg ncgctgccgg cgccggcgtc gctgctgctg 480
tggctgccgg tctggctgct cgaactgctg attgtgcttg ccttcatgcc gacagatgtg 540
cgcagcatgg cctcctctta cttgtaaaaa aatagataaa taggccacct ttggcaattt 600
tctggggttt ggaggtg 617
<210> 288

<211> 638
<212> DNA
<213> Saccharum officinarum
<400> 288
gatttttcc aagccagtga ccggcgattc atcaaccgga gctgagatct tatacaacat 60
tcagaaacac aggagtcctc gcaccgcttt ccactcttgg ctaattttgc aaccactcgc 120
cgtgactgat ctccacgcac accaaagaac acaacacgtg gcaacccgat ctagcgcgat 180
ggaaggaggg gggcaaatac agaggaggaa taatgccgtg aagcggcacc tgcagcagcg 240
gcagcaggag gcggatttcc tcgacaagaa ggtcatcgcg tccacctatt tcagcatcga 300
ggcgttcctc gtgctcgcct gcctcaccgt ctcgctgctg atactgccgc tggtgctgcc 360
gccgctgccg ccgccgccgt cgctgctgct gtggctgccg gtctgcctgc tcatcctgct 420
gattgtgctg gccttcatgc cgacagatgt gcgcagcatg gcctcctctt acttgtaata 480
aatagataaa taggccacct tggtcaatat tctgtgattt ggaggtgatt cgtcctgaga 540
tgagtctcga ttgatgtcag ctactcccaa ggggaaatgc atgtaacact tggtggccga 600
cggtggcaag ataatcatgc taccatgtca gttaaacc 638
<210> 289
<211> 778
<212> DNA
<213> Sorghum bicolor
<400> 289
gctgttggtg ttgttcttga gatctctttc ttgatcttgt gtgggattaa agagggattc 60
ttgccttcct acgggagaaa gagaaaaggg gaagaacgtg cttgttccgg tggttcaaca 120
gtgcggagac ccggagaaca atgcgaggtt caggattaaa ggtgttctgg cttcaggtgc 180
agttcttcca aagcaggtga caggcgattc gatcaccgga gctcagatct gacgaaaaca 240
aaacacagtc ctcctcccac cgcctttcag ttcttgctta cttctgcaac cactcactgc 300
gaccgtacac caaagaacac tgcacatggc aagccgatct agcgcgctgg aaggaggggg 360
ggcagcaata cagcggagga ataatgccgt gaagcggcac ctgcagcagc ggcagcagga 420
ggcggatttc cacgacaaga aggtcatcgc gtccacctac ttcagcatcg gcgcgttcct 480
ggtgctcgcc tgcctcacct tctcgctgct catcctgccg ctggtgctgc cgccgctgcc 540
gccgccgccg tcgctgctgc tgtggctgcc ggtctgcctg ctcgtcctgc tggttgtgct 600
ggccttcatg ccgacagatg tgcgcagcgt ggcggcctct tacttgtaaa tagccagata 660
aataggccac cacctttggc cagttttctg tgttttcggg ggtgattcgt cctaagatga 720
gtcatgatcg agtgtaatgt gaagcatctt ttccaggggt agtagatttc aatgaagt 778
<210> 290
<211> 732
<212> DNA
<213> Arabidopsis thaliana
<400> 290
ttgtcttcct catttcccta ctagtacttg tttcacacag tttcttgatc caaccaaaac 60
caatacacaa agcttctcaa actccttcac ctcaaagctt cttcctttac atctgaatcg 120
ttgagttaac tcggatttgt tctgcatcct ctgtttctga atcgtgggcc atccttattt 180
tgtctcgaat tcttcaccaa ttgcttcgat caagctgcat tggttaacca gttgccctaa 240
agatcagatc tttgagcaaa attttgtcac tgatcttcta aatccaaacc agacacagca 300
aaacaacctc tgtagatgat tcgagaaatc tcaaacttac aaaaagatat tataaacatt 360
caagacagtt attcgaacaa ccgagtcatg gacgtcggaa gaaacaaccg gaaaaacatg 420
agctttcgaa gttcgccgga gaaaagcaag caagagttac ggcggagttt ctcggcgcag 480
aaaaggatga tgatcccggc gaattatttc agtttagagt ctctgttcct attggttggt 540
ctaacggcat ctctgttaat acttccgtta gttttgccgc cgttacctcc gcctccgttt 600
atgctgctat tggttcccat tgggattatg gttttactcg tcgttcttgc cttcatgcct 660
tcttctcatt ctaatgctaa tacagatgta acttgcaatt tcatgtaaat ctgaaattta 720
ttatatgatg at 732
<210> 291
<211> 891
<212> DNA
<213> Arabidopsis thaliana
<400> 291
cccactttat ttcttcttct ctggttttct taccaaaaga aactttcttc gtcttcctct 60
gtatttaagc tttaacaccc tgtttttggt ttccaacgtt caatcttcat cttcttctcg 120
ctgaaggtgt gtttggctct aacggtttaa agctttcttg aaacaccaat tgaatctttt 180
ctctctaccg gcaaaaaaaa aagattagtc cttttaggtc tggaaacgcc aagatcactc 240
gttctaaacc ttagattttg tctgcatttc gggataatca tttcatcgtc agggttcttc 300
aaccaaacta catttacaga agaagaagaa gaagaaaagt tcgttacttt ttatgcgttt 360
ggataaacaa actcaagttt cttcttcata catcgatctg attttccaga tcaaacttcg 420

aaaagagaaa aagccttctt taaatgattc gtgagttctc cagtctacaa aacgacatca 480
taaacattca agaacattat tctctcaaca acaacatgga cgtgagagga gatcataacc 540
ggaaaaacac gagttttcgt ggttcagctc cagctccgat tatggggaag caagaattgt 600
ttcggacatt gtcgtcgcag aacagtccaa ggaggctaat atcagcgagt tacttcagtt 660
tagaatcaat ggttgtgctt gttggtctca cagcatctct cttgatctta ccgttgattc 720
ttccaccatt gcctcctcct ccttttatgc tgcttttgat tcctattggg attatggttt 780
tgcttatggt tcttgctttc atgccttctt ctaattccaa acatgtttct tcttcttcca 840
cttttatgta ataaacgttt ctttaattga agaaagaaat ccttaaacaa a 891
<210> 292
<211> 732
<212> DNA
<213> Arabidopsis thaliana
<400> 292
ttgtcttcct catttcccta ctagtacttg tttcacacag tttcttgatc caaccaaaac 60
caatacacaa agcttctcaa actccttcac ctcaaagctt cttcctttac atctgaatcg 120
ttgagttaac tcggatttgt tctgcatcct ctgtttctga atcgtgggcc atccttattt 180
tgtctcgaat tcttcaccaa ttgcttcgat caagctgcat tggttaacca gttgccctaa 240
agatcagatc tttgagcaaa attttgtcac tgatcttcta aatccaaacc agacacagca 300
aaacaacctc tgtagatgat tcgagaaatc tcaaacttac aaaaagatat tataaacatt 360
caagacagtt attcgaacaa ccgagtcatg gacgtcggaa gaaacaaccg gaaaaacatg 420
agctttcgaa gttcgccgga gaaaagcaag caagagttac ggcggagttt ctcggcgcag 480
aaaaggatga tgatcccggc gaattatttc agtttagagt ctctgttcct attggttggt 540
ctaacggcat ctctgttaat acttccgtta gttttgccgc cgttacctcc gcctccgttt 600
atgctgctat tggttcccat tgggattatg gttttactcg tcgttcttgc cttcatgcct 660
tcttctcatt ctaatgctaa tacagatgta acttgcaatt tcatgtaaat ctgaaattta 720
ttatatgatg at 732
<210> 293
<211> 1130
<212> DNA
<213> Oryza sativa
<400> 293
catgcggcta atgtagatgc tcactgcgct agtagtaagg tactccagta cattatggaa 60
tatacaaagc tgtaatactc gtatcagcaa gagagaggca cacaagttgt agcagtagca 120
caggattaga aaaacgggac gacaaatagt aatggaaaaa caaaaaaaaa caaggaaaca 180
catggcaata taaatggaga aatcacaaga ggaacagaat ccgggcaata cgctgcgaaa 240
gtactcgtac gtaaaaaaaa gaggcgcatt catgtgtgga cagcgtgcag cagaagcagg 300
gatttgaaac cactcaaatc caccactgca aaccttcaaa cgaggccatg gtttgaagca 360
tagaaagcac aggtaagaag cacaacgccc tcgctctcca ccctcccacc caatcgcgac 420
gcacctcgcg gatcggtgac gtggcctcgc cccccaaaaa tatcccgcgg cgtgaagctg 480
acaccccggg cccacccacc tgtcacgttg gcacatgttg gttatggttc ccggccgcac 540
caaaatatca acgcggcgcg gcccaaaatt tccaaaatcc cgcccaagcc cctggcgcgt 600
gccgctcttc cacccaggtc cctctcgtaa tccataatgg cgtgtgtacc ctcggctggt 660
tgtacgtggg cgggttaccc tgggggtgtg ggtggatgac gggtgggccc ggaggaggtc 720
cggccccgcg cgtcatcgcg gggcggggtg tagcgggtgc gaaaaggagg cgatcggtac 780
gaaaattcaa attaggaggt gggggggcggg gcccttggag aataagcgga atcgcagata 840
tgcccctgac ttggcttggc tcctcttctt cttatccctt gtcctcgcaa ccccgcttcc 900
ttctctcctc tcctcttctc ttctcttctc tggtggtgtg ggtgtgtccc tgtctccccct 960
ctccttcctc ctctcctttc ccctcctctc ttccccccctc tcacaagaga gagagcgcca 1020
gactctcccc aggtgaggtg agaccagtct ttttgctcga ttcgacgcgc ctttcacgcc 1080
gcctcgcgcg gatctgaccg cttccctcgg ccttctcgca ggattcagcc 1130

REIVINDICAÇÕES

1. Método para intensificar características relacionadas com o rendimento em plantas, caracterizado pelo fato de que compreende modular a expressão em uma planta de um ácido nucleico que codifica uma quinase semelhante ao receptor de extensina (proteína de ERLK) representada por qualquer um de SEQ ID N° : 2, SEQ ID N° : 12, SEQ ID N° : 14 ou ortólogos ou parálogos de qualquer um dos mencionados acima ou uma porção destes que codifica um polipeptídeo que compreende pelo menos o domínio de transmembrana e o domínio de quinase.

2. Método de acordo com a reivindicação 1, caracterizado pelo fato de que a dita expressão modulada é realizada por qualquer um ou mais de rotulação de ativação de T-DNA, TILLING ou recombinação homóloga.

3. Método de acordo com a reivindicação 1, caracterizado pelo fato de que a dita expressão modulada é realizada por introdução e expressão em uma planta de um ácido nucleico que codifica uma proteína ERLK representada por qualquer um de SEQ ID N° : 2, SEQ ID N° : 12, SEQ ID N° : 14 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima ou uma parte destes que compreende pelo menos o domínio de transmembrana e o domínio de quinase.

4. Método de acordo com qualquer uma das reivindicações de 1 a 3, caracterizado pelo fato de que dita a característica relacionada com o rendimento é um ou mais de biomassa aumentada, rendimento aumentado de semente (peso de semente) ou número aumentado de sementes enchidas.

5. Método de acordo com qualquer uma das reivindicações de 1 a 3, caracterizado pelo fato de que o dito ácido nucleico está operacionalmente ligado a um promotor constituído, preferivelmente a um promotor GOS2, mais preferivelmente ao promotor GOS2 como representado por SEQ ID N° : 58.

6. Método de acordo com qualquer um das reivindicações de 3

a 5, caracterizado pelo fato de que o dito ácido nucleico é uma porção de ou uma variante alélica de ou uma variante de união de ou uma seqüência capaz de hibridizar a uma seqüência de ácido nucleico de acordo com qualquer um de SEQ ID N° : 1, SEQ ID N° : 11, SEQ ID N° : 13, SEQ ID N° : 15, SEQ ID N° : 17, SEQ ID N° : 19, SEQ ID N° : 21, SEQ ID N° : 23, SEQ ID N° : 25, SEQ ID N° : 27, SEQ ID N° : 29, SEQ ID N° : 31, SEQ ID N° : 33, SEQ ID N° : 35, SEQ ID N° : 37, SEQ ID N° : 39, SEQ ID N° : 41, SEQ ID N° : 43, SEQ ID N° : 45, SEQ ID N° : 47, SEQ ID N° : 49, SEQ ID N° : 51, SEQ ID N° : 53 e SEQ ID N° : 55, em que a dita porção, variante alélica, variante de união ou seqüência de hibridização codifica pelo menos o domínio de transmembrana e o domínio de quinase de uma proteína de ERLK.

7. Método de acordo com qualquer uma das reivindicações de 1 a 6, caracterizado pelo fato de que o dito ácido nucleico que codifica uma proteína ERLK ou parte deste é de origem vegetal, preferivelmente de uma planta dicotiledônea, ainda preferivelmente da família Brassicaceae, mais preferivelmente do gênero *Arabidopsis*, mais preferivelmente de *Arabidopsis thaliana*.

8. Planta ou parte da mesma, incluindo sementes, caracterizada pelo fato de ser obtenível por um método de acordo com qualquer uma das reivindicações de 1 a 7, em que a dita planta ou parte da mesma compreendem um ácido nucleico que codifica uma proteína ERLK ou uma parte da mesma que compreende pelo menos o domínio de transmembrana e o domínio de quinase.

9. Construção, caracterizada pelo fato de que compreende:

(i) um ácido nucleico que codifica um proteína de ERLK, representado por qualquer um de SEQ ID N° : 2, SEQ ID N° : 12, SEQ ID N° : 14 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima ou uma parte destes que compreende pelo menos o domínio de transmembrana e o domínio de quinase;

(ii) uma ou mais seqüências de controle capazes de conduzir a expressão da seqüência de ácido nucleico de (i); e opcionalmente

(iii) uma seqüência de terminação de transcrição.

10. Construção de acordo com a reivindicação 9, caracterizada pelo fato de que a dita uma ou mais seqüências de controle é pelo menos um promotor constituído, preferivelmente um promotor GOS2.

11. Uso de uma construção como definida na reivindicação 9 ou 10, caracterizado pelo fato de ser para a fabricação de plantas tendo características relacionadas com o rendimento intensificadas, particularmente biomassa aumentada, rendimento aumentado de semente (peso de semente) e/ou número aumentado de sementes enchidas, com relação às plantas de controle.

12. Planta, parte de planta ou célula de planta, caracterizada pelo fato de ser transformada com uma construção como definida na reivindicação 9 ou 10.

13. Método para a produção de uma planta transgênica tendo características intensificadas relacionadas com o rendimento com relação às plantas de controle, caracterizado pelo fato de compreender:

(i) introdução e expressão em uma planta de um ácido nucleico que codifica uma proteína ERLK representado por qualquer um de SEQ ID N° : 2, SEQ ID N° : 12, SEQ ID N° : 14 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima ou uma parte destes que compreende pelo menos o domínio de transmembrana e o domínio de quinase e

(ii) cultivar a célula de planta sob condições que promovem crescimento e crescimento vegetal.

14. Planta transgênica, caracterizada pelo fato de que tem rendimento aumentado com relação às plantas de controle, o dito rendimento aumentado resultando da expressão aumentada de um ácido nucleico que

codifica uma proteína ERLK representada por qualquer um de SEQ ID N° : 2, SEQ ID N° : 12, SEQ ID N° : 14 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima ou uma parte destes que compreende pelo menos o domínio de transmembrana e o domínio de quinase.

15. Planta de acordo com a reivindicação 8, 12 ou 14, caracterizada pelo fato de que a dita planta é uma planta de safra ou uma monocotiledônea ou um cereal, tal como arroz, milho, trigo, cevada, painço, triticale, centeio, sorgo ou aveias.

16. Partes colhíveis de uma planta como definida em qualquer um das reivindicações 8, 12, 14 ou 15, caracterizadas pelo fato de que as ditas partes colhíveis são sementes.

17. Produtos, caracterizados pelo fato de serem derivados de uma planta como definida na reivindicação 15 e/ou de partes colhíveis de uma planta como definidas na reivindicação 16.

18. Uso de um ácido nucleico que codifica uma proteína ERLK representada por qualquer um de SEQ ID N° : 2, SEQ ID N° : 12, SEQ ID N° : 14 ou ortólogos ou parálogos de qualquer um dos mencionados acima ou uma parte destes que compreende pelo menos o domínio de transmembrana e o domínio de quinase, caracterizado pelo fato de ser empregado na intensificação de características relacionadas com o rendimento em plantas.

19. Uso de acordo com a reivindicação 18, caracterizado pelo fato de que dita característica relacionada com o rendimento é biomassa aumentada, rendimento aumentado de semente (peso de semente) e/ou número aumentado de sementes enchidas.

20. Método para aumentar o rendimento em plantas com relação às plantas de controle, caracterizado pelo fato de que compreende aumentar a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo F-box WD40 (FBXW) e opcionalmente selecionar plantas tendo

rendimento aumentado.

21. Método de acordo com a reivindicação 20, caracterizado pelo fato de que a dita expressão aumentada é realizada por qualquer um ou mais de: ativação de T-DNA, TILLING e recombinação homóloga.

22. Método de acordo com a reivindicação 20, caracterizado pelo fato de que a dita expressão aumentada é realizada por introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo FBXW.

23. Método de acordo com qualquer um das reivindicações de 20 a 22, caracterizado pelo fato de que o dito ácido nucleico é uma porção ou uma variante alélica ou uma variante de união ou uma seqüência capaz de hibridizar a um seqüência de acordo com qualquer um de SEQ ID N° : 59, SEQ ID N° : 60, SEQ ID N° : 61, SEQ ID N° : 63, SEQ ID N° : 65, SEQ ID N° : 67 ou SEQ ID N° : 69, em que a dita porção, variante alélica, variante de união ou a seqüência de hibridização codifica um polipeptídeo FBXW.

24. Método de acordo com a reivindicação 23, caracterizado pelo fato de que a dita porção, variante alélica, variante de união ou seqüência de hibridização codifica um ortólogo ou parálogo de um polipeptídeo FBXW.

25. Método de acordo com qualquer um das reivindicações de 20 a 24, caracterizado pelo fato de que o dito ácido nucleico que codifica um polipeptídeo FBXW é de origem vegetal, preferivelmente de uma planta dicotiledônea, ainda preferivelmente da família Brassicaceae, mais preferivelmente de *Arabidopsis thaliana*.

26. Método de acordo com qualquer um das reivindicações de 20 a 25, caracterizado pelo fato de que o dito ácido nucleico está operacionalmente ligado a um promotor constituído, preferivelmente a um promotor GOS2, ainda preferivelmente a um promotor GOS2 de arroz.

27. Método de acordo com a reivindicação 26, caracterizado pelo fato de que o dito promotor constituído é representado por uma

seqüência de ácido nucleico substancialmente similar à SEQ ID N° : 58, mais preferivelmente o promotor constituído é como representado por SEQ ID N° : 58.

28. Método de acordo com qualquer um das reivindicações de 20 a 27, caracterizado pelo fato de que o dito rendimento aumentado é selecionado de um ou mais de: a) rendimento aumentado de semente; b) número aumentado de sementes (enchidas); c) peso de semente singular aumentado (TKW); d) índice de colheita aumentado e e) taxa de enchimento de semente aumentada.

29. Plantas, partes ou células de tais plantas, incluindo sementes, caracterizadas pelo fato de serem obteníveis por um método como definido em qualquer um das reivindicações de 20 a 28, em que a dita planta, parte de planta ou célula de planta da mesma compreende um ácido nucleico que codifica um polipeptídeo FBXW, cujo ácido nucleico está operacionalmente ligado a um promotor constituído.

30. Construção, caracterizada pelo fato de que compreende:

(i) um ácido nucleico que codifica um polipeptídeo FBXW;

(ii) uma ou mais seqüências de controle operacionalmente ligadas ao ácido nucleico de (i).

31. Construção de acordo com a reivindicação 30, caracterizada pelo fato de que a dita uma ou mais seqüências de controle é de pelo menos um promotor constituído, preferivelmente um promotor GOS2.

32. Uso de uma construção como definida na reivindicação 30 ou 31, caracterizado pelo fato de ser para aumentar o rendimento em plantas.

33. Planta, parte de planta ou célula de planta caracterizadas pelo fato de serem transformadas com uma construção como definida na reivindicação 30 ou 31.

34. Método para a produção de uma planta transgênica tendo rendimento aumentado com relação às plantas de controle, caracterizado pelo

fato de que compreende:

(i) introduzir e expressar a ácido nucleico que codifica um polipeptídeo FBXW em uma célula de planta e

(ii) cultivar a célula de planta sob condições que promovem crescimento e crescimento vegetal.

35. Planta transgênica, caracterizada pelo fato de que tem rendimento aumentado com relação às plantas de controle, o dito rendimento aumentado resultando da expressão aumentada de um ácido nucleico que codifica um polipeptídeo FBXW.

36. Planta de acordo com a reivindicação 29, 33 ou 35, caracterizada pelo fato de que a dita planta é uma planta de safra ou uma monocotiledônea ou um cereal, tal como arroz, milho, trigo, cevada, painço, triticale, centeio, sorgo ou aveias.

37. Partes colhíveis de uma planta como definida em qualquer um das reivindicações 29, 33, 35 ou 36, caracterizadas pelo fato de que as ditas partes colhíveis são sementes.

38. Produtos, caracterizados pelo fato de que são derivados de uma planta como definida na reivindicação 36 ou de partes colhíveis de uma planta como definida na reivindicação 37.

39. Uso de um ácido nucleico que codifica um polipeptídeo FBXW, cujo ácido nucleico está operacionalmente ligado a um promotor constituído, caracterizado pelo fato de ser no aumento de rendimento em uma planta comparado ao rendimento em uma planta de controle.

40. Uso de acordo com a reivindicação 39, caracterizado pelo fato de que o dito rendimento aumentado é selecionado de um ou mais de: a) rendimento aumentado de semente; b) número aumentado de sementes (enchidas); c) peso de semente singular aumentado (TKW); d) índice de colheita aumentado e e) taxa de enchimento de semente aumentada.

41. Método para intensificar características relacionadas com o

rendimento em plantas, caracterizado pelo fato de que compreende preferivelmente modular a expressão em semente ou partes de semente de planta de um ácido nucleico que codifica uma proteína de ligação de RAN (RANBP) que compreende Motivo I: KSC V/L WHAXDF A/S DGELK D/E EXF, onde 'X' é qualquer aminoácido, que permite zero ou uma mudança conservativa em qualquer posição e/ou zero, uma, duas ou três mudanças não conservativas em qualquer posição.

42. Método de acordo com a reivindicação 41, caracterizado pelo fato de que a dita expressão preferencialmente de modulação é realizada por qualquer um ou mais de rotulação de ativação de T-DNA, TILLING ou recombinação homóloga.

43. Método de acordo com a reivindicação 41, caracterizado pelo fato de que a dita expressão preferencialmente de modulação é realizada pela introdução e expressão em semente ou partes de semente de planta um ácido nucleico que codifica um RANBP que compreende o dito Motivo I.

44. Método de acordo com qualquer um das reivindicações 41 a 43, caracterizado pelo fato de que o dito RANBP ainda compreende qualquer um ou mais dos seguintes motivos:

(a) Motivo II como representado por SEQ ID N° : 139 ou 145 ou um motivo tendo em ordem crescente de preferência pelo menos 60%, 70%, 80%, 90% ou mais porcentagem de identidade de seqüência para o Motivo II representado por SEQ ID N° : 139 ou 145;

(b) Motivo III como representado por representado por SEQ ID N° : 140 ou 146 ou um motivo tendo em ordem crescente de preferência pelo menos 70%, 80%, 90% ou mais porcentagem de identidade de seqüência para o Motivo III como representado por SEQ ID N° : 140 ou 146;

(c) Motivo IV como representado por SEQ ID N° : 141 ou 147 seguindo de zero a uma mudanças conservativas em qualquer posição e/ou zero ou uma mudanças não conservativas em qualquer posição;

(d) Motivo V como representado por SEQ ID N° : 142 ou 148 ou um motivo tendo em ordem crescente de preferência pelo menos 70%, 80%, 90% ou mais porcentagem de identidade de seqüência para o Motivo V como representado por SEQ ID N° : 142 ou 148;

(e) Motivo VI como representado por SEQ ID N° : 143 ou 149 seguindo de zero a uma mudanças conservativas em qualquer posição e/ou zero ou uma mudanças não conservativas em qualquer posição;

(f) Motivo VII como representado por SEQ ID N° : 144 ou 150 ou um motivo tendo em ordem crescente de preferência pelo menos 60%, 70%, 80%, 90% ou mais porcentagem de identidade de seqüência para o Motivo VII representado por SEQ ID N° : 144 ou 150.

45. Método de acordo com qualquer um das reivindicações 41 a 44, caracterizado pelo fato de que a dita característica relacionada com o rendimento é rendimento aumentado de semente.

46. Método de acordo com a reivindicação 43, caracterizado pelo fato de que a dita expressão em semente ou partes de semente de planta de um ácido nucleico que codifica um RANBP é realizada pelo dito ácido nucleico sendo operacionalmente ligado a um promotor específico de semente.

47. Método de acordo com a reivindicação 46, caracterizado pelo fato de que o dito promotor específico de semente é um promotor específico de embrião.

48. Método de acordo com a reivindicação 47, caracterizado pelo fato de que o dito promotor específico de embrião é um promotor de prolamina.

49. Método de acordo com qualquer um das reivindicações de 43 a 48, caracterizado pelo fato de que o dito ácido nucleico é uma porção de ou uma variante alélica de ou uma variante de união de ou uma seqüência capaz de hibridizar a uma seqüência de ácido nucleico de acordo com

qualquer um de SEQ ID N° : 113, SEQ ID N° : 116, SEQ ID N° : 119, SEQ ID N° : 121, SEQ ID N° : 123, SEQ ID N° : 125, SEQ ID N° : 127, SEQ ID N° : 129, SEQ ID N° : 131, SEQ ID N° : 133, SEQ ID N° : 135 e SEQ ID N° : 137, em que a dita porção, variante alélica, variante de união ou seqüência de hibridização codifica um RANBP que compreende o dito Motivo I.

50. Método de acordo com qualquer um das reivindicações de 41 a 49, caracterizado pelo fato de que o dito ácido nucleico que codifica um RANBP é de origem vegetal, preferivelmente de uma planta monocotiledônea, ainda preferivelmente da família Poaceae, mais preferivelmente do gênero *Zea*, mais preferivelmente de *Zea mays*.

51. Método de acordo com qualquer um das reivindicações de 41 a 49, caracterizado pelo fato de que o dito ácido nucleico que codifica um RANBP é de origem vegetal, preferivelmente de uma planta dicotiledônea, ainda preferivelmente da família Brassicaceae, mais preferivelmente o ácido nucleico é de *Arabidopsis thaliana*.

52. Planta ou parte da mesma, incluindo sementes, caracterizada pelo fato de ser obtenível por um método como definido em qualquer um das reivindicações de 41 a 51, em que a dita planta ou parte da mesma compreendem um ácido nucleico que codifica um RANBP que compreende o dito Motivo I.

53. Construção, caracterizada pelo fato de que compreende:

(a) ácido nucleico que codifica um RANBP que compreende Motivo I;

(b) um promotor específico de semente operacionalmente ligadas ao ácido nucleico de (a) e opcionalmente

(c) uma seqüência de terminação de transcrição.

54. Construção de acordo com a reivindicação 53, caracterizada pelo fato de que o dito promotor específico de semente é um promotor de prolamina.

55. Uso de uma construção como definida na reivindicação 53 ou 54, caracterizado pelo fato de ser para a fabricação de plantas tendo características relacionadas com o rendimento intensificadas, particularmente rendimento aumentado de semente, com relação às plantas de controle.

56. Planta, parte de planta ou célula de planta, caracterizada pelo fato de ser transformada com uma construção como definida na reivindicação 53 ou 54.

57. Método para a produção de uma planta transgênica tendo características intensificadas relacionadas com o rendimento com relação às plantas de controle, caracterizado pelo fato de que compreende:

(a) introdução e expressão em uma planta de um ácido nucleico que codifica um RANBP que compreende Motivo I: KSC V/L WHAXDF A/S DGELK D/E EXF, onde 'X' é qualquer aminoácido, que permite zero ou uma mudança conservativa em qualquer posição e/ou zero uma, duas ou três mudanças não conservativas em qualquer posição e

(b) cultivar a célula de planta sob condições que promovem crescimento e crescimento vegetal.

58. Planta transgênica, caracterizada pelo fato de que tem rendimento aumentado com relação às plantas de controle, o dito rendimento aumentado resultando da expressão aumentada de um ácido nucleico que codifica um RANBP que compreende Motivo I: KSC V/L WHAXDF A/S DGELK D/E EXF, onde 'X' é qualquer aminoácido, que permite zero ou uma mudança conservativa em qualquer posição e/ou zero uma, duas ou três mudanças não conservativas em qualquer posição.

59. Planta transgênica de acordo com a reivindicação 52, 56 ou 58, caracterizada pelo fato de que a dita planta é uma planta de safra ou uma monocotiledônea ou um cereal, tal como arroz, milho, trigo, cevada, triticale, painço, centeio, sorgo ou aveias.

60. Partes colhíveis de uma planta como definida em qualquer

um das reivindicações 52, 56, 58 ou 59, caracterizadas pelo fato de que ditas partes colhíveis são sementes.

61. Produtos, caracterizados pelo fato de que são derivados de uma planta como definida na reivindicação 59 e/ou de partes colhíveis de uma planta como definidas na reivindicação 60.

62. Uso de um ácido nucleico que codifica um RANBP, caracterizado pelo fato de ser na intensificação de características relacionadas com o rendimento em plantas, o dito RANBP compreendendo Motivo I: KSC V/L WHAXDF A/S DGELK D/E EXF, onde 'X' é qualquer aminoácido, permitindo zero ou uma mudança conservativa em qualquer posição e/ou zero uma, duas ou três mudanças não conservativas em qualquer posição.

63. Uso de acordo com a reivindicação 62, caracterizado pelo fato de que a dita característica relacionada com o rendimento é o rendimento aumentado de semente.

64. Método para intensificar características relacionadas com o rendimento em plantas, caracterizado pelo fato de que compreende modular a expressão em uma planta de um ácido nucleico que codifica uma proteína semelhante a GOLDEN2 (GLK) representada por SEQ ID N° : 157 ou SEQ ID N° : 193 ou que codifica ortólogos ou parálogos da SEQ ID N° : 157 ou SEQ ID N° : 193.

65. Método de acordo com a reivindicação 64, caracterizado pelo fato de que a dita expressão modulada é realizada por qualquer um ou mais de rotulação de ativação de T-DNA, TILLING ou recombinação homóloga.

66. Método de acordo com a reivindicação 64, caracterizado pelo fato de que a dita expressão modulada é realizada por introdução e expressão em uma planta de um ácido nucleico que codifica uma proteína de GLK representado por SEQ ID N° : 157 ou SEQ ID N° : 193 ou ortólogos ou parálogos da SEQ ID N° : 157 ou SEQ ID N° : 193.

67. Método de acordo com qualquer um das reivindicações de 64 a 66, caracterizado pelo fato de que a dita característica relacionada com o rendimento é um ou mais de biomassa aumentada, rendimento aumentado de semente (peso de semente) ou número aumentado de sementes enchidas.

68. Método de acordo com qualquer uma das reivindicações de 64 a 66, caracterizado pelo fato de que o dito ácido nucleico está operacionalmente ligado a um promotor constituído, preferivelmente a um promotor GOS2, mais preferivelmente ao promotor GOS2 como representado por SEQ ID N° : 58.

69. Método de acordo com qualquer um das reivindicações de 66 a 68, caracterizado pelo fato de que o dito ácido nucleico é uma porção de ou uma variante alélica de ou uma variante de união de ou uma seqüência capaz de hibridizar a uma seqüência de ácido nucleico de acordo com qualquer um de SEQ ID N° : 156, SEQ ID N° : 168, SEQ ID N° : 170, SEQ ID N° : 172, SEQ ID N° : 174, SEQ ID N° : 176, SEQ ID N° : 178, SEQ ID N° : 180, SEQ ID N° : 182, SEQ ID N° : 184, SEQ ID N° : 186, SEQ ID N° : 188, SEQ ID N° : 190 e SEQ ID N° : 192, em que a dita porção, variante alélica, variante de união ou seqüência de hibridização codifica uma proteína GLK.

70. Método de acordo com qualquer um das reivindicações de 64 a 69, caracterizado pelo fato de que o dito ácido nucleico que codifica uma proteína de GLK é de origem vegetal, preferivelmente de uma planta monocotiledônea, mais preferivelmente da família Poaceae, mais preferivelmente de *Oryza sativa*.

71. Planta ou parte da mesma, incluindo sementes, caracterizada pelo fato de ser obtenível por um método de acordo com qualquer um das reivindicações de 64 a 70, em que a dita planta ou parte da mesma compreende um ácido nucleico que codifica uma proteína de GLK.

72. Construção, caracterizada pelo fato de que compreende:

(i) ácido nucleico que codifica uma proteína de GLK representado por SEQ ID N° : 157 ou SEQ ID N° : 193 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima;

(ii) uma ou mais seqüências de controle capazes de conduzir a expressão da seqüência de ácido nucleico de (i); e opcionalmente

(iii) uma seqüência de terminação de transcrição.

73. Construção de acordo com a reivindicação 72, caracterizada pelo fato de que a dita uma ou mais seqüências de controle é pelo menos a promotor constituído, preferivelmente um promotor GOS2.

74. Uso de uma construção como definida na reivindicação 72 ou 73, caracterizado pelo fato de ser para a fabricação de plantas tendo características relacionadas com o rendimento intensificadas, particularmente biomassa aumentada, rendimento aumentado de semente (peso de semente) e/ou número aumentado de sementes enchidas, com relação às plantas de controle.

75. Planta, parte de planta ou célula de planta, caracterizada pelo fato de que é transformada com uma construção como definida em qualquer um das reivindicações 72 ou 73.

76. Método para a produção de uma planta transgênica tendo características intensificadas relacionadas com o rendimento com relação às plantas de controle, caracterizado pelo fato de que compreende:

(i) introduzir e expressar em uma célula de planta ao ácido nucleico que codifica uma proteína de GLK representado por SEQ ID N° : 157 ou SEQ ID N° : 193 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima e

(ii) cultivar a célula de planta sob condições que promovem crescimento e crescimento vegetal.

77. Planta transgênica, caracterizada pelo fato de que tem rendimento aumentado com relação às plantas de controle, o dito rendimento

aumentado resultando da expressão aumentada de um ácido nucleico que codifica uma proteína de GLK representada por SEQ ID N° : 157 ou SEQ ID N° : 193 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima.

78. Planta de acordo com a reivindicação 71, 75 ou 77, caracterizada pelo fato de que a dita planta é uma planta de safra ou uma monocotiledônea ou um cereal, tal como arroz, milho, trigo, cevada, painço, triticale, centeio, sorgo ou aveias.

79. Partes colhíveis de uma planta como definida em qualquer um das reivindicações 71, 75, 77 ou 78, caracterizadas pelo fato de que as ditas partes colhíveis são sementes.

80. Produtos, caracterizados pelo fato de que são derivados de uma planta como definida na reivindicação 78 e/ou de partes colhíveis de uma planta como definidas na reivindicação 79.

81. Uso de um ácido nucleico que codifica uma proteína de GLK representada por qualquer um de SEQ ID N° : 157 ou SEQ ID N° : 193 ou ortólogos ou parálogos de qualquer uma das SEQ ID N° mencionadas acima, caracterizado pelo fato de ser empregado na intensificação de características relacionadas com o rendimento em plantas.

82. Uso de acordo com a reivindicação 81, caracterizado pelo fato de que dita característica relacionada com o rendimento é biomassa aumentada, rendimento aumentado de semente (peso de semente) e/ou número aumentado de sementes enchidas.

83. Método para aumentar o rendimento em plantas com relação às plantas de controle, caracterizado pelo fato de ser pela redução da expressão em uma planta de um gene REVOLUTA (REV) endógeno usando-se um domínio de zíper de leucina de homeodomínio REV delta (Δ) (HDZip) / seqüência de ácido nucleico de domínio de transferência (START) de lipídeo relacionado com o regulador agudo esteroidogênico (STAR), e

opcionalmente selecionar plantas tendo rendimento aumentado.

84. Método de acordo com a reivindicação 83, caracterizado pelo fato de que a dita expressão reduzida é realizada por silenciamento mediado por RNA de expressão genética.

85. Método de acordo com a reivindicação 84, caracterizado pelo fato de que o dito silenciamento mediado por RNA é realizado por co-supressão.

86. Método de acordo com a reivindicação 84, caracterizado pelo fato de que o dito silenciamento mediado por RNA é realizado por uso de uma seqüência de ácido nucleico REV ΔHDZip/START anti-sentido.

87. Método de acordo com a reivindicações 83 ou 84, caracterizado pelo fato de que a dita expressão reduzida é realizada usando-se uma repetição invertida de uma seqüência de ácido nucleico REV ΔHDZip/START, preferivelmente capaz de formar uma estrutura de grampo de cabelo.

88. Método de acordo com qualquer um das reivindicações 83 a 87, caracterizado pelo fato de que a dita seqüência de ácido nucleico REV ΔHDZip/START é de uma fonte vegetal ou fonte artificial.

89. Método de acordo com qualquer um das reivindicações 83 a 88, caracterizado pelo fato de que a dita seqüência de ácido nucleico REV ΔHDZip/START é de uma planta monocotiledônea e usada para a transformação de plantas monocotiledôneas ou é de uma planta dicotiledônea e usado para a transformação de plantas dicotiledôneas.

90. Método de acordo com qualquer um das reivindicações de 83 a 89, caracterizado pelo fato de que a dita seqüência de ácido nucleico REV ΔHDZip/START é de uma planta da família Poaceae e usada para transformação de plantas da família Poaceae, mais preferivelmente em que uma seqüência de ácido nucleico REV ΔHDZip/START de arroz é usada para transformar plantas de arroz.

91. Método de acordo com qualquer um das reivindicações 83 a 90, caracterizado pelo fato de que a dita seqüência de ácido nucleico REV ΔHDZip/START compreendem um comprimento suficiente de nucleotídeos substancialmente contíguos da SEQ ID N° : 194 ou SEQ ID N° : 196.

92. Método de acordo com qualquer um das reivindicações 83 a 91, caracterizado pelo fato de que a dita seqüência de ácido nucleico REV ΔHDZip/START que compreende um comprimento suficiente de nucleotídeos substancialmente contíguos é de uma seqüência de ácido nucleico que codifica um polipeptídeo REV ortólogo ou parálogo.

93. Método de acordo com a reivindicação 92, caracterizado pelo fato de que a dita seqüência de ácido nucleico que codifica um polipeptídeo REV ortólogo ou parálogo é representado por SEQ ID N° : 200, SEQ ID N° : 202, SEQ ID N° : 204, SEQ ID N° : 206, SEQ ID N° : 208, SEQ ID N° : 210, SEQ ID N° : 212, SEQ ID N° : 214, SEQ ID N° : 216, SEQ ID N° : 224, SEQ ID N° : 226, SEQ ID N° : 228, e SEQ ID N° : 230.

94. Método de acordo com qualquer um das reivindicações de 83 a 93, caracterizado pelo fato de que o dito rendimento aumentado é rendimento aumentado de semente e/ou biomassa aumentada.

95. Método de acordo com a reivindicação 94, caracterizado pelo fato de que o dito rendimento aumentado de semente é selecionado de um ou mais dos seguintes: (i) peso de semente aumentado; (ii) número aumentado de sementes enchidas; (iii) taxa de enchimento de semente aumentada; (iv) índice de colheita aumentado e (v) comprimento de semente individual aumentado.

96. Método de acordo com a reivindicação 94, caracterizado pelo fato de que a dita biomassa aumentada é biomassa de raiz.

97. Planta, parte de planta ou célula de planta da mesma, caracterizada pelo fato de ser obtenível por um método como definido na qualquer um das reivindicações 83 a 96, em que a dita planta, parte de planta

ou célula de planta da mesma tem expressão reduzida de um gene REV endógeno usando-se uma seqüência de ácido nucleico REV ΔHDZip/START e tendo rendimento aumentado com relação às plantas de controle.

98. Construção para reduzir a expressão em uma planta de um gene REV endógeno caracterizada pelo fato de que compreende uma ou mais seqüências de controle, uma seqüência de ácido nucleico REV ΔHDZip/START e opcionalmente uma seqüência de terminação de transcrição.

99. Construção de acordo com a reivindicação 98, caracterizada pelo fato de que a dita seqüência de ácido nucleico REV ΔHDZip/START é uma repetição invertida, preferivelmente capaz de formar uma estrutura de grampo de cabelo, cuja repetição invertida está sob o controle de um promotor constituído.

100. Construção de acordo com a reivindicação 98, caracterizada pelo fato de que a dita seqüência de controle é um promotor constituído.

101. Construção de acordo com as reivindicações 99 ou 100, caracterizada pelo fato de que o dito promotor constituído é um promotor GOS2.

102. Construção de acordo com a reivindicação 101, caracterizada pelo fato de que o dito promotor GOS2 é substancialmente como representado por SEQ ID N° : 58.

103. Planta, parte de planta ou célula de planta, caracterizada pelo fato de que é transformada com uma construção como definida em qualquer um das reivindicações 98 a 102.

104. Método para a produção de plantas transgênicas tendo rendimento aumentado com relação às plantas de controle, caracterizado pelo fato de que compreende:

(i) introduzir e expressar em uma planta, parte de planta ou

célula de planta uma construção genética que compreende uma ou mais seqüências de controle para reduzir a expressão em uma planta de um gene REV endógeno usando-se uma seqüência de ácido nucleico REV ΔHDZip/START e

(ii) cultivar a planta, parte de planta ou célula de planta sob condições que promovem crescimento e crescimento vegetal.

105. Planta transgênica, tendo rendimento aumentado com relação às plantas de controle, caracterizada pelo fato de que a planta tem expressão reduzida de um gene REV endógeno usando-se uma seqüência de ácido nucleico REV ΔHDZip/START.

106. Planta transgênica de acordo com a reivindicação 97, 103 ou 105, caracterizada pelo fato de que a dita planta é uma planta monocotiledônea, tal como cana de açúcar ou em que a planta é um cereal, tal como arroz, milho, trigo, cevada, triticale, painço, centeio, aveias ou sorgo.

107. Partes colhíveis, caracterizadas pelo fato de que são de uma planta transgênica como definida em qualquer um das reivindicações 97, 103, 105 ou 106.

108. Partes colhíveis de acordo com a reivindicação 107, caracterizadas pelo fato de que as ditas partes colhíveis são sementes.

109. Produtos, caracterizados pelo fato de que são derivados de uma planta como definida em qualquer um das reivindicações 97, 103, 105 ou 106 e/ou de partes colhíveis de uma planta como definidas nas reivindicações 107 ou 108.

110. Uso de uma seqüência de ácido nucleico REV ΔHDZip/START, caracterizado pelo fato de ser para reduzir a expressão em plantas de um gene REV endógeno para aumentar o rendimento nas ditas plantas com relação às plantas de controle.

111. Uso de acordo com a reivindicação 110, caracterizado pelo fato de que o dito rendimento aumentado é rendimento aumentado de

semente e/ou biomassa aumentada.

112. Uso de acordo com a reivindicação 111, caracterizado pelo fato de que o dito rendimento aumentado de semente é selecionado de um ou mais de: selecionado de um ou mais dos seguintes: (i) peso de semente aumentado; (ii) número aumentado de sementes enchidas; (iii) taxa de enchimento de semente aumentada; (iv) índice de colheita aumentado e (v) comprimento de semente individual aumentado.

113. Uso de acordo com a reivindicação 111, caracterizado pelo fato de que a dita biomassa aumentada é biomassa de raiz aumentada.

114. Método para aumentar o rendimento em plantas com relação às plantas de controle, caracterizado pelo fato de ser pela redução da expressão em uma planta de um gene semelhante ao CLE endógeno usando-se uma seqüência de ácido nucleico semelhante a CLE e opcionalmente selecionar plantas tendo rendimento aumentado.

115. Método de acordo com a reivindicação 114, caracterizado pelo fato de que a dita expressão reduzida é realizada por silenciamento mediado por RNA de expressão genética.

116. Método de acordo com a reivindicação 115, caracterizado pelo fato de que o dito silenciamento mediado por RNA é realizada por co-supressão.

117. Método de acordo com a reivindicação 115, caracterizado pelo fato de que o dito silenciamento mediado por RNA é realizado por uso de uma seqüência de ácido nucleico semelhante a CLE anti-sentido.

118. Método de acordo com as reivindicações 114 ou 115, caracterizado pelo fato de que a dita expressão reduzida é realizada usando-se uma repetição invertida de uma seqüência de ácido nucleico semelhante a CLE, preferivelmente capaz de formar uma estrutura de grampo de cabelo.

119. Método de acordo com qualquer um das reivindicações de 114 a 119, caracterizado pelo fato de que o dito ácido nucleico semelhante

a CLE é de uma fonte vegetal ou fonte artificial.

120. Método de acordo com qualquer um das reivindicações de 114 a 119, caracterizado pelo fato de que a dita seqüência de ácido nucleico semelhante a CLE é de uma planta monocotiledônea e usada para transformação de plantas monocotiledôneas ou é de uma planta dicotiledônea e usada para transformação de plantas dicotiledôneas.

121. Método de acordo com qualquer um das reivindicações de 114 a 120, caracterizado pelo fato de que a dita seqüência de ácido nucleico semelhante a CLE é de uma planta da família Poaceae e usada para transformação de plantas da família Poaceae, mais preferivelmente em que a seqüência de ácido nucleico semelhante a CLE de cana de açúcar é usada para transformar plantas de arroz.

122. Método de acordo com qualquer uma das reivindicações de 114 a 121, caracterizado pelo fato de que a dita seqüência de ácido nucleico semelhante a CLE compreende um comprimento suficiente de nucleotídeos substancialmente contíguos da SEQ ID N° : 232.

123. Método de acordo com qualquer um das reivindicações de 114 a 122, caracterizado pelo fato de que a dita seqüência de ácido nucleico semelhante a CLE que compreende um comprimento suficiente de nucleotídeos substancialmente contíguos é de uma seqüência de ácido nucleico que codifica um polipeptídeo semelhante ao CLE ortólogo ou parálogo.

124. Método de acordo com qualquer um das reivindicações de 114 a 123, caracterizado pelo fato de que o dito rendimento aumentado é rendimento aumentado de semente.

125. Método de acordo com a reivindicação 124, caracterizado pelo fato de que o dito rendimento aumentado de semente compreende pelo menos o peso de semente aumentado.

126. Planta, parte de planta ou célula de planta da mesma,

caracterizada pelo fato de ser obtenível por um método como definido em qualquer um das reivindicações 114 a 125, em que a dita planta, parte de planta ou célula de planta da mesma tem expressão reduzida de um gene semelhante ao CLE endógeno usando-se uma seqüência de ácido nucleico semelhante a CLE e tendo rendimento aumentado com relação às plantas de controle.

127. Construção, caracterizada pelo fato de que é para reduzir a expressão em uma planta de gene semelhante ao CLE endógeno que compreende uma ou mais seqüências de controle, a seqüência de ácido nucleico semelhante a CLE, e opcionalmente uma seqüência de terminação de transcrição.

128. Construção de acordo com a reivindicação 127, caracterizada pelo fato de que a dita seqüência de ácido nucleico semelhante a CLE é uma repetição invertida, preferivelmente capaz de formar uma estrutura de grampo de cabelo, cuja repetição invertida está sob o controle de um promotor específico de semente.

129. Construção de acordo com a reivindicação 127, caracterizada pelo fato de que a dita seqüência de controle é um promotor específico de semente.

130. Construção de acordo com a reivindicação 128 ou 129, caracterizada pelo fato de que o dito promotor específico de semente é um promotor específico de endosperma.

131. Construção de acordo com a reivindicação 130, caracterizada pelo fato de que o dito promotor específico de endosperma é substancialmente como representado por SEQ ID N° : 236.

132. Planta, parte de planta ou célula de planta, caracterizada pelo fato de que é transformada com uma construção como definida em qualquer um das reivindicações 127 a 131.

133. Método para a produção de plantas transgênicas tendo

rendimento aumentado com relação às plantas de controle, caracterizado pelo fato de que compreende:

(i) introduzir e expressar em uma planta, parte de planta ou célula de planta uma construção genética que compreende uma ou mais seqüências de controle para reduzir a expressão em uma planta de um gene semelhante a CLE endógeno usado-se uma seqüência de ácido nucleico semelhante a CLE e

(iii) cultivar a planta, parte de planta ou célula de planta sob condições que promovem crescimento e crescimento vegetal.

134. Planta transgênica tendo rendimento aumentado com relação às plantas de controle, caracterizada pelo fato de que a dita planta tem expressão reduzida de um gene semelhante a CLE endógeno usando-se a seqüência de ácido nucleico semelhante a CLE.

135. Planta transgênica de acordo com a reivindicação 126, 132 ou 134, caracterizada pelo fato de que a dita planta é uma planta monocotiledônea, tal como cana de açúcar ou em que a planta é um cereal, tal como arroz, milho, trigo, cevada, painço, centeio, aveias ou sorgo.

136. Partes colhíveis, caracterizadas pelo fato de que são de uma planta transgênica como definida em qualquer um das reivindicações 126, 132, 134 ou 135.

137. Partes colhíveis de acordo com a reivindicação 136, caracterizadas pelo fato de que ditas partes colhíveis são sementes.

138. Produtos, caracterizados pelo fato de que são derivados de uma planta de acordo com qualquer um das reivindicações 126, 132, 134 ou 135 e/ou de partes colhíveis de uma planta como definida na reivindicação 136 ou 137.

139. Uso de uma seqüência de ácido nucleico semelhante a CLE, caracterizado pelo fato de ser para reduzir a expressão em plantas de um gene semelhante ao CLE endógeno para aumentar o rendimento nas ditas

plantas com relação às plantas de controle.

140. Uso de acordo com a reivindicação 139, caracterizado pelo fato de que o dito rendimento aumentado é o rendimento aumentado de semente.

141. Uso de acordo com a reivindicação 140, caracterizado pelo fato de que o dito rendimento aumentado de semente compreendem pelo menos o peso de semente aumentado.

142. Método para aumentar a resistência à tensão abiótica em plantas com relação às plantas de controle, caracterizado pelo fato de que compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo SYR, cujo polipeptídeo SYR compreende um domínio rico em leucina, precedido pelo motivo 1 de tripeptídeo conservado (um de SEQ ID N° : 256, 257, 258 ou 259)) e seguido pelo motivo 2 conservado (SEQ ID N° : 260), em que a dita resistência à tensão abiótica aumentada é a eficiência de captação de nutriente aumentada, com relação às plantas de controle.

143. Método de acordo com a reivindicação 142, caracterizado pelo fato de que o dito polipeptídeo SYR tem, em ordem crescente de preferência, pelo menos 27%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ou mais identidade de seqüência para o polipeptídeo SYR representado por SEQ ID N° : 252.

144. Método de acordo com a reivindicação 142 ou 143, caracterizado pelo fato de que o dito ácido nucleico que codifica um polipeptídeo SYR é representado por qualquer um do ácido nucleico SEQ ID NOs dadas na Tabela II ou uma porção destes ou uma seqüência capaz de hibridizar com qualquer um dos ácidos nucleicos das SEQ ID NOs dadas na Tabela II.

145. Método de acordo com qualquer uma das reivindicações de 142 a 144, caracterizado pelo fato de que a dita seqüência de ácido

nucleico codifica um ortólogo ou parálogo de qualquer uma das SEQ ID NOs dadas na Tabela II.

146. Método de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que a dita proteína SYR além disso compreende o motivo 3 conservado (SEQ ID N° : 261).

147. Método de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que a dita eficiência de captação de nutriente resulta no rendimento aumentado de semente e/ou biomassa aumentada.

148. Método de acordo com a reivindicação 147, caracterizado pelo fato de que o dito rendimento aumentado de semente compreende pelo menos o peso total aumentado de sementes, peso de mil núcleos e/ou número aumentado de sementes enchidas.

149. Método de acordo com a reivindicação 147, caracterizado pelo fato de que a dita biomassa aumentada é biomassa de broto aumentada e/ou biomassa de raiz aumentada.

150. Método de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que a dita eficiência de captação de nutriente aumentada ocorre sob condições de seca brandas.

151. Método de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que a dita expressão modulada é realizada por introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo SYR.

152. Método de acordo com a reivindicação 151, caracterizado pelo fato de que o dito ácido nucleico está operacionalmente ligado a um promotor constituído, preferivelmente a um promotor GOS2.

153. Método de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que o dito ácido nucleico que codifica um polipeptídeo SYR é de origem vegetal, preferivelmente de uma planta

monocotiledônea, ainda preferivelmente da família Poaceae, mais preferivelmente do gênero Oryza, mais preferivelmente de *Oryza sativa*.

154. Uso de uma construção, caracterizado pelo fato de ser em um método para a fabricação de plantas tendo resistência à tensão abiótica aumentada, dita construção compreendendo

(a) ácido nucleico que codifica um polipeptídeo SYR como definido em qualquer um das reivindicações de 142 a 146;

(b) uma ou mais seqüências de controle capazes de conduzir a expressão da seqüência de ácido nucleico de (a); e opcionalmente

(c) uma seqüência de terminação de transcrição,

e em que um das ditas seqüências de controle é um promotor constituído, preferivelmente um promotor GOS2 e em que a dita resistência à tensão abiótica aumentada é a eficiência de captação de nutriente aumentada, com relação às plantas de controle.

155. Uso de um ácido nucleico que codifica um polipeptídeo SYR, caracterizado pelo fato de ser em um método para aumentar a resistência à tensão abiótica em plantas com relação às plantas de controle, em que a dita resistência à tensão abiótica aumentada é a eficiência de captação de nutriente aumentada, com relação às plantas de controle.

156. Uso de acordo com a reivindicação 155, caracterizado pelo fato de que a dita eficiência de captação de nutriente aumentada resulta em rendimento aumentado de semente e/ou biomassa aumentada.

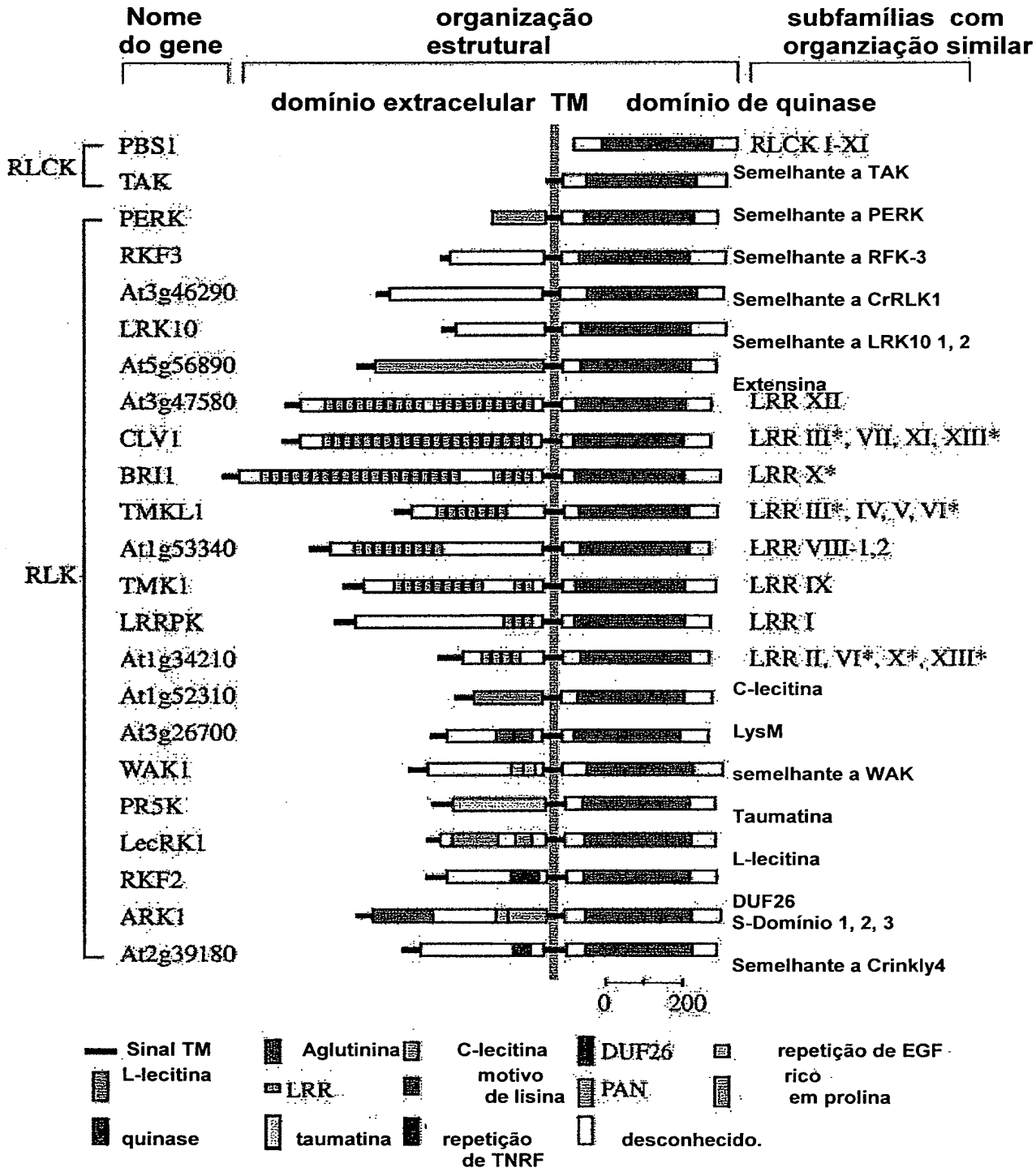

Nome do gene
organização estrutural
subfamílias com organização similar
domínio extracelular
TM
domínio de quinase
RLCK
RLK
PBS1
TAK
PERK
RKF3
At3g46290
LRK10
At5g56890
At3g47580
CLV1
BRI1
TMKL1
At1g53340
TMK1
LRRPK
At1g34210
At1g52310
At3g26700
WAK1
PR5K
LecRK1
RKF2
ARK1
At2g39180
RLCK I-XI
Semelhante a TAK
Semelhante a PERK
Semelhante a RFK-3
Semelhante a CrRLK1
Semelhante a LRK10 1, 2
Extensina
LRR XII
LRR III*, VII, XI, XIII*
LRR X*
LRR III*, IV, V, VI*
LRR VIII-1,2
LRR IX
LRR I
LRR II, VI*, X*, XIII*
C-lecitina
LysM
semelhante a WAK
Taumatina
L-lecitina
DUF26
S-Domínio 1, 2, 3
Semelhante a Crinkly4
0
200
Sinal TM
Aglutinina
C-lecitina
DUF26
repetição de EGF
L-lecitina
LRR
motivo de lisina
PAN
rico em prolina
quinase
taumatina
repetição de TNRF
desconhecido.

FIGURA 1

MENK**SHSNHHRYHHLSSRSH**EQNQRGLISINAIIIIGISIISIFIILAILLIIILLHRLK
SARVKAQELSCKESFNNMNNGGASTNYSYTSSPDDIKRDCLYSRNPTSFRQLPPQTKSCR
RSRAEGVEVYTYKELEIATNN*FSEEKKIGNGDVYKGVLSDGTVAAIKKLHMFNDNASNQK*
HEERSFRLVQRSTSRLQCPYLVELLGYCADQNHRILIYEFMPNGTVEHHLHDHNFKNLKD
RPQPLDWGARLRIALDCARALEFLHENTISTVIHRNFKCTNILLDQNNRAKVSDFGLAKT
GSDKLNGEISTRVIGTTGYLAPEYASTGKLTTKSDVYSYGIVLLQLLTGRTPIDSRRPRG
QDVLVSWALPRLTNREKISEMVDPTMKGQYSQKDLIQVAAIAAVCVQPEASYRPLMTDVV
*HSL*IPLVKAFNKSTDSSRFPSRRESLSFDDIMP

FIGURA 2

```
Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------------------------------
AT2G20300       ------------------------------------------------------------
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       MENLTLLLRICLVSSVLVAASSSGSELLSPLSSPPSPLPETSKGFGQAPSNSPESHKSDN
Osi001397.3p    ------------------------------------------------------------
Osi000142.1p    ------------------------------------------------------------
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       ------------------------------------------------------------
Osi000040.5p    ------------------------------------------------------------

Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------------------------------
AT2G20300       ------------------------------------------------------------
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       VPPSKASSQPSLPPLADLAAPPPSDSVGGKAPAGVPVVFVPNAPAPATIPVKDLPVASPP
Osi001397.3p    ------------------------------------------------------------
Osi000142.1p    ------------------------------------------------------------
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       ------------------------------------------------------------
Osi000040.5p    ------------------------------------------------------------
```

FIGURA 3

```
Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------------------------------
AT2G20300       ------------------------------------------------------------
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       VLQPITPIASPPRFIPGDAPKEPPFSGRVTPAPVSSPVSDIPPIPSVALPPPTPSNVPPR
Osi001397.3p    ------------------------------------------------------------
Osi000142.1p    ---------MRSLACCAVLLLASVLGSTGTDLGPSPVVANSPDAQDQTSSPPEPTIALGP
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       ------------------------------------------------------------
Osi000040.5p    ------------------------------------------------------------

Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------------------------------
AT2G20300       ------------------------------------------------------------
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       NASNNHKPPPIEKSIAPVASPPTISIDIAPPVHPVIPKLTPSSSPVPTSTPTKGSPRRNP
Osi001397.3p    ----------MGRRGGGGRAGGAWIGGCVLALAATLVVCVLVIRGAGGLNVKPPAATSRR
Osi000142.1p    VTLPTGCSEFLMWCGIMALKIITSIDNCPYNCFVSAPSAPSASPPVAKGAVSPAVPTRPQ
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       ------------------------------------------------------------
Osi000040.5p    ------------------------------------------------------------

Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------------------------------
AT2G20300       ------------------------------------------------------------
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       PTTHPVFPIESPAVSPDHAANPVKHPPPSDNGDDSKSPGAAPANETAKPLPVFPHKASPP
Osi001397.3p    VNIQQELYAPSPTISHRGLAIPPLPTTSSPVFPPPIRSYPPLPANKPYPNSPVVAPPKGR
Osi000142.1p    NAPTPVTPPKEYNAPPPVELTPPAPTHVAPVAPPQAAVENPAPVLPGTPALLPSVQAPAP
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       ------------------------------------------------------------
Osi000040.5p    ------------------------------------------------------------
```

FIGURA 3 (Cont.)

```
Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------------------------------
AT2G20300       ------------------------------------------------------------
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       SIAPSAPKFNRHSHHTSPSTTPPPDSTPSNVHHHPSSPSPPPLSSHHQHHQERKKIADSP
Osi001397.3p    HHHSLPVNNTRVKGPAYSPSNSPSIHRKHGIPVAAPPKQHSSNLPPSHHRPHKGSFPVIS
Osi000142.1p    SVARNPNLPIVQPPSVNNPPSGPIGSGNGVPPYPPPQRSLPAIPPSTSGVPRESVKPPVA
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       -------MSIFSLASSSDSQRDDLIMALKAVVIVYCVVSLVSVQLADAQHEGLPVSPTLS
Osi000040.5p    ------------------------------------------------------------

Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------MPPRRAALLLLALAVYVPLGT
AT2G20300       -MRNFAMLLLLILLLHSLASFPICFARLFPMSLPFTRSKAHQMHFFHPYLNPSVAPTPSP
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       APSPLPPHLISPKKSNRKGSMTPPPQSHHAPSPPIPDSLISPAHAPVSFSMKRISPALAP
Osi001397.3p    PTPHKADNASATKHGRSGLHHSPAPAPVGLPPSEGNARGNPAYAPRHPHEYHSPSNSPEP
Osi000142.1p    PPIIAQAPRQQALAPSSDHSNGNSVPPANTSPPHKNSHIPRALPPKESSSQTGTAHKPPI
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       PSTSPVITDLPLPAEFPRFHRKYFAPQQAEAPQHSPPYSRLVASDHPPTSSHFSKPSMKR
Osi000040.5p    ------------------------------------MARIRRASSGRDMSDNLVQHNRRS

Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ASSTTIASYLLGLWSRAHRHSLPAPAPAPAPAPAPETHRPGIRHPVPRHHRKRPHVAPPL
AT2G20300       AFSPNPSRIPPLRHKGHHRHRRWHLRRNATAVSPSSHDCQQTCVEPLTSTPFGSPCGCVF
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       SPTQVFPLRSSSRPSKSRKFPLGPPLQAFPPPPPPNSDCSSTICLEPYTNTPPGSPCGCVW
Osi001397.3p    GLPPVNPPDSHAFKKPKSLAP---APQSFPPPPLSSYCMALNCQDPLTNSLPGTTCLCVW
Osi000142.1p    RGPHISPTMPPIPPQPGPKAPSAHPIWALPPPPPNLDCNSLACPEPLTDPPAGAPCVCVL
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       NAQSPGAGLADIAPAQSSNGVLPDALTQPPLSPSISNCCKSDMVLKRRSIGCHCVYPIKL
Osi000040.5p    EAEISTHVDAAPPDAASNTSAAPSGLVQPPVSPHN-ACCSHNMVQKRGSQDCHCVYPVRV
```

FIGURA 3 (Cont.)

```
Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------------------------------
AT2G20300       ------------------------------------------------------------
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       SIAPSAPKFNRHSHHTSPSTTPPPDSTPSNVHHHPSSPSPPPLSSHHQHHQERKKIADSP
Osi001397.3p    HHHSLPVNNTRVKGPAYSPSNSPSIHRKHGIPVAAPPKQHSSNLPPSHHRPHKGSFPVIS
Osi000142.1p    SVARNPNLPIVQPPSVNNPPSGPIGSGNGVPPYPPPQRSLPAIPPSTSGVPRESVKPPVA
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       -------MSIFSLASSSDSQRDDLIMALKAVVIVYCVVSLVSVQLADAQHEGLPVSPTLS
Osi000040.5p    ------------------------------------------------------------

Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ----------------------------------------MPPRRAALLLLALAVYVPLGT
AT2G20300       -MRNFAMLLLLILLLHSLASFPICFARLFPMSLPFTRSKAHQMHFFHPYLNPSVAPTPSP
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       APSPLPPHLISPKKSNRKGSMTPPPQSHHAPSPPIPDSLISPAHAPVSFSMKRISPALAP
Osi001397.3p    PTPHKADNASATKHGRSGLHHSPAPAPVGLPPSEGNARGNPAYAPRHPHEYHSPSNSPEP
Osi000142.1p    PPIIAQAPRQQALAPSSDHSNGNSVPPANTSPPHKNSHIPRALPPKESSSQTGTAHKPPI
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       PSTSPVITDLPLPAEFPRFHRKYFAPQQAEAPQHSPPYSRLVASDHPPTSSHFSKPSMKR
Osi000040.5p    ------------------------------------MARIRRASSGRDMSDNLVQHNRRS

Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ASSTTIASYLLGLWSRAHRHSLPAPAPAPAPAPAPETHRPGIRHPVPRHHRKRPHVAPPL
AT2G20300       AFSPNPSRIPPLRHKGHHRHRRWHLRRNATAVSPSSHDCQQTCVEPLTSTPFGSPCGCVF
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       SPTQVFPLRSSSRPSKSRKFPLGPPLQAFPPPPPNSDCSSTICLEPYTNTPPGSPCGCVW
Osi001397.3p    GLPPVNPPDSHAFKKPKSLAP---APQSFPPPPLSSYCMALNCQDPLTNSLPGTTCLCVW
Osi000142.1p    RGPHISPTMPPIPPQPGPKAPSAHPIWALPPPPPNLDCNSLACPEPLTDPPAGAPCVCVL
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       NAQSPGAGLADIAPAQSSNGVLPDALTQPPLSPSISNCCKSDMVLKRRSIGCHCVYPIKL
Osi000040.5p    EAEISTHVDAAPPDAASNTSAAPSGLVQPPVSPHN-ACCSHNMVQKRGSQDCHCVYPVRV
```

FIGURA 3 (Cont.)

```
Osi009054.1p      --------------------GRSTLSVSRVSSASASMLSTVATCTTSVKTFSLSQLEKAT
Osi001078.4p      RGTGR------VSNSSATRKTGMRSSFSSMTSSTASFVSTIATCPPTVKTFSISELEKAT
AT2G20300         SSNAVGPALAPSINKRPGAGS--MFSSSARSSGSDSLMSSMATCALSVKTFTLSELEKAT
ABE82646.1        ------------------------------------------------------------
ABB36644.1        PSNAVGPVFTPSMHKRSGKGIGSTISSSPTSSTSISLISAMPASILSVKTFTLAELERAT
At5g56890         LSKRTPLARPSLPSLSKPSGSARSLTGSRFSSTSLSFESSIAPFTLSAKTFTASEIMKAT
Osi001397.3p      TEESPMPPKPAGPGSAVVGSRLGS----RPISASPSFSSSIVTYKGTAKTFSLIEMERAT
Osi000142.1p      AQRFPHSALPKFSRSSGTGQTLLAGRYSSPSGPSGSLGSSIATYAGQAKTFKFAEIEKAT
CDS845            NGGASTNYSYTSSPDDIKRDCLYSRNPTSFRQLPPQTKSCRRSRAEGVEVYTYKELEIAT
NP_175879.2       NGGASTNYSYTSSPDDIKRDCLYSRNPTSFRQLPPQTKSCRRSRAEGVEVYTYKELEIAT
Popsp             ----------------------------------MGHKPPDKYKGVQVFTYKELEIAT
Osi093820.1p      ------------------------------------------------------------
Osi015947.1p      PGGAPASSCGSSVRG--YNNSRYYAAAAAGCIYGGRLGFSVQPRNRGAQVFTYRELESAT
AT3G58690         ---------------------------------------VQDVTENGLQIFTFKQLHSAT
At3g58690sv       ---------------------------------------VQDVTENGLQIFTFKQLHSAT
TC196912          ------------------------------------------------------------
TC81885           ------------------------------------------------------------
Osi003977.2p      PMPPPLPPPPPSANAPTLGKESPSSNSASDG---AAAAVVVGGERGAVQVFSYRQLHAAT
TC15181           ----PPPPPPRAPPRQQQGKDSPSSNSASDGGGAAAAMSVVVAGERGVQVFSYRQLHAAT
AT4G02010         ILAIITVLVICSRALREEKAPDPHKEAVKPRNLDAGSFGGSLPHPASTRFLSYEELKEAT
Osi000040.5p      IAVLVIAMFICFCKLRKGKRKVPPVETPKQRTPDAVSAVDSLPRPTSTRFLAYDELKEAT

Osi009054.1p      DGFDSKRVLGQGGFGRVYHGTMDGGDEIAVKLLTREDRSGDRE---------FIAEVEML
Osi001078.4p      ENFSFNKIIGEGGYGRVYRGTIDDEVDVAVKLLTRKHQNRDRE---------FIAEVEML
AT2G20300         DRFSAKRVLGEGGFGRVYQGSMEDGTEVAVKLLTRDNQNRDRE---------FIAEVEML
ABE82646.1        ------------------------------------------------------------ML
ABB36644.1        DKFSLKRVLGEGGFGRVYHGILEDRTEVAVKVLTRDNQNGDRE---------FIAEVEML
At5g56890         NNFDESRVLGEGGFGRVYEGVFDDGTKVAVKVLKRDDQQGSRE---------FLAEVEML
Osi001397.3p      QRFDNSRIIGEGGFGRVYEGILEDGERVAVKILKRDDQQGTRE---------FLAEVEML
Osi000142.1p      NSFDDSTVLGEGGFGCVYQGTLEDGTRVAVKVLKRYDGQGERE---------FLAEVEML
CDS845            NNFSEE---KKIGNGDVYKGVLSDGTVAAIKKLHMFNDNASNQKHEERS---FRLVQRST
NP_175879.2       NNFSEE---KKIGNGDVYKGVLSDGTVAAIKKLHMFNDNASNQKHEERS---FRLEVDLL
Popsp             NKFSEANVTWNEGYGVVYGGTLSDGTVAAIKMLHRAGK------QGERA---FRIEVDLL
Osi093820.1p      ------------------------------------------------------MPTSRVDLL
Osi015947.1p      DGFSECNVVGRGAYGVVFRGRLGDGTTAAIKRLKMDGRREGEREFRIEMGVAITAQVDLL
AT3G58690         GGFSKSNVVGNGGFGLVYRGVLNDGRKVAIKLMDHAGKQGEEE---------FKMEVELL
At3g58690sv       GGFSKSNVVGNGGFGLVYRGVLNDGRKVAIKLMDHAGKQGEEE---------FKMEVELL
TC196912          ---SKSNVIGHGGFGLVYRGVLNDGRKVAIKFMDQAGKQGEEE---------FKVEVELL
TC81885           ------------------------------------QGEEE---------FKVEVELL
Osi003977.2p      GGFGRAHVVGQGSFGAVYRGVLPDGRKVAVKLMDRPGKQGEEE---------FEMEVELL
TC15181           GGFGRAHMVGQGSFGAVYRGVLPDGRKVAVKLMDRPGKQGEEE---------FEMEVELL
AT4G02010         SNFESASILGEGGFGKVYRGILADGTAVAIKKLTSGGPQGDKE---------FQVEIDML
Osi000040.5p      NNFDPSSMLGEGGFGRVFKGVLTDGTAVAIKKLTSGGHQGDKE---------FLVEVEML

Osi009054.1p      SRLHHRNLVKLIGICIEH--NKRCLVYELIRNGSVESHLHGAD----------------
Osi001078.4p      SRLHHRNLVKLIGICIER--STRCLVFELVPNGSVESHLHGSD----------------
AT2G20300         SRLHHRNLVKLIGICIEG--RTRCLIYELVHNGSVESHLHEG----------------
ABE82646.1        SRLHHRNLVKLIGICIEG--RRRCLVYELVPNGSVESHLHGDD----------------
ABB36644.1        SRLHHRNLVKLIGICSEE--RTRSLVYELVRNGSVESHLHGRD----------------
At5g56890         SRLHHRNLVNLIGICIED--RNRSLVYELIPNGSVESHLHGID----------------
Osi001397.3p      SRLHHRNLVKLIGICTEE--HIRCLVYELVPNGSVESHLHGSD----------------
Osi000142.1p      GRLHHRNLVKLLGICVEE--NARCLVYELIPNGSVESHLHGVD----------------
CDS845            SRLQCPYLVELLGYCADQ--NHRILIYEFMPNGTVEHHLHDH----------------
NP_175879.2       SRLQCPYLVELLGYCADQ--NHRILIYEFMPNGTVEHHLHDH----------------
Popsp             SRLHSPYLVELLGYCADQ--NHRLLVFEFMPNGTLQHHLHHK----------------
Osi093820.1p      SRMHSPYLVGLLGYCADQ--SHRLLVFEFMPNGSLKSHLHRR----------------
Osi015947.1p      SRMHSPYLVGLLGYCADQ--SHRLLVFEFMPNGSLKSHLHRR----------------
AT3G58690         SRLRSPYLLALLGYCSDN--SHKLLVYEFMANGGLQEHLYLP----------------
At3g58690sv       SRLRSPYLLALLGYCSDN--SHKLLVYEFMANGGLQEHLYLP----------------
TC196912          SRLHSPYLLALLGYCSDS--NHKLLVYEFMANGGLQEHLYPV----------------
TC81885           CRLRSPYLLSLIGYCSES--SHKLLVYEFMANGGLQEHLYPI----------------
Osi003977.2p      SRLRSPYLLGLIGHCSEG--GHRLLVYEFMANGGLQEHLYPNGAFEKIETFSIYLVKQRP
TC15181           SRLRSPYLLGLIGHCSEG--GHRLLVYEFMANGGLQEHLYPN----------------
AT4G02010         SRLHHRNLVKLVGYYSSRDSSQHLLCYELVPNGSLEAWLH----------------
Osi000040.5p      SRLHHRNLVKLIGYYSNR------------------------------------------
                   *::   *: *:*   .
```

FIGURA 3 (Cont.)

```
Osi009054.1p    ------------------------------KAKGMLNWDVRMKIALGAARGLAYLHEDS
Osi001078.4p    ------------------------------KIYGPLDFDTRMKIALGAARGLAYLHEDA
AT2G20300       -----------------------------------TLDWDARLKIALGAARGLAYLHEDS
ABE82646.1      ------------------------------KNRGPLDWEARMKIALGAARGLAYLHEDS
ABB36644.1      -------------------------------GRKEPLDWDVRLKIALGAARGLAYLHEDS
At5g56890       ------------------------------KASSPLDWDARLKIALGAARGLAYLHEDS
Osi001397.3p    ------------------------------KGTAPLYWDARLKIALGAARALAYLHEDS
Osi000142.1p    -----------------------------LETAPLDWNARMKIALGAARALAYLHEDS
CDS845          ------------------------NFKNLKDRPQPLDWGARLRIALDCARALEFLHENT
NP_175879.2     ------------------------NFKNLKDRPQPLDWGARLRIALDCARALEFLHENT
Popsp           ------------------------QYR-------PLDWGTRLRIALDCARALEFLHELT
Osi093820.1p    ------------------------ALAPAE-QPPPLDWQTRLGIALDCARALEFLHEHS
Osi015947.1p    ------------------------ALAPAE-QPPPLDWQTRLGIALDCARALEFLHEHS
AT3G58690       -------------------------NRSGSVPPRLDWETRMRIAVEAAKGLEYLHEQV
At3g58690sv     -------------------------NRSGSVPPRLDWETRMRIAVEAAKGLEYLHEQV
TC196912        -------------------------SNSIITPVKLDWETRLRIALEAAKGLEYLHEHV
TC81885         -------------------------KGSNNCCPKLDWKTRLRIALEAAKGLEYLHEHV
Osi003977.2p    IFDKNIGHPICRRAHFSRQLISHKADIVGSCGGISKLDWPTRMRIALEAAKGLEYLHERV
TC15181         -------------------------RGSCGGISKLDWDTRMRIALEAAKGLEYLHERV
AT4G02010       ----------------------------GPLGLNCPLDWDTRMKIALDAARGLAYLHEDS
Osi000040.5p    ----------------------------TLGASRPLDWDTRMRIALDAARGLAYLHEDS
                                             * : .*: **: .*:.* :***

Osi009054.1p    NPHVIHRDFKGSNILLEEDFTPKVTDFGLAREATNGIQ--PISTRVMGTFGYVAPEYAMT
Osi001078.4p    NPHVIHRDFKASNVLLENDFTPKVADFGLAKEASEGMD--HISTQVMGTFGYVAPEYAMT
AT2G20300       NPRVIHRDFKASNVLLEDDFTPKVSDFGLAREATEGSQ--HISTRVMGTFGYVAPEYAMT
ABE82646.1      NPRVIHRDFKASNVLLEDDFTPKVSDFGLAREATEGSN--HISTRVMGTFGYVAPEYAMT
ABB36644.1      NPRVIHRDFKASNVLLEDDFTPKVADFGLAREATEGSH--HISTRVMGTFGYVAPEYAMT
At5g56890       SPRVIHRDFKSSNILLENDFTPKVSDFGLARNALDDEDNRHISTRVMGTFGYVAPEYAMT
Osi001397.3p    SPRVIHRDFKSSNILLEHDFTPKVSDFGLARTAIGEGN-EHISTRVMGTFGYVAPEYAMT
Osi000142.1p    SPCVIHRDFKSSNILLEHDFTPKVSDFGLARTARGEGN-QHISTRVMGTFGYVAPEYAMT
CDS845          ISTVIHRNFKCTNILLDQNNRAKVSDFGLAKTGSDKLN-GEISTRVIGTTGYLAPEYAST
NP_175879.2     ISTVIHRNFKCTNILLDQNNRAKVSDFGLAKTGSDKLN-GEISTRVIGTTGYLAPEYAST
Popsp           IPAVIHRDFKCSNILLDQNFRAKVSDFGSAKMGSERIN-ARNSMCLPSTTGYLAPEHAST
Osi093820.1p    SPAVIHRDFKCSNILLDHNYRARVSDFGMGKLGSNKAN-GQ------------------
Osi015947.1p    SPAVIHRDFKCSNILLDHNYRARVSDFGMAKLGSNKAN-GQVAAITAMCIQTKADYRP--
AT3G58690       SPPVIHRDFKSSNILLDRNFNAKVSDFGLAKVGSDKAG-GHVSTRVLGTQGYVAPEYALT
At3g58690sv     SPPVIHRDFKSSNILLDRNFNAKVSDFGLAKVGSDKAG-GHVSTRVLGTQGYVAPEYALT
TC196912        SPPVIHRDFKSSNILLDKKFHAKVSDFGLAKLGPDRAG-GHVSTRVLGTQGYVAPEYALT
TC81885         NPPIIHRDLKSSNILLDKNFHAKVSDFGLAKLGSDKAG-GHVSTRVLGTQGYVAPEYALT
Osi003977.2p    NPPVIHRDFKSSNILLDKDFRARVSDFGLAKLGSDRAG-GHVSTRVLGTQGYVAPDYG--
TC15181         NPPVIHRDFKSSNILLDKDFHARVSDFGLTKLGSDRAG-GHVSTRVLGTQGYVAPEYALT
AT4G02010       QPSVIHRDFKASNILLENNFNAKVADFGLAKQAPEGRG-NHLSTRVMGTFGYVAPEYAMT
Osi000040.5p    QPCVIHRDFKASNILLEDDFHAKVSDFGLAKQAPEGRT-NYLSTRVMGTFGYVAPEYAMT
                 . :***::* :*:**: .  .:*:***  : .

Osi009054.1p    GHLLVKSDVYSYGVVLLELLSGRKPVCMSDTNGPQNLVTWAR------------------
Osi001078.4p    GHLLVKSDVYSYGVVLLELLSGRKPVDMTQPPGSENLVTWAR------------------
AT2G20300       GHLLVKSDVYSYGVVLLELLTGRRPVDMSQPSGEENLVTWAR------------------
ABE82646.1      GHLLVKSDVYSYGVVLLELLTGRKPVDMSQPQGQENLVTWAR------------------
ABB36644.1      GHLLVKSDVYSYGVVLLELLSGRKPVDMSQPPGEENLVTWAR------------------
At5g56890       GHLLVKSDVYSYGVVLLELLTGRKPVDMSQPPGQENLVSWTR------------------
Osi001397.3p    GHLLVKSDVYSYGVVLLELLTGRKPVDILRPPGQENLVAWAC------------------
Osi000142.1p    GHLLVKSDVYSYGVVLLELLTGRKPVDMSRPGGQENLVSWAR------------------
CDS845          GKLTTKSDVYSYGIVLLQLLTGRTPIDSRRPRGQDVLVSWAL------------------
NP_175879.2     GKLTTKSDVYSYGIVLLQLLTGRTPIDSRRPRGQDVLVSWAL------------------
Popsp           GKLTTKSDVYSYGVVLLQLLTGRKPVDTKQPSGEHVLVSWAL------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ----LMTDVVQSLIPIVKSPLMSCTSTPLRPAHGHHHVVYMS------------------
AT3G58690       GHLTTKSDVYSYGVVLLELLTGRVPVDMKRATGEGVLVSWAL------------------
At3g58690sv     GHLTTKSDVYSYGVVLLELLTGRVPVDMKRATGEGVLVSWAL------------------
TC196912        GHLTTKSDVYSYGVVLLELLTGRVPVDMKRPPGEGVLVSWAL------------------
TC81885         GHLTTKSDVYSYGVVLLELLTGRVPVDMKRSPGEGVLVSWAL------------------
Osi003977.2p    -------------VVLLELLTGRVPVDMKRPPGEGVLVNWAL------------------
TC15181         GHLTTKSDVYSYGVVLLELLTGRVPVDMKRPPGEGVLVNWAL------------------
AT4G02010       GHLLVKSDVYSYGVVLLELLTGRKPVDMSQPSGQENLVTWT------------------
Osi000040.5p    GHLLVKSDVYSYGVVLLELLTGRRPVDMSQPSGQENLVTWLTQMFPDLSTKSLVSHQPLL
```

FIGURA 3 (Cont.)

```
Osi009054.1p      ------------------PLLCHKEGLERLIDPSLNGN-FNFDDVAKVASIASMCVHND
Osi001078.4p      ------------------PLLTDRDGLQQLVDPSMPAASYGFEKLAKAAAIASMCVHVE
AT2G20300         ------------------PLLANREGLEQLVDPALAGT-YNFDDMAKVAAIASMCVHQE
ABE82646.1        ------------------ALLTSREGLEQLVDPSLAGG-YNFDDMAKVAAIASMCVHSE
ABB36644.1        ------------------PLLTTREGLEQLVDPSLAGS-YDFDDMAKVAAIASMCVHPE
At5g56890         ------------------PFLTSAEGLAAIIDQSLGPE-ISFDSIAKVAAIASMCVQPE
Osi001397.3p      ------------------PFLTSRDGLETIIDPSLGNS-ILFDSIAKVAAIASMCVQPE
Osi000142.1p      ------------------PLLTNVVSLRQAVDPLLGPN-VPLDNVAKAAAIASMCVQPE
CDS845            ------------------PRLTNREKISEMVDPTMKGQ-YSQKDLIQVAAIAAVCVQPE
NP_175879.2       ------------------PRLTNREKISEMVDPTMKGQ-YSQKDLIQVAAIAAVCVQPE
Popsp             ------------------PRLTNRDKVVEMVDPAMQDQ-YSKKDLIQVAAIAAVCVQPE
Osi093820.1p      ------------------------------------------------------------
Osi015947.1p      ------------------PSRGSSNGGALETRCVMHGL-D------------------
AT3G58690         ------------------PQLADRDKVVDIMDPTLEGQ-YSTKEVVQVAAIAAMCVQAE
At3g58690sv       ------------------PQLADRDKVVDIMDPTLEGQ-YSTKEVVQVAAIAAMCVQAE
TC196912          ------------------PLLTDREKVVKIMDPSLEGQ-YSMKEVVQVAAIAAMCVQPE
TC81885           ------------------PRLTDREKVVEIMDPALEGQ-YSMKEVIQVAAIAAMCVQPE
Osi003977.2p      ------------------PMLTDREKVVQILDPALEGQ-YSLKDAVQVAAIAAMCVQQE
TC15181           ------------------PMLTDREKVVRILDPALEGQ-YSLKDAVQVAAIAAMCVQPE
AT4G02010         ------------------RPVLRDKDRLEELVDSRLEGK-YPKEDFIRVCTIAAACVAPE
Osi000040.5p      AKLSGVEILICSILTTQARPILRDKDTLEELADPKLGGQ-YPKDDFVRVCTIAAACVSPE

Osi009054.1p      PSQRPFMGEVVQALKLIYNDAEAACDDSYSHRDSSCDQYDDYHGALALDSGSGSWWNRSS
Osi001078.4p      ASHRPFMGEVVQALKLIYNGNNDDTCTSGSFGGGGGEEYEDEEASSPWNNRSWSHDFAAT
AT2G20300         VSHRPFMGEVVQALKLIYNDADETCGDYCSQKDSSVPDSADFKGDLAPSDSSWWN---LT
ABE82646.1        VTQRPFMGEVVQALKLIYNDTDETGGDYCSQKDSSAQE-SDFRGELAPSDSSWWNGGGLT
ABB36644.1        VTQRPFMGEVVQALKLIYNDNDETCADGCSQKESSLPDSDFKG---VPSDSSWWNAGGVT
At5g56890         VSHRPFMGEVVQALKLVSNECDEAKELNSLTSISKDDFRDDTQAESSCGDSSARMARYPL
Osi001397.3p      VDQRPFMGEVVQALKLVCDEGSEFNESRSFSQDLHIQDSGIISRASLDVDVEPVVSAELF
Osi000142.1p      VAHRPSMGEVVQALKLVCSDGDEGLGSGSFSQELAAQAAAIYDVT--GMEAERVLLSEMF
CDS845            ASYRPLMTDVVHSLIPLVKAFNKSTDSSRFPSRRESLSFDDIMP----------------
NP_175879.2       ASYRPLMTDVVHSLIPLVKAFNKSTDSSRFPSRRESLSFDDIMP----------------
Popsp             ADYRPLMTDVVQSLIPLVKNLSSVSSSCSSKFMNQMLCSPRPM----------------
Osi093820.1p      ------------------------------------------------------------
Osi015947.1p      ------------------------------------------------------------
AT3G58690         ADYRPLMADVVQSLVPLVRNRRSASKLSGCSSSFSLARSPNSPGKASIGSQ---------
At3g58690sv       ADYRPLMADVVQSLVPLVRNRRSASKLSGCSSSFSLARSPNSPGKASIGSQ---------
TC196912          ADYRPLMADVVQSLVPLVKTQRSPSKVGSSSSFNSPKLSPGPTF----------------
TC81885           ADYRPLMADVVQSLVPLVKQPRPTVKPGSSSSFHATQSPSPHVTQSPKA-----------
Osi003977.2p      ADYRPLMADVVQSLVPLVKNRSTPKTCNPSVQA----------------------------
TC15181           ADYRPLMADVVQSLVPLVKNRSNQKACNPNVQSSKPLD----------------------
AT4G02010         ASQRPTMGEVVQSLKMVQR-----VVEYQDPVLNTSNKARPNRRQSSATFESEVTSSMFS
Osi000040.5p      ASQRPTMGEVVQSLKMVQRSEFQESIPTPPARPNASAPLRLRTYRERLSLKIFTKGGNSG

Osi009054.1p      NPSG---FFDNRNPLPVITMEYSSGRIEGARDPRFALSTG---------GHAQSPALQNR
Osi001078.4p      PPPASRRLAFPRAPARPTTMDYSSDPADGAAGTSSASARRQRSTSSLVLDKIESLAAYDW
AT2G20300         PRLRYG------QASSFITMDYSSGPLEDMENRPHSASSIPRVGG---------LILPNR
ABE82646.1        PRLTYG------QASSFITMEYSSGPLEDMENRPFSTSSFNGDELS------LPIRHGNR
ABB36644.1        PRLTYG------QASTFMTMDYSSGPLEEFENRPFSASSFNLGGGAG-----LTISHGNR
At5g56890         LPNYDS--------EPDTERGLSTSEMYTGSGR--------------------FERQSN
Osi001397.3p      NASAHYDTLDA-------------------------------------------SGSFRRYSS
Osi000142.1p      GSTPVFTPAAD-------------------------------------------SGSFRKQSS
CDS845            ------------------------------------------------------------
NP_175879.2       ------------------------------------------------------------
Popsp             ------------------------------------------------------------
Osi093820.1p      ------------------------------------------------------------
Osi015947.1p      ------------------------------------------------------------
AT3G58690         ------------------------------------------------------------
At3g58690sv       ------------------------------------------------------------
TC196912          ------------------------------------------------------------
TC81885           ------------------------------------------------------------
Osi003977.2p      ------------------------------------------------------------
TC15181           ------------------------------------------------------------
AT4G02010         SGPYSGLSAFDHENITRTTVFSEDLHEGR--------------------------------
Osi000040.5p      EQDPATMVGFADQTADVSAAAFQTMYRLNVGGAYIPPSNDSGLTRPWYDDTPFVQGPLRG
```

FIGURA 3 (Cont.)

```
Osi009054.1p    SGPIRMKKKLASFYRSRGSFSEHGQLPRH----------------------------
Osi001078.4p    SGPLRASRGRN-FYRLRGSMSEHGGHPSEDCSMEGYWM-------------------
AT2G20300       SGPLRPMRSRRNFFRLRGSMSEHGGPSSSRHLWSGNGDWL-----------------
ABE82646.1      SGPLRTTRSKLSLYRFSGSRSEHGEVSSKRNWI-------------------------
ABB36644.1      SGPLRTVRSKPALYRLRGSMSEHGALLPRHDWRDGTNYDASF----------------
At5g56890       SGPLTSGRGKSFWQKMRRLSTGSLSEHGTPTVMLRSGSR-------------------
Osi001397.3p    SGPLRVGRTG--------HNRGSSSEHCGTQRFRIDSE--------------------
Osi000142.1p    SGPLMTGKNRKFWQRLRSLSRGSMSEHGASPDFETRSQCSNR----------------
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       ------------------------------------------------------------
Osi000040.5p    LVYNAGPHFHIKYPSDAAEYAAPPEVYLGGRSMGRDQRLNQNSNLTWSLHVECNFTYVVR

Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------------------------------
AT2G20300       ------------------------------------------------------------
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       ------------------------------------------------------------
Osi001397.3p    ------------------------------------------------------------
Osi000142.1p    ------------------------------------------------------------
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       ------------------------------------------------------------
Osi000040.5p    LHFCELQLIHGNQRVFDIYINNRTAQTDVDVLEMATERGVPVYKDYAVRLSNDTADEHLW

Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------------------------------
AT2G20300       ------------------------------------------------------------
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       ------------------------------------------------------------
Osi001397.3p    ------------------------------------------------------------
Osi000142.1p    ------------------------------------------------------------
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       ------------------------------------------------------------
Osi000040.5p    VAVHPSVMLRPQFYDAILNGLEVFKVNNTGGSLASPDPVPYKLLAEKELGWGGPPEFSTD
```

FIGURA 3 (Cont.)

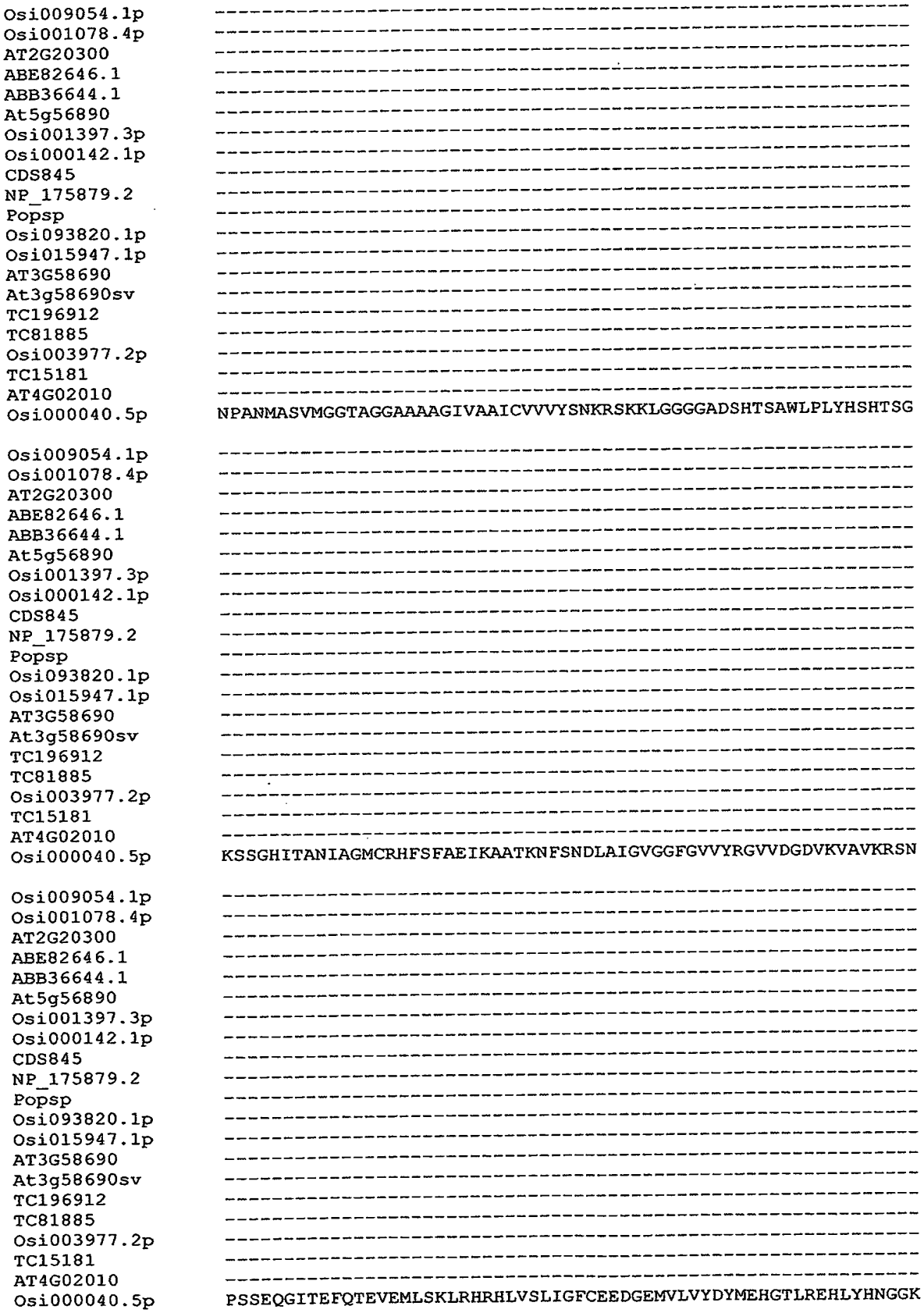


FIGURA 3 (Cont.)

```
Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------------------------------
AT2G20300       ------------------------------------------------------------
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       ------------------------------------------------------------
Osi001397.3p    ------------------------------------------------------------
Osi000142.1p    ------------------------------------------------------------
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       ------------------------------------------------------------
Osi000040.5p    PTLSWRHRLDICIGAARGLHYLHTGESHVSTVVKGSFGYLDPEYYRRQQLTDKSDVYSFG

Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------------------------------
AT2G20300       ------------------------------------------------------------
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       ------------------------------------------------------------
Osi001397.3p    ------------------------------------------------------------
Osi000142.1p    ------------------------------------------------------------
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       ------------------------------------------------------------
Osi000040.5p    VVLFEVLMARPALDPALPRDQVSLADYALACKRGGALPDVVDPAIRDQIAPECLAKFADT

Osi009054.1p    ------------------------------------------------------------
Osi001078.4p    ------------------------------------------------------------
AT2G20300       ------------------------------------------------------------
ABE82646.1      ------------------------------------------------------------
ABB36644.1      ------------------------------------------------------------
At5g56890       ------------------------------------------------------------
Osi001397.3p    ------------------------------------------------------------
Osi000142.1p    ------------------------------------------------------------
CDS845          ------------------------------------------------------------
NP_175879.2     ------------------------------------------------------------
Popsp           ------------------------------------------------------------
Osi093820.1p    ------------------------------------------------------------
Osi015947.1p    ------------------------------------------------------------
AT3G58690       ------------------------------------------------------------
At3g58690sv     ------------------------------------------------------------
TC196912        ------------------------------------------------------------
TC81885         ------------------------------------------------------------
Osi003977.2p    ------------------------------------------------------------
TC15181         ------------------------------------------------------------
AT4G02010       ------------------------------------------------------------
Osi000040.5p    AEKCLSENGTERPTMGDVLWNLESAMHFQDAFDAAAGRPVPALDAAAGSSSHLDDGSTAS
```

FIGURA 3 (Cont.)

```
Osi009054.1p      ----------------------------------------
Osi001078.4p      ----------------------------------------
AT2G20300         ----------------------------------------
ABE82646.1        ----------------------------------------
ABB36644.1        ----------------------------------------
At5g56890         ----------------------------------------
Osi001397.3p      ----------------------------------------
Osi000142.1p      ----------------------------------------
CDS845            ----------------------------------------
NP_175879.2       ----------------------------------------
Popsp             ----------------------------------------
Osi093820.1p      ----------------------------------------
Osi015947.1p      ----------------------------------------
AT3G58690         ----------------------------------------
At3g58690sv       ----------------------------------------
TC196912          ----------------------------------------
TC81885           ----------------------------------------
Osi003977.2p      ----------------------------------------
TC15181           ----------------------------------------
AT4G02010         ----------------------------------------
Osi000040.5p      INTLATSSTSHPHEPCVDVVLEPDDVVAERATFSQLVQPTGR
```

FIGURA 3 (Cont.)

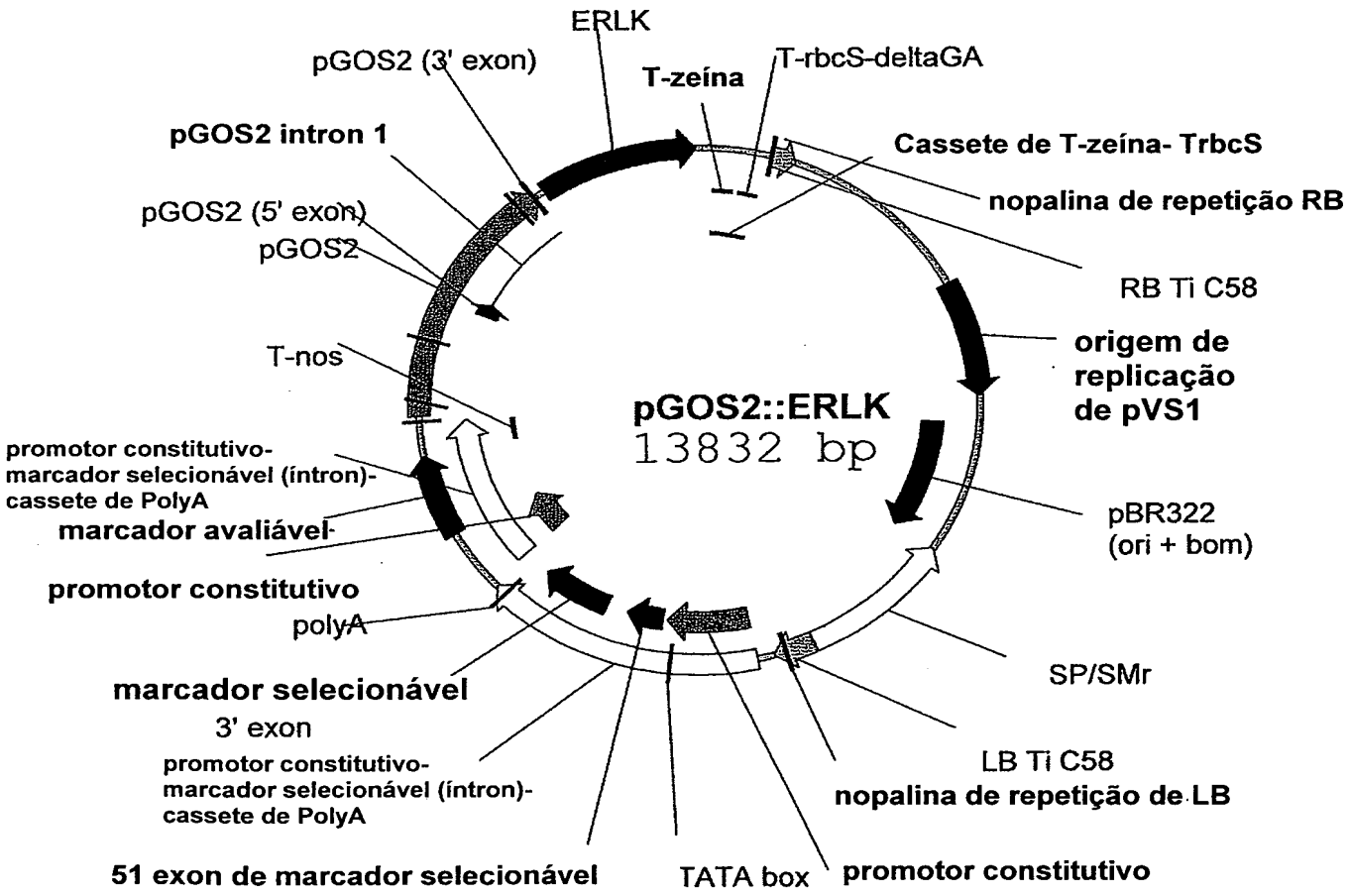


FIGURA 4

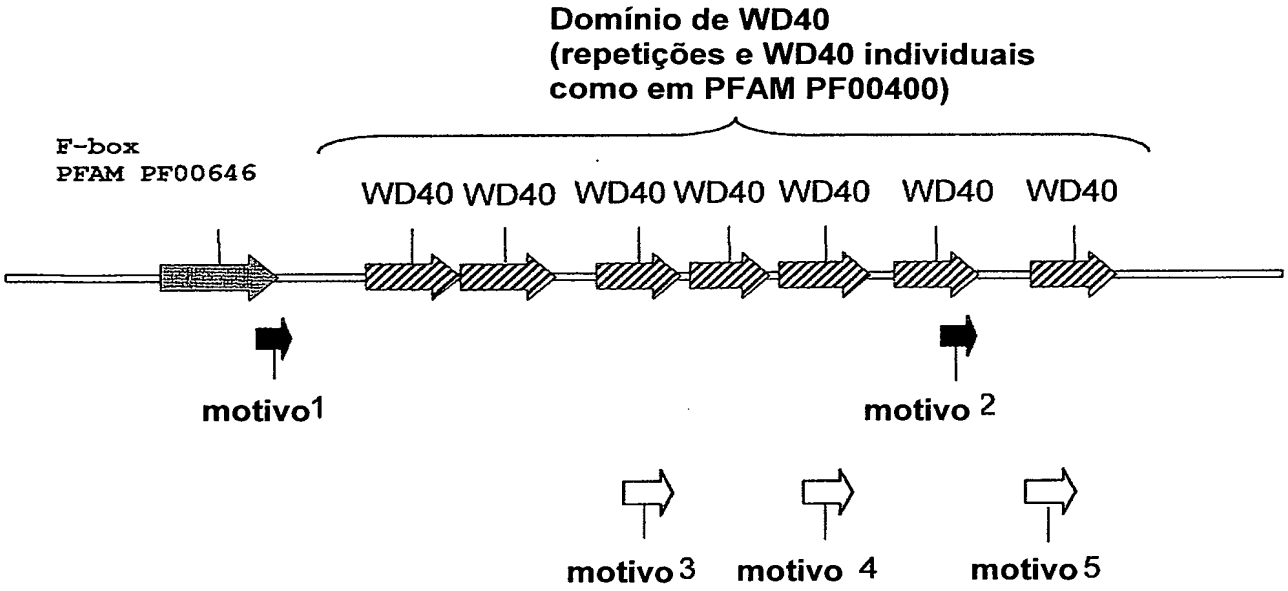


FIGURA 5

CLUSTAL W (1.83) alinhamento de seqüência múltiplo

```
Arath_FBXW      MEFECQERLEAAN----------------------DQRSESGLLNTDLRIGNDGSVEIPNV
Poptr_FBXW      MAFECQKSTEVS-------------------------KSLAICVEKSNSNTEHLEISKS
Medtr_FBXW      MAFDCPQSIKVSQNLVENPGISIDIVELNPKPNSKAISISFPDFVTNTKSESGKGEIFKG
Orysa_FBXW      MDFDCKTARGDSSS-----------VNRSCIVTEGTVVQAKPVSHNGKAKHWNSLSTLN
Zeama_FBXW      MAFDCNKERGVSSP-----------DSCSRICTEGTLVQANPLSHYWKPKRLENFNKLE
Triae_FBXW      MDFDCN-KAGESSA-----------KHCSSICNEGTLIQANTLIHCGKAKKWNSLNKFN
                * *:*      :                                .          :

                                                                F-box

Arath_FBXW      KLCCQKKKGTLVPSG---SKQLLSDKDL---STTIIDLPQALISEILNCLDPKELGLVSC
Poptr_FBXW      PLDCYPKKGILSSLS---SKQLSCNDVLLNCRRSITDLPPALISEILSCLDPQELGIVSC
Medtr_FBXW      KSKLNSKKGIKAASAGALNQSSVPGDVVIGNCRSITDLPPVLISEILNCLDPKELGIVSC
Orysa_FBXW      NQKCSYELLS---DPKKNVETSDGETASKCDSWCFTDLPSALVCEVLEHLDPKELGIVSC
Zeama_FBXW      NQKSSYESAPSGSDPKQDAEASDEATASLCGFRCFTDLPASLVCEVLARLDAKDLGIVSC
Triae_FBXW      NPESSHGSLPRVNDPKEDGETGNDATASECSVMCFTDLPSALVCEVLARLDPKGLGVVSC
                                :                : *** *:.*:* **.: **:***

                F-box (cont'd)  motivo 1

Arath_FBXW      VSTYLHRLASEHHAWKEFYRERWGLPVVFGAASSG-------LSDERSWKDLFVEREFRS
Poptr_FBXW      VSRVLNSLATENSVWKEVYGERWGLPIVPAPLGAG-------VSNEKSWKELFVEREYRS
Medtr_FBXW      VSLILRSLASEHHAWKEFYCERWGLPAVPEAVLDSDSGVGDSVKDEKSWKDIFVERDYRS
Orysa_FBXW      VSTLLHTLATDHQGWKKFYCERWGIPTPPVTLNGPLVPG--GTSDWKSWKTLFVEREFRS
Zeama_FBXW      VSTLLHALATDHQGWKKLYCERWGLPDLPTTLNGPLLPG--VPLDGKFWKTFFVEREFRS
Triae_FBXW      VSTVLQTLATDHQGWKKFYCERWGLPNAP---IGPLVPG--GTPDGRSWKALFVDREFQS
                **  *. **::: **:.* ****:*                  : : ** :**:*:::*
```

FIGURA 6

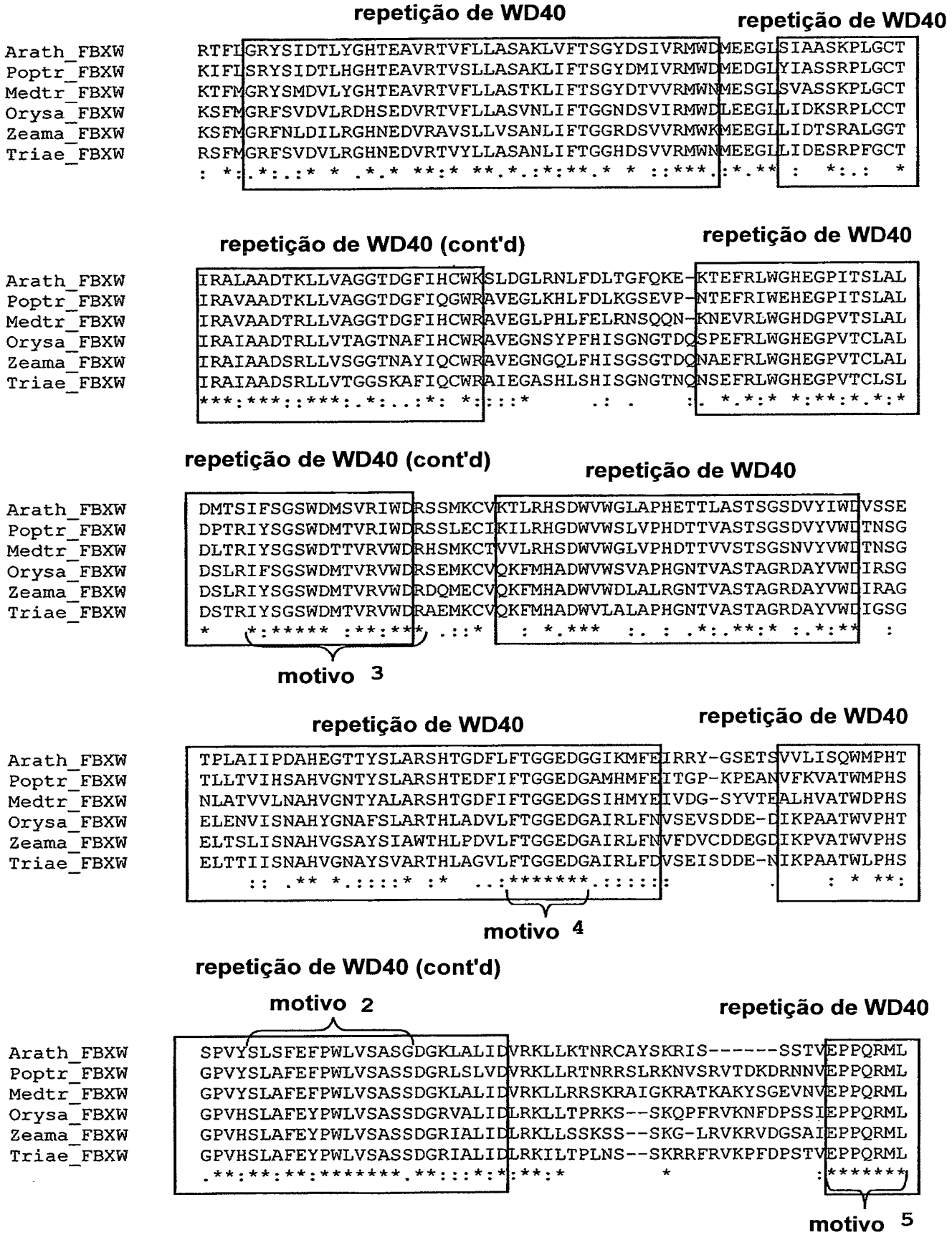


FIGURA 6 (Cont.)

repetição de WD40 (cont'd)

```
Arath_FBXW      HGFGSNLFSVDVGYDRIVCGGEEGTVRIWNFTQALEIERRTRALKGMRHENRMRRRRMQM
Poptr_FBXW      HGFGPNLFSVDVGADRIVCGGEEGVVRIWNFSKALERERTARAMRGIRLENRMRRRRLQI
Medtr_FBXW      HGFKSNLFSVGIGADRIVCGGEEGVVRIWNFTEALEIESRVRALRGMRLENRMRRRKNQT
Orysa_FBXW      HGFGCDLFSVAIGADRIVCGGEDGAVKVWNFSEALEIEKRAQALRSMRQENRMRRKKAQV
Zeama_FBXW      HGVGCDLFSIAIGADRIVCAGEDGSVRVWNFSKALEIERRAQALRSLRQENRIRRRKSQA
Triae_FBXW      HGFGCYLFSVGIGADRIICGGEDGSVRVWNFSEALEIEKKAQALRSLRQENRMRRRKAQV
                **.   ***: :* ***:*.**:* *::***::*** *  .:*::.:* ***:**:: *

Arath_FBXW      EMNAKNGRPDQC-SIAAHKNPINGERNRAWHSKRRASGKAKA
Poptr_FBXW      EMSSKGGRTDQC-SVAAKKNPMHVDKSGVWRNKHGMSGKLKA
Medtr_FBXW      ELNSKGGKSDQC-SAAAKKS----SVTCIWPSKRGMGGKTKA
Orysa_FBXW      EMNANGRRSDQCGSIAMKRNQLKGDKSVTWHSKRAINDKVKS
Zeama_FBXW      EMNTNTRRPDQC-SIAMKKNQLKGDKSATWHNKRAINEKAKS
Triae_FBXW      EMNANGRRVDHC-SVAMKRNPLKGDKSVTWQIKRPISDKVKS
                *:.::  : *:*.  * ::.    . .  *  *:  . * *:
```

FIGURA 6 (Cont.)

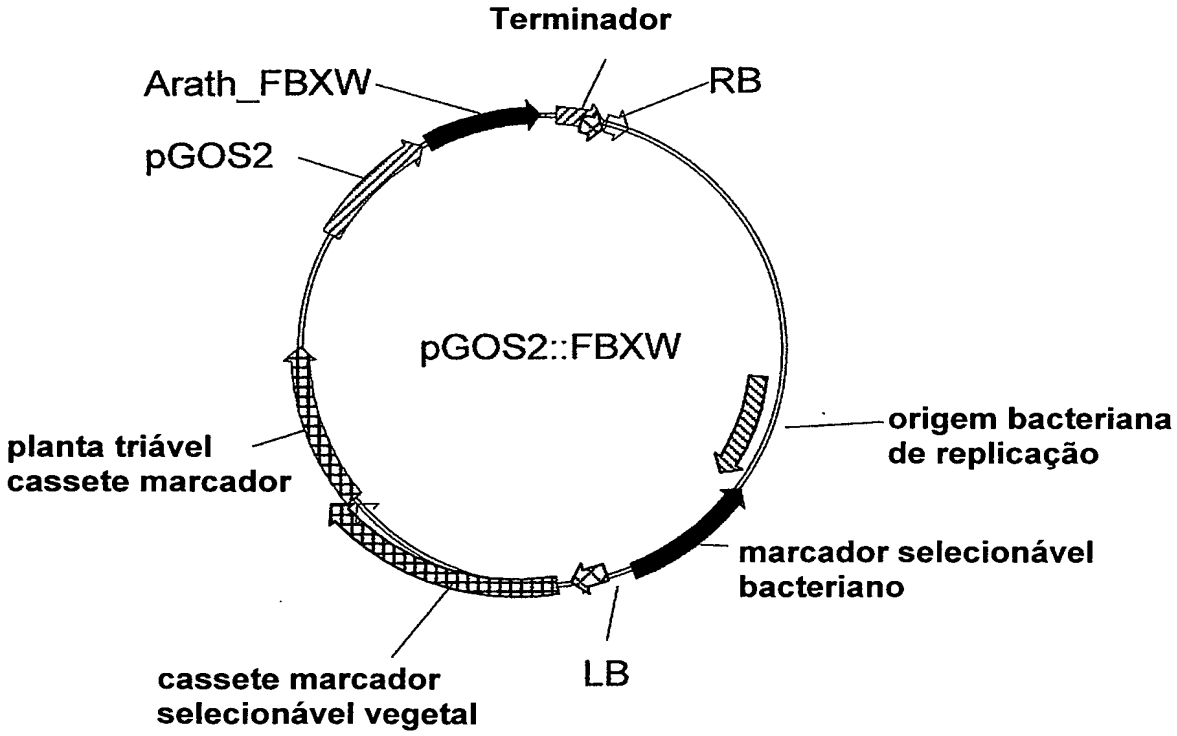
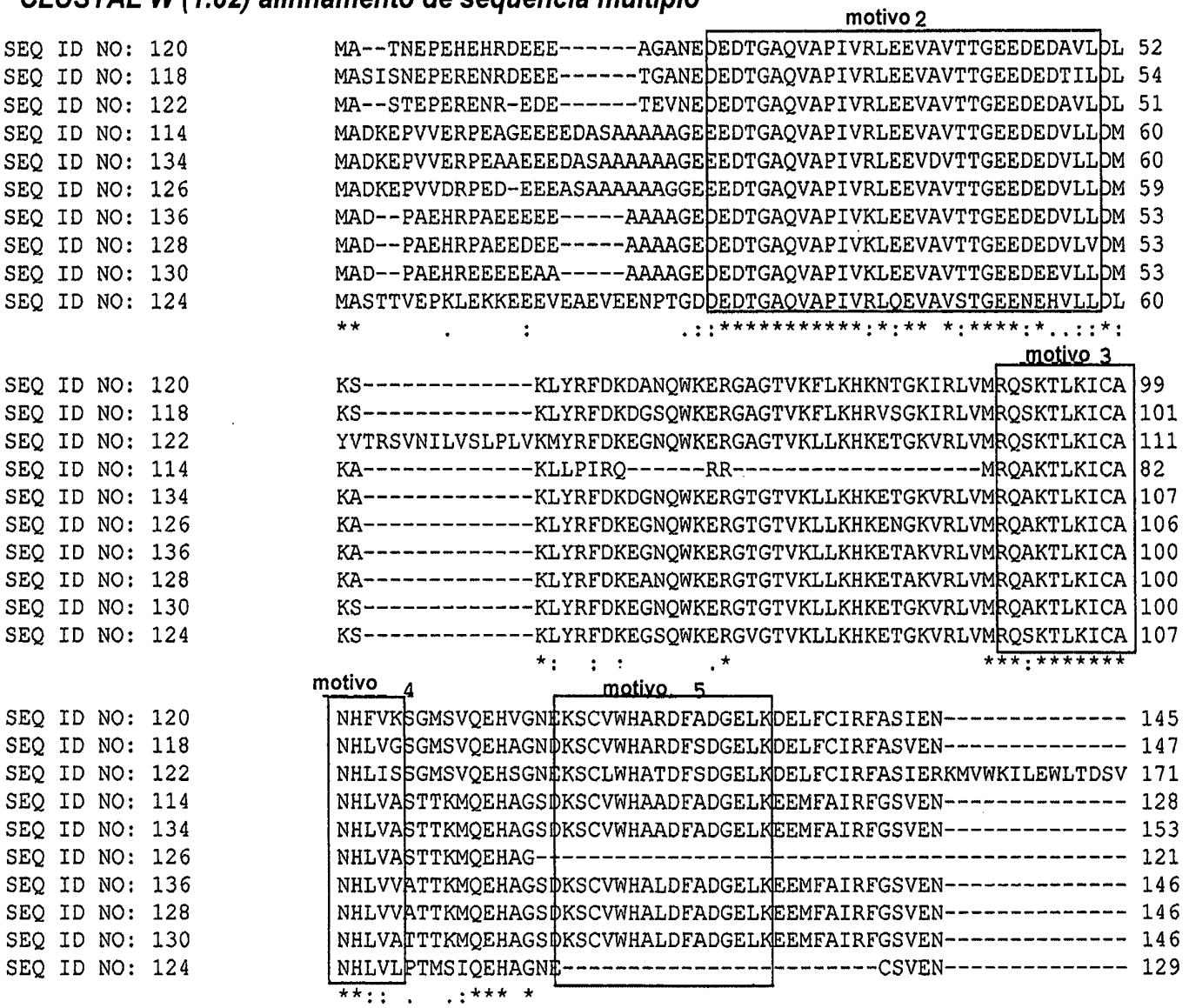


FIGURA 7

FIGURA 8

CLUSTAL W (1.82) alinhamento de seqüência múltiplo



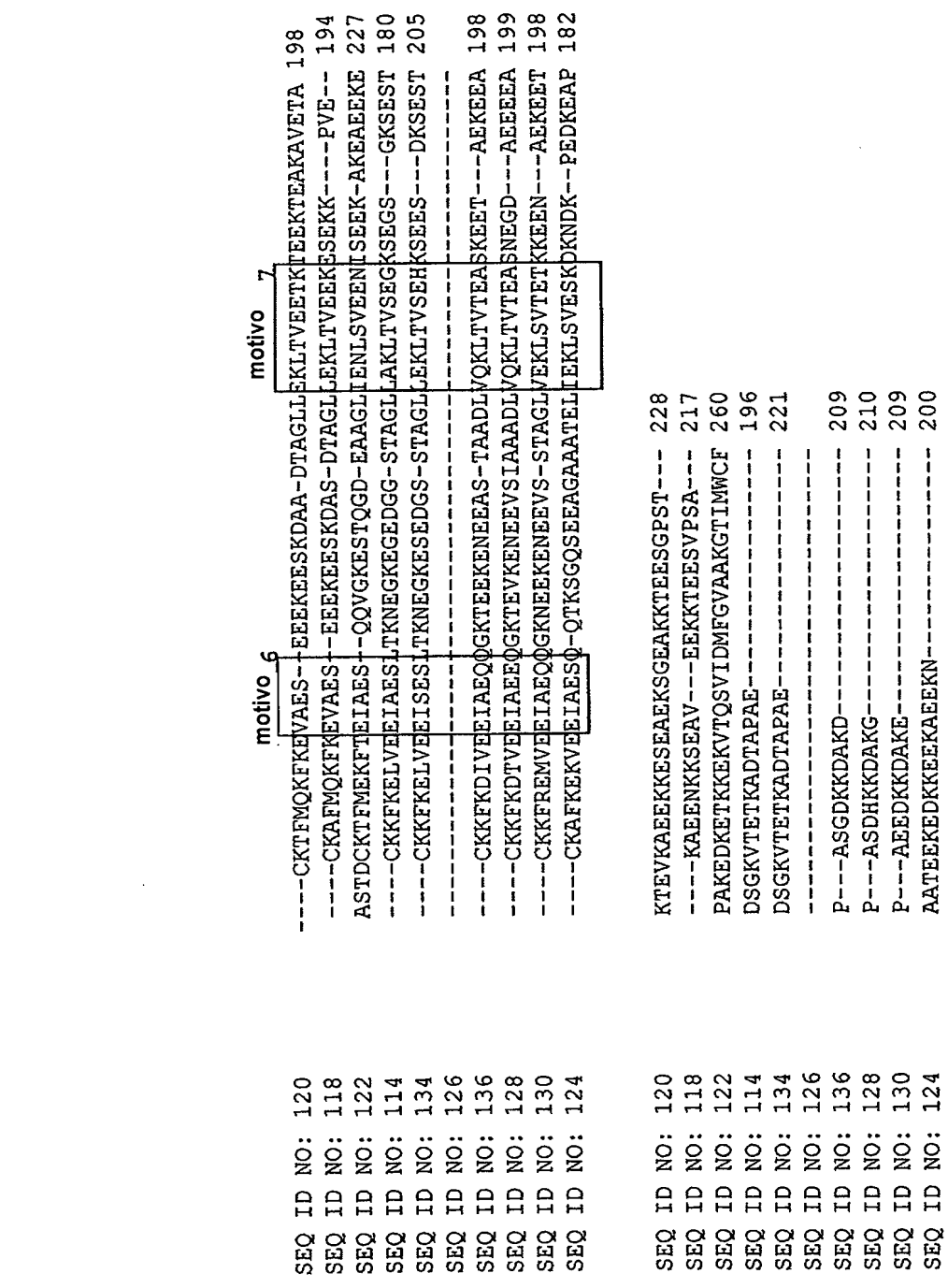


FIGURA 8 (Cont.)

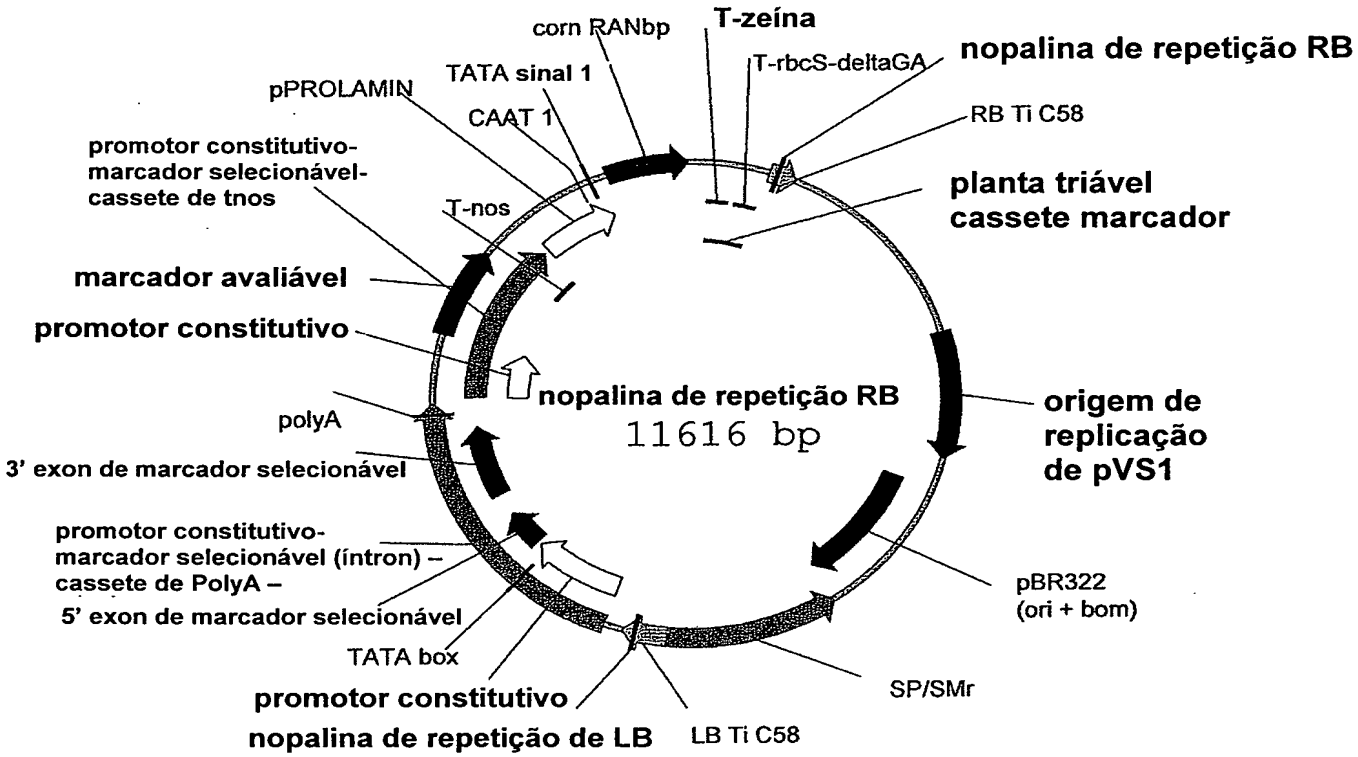

corn RANbp
T-zeína
nopalina de repetição RB
TATA sinal 1
T-rbcS-deltaGA
pPROLAMIN
CAAT 1
RB Ti C58
promotor constitutivo-
marcador selecionável-
cassete de tnos
T-nos
planta triável
cassete marcador
marcador avaliável
promotor constitutivo
nopalina de repetição RB
11616 bp
origem de
replicação
de pVS1
polyA
3' exon de marcador selecionável
promotor constitutivo-
marcador selecionável (íntron) –
cassete de PolyA –
5' exon de marcador selecionável
pBR322
(ori + bom)
TATA box
SP/SMr
promotor constitutivo
nopalina de repetição de LB
LB Ti C58

FIGURA 9

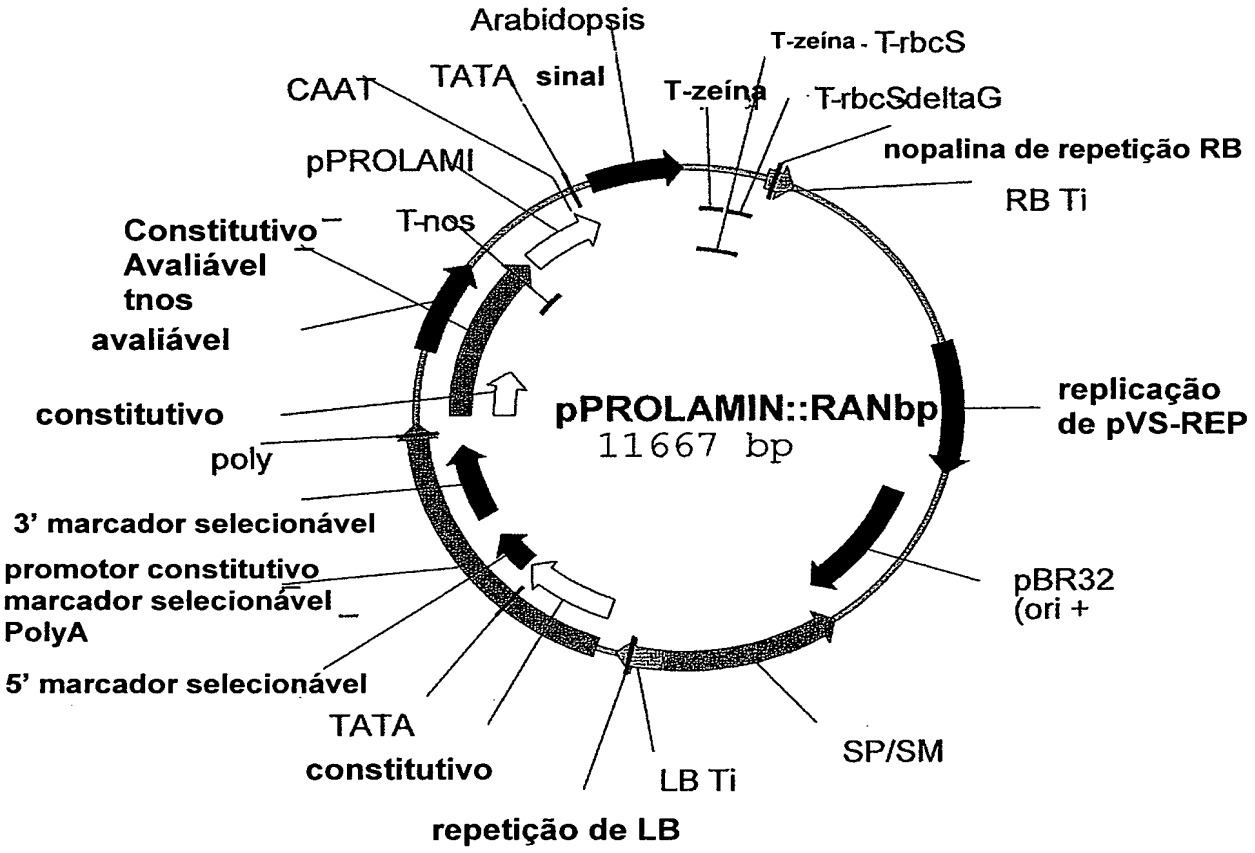

Arabidopsis
T-zeína - T-rbcS
CAAT
TATA
sinal
T-zeína
T-rbcSdeltaG
nopalina de repetição RB
pPROLAMI
RB Ti
T-nos
Constitutivo
Avaliável
tnos
avaliável
constitutivo
pPROLAMIN::RANbp
11667 bp
replicação
de pVS-REP
poly
3' marcador selecionável
promotor constitutivo
marcador selecionável
PolyA
pBR32
(ori +
5' marcador selecionável
TATA
constitutivo
SP/SM
LB Ti
repetição de LB

FIGURA 10

20/124

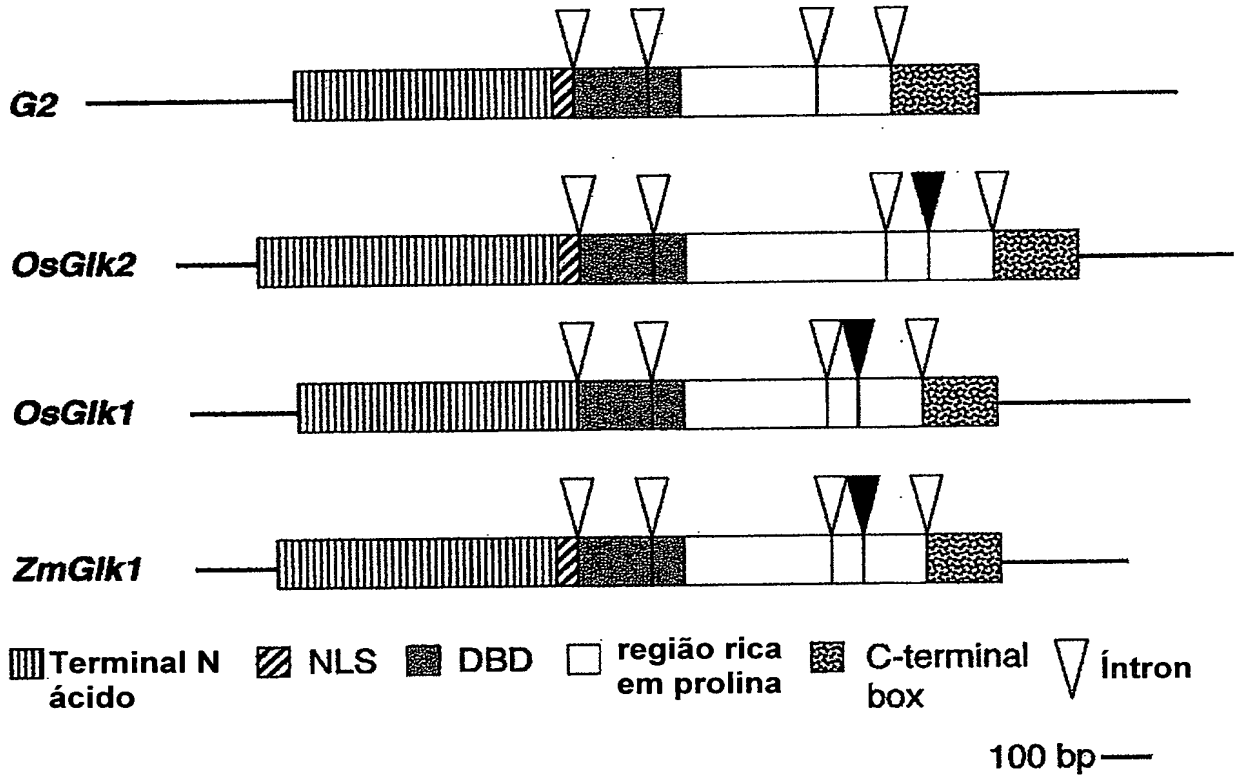


FIGURA 11

MLAVSPAMCPDIEDRAAVAGDAGMEVVGMSSDDMDQFDFSVDDIDFGDFFLRLEDGDVLP
DLEVDPAEIFTDFEAIATSGGEGVQDQEVPTVELLAPADDVGVLDPCGDVVVGKENAAFA
GAGEEKGGCNQDDDAGEANADDGAAAVEAKSSSPSSSTSSSQEAESRHKSSSKSSHGKK**K**
AKVDWTPELHRRFVQAVEQLGIDKAVPSRILEIMGIDSLTRHNIASHLQKYRSHRKHMIA
REAEAASWTQRRQIYAAGGGAVAKRPESNAWTVPTIGFPPPPPPPPSPAPMQHFARPLHV
WGHPTMDPSRVPVWPPRHLVPRGPAPPWVPPPPPSDPAFWHHPYMRGPAHVPTQGTPCMA
MPMPAARFPAPPVPGVVPCPMYRPLTPPALTSKNQQDAQLQLQVQPSSESIDAAIGDVLS
KPWLPLPLGLKPPSVDSVMGELQRQGVANVPPACG

FIGURA 12

CLUSTAL W (1.83) alinhamento de seqüência múltiplo

```
ZmG2          ---------------------------------------------MLEVSTLRGPTSSG
SoGLK2        ------------------------------------------------------------
ZmGLK1        ----------------------------------MLAVSPSPVRCADAEECGGGGAS
SbGLK1        ------------------------------------------------------------
OsGLK1        ---------------------------------------MLAVSPAMCPDIEDRAAVA
OsGLK1var     ---------------------------------------MLAVSPAMCPDIEDRAAVA
TaGLK1        ------------------------------------------------------------
OsGLK2        --------------MLEVSTLRSPKADQRAGVGGHHVVGFVPAPPSPADVADEVDAFIV
HvGLK2        ------------------------------------------------------------
AtGLK2        ------------------------------------------------------------
AtGLK1        ------------------------------------------------------------
PpGLK2        MAMDMARIDESTAVEVN---SLSLVHCVLDGLPDS-PCLKSSPTSFEEAVAEGRSVFGDE
PpGlk1        MAMDMARIEVEPLVEVHNNTNNLLVHCVLDALPDSSPCLKSSETSFEAVVVKEDEEGG--
AcGLK1        ------------------------MLTISPLESSNPDAKDEGDFVLEDISFDDLFVG--

ZmG2          SKAEQHCGGGGGFVGDHHVVFPTSGDCFAMVDDNLLDYIDFSCDVPFFDADGDILPDLEV
SoGLK2        ------------------------------------------------------------
ZmGLK1        KEMEETAVGPVSDSDLDFDFTVDDIDFGDFFLRLDDGDDALPGLEVDPAEIVFADFEAIA
SbGLK1        ------------------------------------------------------------
OsGLK1        GDAGMEVVGMSSDDMDQFDFSVDDIDFGDFFLRLEDGDVLPDLEVDPAEIFTDFEAIATS
OsGLK1v       GDAGMEVVGMSSDDMDQFDFSVDDIDFGDFFLRLEDGDVLPDLEVDPAEIFTDFEAIATS
TaGLK1        ------------------------------------------------------------
OsGLK2        DDSCLLEYIDFSCCDVPFFHADDGDILPDLEVDPTELLAEFASSPDDEPPPTTSAPGPGE
HvGLK2        ------------------------------------------------------------
AtGLK2        --------------MLTVSPAPVLIGNNSKDTYMAADFADFTTEDLPDFTTVGDFSDDLL
AtGLK1        --------------MLALSPATRDGCDGASEFLDTSCGFTIINPEEEEEFPDFADHGDLL
PpGLK2        EDIINNSNDQDNSSSCGAVVTTHEDFAECLNFVTEAECGDVGVRCFEDFDKLPDCGDEGE
PpGlk1        RPLFGKPELHPASPSSDTAAAANGEFARCLNFVTEADCGDVGVQCFEDFGTLPDCGEAGI
AcGLK1        ---------------------------------------------FDELPDLEIDP-

ZmG2          DTTELLAEFSSTPPADDLLAVAVFGADDQPAAAVAQEKPSSSLEQTCGDDKGVAVAAARR
SoGLK2        ------------------------------------------------------------
ZmGLK1        TAGGDGGVTDQEVPSVLPFADAAHIGAVDPCCGVLGEDNDAACADVEEGKGECDHADEVA
SbGLK1        ------------------------------------------------------------
OsGLK1        GGEGVQDQEVPTVELLAPADDVGVLDPCG---------DVVVGKENAAFAGAGEEKGGCN
OsGLK1var     GGEGVQDQEVPTVELLAPADDVGVLDPCG---------DVVVGEENAAFAGAGEEKVGCN
TaGLK1        ------------------------------------------------------------
OsGLK2        PAAAAGAKEDVKEDGAAAAAAAAAADYDGSPPPPRGKKKKDDEERSSSLPEEKDAKNGGG
HvGLK2        ------------------------------------------------------SLSIVG
AtGLK2        DGIDYYDDLFIGFDGDDVLPDLEIDSEILG---------EYSGSGRDEEQEMEGNTSTAS
AtGLK1        DIIDFDDIFGVAG---DVLPDLEIDPEILSGDFSNHMNASSTITTTSDKTDSQGETTKGS
PpGLK2        TSKAEEEGCVRIGGGGEQGELLESVSLDCSRNSENLELRDLGELWEGSERPDSVPGNEVG
PpGlk1        SREEGG------GRDGEQVELLESMSLDGSRNSENLELGELGELLQGSETLDSVPGNEVG
AcGLK1        ------------------CDIFADFSFNSSEEGDGNNKGSTSEENDVQHKRDGYYKEHSG
```

FIGURA 13

```
ZmG2          KLQTTTTTTTTEEEDSSPAGSGANKSSASAEGHSSKKKSAGKNSNGGKRKVKVDWTPELH
SoGLK2        ------------------------------------------------------------
ZmGLK1        AAGNNNSDSGEAGCGGAFAGEKSPSSTASSSQEAESRRKVSKKHSQGKKKAKVDWTPELH
SbGLK1        ------------------------------------------------------------
OsGLK1        QDDDAGEANADDGAAAVEAKSSSPSSSTSSSQEAESRHKSSSKSSHGKKKAKVDWTPELH
OsGLK1var     QDDDAGEANVDDGAAAVEAKSSSPSSTTSSSQEAESRHKSSSKSSHGKKKAKVDWTPELH
TaGLK1        ------------------------------------------------------------
OsGLK2        DEVLSAVTTEDSSAGAAKSCSPSAEGHSKRKPSSSSSSAAAGKNSHGKRKVKVDWTPELH
HvGLK2        DEVCSAVTTDDSSAAVGSENSKSSASAEG--HSKRTSAAAATKSSHGRRKVKVDWTPELH
AtGLK2        ETSERDVGVCKQEGGGGGDGGFRDKTVRRGKRKGKKSKDCLSDENDIKKKPKVDWTPELH
AtGLK1        SGKGEEVVSKRDDVAAETVTYDGDSDRKRKYSSSASSKNNRISNNEGKRKVKVDWTPELH
PpGLK2        EEEALLLAEAAKATGDVVSASDSGECSSVDRKDNQASPKSSKNAAPGKKKAKVDWTPELH
PpGlk1        EEEALLLAKAAKATGVVVSASDSGECSSVDRKENQQSPKSCKSAAPGKKKAKVDWTPELH
AcGLK1        KEETEVVSARTREDEKKPPSSKSENVKSLKS--------SGKKSSQSKKKAKVDWTPELH

ZmG2          RRFVQAVEQLGIDKAVPSRILEIMG-TDCLTRHNIASHLQKYRSHRKHLMAREAEAATWA
SoGLK2        -----------IDKAVPSRILEIMG-MDCLTRHNIASHLQKYRSHRKHLMAREAEAATWA
ZmGLK1        RRFVQAVEELGIDKAVPSRILEIMG-IDSLTRHNIASHLQKYRSHRKHMLAREVEAATWT
SbGLK1        ------------------------------------------------------------
OsGLK1        RRFVQAVEQLGIDKAVPSRILEIMG-IDSLTRHNIASHLQKYRSHRKHMIAREAEAASWT
OsGLK1var     RRFVQAVEQLGIDKAVPSRILEIMG-IDSLTRHNIASHLQKYRSHRKHMIAREAEAASWT
TaGLK1        ----------------------MG-INSLTRHNIASHLQKYRSHRKHMIAREAEAASWT
OsGLK2        RRFVQAVEQLGIDKAVPSRILELMG-IECLTRHNIASHLQKYRSHRKHLMAREAEAASWT
HvGLK2        RRFVQAGEQLGLDKAVPSRILELMGNEYRLTRHNIASHLQKYRSHRKHLMAREGEAGSWT
AtGLK2        RKFVQAVEQLGVDKAVPSRILEIMN-VKSLTRHNVASHLQKYRSHRKHLLAREAEAASWN
AtGLK1        RRFVEAVEQLGVDKAVPSRILELMG-VHCLTRHNVASHLQKYRSHRKHLLAREAEAANWT
PpGLK2        RRFVHAVEQLGVEKAYPSRILELMG-VQCLTRHNIASHLQKYRSHRRHLAAREAEAASWT
PpGlk1        RRFVHAVEQLGVEKAFPSRILELMG-VQCLTRHNIASHLQKYRSHRRHLAAREAEAASWT
AcGLK1        RRFVQAVEQLGVDKAVPSRILELMG-IDCLTRHNIASHLQKYRSHRKHLLAREAEAASWS

ZmG2          QKRHMYAPPAPRTTTTTDAARPPWVVPTTIGFP----------------------PPR---
SoGLK2        QKRHMYAAAGGVAPRTDAPHGSRPWVVPTIGFP----------------------PPAPPP
ZmGLK1        THRRPMYAAP-SGAVKRPDSN--AWTVPTIGFP----------------PPAGTPPRPVQH
SbGLK1        ------------------ARGL--KFPPPPPPPP----------------PPPPPSPHPMQH
OsGLK1        QRRQIYAAGG-GAVAKRPESN--AWTVPTIGFP--------------PPPPPPPSPAPMQH
OsGLK1var     QRRQIYAAGG-GAVAKRPESN--AWTVPTIGFP--------------PPPPPPPSPAPIQH
TaGLK1        QRRQMYAAGGPAAAVKRQDSN--MWTVPTIGFAPAHP----------PPPPPPPSPAAMQH
OsGLK2        QKRQMYTAAAAAAAVAAGGGPRKDAAAATAAVAPWVMPTIGFPPPHAAAMVPPPPHPPPF
HvGLK2        PQPQMCWGAKVVRVGRALTL-------------------------------------------
AtGLK2        LRRHATVAVPGVGGGGKKP----WTAPALGYPP--------------------HVAPMHHG
AtGLK1        RKRHIYGVDTGANLNGRTKNGWLAPAPTLGFPPP--------------PPVAVAPPPVHHH
PpGLK2        HRRTYTQAPWPRSSRRDGLPYLVPIHTPHIQPRPSMA-----------MAMQPQLQTPHHP
PpGlk1        HRRAYTQMPWSRSSRRDGLPYLVPLHTPHIQPRPSMV-----------MAMQPQLQTQHTP
AcGLK1        HRRQLYVSSTNNGKRR-----------------------------------------------
```

FIGURA 13 (Cont.)

23/124

```
ZmG2          FCRPLHVWGHPP----PHAAAAEAAAATPMLPVWP-RHLAPPRHLAPWAHPTP--VDPAF
SoGLK2        FCRPLHVWGHPPHAAAAEAAAAVAATPTPMLPVWP-RHLAPPRPLPPWAHPHPPAVDPAF
ZmGLK1        FGRPLHVWGHPS---------PTPAVESPRVPMWP-RHLAPRAPPPPPWAPPPPADPASF
SbGLK1        FGRPLHAWGHP----------TPTVESPRVPMWP-RHLVPRTPPPP--WAPPPPSDPAF
OsGLK1        FARPLHVWGHP----------TMDPSRVPVWPPRHLVPRGPAPPWVPPPPP-SDPAF
OsGLK1var     FARPLHVWGHP----------TMDPSRVPVWPPRHLVPRGPAPPWVPPPPP-SDPAF
TaGLK1        YARPLHVWGHP----------TMDSPRMPMWP-RHTISRAPMPAWAPPPPPPSDPAF
OsGLK2        CRPPLHVWGHPTAG--------VEPTTAAAPPPPSPHAQPPLLPVWPRHLAPPPPPLPAA
HvGLK2        ------------------------------------------------------------
AtGLK2        HFRPLHVWGHP----------TWPKHKPNTPASAHRTYPMPAIAAAPASWPGHPPY
AtGLK1        HFRPLHVWGHP----------TVDQSIMPHVWPKHLPPPS-TAMPNPPFWVSDSPY
PpGLK2        ISTPLKVWGYPTVDHSNVHMWQQPAVATPSYWQAADGSYWQHPATGYDAFSARACYSHPM
PpGlk1        VSTPLKVWGYPTVDHSSVHMWQQPAVATPSYWQAPDGSYWQHPATNYDAYSARACYPHPM
AcGLK1        ----------------------------------------------------RKNNEHET

ZmG2          WHQ----------QYSAARKWGPQ---------AAAVTQGTPCVPLPRFPVPHPIYSRP
SoGLK2        WHQ----------QYNAARKWGPQ---------ATAVTQGTPACVPXXAAXMLPRFSRV
ZmGLK1        WHH----------AYMRGPAAHMPDQVAVTPCVAVPMAAARFPAPHVRGSLPWPPPMYR
SbGLK1        WHH----------AYMRGPAHMP---GQVTPCVAVPMPAARFPAPPVRGALPCPPPMYR
OsGLK1        WHH----------PYMRG-PAHVP--TQGTPCMAMPMPAARFPAPPVPGVVPCPMYR--
OsGLK1var     WHH----------PYMRG-PAHVP--TQGTPCMAMPMPAARFPAPPVPGVVPCPMYR--
TaGLK1        WHH----------PYMRGPAAYMP--THGTPCMAMPMTPK-FPAAPVPVAMPCPVYAS-
OsGLK2        WAHGHQPAPVDPAAY----------------WQQQYNLQRFPVPPVPGMVPHPMYRP-
HvGLK2        ------------------------------------------------------------
AtGLK2        WHQ--------------------------QPLYPQGYG---------------MASS
AtGLK1        WHP--------------------------MHNGTTPYLPTVATRFRAPPVAGIPHALP
PpGLK2        QRVPVTTTHAG--------------------LPIVAPGFPDESCYYGDDMLAGSMYLCNQ
PpGlk1        R-VSLGTTHAG--------------------SPMMAPGFPDE-SYYGEDVLAATMYLCNQ
AcGLK1        ------------------------------------------------------------

ZmG2          AMVPPPPSTTKLAQLHLELQAHPSKESIDAAIGDVLVKPWLPLPLGLKPPSLDSVMSELH
SoGLK2        GPGGGGGKKIF--------------------------------------------------
ZmGLK1        PLVPPALAGKSQQDALFQLQIQPSSESIDAAIGDVLTKPWLPLPLGLKPPSVDSVMGELQ
SbGLK1        PLVPPTLDT------QFQLQTQPSSESIDAAIGDVLTKPWLPLPLGLKPPSVDSVMGELQ
OsGLK1        PLTPPALTSKNQQDAQLQLQVQPSSESIDAAIGDVLSKPWLPLPLGLKPPSVDSVMGELQ
OsGLK1var     PLTPPALASKNQQDAQLQLQVQPSSESIDAAIGDVLSKPWLPLPLGLKPPSVDSVMGELQ
TaGLK1        PSPAPALASKSQQDSQLQLQAQPSNESIDAAIGDVLSKPWLPLPLGLKPPSLGSVMGELE
OsGLK2        IPPPSPPQGNKLAALQLQLDAHPSKESIDAAIGDVLVKPWLPLPLGLKPPSLDSVMSELH
HvGLK2        ------------------------------------------------------------
AtGLK2        NHSSIGVPTRQLGPTNPPIDIHPSNESIDAAIGDVISKPWLPLPLGLKPPSVDGVMTELQ
AtGLK1        PHHTMYKPNLGFGGARPPVDLHPSKESVDAAIGDVLTRPWLPLPLGLNPPAVDGVMTELH
PpGLK2        SYDSEIGRAAGVAACSKPIETHLSKEVLDAAIGEALANPWTPPPLGLKPPSMEGVIAELQ
PpGlk1        SYDSELGRAAGVAACSKPPETHLSKEVLDAAIGEALANPWTPPPLGLKPPSMEGVIAELQ
AcGLK1        ------------------------------------------------------------
```

FIGURA 13 (Cont.)

```
ZmG2          KQGVPKIPPAAATTTGATG
SoGLK2        -------------------
ZmGLK1        RQGVANVPQACG-------
SbGLK1        RQGVANVPPACG-------
OsGLK1        RQGVANVPPACG-------
OsGLK1var     RQGVANVPPACG-------
TaGLK1        RQGVANVPQACG-------
OsGLK2        KQGIPKVPPAASGAAG---
HvGLK2        -------------------
AtGLK2        RQGVSNVPPLP--------
AtGLK1        RHGVSEVPPTASCA-----
PpGLK2        RQGINTVPPSTC-------
PpGlk1        RQGINTVPPSTC-------
AcGLK1        -------------------
```

FIGURA 13 (Cont.)

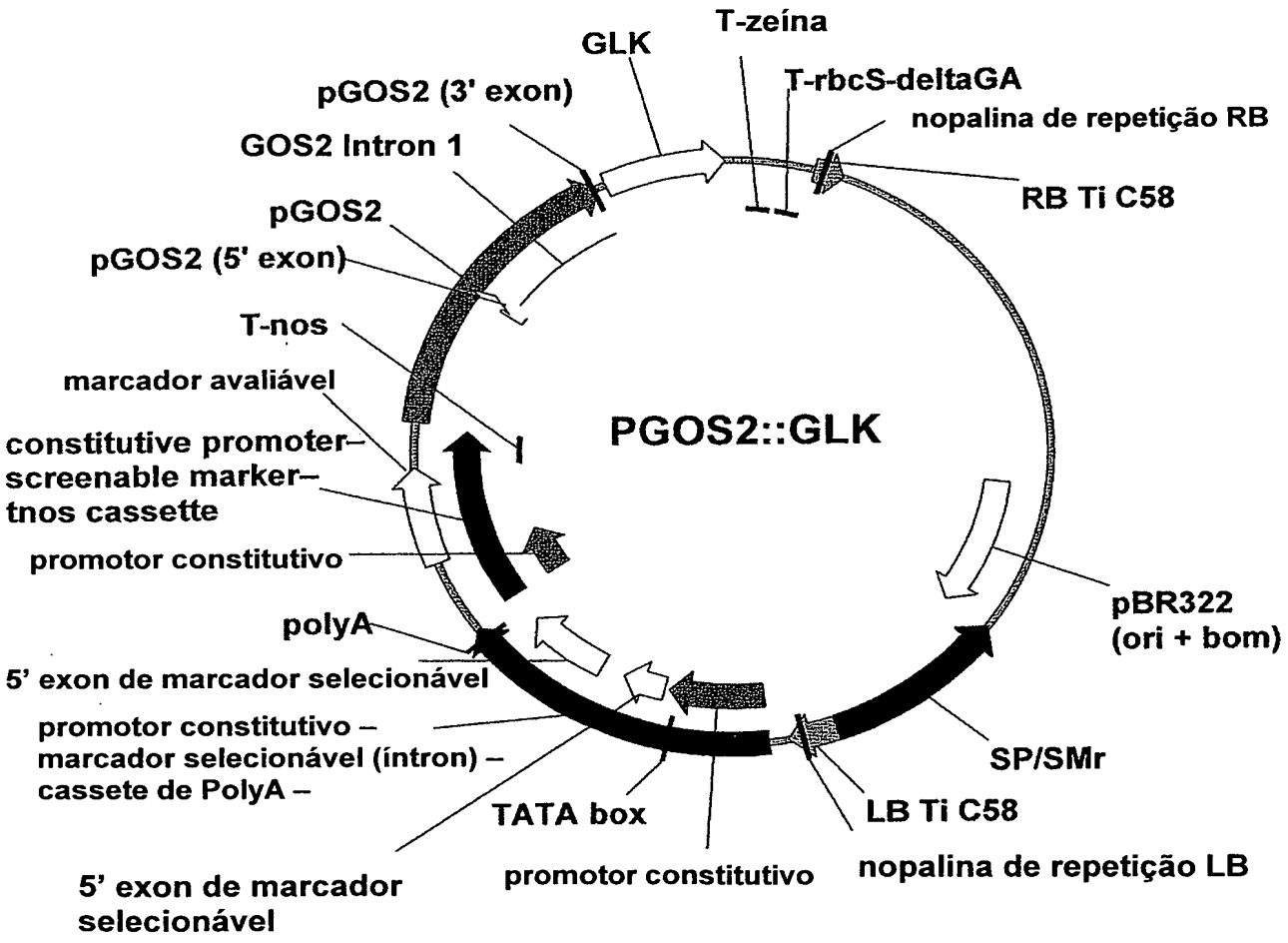
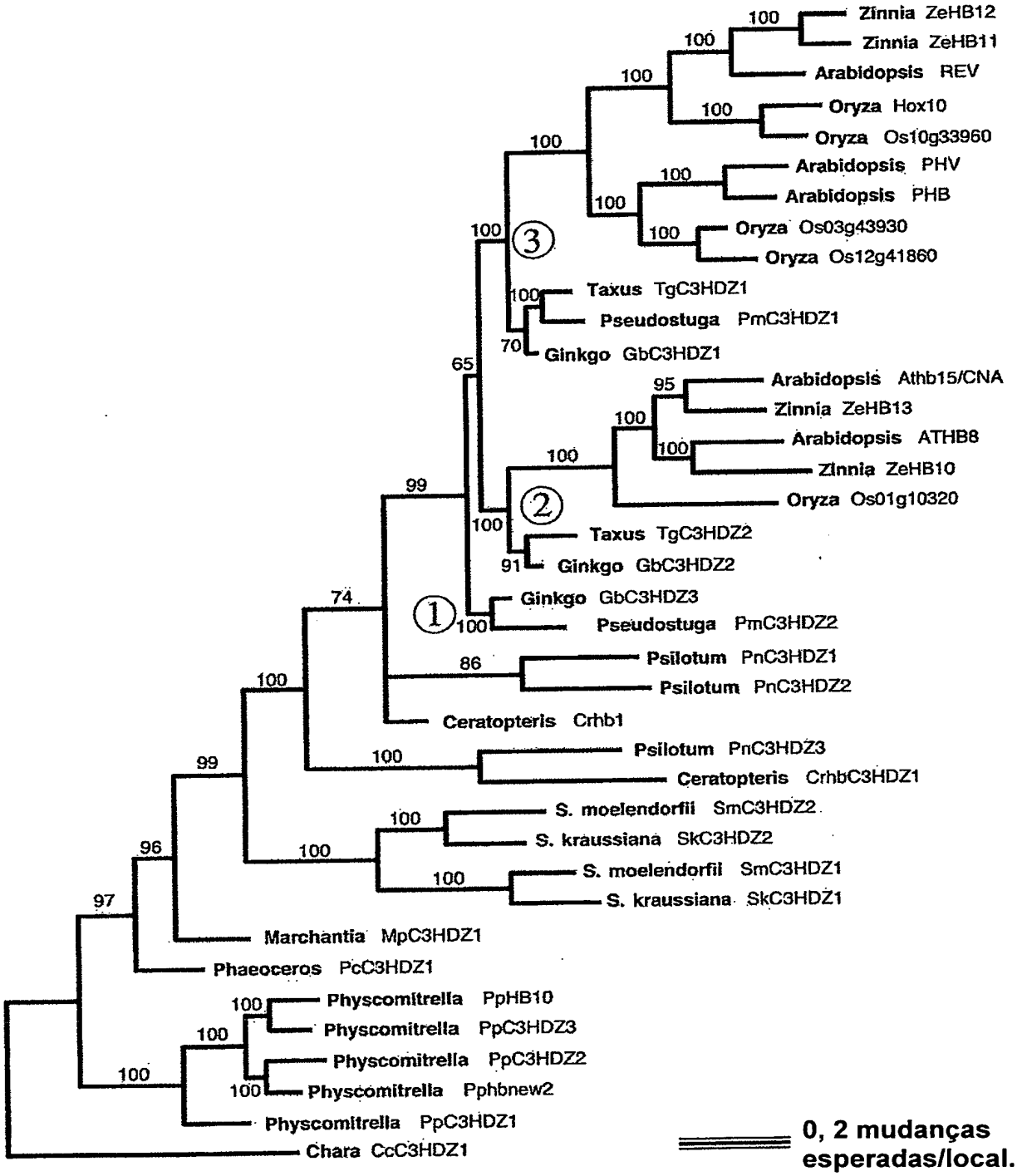


FIGURA 14



After Floyd *et al.* (2006) Genetics 173(1): 373-88

FIGURA 15

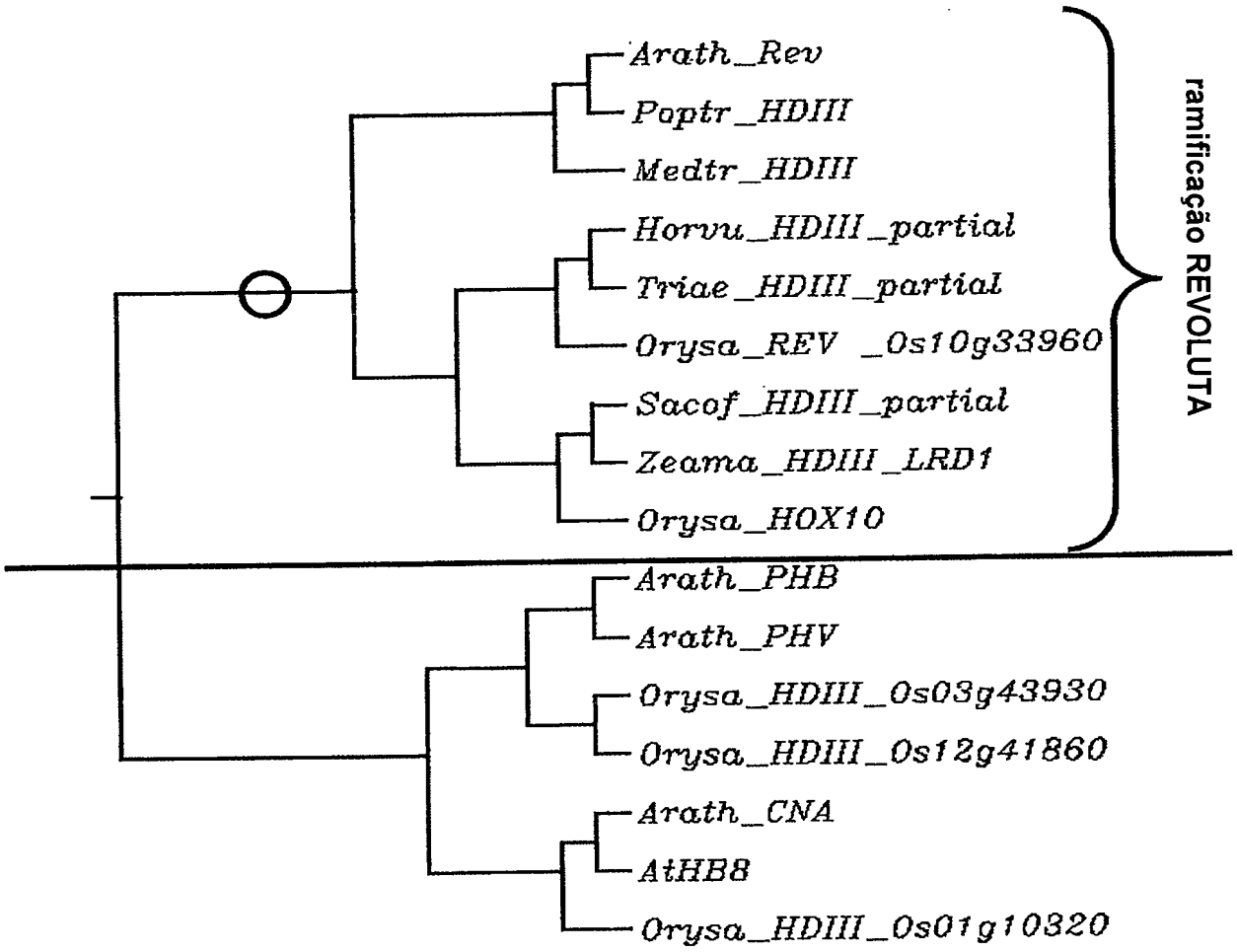


FIGURA 16

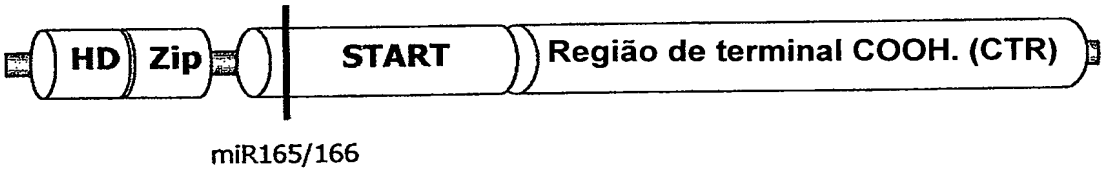


FIGURA 17

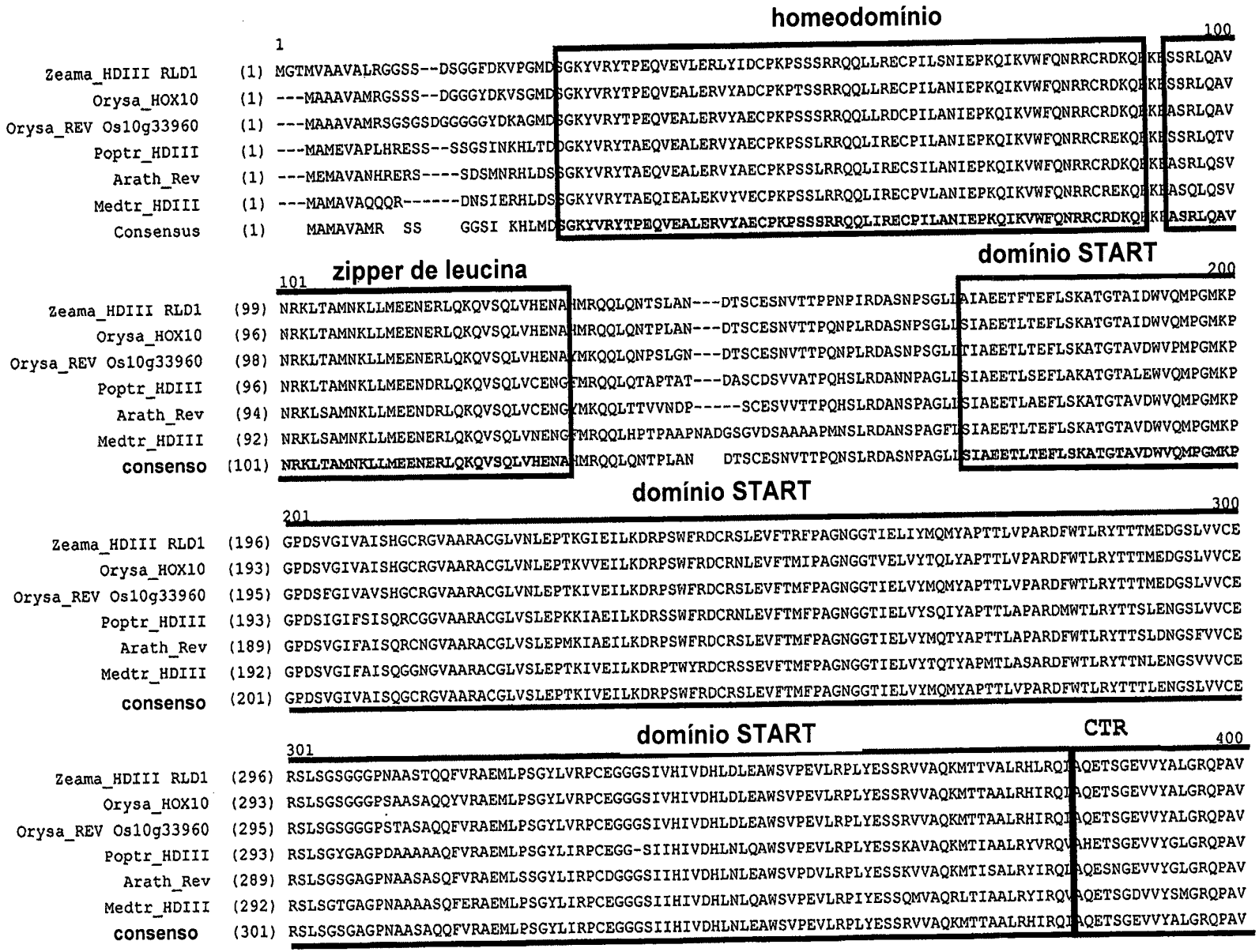

homeodomínio
zipper de leucina
domínio START
domínio START
domínio START
CTR
Zeama_HDIII RLD1
Orysa_HOX10
Orysa_REV Os10g33960
Poptr_HDIII
Arath_Rev
Medtr_HDIII
consenso
Consensus

FIGURA 18

FIGURA 18 (Cont.)

```
CTR_Poptr_HDIII AHETSGEVVYGLGRQPAVLRTFNQRLSRGFNDAINGFNDDGWSLMNADGAEDVIIAVNST
CTR_Arath_REV   AQESNGEVVYGLGRQPAVLRTFSQRLSRGFNDAVNGFGDDGWSTMHCDGAEDIIVAINST
CTR_Medtr_REV   AQETSGDVVYSMGRQPAVLRTFSQRLSRGFNDAVNGFNDNGWSVLNCDGAEGVTISVNSI
CTR_Triae_HDIII AQETSGEVVYALGRQPAVLRTFSQRLSRGFNDAISGFNDDGWSVMAGDGIEDVIIACNS-
CTR_Horvu_HDII  ----------ALGRQPAVLRTFSQRLSRGFNDAISGFNDDGWSVMAGDGIEDVIIACNS-
CTR_Orysa_REV   AQETSGEVVYALGRQPAVLRTFSQRLSRGFNDAISGFNDDGWSVMGGDGIEDVIIACNA-
CTR_Phypr_HDIII AQETSGEVVYALGRQPAVLRTFSQRLSRGFNDAISGFNDDGWSVMGGDGIEDVIIACNS-
CTR_Orysa\HOX10 AQETSGEVVYALGRQPAVLRTFSQRLSRGFNDAISGFNDDGWSIMGGDGVEDVVIACNST
CTR_Zeama_HDIII AQETSGEVVYALGRQPAVLRTFSQRLSRGFNDAISGFNDDGWSVMGGDGIEDVVIACNST
CTR_Sacof_HDIII AQETSGEVVYALGRQPAVLRTFSQRLSRGFNDAISGFNDDGWSVMGGDGIEDVAVACTST
                .:**********.**********:.**.*:*** :  ** *.: :: .:

CTR_Poptr_HDIII KNLIGANNSAHSLSFLGGILCAKASMLLQNVHPAVLVCFLREHHAEWADFSVDAYSAALW
CTR_Arath_REV   KHLNNISN---SLSFLGGVLCAKASMLLQNVPPAVLIRFLREHRSEWADFNVDAYSAATL
CTR_Medtr_REV   KNLSGTSNPASSLSLLGGIVCAKASMLLQNTTPAVLVRFLREHRSEWADFSVDAFSAASL
CTR_Triae_HDIII KKIRSNNTAPNAFIAPGGVICAKASMLLQSVPPAVLVRFLREHRSEWADYNFDAYSASAL
CTR_Horvu_HDIII KKIRSNNTAPNAFIAPGGVICAKASMLLQSVPPAVLVRFLREHRSEWADYNFDAYSASAL
CTR_Orysa_REV   KKVRNTSTSANAFVTPGGVICAKASMLLQSVPPAVLVRFLREHRSEWADYNFDAYSASSL
CTR_Phypr_HDIII KKIRNNSTAANAFGAPGGVICAKASMLLQSVPPAVLVRFLREHRSEWADYNFDAYSALAL
CTR_Orysa\HOX10 KKIRSNSNAGIAFGAPGGIICAKASMLLQSVPPAVLVRFLREHRSEWADYNIDAYLASTL
CTR_Zeama_HDIII KKIRNTSNAGITFGAPGGIICAKASMLLQSVPPAVLVRFLREHRSEWADYNIDAYLASSL
CTR_Sacof_HDIII KKIRNNSNAGITFGAPGGIICAKASMLLQSVPPAVLVRFLREHRSEWADYNIDAYLASSL
                *:: . ..   ::   **::**********.. ****: *****::****:..**: *

CTR_Poptr_HDIII KAGSYAYPGMRPMRFTGSQITMPLGHTIEQEDLLEVIRLEGHSFAQEDAFVSQDIHLLQI
CTR_Arath_REV   KAGSFAYPGMRPTRFTGSQIIMPLGHTIEHEEMLEVVRLEGHSLAQEDAFMSRDVHLLQI
CTR_Medtr_REV   KAGSYGYPGMRSTKFTGNQAIMPLGHTIEHEEMLEIIRLEG--LAQDDSFVSRDVHLLQV
CTR_Triae_HDIII KSSSCSLPGLRPMRFSGSQIIMPLAHTVENEEILEVVRLEGQAL--DEGLLSRDIHLLQF
CTR_Horvu_HDIII KSSSCSLPGLRPMRFSGSQIIMPLAHTVENEEILEVVRLEGQAL--DEGLLSRDIHLLQF
CTR_Orysa_REV   KTSSCSLPGLRPMRFSGSQIIMPLAHTVENEEILEVVRLEGQALTHDDGLMSRDIHLLQL
CTR_Phypr_HDIII KTSSCSLPGLRPTRFSGSQIIMPLAHTVENEEILEVIRLEGQALTHDEGLLSRDIHLLQL
CTR_Orysa\HOX10 KTSACSLTGLRPMRFSGSQIIIPLAHTVENEEILEVVRLEGQPLTHDEALLSRDIHLLQL
CTR_Zeama_HDIII KTSACSLPGLRPMRFSEGQMIMPLAHTVENEEILEVVRLEGQPLTHDEALLSRDIHLLQL
CTR_Sacof_HDIII KTSACSLPXLQPMRXSGGQMIMPLAHTVENEEILEVIRLEGHPLTHDEALLSRDIHLLQL
                *:.: . . ::. : : .*  :**.**:*:*::**::****  :  ::.::*:*:****.

CTR_Poptr_HDIII --------CSGIDENAVGACSELVFAPIDETFPDDAPLLPSGFRIISLESKAKDTQEVLT
CTR_Arath_REV   --------CTGIDENAVGACSELIFAPINEMFPDDAPLVPSGFRVIPVDAKTGDVQDLLT
CTR_Medtr_REV   LPLTLVMLCTGIDENAVGACSELIFAPIDDMFPEDAPLVPSGFRIVLLNSQPGDTKNTTT
CTR_Triae_HDIII --------CTGLDEKSMGSCFQLVFAPIDELFPDDAPLISSGFRVIPLDMKT----DGAP
CTR_Horvu_HDIII --------CTGIDEKSMGSCFQLVFAPIDELFPDDAPLISSGFRVIPLDMKT----DGAP
CTR_Orysa_REV   --------CTGIDEKSMGSCFQLVSAPIDELFPDDAPLISSGFRVIPLDMKT----DGTP
CTR_Phypr_HDIII --------CTGIDEKSMGSCFQLVFAPIDELFPDDAPLISSGFRVIPLDMKT----DGAP
CTR_Orysa\HOX10 --------CTGIDEKSVGSSFQLVFAPIDD-FPDETPLISSGFRVIPLDMKT----DGAS
CTR_Zeama_HDIII --------CTGIDEKSVGSSFQLVFAPIDEHFPDDAPLISSGFRVIPLDVKT----DGVS
CTR_Sacof_HDIII --------CTGIDEKSVGSSFQLVFAPIDEHFPDDAPLISSGFRVIPLDMKT----DGVS
                *:*:**:::*:. :*: ***:: **:::**:.****:: :: :.    :  .

CTR_Poptr_HDIII TN-CTLDLTSSLEAGLAINHTAVDGS--SCHSLRSVLTIAFQFPFESNLQDNVATMARQY
CTR_Arath_REV   ANHRTLDLTSSLEVGPSPENASGNSF--SSSSSRCILTIAFQFPFENNLQENVAGMACQY
CTR_Medtr_REV   AN-RTLDLTSGLEVSPATAHANGDA---SCPNNRCVLTVAFQFPFESGLQDNVAAMARQY
CTR_Triae_HDIII AG-RTLDLASSLEAGSTTLQASGGA--DDC-NLRSVLTIAFQFPYEMHLQDSVATMARQY
CTR_Horvu_HDIII TG-RTLDLASSLEAGSTTLQASGNA--DDC-NLRSVLTIAFQFPYEMHLQDSVATMARQY
CTR_Orysa_REV   AG-RTLDLASSLEVGSTA-QPTGDASMDDC-NLRSVLTIAFQFPYEMHLQDSVATMARQY
CTR_Phypr_HDIII TG-RTLDLASSLEVGSTTQQATGDASLDDCRNLRSVLTIAFQFPYEIHLQDSVATMARQY
CTR_Orysa\HOX10 SG-RTLDLASSLEVGSATAQASGDASADDC-NLRSVLTIAFQFPYELHLQDSVAAMARQY
CTR_Zeama_HDIII SG-RTLDLASSLDVGSAAPQASGDASPDDC-SLRSVLTIAFQFPYEMHLQDSVAAMARQY
CTR_Sacof_HDIII SG-RTLDLASSLDVGSAAPQASGDASPDDC-NLRSVLTIAFQFPYEMHLQDSVATMARQY
                :.  ****:*.*:.. :  :.  ..   .. . *.:**:*****:*  **:.** ** **
```

FIGURA 19

```
CTR_Poptr_HDIII VRSVISSVQRVAMAISPSGLSPVLGPKLSAGSPEALTLAHWICQSHRQVLLKFSSCYHLG
CTR_Arath_REV   VRSVISSVQRVAMAISPSGISPSLGSKLSPGSPEAVTLAQWISQSYS---------HHLG
CTR_Medtr_REV   VRRVVSAVQAVATAISPSSVNTSGGAKLSPGTPEALTLAQWICQSYS---------HHLG
CTR_Triae_HDIII VRSIVSAVQRVSMAISPSRSGLNAEQKIISGFPEAATLARWICQSYR---------FHLG
CTR_Horvu_HDIII VRSIVSAVQRVSMAISPSRSGLNAEQKIISGFPEAATLARWICQSYR---------FHLG
CTR_Orysa_REV   VRSIVSSVQRVSMAISPSRSGLNAGQKIISGFPEAPTLARWICQSYQ---------FHLG
CTR_Phypr_HDIII VRSVVSAVQRVSMAISPPRSGVNAGQKIFSGFPEAATLARWICQSYQ---------FHLG
CTR_Orysa\HOX10 VRSIVSAVQRVSMAISPSQTGLNAGQRIISGFPEAATLARWVCQSYH---------YHLG
CTR_Zeama_HDIII VRSVISAVQRVSMAISPSQSGLNAGHRMLSGFPEAATLARWVCQSYH---------YHLG
CTR_Sacof_HDIII VRSVVSAVQRVSMAISPSQSGLNA-QRTLSGFPETATLARWVCQSYH---------YHLG
                ** ::*:** *: ****.  .     :  .* **: ***:*:.**:          .***

CTR_Poptr_HDIII AELLRSDSV-GGDSVLKHLWHHPDAILCCSLKSLPVFIFANQAGLDMLETTLVALQDITL
CTR_Arath_REV   SELLTIDSLGSDDSVLKLLWDHQDAILCCSLKPQPVFMFANQAGLDMLETTLVALQDITL
CTR_Medtr_REV   AQLLRSDSL-IGDMLLKHLWHHPDAILCCSLKQVPVFIFANQAGLDMLETTLVALQDITL
CTR_Triae_HDIII VELFRQADE-AGESLLRMLWDHEDAILCCSFKEKPVFTFANEMGINMLETSFVALQDLSL
CTR_Horvu_HDIII VELFRQADE-AGESLLRMLWDHEDAILCCSFKEKPVFTFANEMGINMLETSFVALQDLSL
CTR_Orysa_REV   VELLRQADD-AGEALLKMLWDYEDAILCCSFKEKPVFTFANEMGLNMLETSLVALQDLSL
CTR_Phypr_HDIII VELLRQADE-AGESLLRMLWDYEDAILCCSFKEKPVFTFANEMGLNMLETSLVALQDLSL
CTR_Orysa\HOX10 VELLSQSDG-DAEQLLKMLWHYQDAILCCSFKEKPVFTFANKAGLDMLETSLVALQDLTL
CTR_Zeama_HDIII MELLNQSDG-AGEALLKMLWHHPDAVLCCSFKEKPMFTFANKAGLDMLETSLVALQDLTL
CTR_Sacof_HDIII VELLNQSDE-AGEALLKMLWHHPDAVLCCSFKEKPMFTFANKAGLDMLETSLIALQDLTL
                 :*:   .    : :*: **.: **:****:*  *:* ***: *::****:::****::*

CTR_Poptr_HDIII DKIFNESGRQALYTEFAKLMQQGFACLPAGICMSTMGRNVSYEQAVAWKVLSA---EENA
CTR_Arath_REV   EKIFDESGRKAICSDFAKLMQQGFACLPSGICVSTMGRHVSYEQAVAWKVFAASEENNNN
CTR_Medtr_REV   DKIFDESARKNLIAYFAKLMQQGFACMPAGICMSTMGRHASYDQAVAWKVHA----EDNS
CTR_Triae_HDIII DKIFDEAGRKALYSEIPKLMEQGFVYLPGGVCLSGMGRHVSFESAVAWKVVG----EDNN
CTR_Horvu_HDIII DKIFDEAGRKALYSEIPKLMEQGFVYLPGGVCLSGMGRHVSFENAIAWKVVG----EDNN
CTR_Orysa_REV   DKIFDEAGRKALYNEIPKLMEQGYVYLPGGVCLSGMGRHVSFEQAVAWKVLG----EDNN
CTR_Phypr_HDIII DKIFDETGRKALHSEIPKLMEQGYVYLPAGVCLSGMGRHVSFEQAVAWKVLG----EDNN
CTR_Orysa\HOX10 DRIFDEPGKEALFSNIPKLMEQGHVYLPSGVCMSGMGRHVSFDQAVAWKVLA----EDSN
CTR_Zeama_HDIII DKIFDESGRKALFSDISKLMEQGYAYLPSGVCMSGMGRHVSFDQAVAWKVLG----EDSN
CTR_Sacof_HDIII DKIFDESGRKAIFSDISKLMEQGYAYLPSGVCMSGMGRHVSFDQAVAWKVLG----EDSN
                .::**:*..:: :   :.***:**.. :*.*:*:* ***:.*::.*:**** .    ::.

CTR_Poptr_HDIII VHCIAFSFVNWSFL
CTR_Arath_REV   LHCLAFSFVNWSF-
CTR_Medtr_REV   VHCLAFSFINWSFI
CTR_Triae_HDIII VHCLAFCFVNWSF-
CTR_Horvu_HDIII VHCLAFCFVNWSFV
CTR_Orysa_REV   VHCLAFCFVNWSF-
CTR_Phypr_HDIII VHCLAFCFVNWSF-
CTR_Orysa\HOX10 VHCLAFCFVNWSFV
CTR_Zeama_HDIII IHCLAFCFVNWSFV
CTR_Sacof_HDIII VHCLAFCFVNWSF-
                :**:**.*:****
```

FIGURA 19 (Cont.)

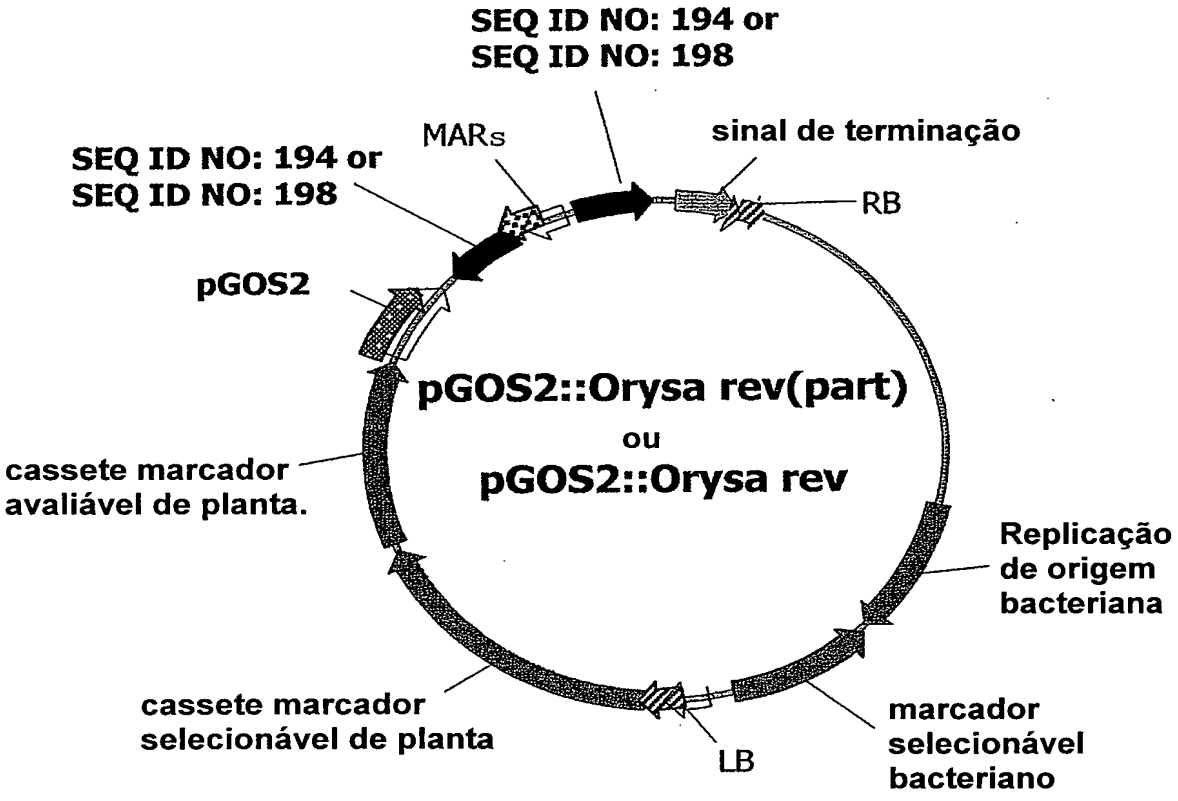

SEQ ID NO: 194 or
SEQ ID NO: 198
MARs
sinal de terminação
SEQ ID NO: 194 or
SEQ ID NO: 198
RB
pGOS2
pGOS2::Orysa rev(part)
ou
pGOS2::Orysa rev
cassete marcador
avaliável de planta.
Replicação
de origem
bacteriana
cassete marcador
selecionável de planta
LB
marcador
selecionável
bacteriano

FIGURA 20

*MRMFFRHLLSCILLLLLLMSHLPSPILG*LRTLRGEEAKSDLRRHEHELPP
AVSPSKEVDNDDAAASSKFTVSRRMVPQGPNPLHNR

FIGURA 21

```
SoCLE       MRMFFRHLLS--CILLLLLMSHLPSPILGLRTLRGEEAKSDLRRHEHELPPAVSPSKEVD
SoCLEb      -------------------MSHLPSSILGLRTSRGEEAKSDLRRHEHELPPAVSPSKEVD
OsCLE       MRRFSKQHLVPFILLLLLVMSHLPISSLGSRRAFREEAVSGFR--SHELAPTMAPSQEKE
PtCLE       -MRNKNPQLFLIFLIVFLVLVHGTTCRDAKRSTSNGETVQGSK-TKHSSMFLQALSSIFK
BnCLE       --MKIKSLILASSFLILAFIHHSESASFRSLLMKNGLYE------EEEAKILLGDSKETI
AtCLE       ----MAKLSFTFCFLLFLLLSSIAAGSRPLEGARVGVKVRGLS------PSIEATSPTVE
                                  :                                . *

SoCLE       NDDAAASS-----KFTVSRRMVPQGPNPLHNR
SoCLEb      NDDAAAAS-----KLTVSRRMVPQGPNPLHNR
OsCLE       AGVVAGAGSICGQKYAVSRRMVPQGPNPLHN-
PtCLE       ASESSTNN--IKALHTVSRRLVPCGPNPLHN-
BnCLE       TNSTALE----------SKRIIPTGPNPLHNR
AtCLE       DDQAAGSHG------KSPERLSPGGPDPQHH-
             .  :              .*: * **:* *:
```

FIGURA 22

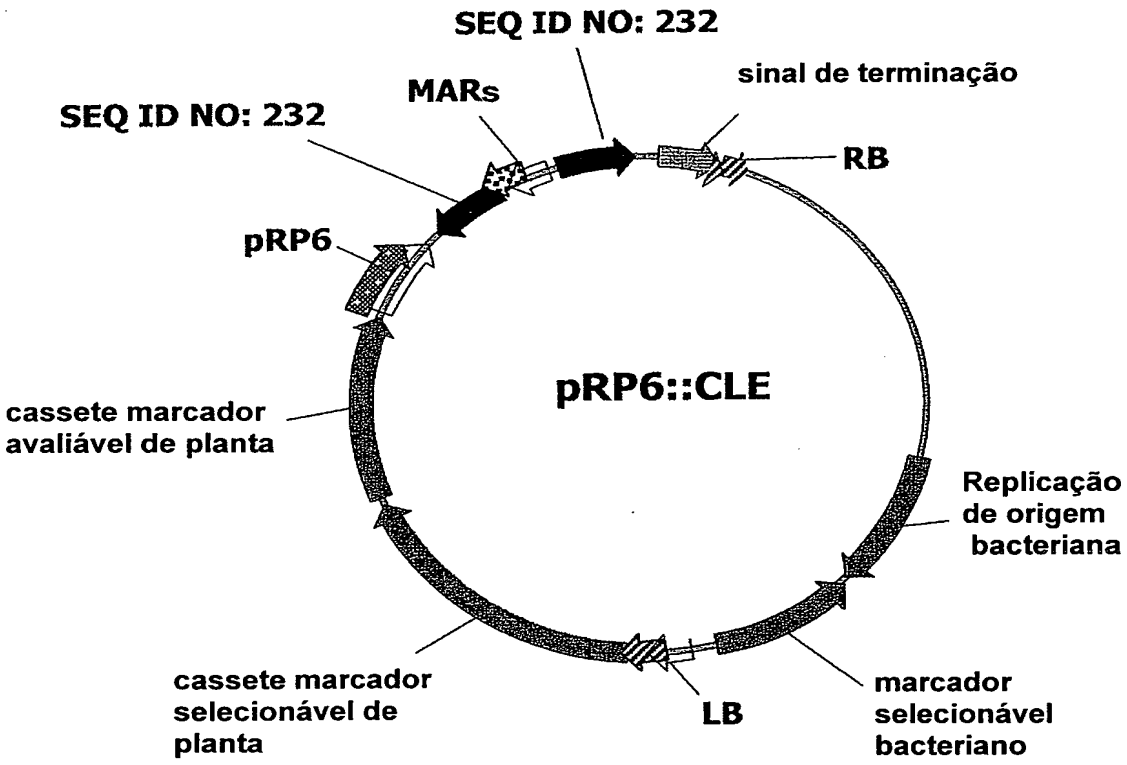


FIGURA 23

MEGVGARQRRNPLIPRP*NGSKR*HLQHQHQPNAAEKKTAATSNYFS**IEAFLVLVFLT**

MSLLILPLVLPPLPPPPSLLLLLPVCLLILLVVLAFMPT**DVRSMAS**SYL

FIGURA 24

CLUSTAL W (1.83) alinhamento de seqüência múltiplo

```
SEQID284    ------------------------------------------------------------
SEQID285    ------------------------------------------------------------
SEQID270    ----------MIREISNLQKD-----IINIQDS-YSNNRVMDV-GRNNR---KNMSFR
SEQID271    ----------MIREFSSLQND-----IINIQEH-YSLNNNMDVRGDHNR---KNTSFR
SEQID272    ----------MSIEQPEADSRLSEGPLINLQDR--YLSGIMEARGRRNSAPLQVERKN
SEQID273    ----------MNSDNSESRQRLSKG-IINLQDR--YPTSIMD-RG---------VRKI
SEQID274    ----------MNMDMESSEAKLRSSKG-FINLEEHQQYFNNIME--G----------
SEQID275    ------------------------------------------ME--G----------
SEQID265    ------------------------------MDSQFGALERG----GSRQRR------SP
SEQID266    ------------------------------MDSQFGAMDRG----GSRQRS------SP
SEQID252    ------------------------------------MEGV----GARQRR------NP
SEQID264    ------------------------------MASRSSAMEGG----AAIQR---------
SEQID269    ------------------------------MASRSSALEGGG---AAIQR---------
SEQID267    ------------------------------------MEGG----GQIQR---------
SEQID268    ------------------------------------MEGG----GQIQR---------
SEQID263    ----------------------------MLLEHLMITMEEQMFREQQMQRG------GR
SEQID262    MYLLSPRNGDEEDEQEEIQELISDDEPPNLKLASCATAASSSSSSGSDMEKGRGKACGGG


SEQID284    ----------MVR----------------------CFSLGSVLILIALAASMVVLPL
SEQID285    ----------MIMVAS------KEKTNSG--------GCMFRYSVLILSLLALSILVLPL
SEQID270    -SSPE--KSKQELRRSFSAQK--RMMIPA--------NYFSLESLFLLVGLTASLLILPL
SEQID271    GSAPAPIMGKQELFRTLSSQNSPRRLISA--------SYFSLESMVVLVGLTASLLILPL
SEQID272    PTPPMAEGKKMEYNRTPLSRENSRRLIPA--------SYFSLESLLLLICLTASLLILPL
SEQID273    ATPPVEK-RKVEYHRS-YSQGASRKLFSA--------SYFTLESLLLLVCLTASLLILPL
SEQID274    ---------NKMEHKRS-FTQGHGKKMLSM--------NYFSLESIILLLGLTASLLLLPL
SEQID275    ---------NKMEHKRS-FTQGHGKKMLSM--------NYFSLESIILLLGLTASLLLLPL
SEQID265    VLARPNTTKRHIQQQ--RANAADKKVVMP--------NYFSIEAFFVLACLTVSLLILPL
SEQID266    VLARPNTAKRQMQQQ--RANAADKKVVIP--------NYFGVEAFFVLACLTVSLLILPL
SEQID252    LIPRPNGSKRHLQHQH-QPNAAEKKTAATS--------NYFSIEAFLVLVFLTMSLLILPL
SEQID264    ---RN-AVKRHLQQRQQEADFLDKKVIAS--------TYFSIGAFLVLACLTVSLLILPL
SEQID269    ---RNNAVKRHLQQRQQEADFHDKKVIAS--------TYFSIGAFLVLACLTFSLLILPL
SEQID267    ---RNNAVKRHLQQRQQEADFLDKKVIAS--------TYFSIEAFLVLACLTVSLLILPL
SEQID268    ---RNNAVKRHLQQRQQEADFLDKKVIAS--------TYFSIEAFLVLACLTVSLLILPL
SEQID263    HHQHHTTREQEQQQKQQQRRRLMNNATNGGGGDGGSRCYFSTEAILVLACVTVSLLVLPL
SEQID262    STAPPPPPPSSSGKSGGGGGSNIREAAASGGGGGVWGKYFSVESLLLLVCVTASLVILPL
                                                    :   :..:*  :: *:::***
```

FIGURA 23

```
SEQID284    MLPPLPPPPLALLFFPVGIMAALVVLAFSPSENV----KNVVV----------------
SEQID285    VMPPLPPPPLLLLLVPVFIMLLLFFIAFSPSKKV----PNKAS----------------
SEQID270    VLPPLPPPPFMLLLVPIGIMVLLVVLAFMPSSHSNANTDVTCN----------------
SEQID271    ILPPLPPPPFMLLLIPIGIMVLLMVLAFMPSSNS-KHVSSSST----------------
SEQID272    ILPPLPPPPFMLLLLPIGILAVLMILAFMPSNVR----DLTYT----------------
SEQID273    VLPPLPPPPFLLLLVPIXILAVLLVLAFMPSNVR----DITST----------------
SEQID274    MLPPLPPPPFMLLLVPIFILVVLMILAFMPSNVR----NVTCS----------------
SEQID275    MLPPLPPPPFMLLLVPIFILVVLMILAFMPSNVR----NVTCS----------------
SEQID265    VLPPLPPPPSLLLFVPVCLLILLMVLAFMPTDMR----SMATS----------------
SEQID266    VLPPLPPPPSLLLLLPVCLLILLMVLAFMPTDVR----SMATS----------------
SEQID252    VLPPLPPPPSLLLLLPVCLLILLVVLAFMPTDVR----SMASS----------------
SEQID264    XXXXXXXXXXXXLLWLPVCLLVLLVVLAFMPTDVR----SMASS----------------
SEQID269    VLPPLPPPPSLLLWLPVCLLVLLVVLAFMPTDVR----SVAAS----------------
SEQID267    VLPXLPAPASLLLWLPVWLLELLIVLAFMPTDVR----SMASS----------------
SEQID268    VLPPLPPPPSLLLWLPVCLLILLIVLAFMPTDVR----SMASS----------------
SEQID263    ILPPLPPPPTLLLLLPVCLLALLVVLAFMPTDMR----TMASS----------------
SEQID262    VLPPLPPPPSMLMLVPVAMLVLLLALAFMPTTTS----SSSSAGGGGGGGRNGATTGHAP
                       *: .*: ::  *. :** *:
```

```
SEQID284    YSSSSSGIANSKR
SEQID285    FVS----------
SEQID270    FM-----------
SEQID271    FM-----------
SEQID272    YV-----------
SEQID273    YV-----------
SEQID274    YL-----------
SEQID275    YL-----------
SEQID265    YL-----------
SEQID266    YL-----------
SEQID252    YL-----------
SEQID264    YL-----------
SEQID269    YL-----------
SEQID267    YL-----------
SEQID268    YL-----------
SEQID263    YL-----------
SEQID262    YL-----------
            :
```

FIGURA 25 (Cont.)

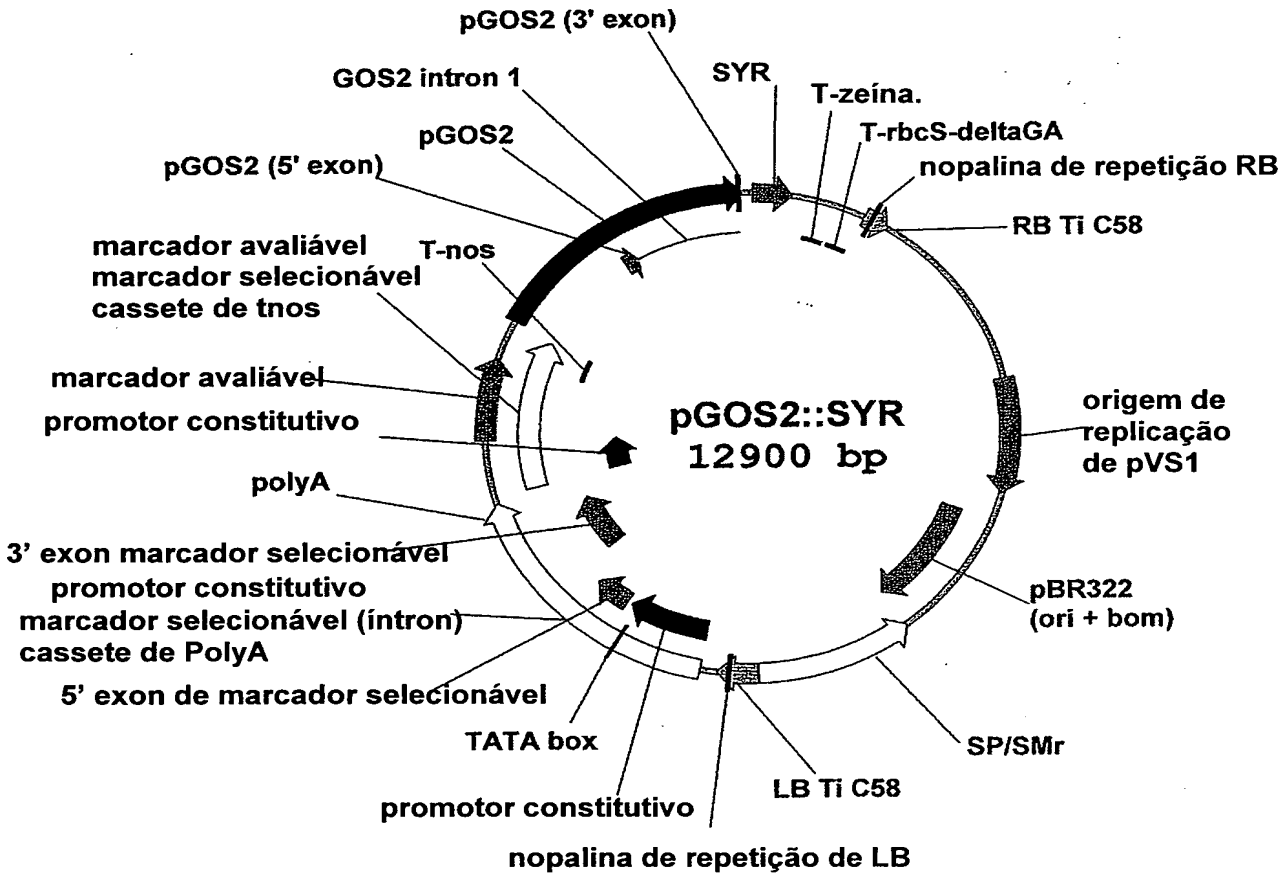

pGOS2 (3' exon)
GOS2 intron 1
SYR
T-zeína.
pGOS2
T-rbcS-deltaGA
pGOS2 (5' exon)
nopalina de repetição RB
RB Ti C58
marcador avaliável
marcador selecionável
cassete de tnos
T-nos
marcador avaliável
promotor constitutivo
pGOS2::SYR
12900 bp
origem de
replicação
de pVS1
polyA
3' exon marcador selecionável
promotor constitutivo
marcador selecionável (íntron)
cassete de PolyA
pBR322
(ori + bom)
5' exon de marcador selecionável
TATA box
SP/SMr
LB Ti C58
promotor constitutivo
nopalina de repetição de LB

FIGURA 26

SEQ ID N° : 1, CDS845, seqüência de clone de entrada
ATGGAAAACAAAAGCCATAGCAACCACCACCGGTACCATCATCTATCATCCCGGAGCCACGA
GCAAAACCAGCGTGGCCTAATTTCGATAAATGCCATAATCATCATTGGCATCTCCATTATCTC
CATATTCATAATCCTTGCAATTCTCCTAATAATCATTTTACTTCATCGACTTAAATCCGCAAGA
GTCAAAGCACAAGAATTATCTTGCAAAGAAAGCTTCAACAACATGAACAATGGTGGAGCTTC
CACCAACTACAGCTACACTTCTAGCCCTGATGACATCAAGAGAGATTGTCTATACTCAAGAAA
CCCAACATCGTTCAGACAATTACCCCCACAGACAAAAAGTTGTAGGAGAAGCCGAGCCGAAG
GAGTAGAAGTGTACACATACAAGGAATTAGAGATTGCAACAAATAATTTCAGCGAGGAGAAG
AAAATCGGAAATGGAGATGTGTACAAAGGAGTACTAAGTGATGGAACTGTTGCGGCCATTAA
AAAGCTTCATATGTTTAATGATAATGCTAGTAACCAAAAGCATGAAGAACGGTCCTTTAGGCT
TGTACAAAGGAGTACTAGTCGATTACAATGCCCATATTTGGTGGAATTACTTGGATATTGTGC
AGATCAAAACCATAGGATCTTGATATATGAATTTATGCCTAATGGTACTGTGGAACATCATCT
CCACGATCATAATTTTAAGAATCTAAAGGATCGACCTCAACCATTAGATTGGGGAGCTCGGCT
TAGGATCGCCCTTGATTGCGCCAGAGCCCTAGAGTTTCTTCATGAGAATACAATCTCTACTGTT
ATCCACCGGAACTTCAAGTGTACCAACATTTTGTTGGATCAAAATAATCGAGCTAAAGTTTCA
GATTTCGGATTAGCCAAAACCGGATCCGATAAACTAAACGGTGAGATTTCAACAAGGGTTATT
GGCACCACAGGATATCTTGCTCCAGAGTATGCCTCTACTGGAAAGCTTACTACAAAATCAGAT
GTTTATAGTTATGGTATTGTGTTACTACAACTCTTAACGGGTCGCACACCGATCGATTCAAGAC
GTCCACGGGGACAAGATGTCCTTGTCTCTTGGGCTCTTCCAAGACTAACCAATAGAGAGAAAA
TTAGTGAAATGGTGGATCCAACCATGAAAGGTCAATACTCACAAAAAGATCTTATTCAGGTAG
CAGCCATAGCAGCAGTGTGTGTGCAACCAGAAGCAAGTTATAGACCGTTGATGACGGATGTT
GTTCACTCGCTCATTCCCCTTGTTAAAGCTTTCAACAAAAGTACTGATTCTTCTCGGTTTCCTA
GCCGTAGAGAAAGCTTGTCATTTGATGATATTATGCCATGA

SEQ ID N° : 2, CDS845, traduzido
MENKSHSNHHRYHHLSSRSHEQNQRGLISINAIIIIGISIISIFIILAILLIIILLHRLKSARVKAQELSCKE
SFNNMNNGGASTNYSYTSSPDDIKRDCLYSRNPTSFRQLPPQTKSCRRSRAEGVEVYTYKELEIAT
NNFSEEKKIGNGDVYKGVLSDGTVAAIKKLHMFNDNASNQKHEERSFRLVQRSTSRLQCPYLVEL
LGYCADQNHRILIYEFMPNGTVEHHLHDHNFKNLKDRPQPLDWGARLRIALDCARALEFLHENTI
STVIHRNFKCTNILLDQNNRAKVSDFGLAKTGSDKLNGEISTRVIGTTGYLAPEYASTGKLTTKSD
VYSYGIVLLQLLTGRTPIDSRRPRGQDVLVSWALPRLTNREKISEMVDPTMKGQYSQKDLIQVAAI
AAVCVQPEASYRPLMTDVVHSLIPLVKAFNKSTDSSRFPSRRESLSFDDIMP

SEQ ID N° : 3, iniciador prm02500
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACAATGGAAAACAAAAGCCATAGC

SEQ ID N° : 4, iniciador prm02501

GGGGACCACTTTGTACAAGAAAGCTGGGTAAACAAAAGAGTGTCATGGCA

SEQ ID N° : 5, promotor GOS2 de arroz
AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAA
TATAAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCA
CCTACTTTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGT
AATTAAGTGGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGT
TAAATTATTCGAGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTTCTAGCT
GAACTCAATGGGTAAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTCTGAACGT
ATTGGCAAAGATTTAAACATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACA
TCATTAAGGACATGTCTTACTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATT
TTTATTATTTATCTTTTTTCGATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCA
TGTGTGAGTGCAC

FIGURA 27

CTCCTCAATACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTAG
GTAGCAATATCTGAATTCAAGCACTCCACCATCACCAGACCACTTTTAATAATATCTAAAATA
CAAAAAATAATTTTACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTCACATA
CAAAAAAAAAAAGAATTTTGCTCGTGCGCGAGCGCCAATCTCCCATATTGGGCACACAGGCA
ACAACAGAGTGGCTGCCCACAGAACAACCCACAAAAAACGATGATCTAACGGAGGACAGCA
AGTCCGCAACAACCTTTTAACAGCAGGCTTTGCGGCCAGGAGAGAGGAGGAGAGGCAAAGAA
AACCAAGCATCCTCCTCCTCCCATCTATAAATTCCTCCCCCCTTTTCCCCTCTCTATATAGGAG
GCATCCAAGCCAAGAAGAGGGAGAGCACCAAGGACACGCGACTAGCAGAAGCCGAGCGACC
GCCTTCTTCGATCCATATCTTCCGGTCGAGTTCTTGGTCGATCTCTTCCCTCCTCCACCTCCTCC
TCACAGGGTATGTGCCCTTCGGTTGTTCTTGGATTTATTGTTCTAGGTTGTGTAGTACGGGCGT
TGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCCTGTTCTTGGATTTGGGATAGAGGGG
TTCTTGATGTTGCATGTTATCGGTTCGGTTTGATTAGTAGTATGGTTTTCAATCGTCTGGAGAG
CTCTATGGAAATGAAATGGTTTAGGGTACGGAATCTTGCGATTTTGTGAGTACCTTTTGTTTGA
GGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAAGTACGGTTGTTTGGTCCTCGA
TTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTCCCTATTGAACAAAAA
TAATCCAACTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTTAAGCCTGTCCAA
AATTTCGCAGCTGGCTTGTTTAGATACAGTAGTCCCCATCACGAAATTCATGGAAACAGTTAT
AATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTTTTTTCCC
AAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATGCTTTT
ATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGAAG
AACTTATCCGATTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATT
CATTTGGATTATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGA
ACTGTCCTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCT
GTAGAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTA
ATCGGGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCTTG
CCACTTTCACCAGCAAAGTTC

SEQ ID N° : 6, motivo 1, X pode ser qualquer aminoácido, preferivelmente um de K, N, A, S, G, E
(M/L)L(S/G/R)R(L/M)(H/R/Q)(H/S/C)(R/P)(N/Y)L(L/V)XL(I/L/V)G

SEQ ID N° : 7, motivo 2, X pode ser qualquer aminoácido, preferivelmente um de D, N, E, K, P, Q ou G
L(Y/D/N)(W/F)X(A/V/T)R(L/M)(L/R/G)IA(L/V)

SEQ ID N° : 8, motivo 3
A(R/K)(A/G)L(A/E)(Y/F)LHE

SEQ ID N° : 9, motivo 4, aminoácido na posição 6 é preferivelmente not I, V ou M
(V/I)IHR(D/N)(F/L)K(S/A/G/C)(S/T)N(I/V)LL(E/D)

SEQ ID N° : 10, motivo 5
(K/R)V(S/A/T)DFG(L/M/S)

SEQ ID N° : 11, variante de união de SEQ ID N° : 1 seqüência codificadora de Q9ZVL4 (AAC64891; Arabidopsis thaliana)
ATGGAAAACAAAAGCCATAGCAACCACCACCGGTACCATCATCTATCATCCCGGAGCCACGA
GCAAAACCAGCGTGGCCTAATTTCGATAAATGCCATAATCATCATTGGCATCTCCATTATCTC
CATAT

FIGURA 27 (continuação)

TCATAATCCTTGCAATTCTCCTAATAATCATTTTACTTCATCGACTTAAATCCGCAAGAGTCAA
AGCACAAGAATTATCTTGCAAAGAAAGCTTCAACAACATGAACAATGGTGGAGCTTCCACCA
ACTACAGCTACACTTCTAGCCCTGGTAAAACACCTAAACATTCATTCCATAATTTTGTAAAGA
ATCATAAGTTTTTAATTCAAAATCAATCAGCAATGAGTGGAGATTCTCATATCTTTATATCAGT
ATTGAATTACTTCCACATCCAATGTGTATTTCTTATTCGAACTTTAATGTTCTTGTTAGATGAC
ATCAAGAGAGATTGTCTATACTCAAGAAACCCAACATCGTTCAGACAATTACCCCCACAGACA
AAAAGTTGTAGGAGAAGCCGAGCCGAAGGAGTAGAAGTGTACACATACAAGGAATTAGAGA
TTGCAACAAATAATTTCAGCGAGGAGAAGAAAATCGGAAATGGAGATGTGTACAAAGGAGTA
CTAAGTGATGGAACTGTTGCGGCCATTAAAAAGCTTCATATGTTTAATGATAATGCTAGTAAC
CAAAAGCATGAAGAACGGTCCTTTAGGCTTGAGGTTGATCTTCTTAGTCGATTACAATGCCCA
TATTTGGTGGAATTACTTGGATATTGTGCAGATCAAAACCATAGGATCCTGATATATGAATTT
ATGCCTAATGGTACTGTGGAACATCATCTCCACGATCATAATTTTAAGAATCTAAAGGATCGA
CCTCAACCATTAGATTGGGGAGCTCGGCTTAGGATCGCCCTTGATTGCGCCAGAGCCCTAGAG
TTTCTTCATGAGAATACAATCTCTACTGTTATCCACCGGAACTTCAAGTGTACCAACATTTTGT
TGGATCAAAATAATCGAGCTAAAGTTTCAGATTTCGGATTAGCCAAAACCGGATCCGATAAAC
TAAACGGTGAGATTTCAACAAGGGTTATTGGCACCACAGGATATCTTGCTCCAGAGTATGCCT
CTACTGGAAAGCTTACTACAAAATCAGATGTTTATAGTTATGGTATTGTGTTACTACAACTCTT
AACGGGTCGCACACCGATCGATTCAAGACGTCCACGGGGACAAGATGTCCTTGTCTCTTGGGC
TCTTCCAAGACTAACCAATAGAGAGAAAATTAGTGAAATGGTGGATCCAACCATGAAAGGTC
AATACTCACAAAAAGATCTTATTCAGGTAGCAGCCATAGCAGCAGTGTGTGTGCAACCAGAA
GCAAGTTATAGACCGTTGATGACGGATGTTGTTCACTCGCTCATTCCCCTTGTTAAAGCTTTCA
ACAAAAGTACTGATTCTTCTCGGTTTCCTAGCCGTAGAGAAAGCTTGTCATTTGATGATATTAT
GCCATGA

SEQ ID N° : 12, variante de união de SEQ ID N° : 2 (Q9ZVL4|Q9ZVL4_ARATH T22H22.21 proteína - Arabidopsis thaliana). variante de união de NP_175879

MENKSHSNHHRYHHLSSRSHEQNQRGLISINAIIIIGISIISIFIILAILLIIILLHRLKSARVKAQELSCKE
SFNNMNNGGASTNYSYTSSPGKTPKHSFHNFVKNHKFLIQNQSAMSGDSHIFISVLNYFHIQCVFLI
RTLMFLLDDIKRDCLYSRNPTSFRQLPPQTKSCRRSRAEGVEVYTYKELEIATNNFSEEKKIGNGDV
YKGVLSDGTVAAIKKLHMFNDNASNQKHEERSFRLEVDLLSRLQCPYLVELLGYCADQNHRILIY
EFMPNGTVEHHLHDHNFKNLKDRPQPLDWGARLRIALDCARALEFLHENTISTVIHRNFKCTNILL
DQNNRAKVSDFGLAKTGSDKLNGEISTRVIGTTGYLAPEYASTGKLTTKSDVYSYGIVLLQLLTGR
TPIDSRRPRGQDVLVSWALPRLTNREKISEMVDPTMKGQYSQKDLIQVAAIAAVCVQPEASYRPL
MTDVVHSLIPLVKAFNKSTDSSRFPSRRESLSFDDIMP

SEQ ID N° : 13, >gi|42562757|ref|NM_104355.2| ligação de ATP de Arabidopsis thaliana / quinase / proteína quinase / proteína serina / treonina quinase / proteína-tirosinaquinase AT1G54820 mRNA, cds completos

CAAAAATGAACAACACCAAAATGGAAAACAAAAGCCATAGCAACCACCACCGGTACCATCAT
CTATCATCCCGGAGCCACGAGCAAAACCAGCGTGGCCTAATTTCGATAAATGCCATAATCATC
ATTGGCATCTCCATTATCTCCATATTCATAATCCTTGCAATTCTCCTAATAATCATTTTACTTCA
TCGACTTAAATCCGCAAGAGTCAAAGCACAAGAATTATCTTGCAAAGAAAGCTTCAACAACA
TGAACAATGGTGGAGCTTCCACCAACTACAGCTACACTTCTAGCCCTGATGACATCAAGAGAG
ATTGTCTATACTCAAGAAACCCAACATCGTTCAGACAATTACCCCCACAGACAAAAAGTTGTA
GGAGAAGCCGAGCCGAAGGAGTAGAAGTGTACACATACAAGGAATTAGAGATTGCAACAAA
TAATTTCAGCGAGGAGAAGAAAATCGGAAATGGAGATGTGTACAAAGGAGTACTAAGTGATG
GAACTGTTGCGGCCATTAAAAAGCTTCATATGTTTAATGATAATGCTAGTAACCAAAAGCATG
AAGAACGGTCCTTTAGGCTTGAGGTTGATCTTCTTAGTCGATTACAATGCCCATATTTGGTGGA
ATTACTTGGATATTGTGCAGAT

FIGURA 27 (continuação)

CAAAACCATAGGATCCTGATATATGAATTTATGCCTAATGGTACTGTGGAACATCATCTCCAC
GATCATAATTTTAAGAATCTAAAGGATCGACCTCAACCATTAGATTGGGGAGCTCGGCTTAGG
ATCGCCCTTGATTGCGCCAGAGCCCTAGAGTTTCTTCATGAGAATACAATCTCTACTGTTATCC
ACCGGAACTTCAAGTGTACCAACATTTTGTTGGATCAAAATAATCGAGCTAAAGTTTCAGATT
TCGGATTAGCCAAAACCGGATCCGATAAACTAAACGGTGAGATTTCAACAAGGGTTATTGGC
ACCACAGGATATCTTGCTCCAGAGTATGCCTCTACTGGAAAGCTTACTACAAAATCAGATGTT
TATAGTTATGGTATTGTGTTACTACAACTCTTAACGGGTCGCACACCGATCGATTCAAGACGT
CCACGGGGACAAGATGTCCTTGTCTCTTGGGCTCTTCCAAGACTAACCAATAGAGAGAAAATT
AGTGAAATGGTGGATCCAACCATGAAAGGTCAATACTCACAAAAAGATCTTATTCAGGTAGC
AGCCATAGCAGCAGTGTGTGTGCAACCAGAAGCAAGTTATAGACCGTTGATGACGGATGTTG
TTCACTCGCTCATTCCCCTTGTTAAAGCTTTCAACAAAAGTACTGATTCTTCTCGGTTTCCTAG
CCGTAGAGAAAGCTTGTCATTTGATGATATTATGCCATGACACTCTTTTGTTTAATATGTATTA
ATTTCTCTCTTTTAAGTTGAGATTCTCGTTGTTATTGTATATTTGTATAGGTAAATGAATCCATT
TATCTTGTACT

SEQ ID N° : 14, >gi|42562758|ref|NP_175879.2| ligação de ATP / quinase/ proteína quinase/ proteína serina/treonina quinase/ proteína-tirosinaquinase [Arabidopsis thaliana]
MNNTKMENKSHSNHHRYHHLSSRSHEQNQRGLISINAIIIIGISIISIFIILAILLIIILLHRLKSARVKAQ
ELSCKESFNNMNNGGASTNYSYTSSPDDIKRDCLYSRNPTSFRQLPPQTKSCRRSRAEGVEVYTYK
ELEIATNNFSEEKKIGNGDVYKGVLSDGTVAAIKKLHMFNDNASNQKHEERSFRLEVDLLSRLQCP
YLVELLGYCADQNHRILIYEFMPNGTVEHHLHDHNFKNLKDRPQPLDWGARLRIALDCARALEFL
HENTISTVIHRNFKCTNILLDQNNRAKVSDFGLAKTGSDKLNGEISTRVIGTTGYLAPEYASTGKLT
TKSDVYSYGIVLLQLLTGRTPIDSRRPRGQDVLVSWALPRLTNREKISEMVDPTMKGQYSQKDLIQ
VAAIAAVCVQPEASYRPLMTDVVHSLIPLVKAFNKSTDSSRFPSRRESLSFDDIMP

SEQ ID N° : 15, NM_115731.3| Arabidopsis thaliana quinase AT3G58690 mRNA, cds completos
CAAAAACCAAAAACTAATTTTCACGAAATCCGAAGCTTAAAAATTATTCCTCTACTGTTTCAA
TTTCATCTCTTCTTCCTCTTGAGTTGGCAATTTCTAGGGTTTTCTCATTTTCCTTTTTTTCCATTT
CTGGAAACTGGTAAAAGGCTTTTTCTTTGTCTGTGGGGGAGAAACAAAAATGGAGACAGACG
AAGCATACCAAAAGAAAGAGCGAGCTGCATTAGTAGCCATCGTTGTCCTTGCTTGTCTTGCTT
TATCTTCTTTGTTCGTCGCCTTTAGCTACTACTGCTATATTCGTAACAAGGTTTCTAAGCGTCAC
AGAATTAGCAAGAGGTTTGATTGTGAGGAAAAAGGCGATTGTCAGAAAGTACAAGATGTTAC
TGAAAATGGGTTGCAAATTTTCACCTTTAAGCAATTGCATTCAGCAACTGGTGGATTTAGCAA
GTCTAATGTGGTTGGGAATGGTGGGTTTGGCTTGGTTTATCGTGGTGTGCTTAACGATGGCAG
AAAAGTTGCTATCAAGCTCATGGATCATGCAGGAAAGCAAGGGGAAGAAGAATTCAAGATGG
AGGTTGAGTTACTGAGCCGTCTGCGTTCACCATACTTGTTGGCACTTCTTGGTTATTGCTCAGA
CAATAGTCATAAGCTGCTTGTGTATGAGTTCATGGCAAATGGTGGTCTGCAGGAACACCTGTA
TCTTCCCAACAGGTCTGGTTCTGTCCCACCAAGATTAGATTGGGAAACTAGGATGAGAATAGC
TGTGGAAGCTGCAAAAGGCTTGGAATATCTCCATGAACAAGTATCACCCCCGGTGATTCATAG
AGACTTTAAGAGCAGCAATATTCTGCTGGACAGAAACTTCAATGCCAAAGTTTCAGATTTTGG
ATTGGCTAAAGTCGGATCAGATAAAGCTGGTGGACACGTTTCTACCCGTGTATTAGGCACACA
AGGATATGTAGCTCCTGAGTACGCTTTAACTGGTCACTTGACAACAAAGTCGGATGTCTATAG
CTACGGTGTTGTCCTGTTAGAGTTACTAACCGGTAGAGTTCCAGTAGATATGAAGAGAGCTAC
AGGAGAAGGTGTTCTTGTCTCTTGGGCTTTGCCTCAATTGGCTGATAGAGACAAAGTAGTAGA
CATAATGGATCCAACACTCGAAGGACAATACTCTACGAAAGAAGTTGTTCAAGTTGCAGCAA
TAGCAGCAATGTGTGTTCAAGCAGAAGCTGACTACAGACCCTTAATGGCAGATGTGGTGCAGT
CACTGGTTCCATTAGTGAGAAACCGTAGGTCAGCTTCA

FIGURA 27 (continuação)

AAGCTTAGTGGCTGCTCTAGTAGCTTTAGCTTGGCTCGGTCTCCTAACTCTCCGGGTAAAGCA
AGCATAGGTTCTCAATGAATAAATCAATCAGAGAAGACGAAATGTGAGGCTTTTATTATGTAA
CGGTATTTACTAAAACAGACAATGATGTAAACACTTGAGTGATCAAACATGGAGTCATAGAG
ATAATCATCAAATAAACTTTTTCGAATC

SEQ ID N° : 16, NP_191428.3| quinase [Arabidopsis thaliana] AT3G58690

METDEAYQKKERAALVAIVVLACLALSSLFVAFSYYCYIRNKVSKRHRISKRFDCEEKGDCQKVQ
DVTENGLQIFTFKQLHSATGGFSKSNVVGNGGFGLVYRGVLNDGRKVAIKLMDHAGKQGEEEFK
MEVELLSRLRSPYLLALLGYCSDNSHKLLVYEFMANGGLQEHLYLPNRSGSVPPRLDWETRMRIA
VEAAKGLEYLHEQVSPPVIHRDFKSSNILLDRNFNAKVSDFGLAKVGSDKAGGHVSTRVLGTQGY
VAPEYALTGHLTTKSDVYSYGVVLLELLTGRVPVDMKRATGEGVLVSWALPQLADRDKVVDIM
DPTLEGQYSTKEVVQVAAIAAMCVQAEADYRPLMADVVQSLVPLVRNRRSASKLSGCSSSFSLAR
SPNSPGKASIGSQ

SEQ ID N° : 17, nucleotídeo NP209041, Arabidopsis thaliana, At3g58690

ATGGAGACAGACGAAGCATACCAAAAGAAAGAGCGAGCTGCATTAGTAGCCATCGTTGTCCT
TGCTTGTCTTGCTTTATCTTCTTTGTTCGTCGCCTTTAGCTACTACTGCTATATTCGTAACAAGG
AAAAAGGCGATTGTCAGAAAGTACAAGATGTTACTGAAAATGGGTTGCAAATTTTCACCTTTA
AGCAATTGCATTCAGCAACTGGTGGATTTAGCAAGTCTAATGTGGTTGGGAATGGTGGGTTTG
GCTTGGTTTATCGTGGTGTGCTTAACGATGGCAGAAAAGTTGCTATCAAGCTCATGGATCATG
CAGGAAAGCAAGGGGAAGAAGAATTCAAGATGGAGGTTGAGTTACTGAGCCGTCTGCGTTCA
CCATACTTGTTGGCACTTCTTGGTTATTGCTCAGACAATAGTCATAAGCTGCTTGTGTATGAGT
TCATGGCAAATGGTGGTCTGCAGGAACACCTGTATCTTCCCAACAGGTCTGGTTCTGTCCCAC
CAAGATTAGATTGGGAAACTAGGATGAGAATAGCTGTGGAAGCTGCAAAAGGCTTGGAATAT
CTCCATGAACAAGTATCACCCCCGGTGATTCATAGAGACTTTAAGAGCAGCAATATTCTGCTG
GACAGAAACTTCAATGCCAAAGTTTCAGATTTTGGATTGGCTAAAGTCGGATCAGATAAAGCT
GGTGGACACGTTTCTACCCGTGTATTAGGCACACAAGGATATGTAGCTCCTGAGTACGCTTTA
ACTGGTCACTTGACAACAAAGTCGGATGTCTATAGCTACGGTGTTGTCCTGTTAGAGTTACTA
ACCGGTAGAGTTCCAGTAGATATGAAGAGAGCTACAGGAGAAGGTGTTCTTGTCTCTTGGGCT
TTGCCTCAATTGGCTGATAGAGACAAAGTAGTAGACATAATGGATCCAACACTCGAAGGACA
ATACTCTACGAAAGAAGTTGTTCAAGTTGCAGCAATAGCAGCAATGTGTGTTCAAGCAGAAG
CTGACTACAGACCCTTAATGGCAGATGTGGTGCAGTCACTGGTTCCATTAGTGAGAAACCGTA
GGTCAGCTTCAAAGCTTAGTGGCTGCTCTAGTAGCTTTAGCTTGGCTCGGTCTCCTAACTCTCC
GGGTAAAGCAAGCATAGGTTCTCAATGA

SEQ ID N° : 18, NP209041, seqüência de proteína deduzida, variante de união de At3g58690

METDEAYQKKERAALVAIVVLACLALSSLFVAFSYYCYIRNKEKGDCQKVQDVTENGLQIFTFKQ
LHSATGGFSKSNVVGNGGFGLVYRGVLNDGRKVAIKLMDHAGKQGEEEFKMEVELLSRLRSPYL
LALLGYCSDNSHKLLVYEFMANGGLQEHLYLPNRSGSVPPRLDWETRMRIAVEAAKGLEYLHEQ
VSPPVIHRDFKSSNILLDRNFNAKVSDFGLAKVGSDKAGGHVSTRVLGTQGYVAPEYALTGHLTT
KSDVYSYGVVLLELLTGRVPVDMKRATGEGVLVSWALPQLADRDKVVDIMDPTLEGQYSTKEV
VQVAAIAAMCVQAEADYRPLMADVVQSLVPLVRNRRSASKLSGCSSSFSLARSPNSPGKASIGSQ

FIGURA 27 (continuação)

SEQ ID N° : 19, NM_116432.2| Arabidopsis thaliana quinase AT4G02010 mRNA, cds completos

AACGGAATATTTTTTTTCTGGGAGCTGAAAATCGTTGGTGACGATGGCGGATTCTCTTTCTAGG
GTTTACGGAACTCTGTTTAGTGGTGTTAATATATCTCTCTTCTCGGATCCACCGTTGAAAACCC
CCTGAAGAGTGTTTTTTTTTCCGGGTTTAAGGTCTTTTACGTTTCGTTTTTTCTTGAGATTTCGT
CTTCCGCTGTGAGTTTCTGTTCTCGCGGCGAGTTTTTGTGTAGATCTGAGGGTTTTCAAGGCAA
GAGAAAGTGTCGGTGATGTCGATTTTCTCACTCGCGAGCTCTTCTGATTCACAAAGAGATGAC
CTAATAATGGCGTTGAAGGCGGTGGTTATTGTATATTGTGTGGTTTCTCTCGTCAGTGTTCAAT
TAGCTGATGCTCAACATGAAGGACTGCCAGTTTCACCAACGTTATCACCTTCAACTTCACCAG
TTATCACTGATCTGCCCCTACCAGCTGAATTTCCGCGGTTTCACAGAAAGTATTTCGCACCACA
ACAAGCAGAAGCACCTCAGCATTCTCCCCCTTATAGTCGTTTGGTCGCTTCTGATCATCCCCCT
ACCAGCTCACATTTCTCCAAACCTTCCATGAAAAGGAATGCTCAGTCTCCTGGAGCCGGCTTG
GCTGATATTGCTCCAGCACAATCTAGCAATGGTGTTCTTCCTGATGCCTTAACTCAGCCACCTT
TGTCGCCCTCCATTTCAAATTGTTGCAAATCAGATATGGTGCTTAAACGAAGAAGTATTGGTT
GCCACTGCGTGTATCCTATAAAACTGGACATCCTTCTCTTGAATGTTTCAGAAACTCCTAGTTG
GAACATGTTCTTGAACGAATTTGCTACCCAGCTTGGTCTCCTACCTCACCAAATCGAGCTGATT
AACTTCTATGTGCTAAGCTTATCAAGGATGAACATATCGATGGATATCACCCCTCATTCTGGA
ATTAGTTTCTCAGCTAGTCAGGCATCCGCAATAAACTCTTCCCTTATCAGCCACAAGATTCAAT
TTAGCCCTACTTTGGTGGGAGATTACAAACTTCTAAACCTTACTTGGTTTGAGGCCCCTGCACC
TTCGCAAGCACCTCTAGTGGCTTCTTCACCTCATAAAGCACCATCACAAGGATCCTCAGCAAC
TACGTCAGTAAGATCTCCAGGGAAAAAGAGGCATCCCAATCTTATTCTTATCTTTTCTATAGCC
GCTGGTGTGCTTATACTTGCCATAATCACTGTACTTGTTATTTGTTCCCGCGCACTCCGAGAAG
AGAAAGCTCCAGATCCTCACAAAGAAGCTGTAAAACCAAGGAACCTGGACGCTGGTTCATTT
GGGGGATCTCTTCCTCACCCAGCAAGTACACGGTTTCTGTCATATGAAGAACTCAAAGAGGCA
ACTAGCAATTTTGAATCTGCTAGCATTCTAGGAGAAGGTGGGTTTGGCAAGGTTTACAGAGGC
ATCTTAGCCGATGGTACTGCTGTAGCGATTAAGAAGCTCACAAGTGGTGGGCCACAAGGTGAT
AAAGAATTCCAGGTGGAGATTGATATGCTTAGCCGTCTTCATCATCGTAATCTTGTGAAACTT
GTGGGTTACTATAGTAGTCGAGATTCTTCTCAGCACCTACTTTGTTATGAGCTTGTTCCAAATG
GCAGCCTCGAGGCTTGGCTCCATGGGCCTCTCGGGTTGAACTGTCCTCTTGATTGGGACACCA
GAATGAAGATTGCACTTGATGCTGCAAGAGGACTTGCATACCTTCATGAAGACTCGCAACCCT
CCGTTATACACAGAGATTTTAAAGCCTCTAATATACTCCTTGAAAACAACTTCAACGCCAAAG
TTGCAGATTTTGGCCTAGCCAAACAAGCTCCTGAAGGCAGGGGTAATCACTTATCTACTCGTG
TTATGGGCACATTTGGATATGTTGCGCCTGAATATGCAATGACGGGACACCTACTCGTCAAGA
GTGATGTTTATAGTTACGGTGTGGTCCTTCTCGAATTGTTAACTGGTAGAAAACCTGTGGATAT
GTCACAACCTTCAGGCCAAGAAAATCTCGTCACTTGGACAAGGCCAGTTTTAAGAGACAAAG
ACCGGTTAGAAGAACTAGTCGATTCAAGACTTGAAGGAAAATACCCGAAAGAAGATTTCATA
AGAGTATGCACAATCGCTGCAGCTTGTGTTGCACCTGAAGCTAGCCAGAGACCAACGATGGG
CGAAGTGGTTCAGTCACTTAAAATGGTTCAACGGGTGGTTGAGTATCAAGACCCGGTTTTAAA
CACTTCAAATAAAGCTCGTCCTAACCGGAGACAATCATCAGCTACGTTCGAGTCAGAAGTAAC
CTCTTCTATGTTCTCTTCTGGTCCTTATTCTGGTCTAAGCGCTTTTGATCATGAAAATATTACAC
GAACAACTGTTTTCTCAGAAGATCTTCACGAAGGCCGATGATAGAAGCCAGGGTTTTCTTCTT
TTTATTTGTTTTTCTCCCACTTACGGTGAGAAAATTTCCACCAGAGAAGCTTTTGAGTTTGGGA
CATTATTACAGCTCTTTGGATTTTGGATTCTCCTTTATTGGAGGATAGAATTTTGTATATATTTT
TGCGTTAATTAATTAATATTTGCCAC

SEQ ID N° : 20, NP_192110.2| quinase [Arabidopsis thaliana], AT4G02010

MSIFSLASSSDSQRDDLIMALKAVVIVYCVVSLVSVQLADAQHEGLPVSPTLSPSTSPVITDLPLPAE
FPRFHRKYFAPQQAEAPQHSPPYSRLVASDHPPTSSHFSKPSMKRNAQSPGAGLADIAPAQS

FIGURA 27 (continuação)

SNGVLPDALTQPPLSPSISNCCKSDMVLKRRSIGCHCVYPIKLDILLLNVSETPSWNMFLNEFATQL
GLLPHQIELINFYVLSLSRMNISMDITPHSGISFSASQASAINSSLISHKIQFSPTLVGDYKLLNLTWFE
APAPSQAPLVASSPHKAPSQGSSATTSVRSPGKKRHPNLILIFSIAAGVLILAIITVLVICSRALREEK
APDPHKEAVKPRNLDAGSFGGSLPHPASTRFLSYEELKEATSNFESASILGEGGFGKVYRGILADGT
AVAIKKLTSGGPQGDKEFQVEIDMLSRLHHRNLVKLVGYYSSRDSSQHLLCYELVPNGSLEAWLH
GPLGLNCPLDWDTRMKIALDAARGLAYLHEDSQPSVIHRDFKASNILLENNFNAKVADFGLAKQA
PEGRGNHLSTRVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLELLTGRKPVDMSQPSGQENLV
TWTRPVLRDKDRLEELVDSRLEGKYPKEDFIRVCTIAAACVAPEASQRPTMGEVVQSLKMVQRV
VEYQDPVLNTSNKARPNRRQSSATFESEVTSSMFSSGPYSGLSAFDHENITRTTVFSEDLHEGR

SEQ ID N° : 21, Arabidopsis Extensina RLK At5g56890 mRNA

CTTCTTCTTCTGCCACTTCGATTGCTTTAAACTTGGAGATTAAGATTACACATCAAATAATTCT
TCTTTGTCTTCCTCGTGATCATCGAGGAAGCTTGATCTCTTCTATGGAATGGAGAATTTGACTC
TCCTCCTTAGGATCTGTCTTGTTTCGTCGGTTTTAGTTGCTGCTTCTTCCTCAGGATCGGAATTG
TTATCTCCGTTATCATCACCACCGTCTCCATTACCCGAAACTTCAAAAGGGTTTGGTCAAGCAC
CGAGTAATTCACCTGAATCTCACAAATCCGACAATGTGCCACCGTCTAAAGCTTCTTCTCAAC
CGTCTCTTCCTCCTTTAGCTGATTTGGCAGCTCCACCACCGTCAGATTCCGTCGGAGGTAAAGC
ACCGGCCGGTGTGCCTGTTGTCTTTGTTCCTAATGCTCCAGCTCCAGCTACAATACCCGTTAAG
GATTTGCCCGTAGCTTCGCCTCCGGTTCTTCAACCCATTACACCGATTGCTTCACCACCTCGAT
TCATTCCAGGAGATGCACCAAAGGAGCCTCCTTTCTCAGGAAGAGTTACTCCAGCCCCGGTTT
CATCACCTGTTTCGGATATTCCTCCAATTCCATCAGTGGCTCTGCCTCCGCCTACACCGAGCAA
TGTACCTCCTAGAAATGCTTCTAACAATCACAAGCCTCCTCCTATCGAAAAGTCTATTGCTCCT
GTTGCATCACCACCAACAATATCAATTGACATTGCACCACCAGTACATCCCGTCATACCTAAG
TTAACACCCTCGAGCTCACCTGTACCTACCTCGACTCCAACTAAGGGATCTCCGAGAAGAAAT
CCGCCAACTACCCACCCAGTTTTCCCCATAGAATCTCCTGCTGTCTCACCAGATCATGCTGCAA
ATCCAGTAAAACATCCGCCCCCCAGCGATAACGGAGATGATAGTAAAAGTCCGGGTGCTGCA
CCGGCAAATGAAACTGCTAAACCCTTACCTGTCTTCCCACATAAAGCTTCTCCTCCTTCGATTG
CTCCCTCAGCCCCAAAATTTAACAGACACTCTCATCATACATCTCCATCTACTACTCCACCGCC
AGATTCTACTCCTAGTAATGTACACCACCATCCTTCATCACCATCTCCTCCTCCTCTATCTTCAC
ACCATCAACATCATCAAGAACGAAAGAAAATCGCAGATAGTCCAGCTCCTTCACCACTGCCA
CCGCATCTCATATCTCCAAAGAAGTCAAACCGTAAAGGTTCAATGACTCCTCCTCCACAATCG
CACCATGCTCCTTCTCCTCCCATTCCTGATTCTTTGATTTCTCCTGCTCACGCTCCAGTCTCTTT
CTCTATGAAACGCATCTCCCCGGCTCTAGCACCTTCTCCAACCCAAGTGTTTCCTCTCAGGTCC
TCTTCAAGACCTTCAAAATCTCGGAAATTCCCTCTAGGTCCACCTCTTCAAGCATTTCCTCCTC
CTCCCCCTAATTCAGATTGTTCTTCAACTATTTGTTTGGAACCGTACACAAACACACCTCCGGG
ATCTCCATGTGGCTGTGTCTGGCCAATTCAAGTTGAGCTACGCCTTAGCATGGCGCTCTACGA
CTTCTTCCCAATGGTTTCAGAATTTGCTCGGGAAATCAGCGCCGGAGTTTTCATGAAACAGAG
CCAAGTCCGTATAATGGGAGCTAACGCGGCCAGCGAACAACCTGAGAAATCCATTGTTCTAAT
CGATTTAGTACCGCTTGGAGATAAATTCGATAACATGACTGCAATGCTGACTTACCAGAGATT
CTGGAGTAAAAAAGTCTATATAGATGAACCAATCTTTGGCGGATACGACGTGATTTACGTGCG
TTATCCTGGTTTACCCGCTTCCCCGCCAACTTCTGGTATGACCATTATAGATCAAGGACCGTAC
TCTGGTAATAATAACGGAAGGGCGGTCAAACCGCTCGGAGTTGATGTTCCAAGGAAGCCGCG
CAAGAAAGAGCTCAATGGTGGAAGTATTGCTGTGATTGTTTTATCCGCAGCTGCTTTTATCGG
TTTATGTTTCGTTATCGTTTGGTTCTTGGTTTTCCGGCGACAAAGAGATCGACGACGTCTTTCG
AAAAGAACACCACTTGCTCGCCCTTCACTGCCTTCCCTCTCGAAACCATCAGGCTCGGCGAGA
TCTTTAACTGGAAGCCGATTTAGCTCAACTTCATTGTCATTTGAATCCAGCATTGCTCCGTTTA
CTCTCTCTGCAAAGACTTTCACCGCAAGTGAAATAATGAAAGCTACAAATAACTTTGATGAAT
CGAGGGTTCTTGGTGAAGGCGGATTTGGACGAGTTTACGAAGGTGTTTTCGA

FIGURA 27 (continuação)

TGACGGAACTAAAGTAGCAGTTAAGGTATTGAAGAGAGATGACCAGCAAGGTAGCAGGGAGT
TCTTAGCCGAAGTGGAAATGCTTAGTCGTCTTCATCACAGAAACCTCGTGAATTTAATTGGTA
TTTGTATCGAAGATCGTAACCGTTCCCTGGTTTATGAACTCATACCAAATGGCAGCGTTGAAT
CTCACCTCCACGGGATTGATAAAGCATCTTCGCCTTTAGATTGGGATGCTCGGTTGAAGATAG
CTCTTGGTGCGGCCCGTGGTTTAGCGTATTTACACGAAGACTCGAGCCCGCGAGTTATACACA
GAGACTTCAAGTCCAGCAACATCTTGCTGGAAAACGATTTCACACCTAAAGTATCTGACTTTG
GATTGGCTCGTAACGCCCTTGATGATGAAGATAACAGACATATATCGACACGCGTTATGGGAA
CGTTTGGGTATGTGGCACCGGAATACGCAATGACAGGGCATCTTTTGGTGAAGAGTGATGTGT
ATAGTTACGGAGTTGTGCTTCTCGAGCTACTTACAGGGCGGAAACCAGTGGATATGTCTCAAC
CGCCCGGACAAGAGAACTTAGTTTCATGGACTCGGCCTTTTCTTACGAGTGCAGAAGGGCTTG
CAGCTATTATAGATCAGTCTTTAGGACCCGAGATCTCATTCGATAGCATAGCTAAAGTCGCGG
CAATAGCTTCTATGTGCGTTCAACCAGAAGTATCACACCGTCCTTTCATGGGAGAAGTAGTGC
AAGCTTTGAAACTTGTCAGCAACGAATGTGATGAAGCTAAAGAACTCAATTCTTTAACTTCTA
TCTCTAAAGACGATTTTAGAGATGACACTCAAGCTGAGAGCTCTTGTGGTGACAGCTCAGCAA
GGATGGCTCGGTATCCATTGCTACCAAATTACGACTCTGAGCCTGATACAGAGAGAGGATTAT
CTACATCGGAAATGTACACGGGGTCAGGGAGGTTCGAGAGACAGTCTAACTCGGGTCCCTTG
ACCTCGGGTCGAGGCAAGAGTTTCTGGCAAAAGATGAGGAGATTATCGACCGGAAGCTTGAG
TGAGCATGGAACACCAACGGTCATGTTACGATCCGGTTCTCGCTAAACAATCTTTTGCCTACA
GAACACTGAAGAGTTTTTGATTGCCCGTTTAAGTTAAGAGACTGAAAGGAAAGAAGGTGCCTT
CTCCTACAAGCAAAACTCTTTGCCTCATGGAAACCGGTTAAGATAAAACCAGGAGTGATCATT
TCCCGGTTCAAAATTTTTAGTTGAATAGATCTGTTTATCTGAGAAGAAGAAACAAGAGATTCT
GTTAAATCTAATTTGAATGTACATATCACGTTTTAAAGTAACAGGCTTAAGCATTAAGTATCA
TCATCCTTC

SEQ ID N° : 22, Arabidopsis Extensina RLK At5g56890, seqüência de proteína (NP_680446)
MENLTLLLRICLVSSVLVAASSSGSELLSPLSSPPSPLPETSKGFGQAPSNSPESHKSDNVPPSKASSQ
PSLPPLADLAAPPPSDSVGGKAPAGVPVVFVPNAPAPATIPVKDLPVASPPVLQPITPIASPPRFIPGD
APKEPPFSGRVTPAPVSSPVSDIPPIPSVALPPPTPSNVPPRNASNNHKPPPIEKSIAPVASPPTISIDIAP
PVHPVIPKLTPSSSPVPTSTPTKGSPRRNPPTTHPVFPIESPAVSPDHAANPVKHPPPSDNGDDSKSPG
AAPANETAKPLPVFPHKASPPSIAPSAPKFNRHSHHTSPSTTPPPDSTPSNVHHHPSSPSPPPLSSHHQ
HHQERKKIADSPAPSPLPPHLISPKKSNRKGSMTPPPQSHHAPSPPIPDSLISPAHAPVSFSMKRISPA
LAPSPTQVFPLRSSSRPSKSRKFPLGPPLQAFPPPPPNSDCSSTICLEPYTNTPPGSPCGCVWPIQVEL
RLSMALYDFFPMVSEFAREISAGVFMKQSQVRIMGANAASEQPEKSIVLIDLVPLGDKFDNMTAM
LTYQRFWSKKVYIDEPIFGGYDVIYVRYPGLPASPPTSGMTIIDQGPYSGNNNGRAVKPLGVDVPR
KPRKKELNGGSIAVIVLSAAAFIGLCFVIVWFLVFRRQRDRRRLSKRTPLARPSLPSLSKPSGSARSL
TGSRFSSTSLSFESSIAPFTLSAKTFTASEIMKATNNFDESRVLGEGGFGRVYEGVFDDGTKVAVKV
LKRDDQQGSREFLAEVEMLSRLHHRNLVNLIGICIEDRNRSLVYELIPNGSVESHLHGIDKASSPLD
WDARLKIALGAARGLAYLHEDSSPRVIHRDFKSSNILLENDFTPKVSDFGLARNALDDEDNRHIST
RVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLELLTGRKPVDMSQPPGQENLVSWTRPFLTSA
EGLAAIIDQSLGPEISFDSIAKVAAIASMCVQPEVSHRPFMGEVVQALKLVSNECDEAKELNSLTSIS
KDDFRDDTQAESSCGDSSARMARYPLLPNYDSEPDTERGLSTSEMYTGSGRFERQSNSGPLTSGRG
KSFWQKMRRLSTGSLSEHGTPTVMLRSGSR

FIGURA 27 (continuação)

SEQ ID N° : 23, NM_179667.2| ligação de ATP de Arabidopsis thaliana / quinase/ proteína quinase/ proteína serina/treonina quinase/ proteína-tirosinaquinase AT2G20300 mRNA, cds completos

TTCTGGTTTCGAATTCACTTTACTCTCTTCGCTTCCAATCTCTCTCTACCACTCTGATACCTTCG
CCGGAGAAGAAGCTCTCGTTCTCTCTATCGTCGGAGAAGCTCTCGTTCTCTGCTTCCCTGCTCT
GTACAGATCTCGAGCCGCATTTCGTGGATCTGTCAAAATCGCGTTAAAACCTACTCGCTTCTCT
CTAACCCTAGGGTTTTTGAATTTTTTTCGGGACGAGGGAGAAATCAATGTCTGTGAAGCTTTG
AGGAAGCTATATGGTTACACCTTTGGTTCTGGTTAGTCTTCGTTTTGATCGGCGGAGAGTGTTC
GCCGAGGCTTGGTGTCTCCGTTATTGACTAATGGTGGTTTTTGATTTGAGATAAGAGATGCGG
AACTTTGCGATGCTTCTGCTGTTAATTCTTCTTCTTCACTCTCTCGCCTCGTTTCCTATATGCTTT
GCTCGGTTATTTCCGATGAGTTTGCCATTTACCCGATCAAAGGCGCATCAAATGCATTTCTTTC
ATCCTTATCTTAATCCATCAGTTGCTCCTACACCTTCACCAGCTTTTTCCCCCAACCCAAGTCG
CATTCCACCACTAAGGCACAAGGGTCATCACCGTCATCGTCGCTGGCATTTAAGACGCAATGC
CACTGCAGTGTCACCTTCGTCACATGACTGTCAGCAGACATGTGTGGAGCCTCTCACTTCAAC
TCCCTTTGGTTCACCTTGTGGTTGTGTCTTCCCCATGAAAGTTCAGCTTCTACTAAGTGTCGCA
CCTTTTTCTATTTTCCCCGTGACCAACGAGTTAGAAATTGAGGTTGCTGCTGGAACATACTTGG
AACAAAGCCAAGTGAAAATAATGGGTGCCAGTGCCGACAGTGAAAATCAAGGGAAAACAGT
GGTGGATATTAACCTTGTTCCACTTGGGGAAAAGTTCGACAACACTACAGCAACACTTATTTA
TCAGAGATTCCGGCACAAGAAAGTACCTCTAAATGAGACTGTTTTTGGTGATTATGAGGTTAC
GCACATAAGTTATCCAGGAATTCCTTCTTCTTCGCCAAATGGAGATGTTACTGGAGATGCTCC
AGGAGGCCTTCCAATTCCGATCAATGCAACAACTTTTGCCAACAAAAGCCAGGGAATAGGTTT
TAGAACGATTGCAATTATTGCGTTGTCAGGTTTTGTCCTCATTCTGGTTTTGGTTGGAGCTATT
TCTATAATCGTGAAATGGAAGAAAATTGGGAAGTCATCTAATGCTGTTGGTCCGGCTTTGGCT
CCATCGATTAACAAAAGGCCTGGTGCTGGATCTATGTTTTCCAGTAGCGCCAGAAGCTCAGGC
TCAGATTCTTTGATGTCTTCCATGGCTACATGTGCTTTATCAGTTAAGACCTTCACACTCTCCG
AGCTTGAGAAGGCTACTGATAGATTCAGTGCCAAGAGGGTTTTGGGCGAGGGCGGATTTGGT
CGTGTTTATCAGGGGAGTATGGAAGATGGAACTGAGGTTGCAGTGAAACTGCTAACTAGGGA
CAATCAGAATCGAGACCGTGAATTTATTGCCGAAGTCGAGATGCTAAGCCGTTTGCACCACCG
CAATCTTGTGAAACTGATTGGAATATGCATTGAAGGTCGTACACGTTGCTTGATTTATGAGCT
CGTCCACAATGGGAGTGTTGAGTCCCACTTACACGAAGGAACGCTTGATTGGGATGCGCGGTT
GAAGATTGCACTTGGAGCAGCGAGAGGACTGGCTTATCTCCATGAAGATTCAAATCCCCGAGT
AATCCACCGTGATTTCAAGGCTAGCAATGTTCTCCTAGAAGATGACTTTACCCCAAAAGTCTC
AGACTTTGGACTGGCAAGAGAAGCAACCGAAGGAAGTCAACATATTTCAACTCGAGTCATGG
GGACCTTTGGGTACGTGGCTCCTGAGTATGCAATGACGGGGCATCTCCTTGTTAAAAGCGATG
TCTATAGTTATGGTGTAGTTCTACTAGAGCTTCTAACCGGTAGAAGACCAGTAGACATGTCTC
AACCTTCGGGAGAGGAAAACCTTGTGACATGGGCGAGACCCTTACTAGCCAACAGAGAGGGC
CTGGAGCAACTGGTGGACCCTGCATTGGCTGGAACCTACAACTTTGATGACATGGCGAAAGTG
GCTGCGATTGCCTCCATGTGCGTCCACCAAGAGGTCTCACACAGACCATTCATGGGTGAAGTT
GTGCAGGCACTAAAACTTATATACAACGATGCAGATGAGACCTGTGGTGATTATTGCAGCCAA
AAGGACTCATCAGTACCAGACTCTGCTGACTTCAAAGGCGATCTGGCTCCTTCGGATAGCAGC
TGGTGGAATTTGACACCTCGGTTAAGGTATGGTCAAGCTTCGTCGTTCATAACAATGGATTAT
AGCTCAGGACCACTAGAGGACATGGAGAACCGGCCTCACTCTGCCTCTAGCATTCCAAGAGT
AGGAGGTCTTATTCTACCAAACAGATCAGGTCCGTTAAGACCCATGCGTAGTAGAAGAAACTT
CTTTAGACTGAGAGGAAGCATGAGCGAACACGGTGGACCGTCGTCGTCTAGACATCTCTGGTC
CGGCAATGGTGACTGGCTCTGAGAAACCAGAAAATGTACAGTGAAGAAAAGGGAAAAGGAG
GCTCACAAACTCTTGAGCTGCATGGTTAGGTTCTGGTTAATCCGGCCAGGTTCGGTTCATCTGG
GTAAAGATTTCCGGTTTGATTTCTTTTAGGAGCTGTTTGTTGTTATGATTATCCTCAATAGAGA
TTTAATTGTTTGTTTTTGCGAGGATGCAAAAGCACAGGGAGGGTTGTATTTTGAAGGACGCCT
GTATTTAGTTCTCTCCTTTTCTCTCTCACCCACAAAAAAACATTTTTATATATTTTATTTTAGTG
AAGTTAGTAATTACCTAA

FIGURA 27 (continuação)

TTTTTTTTAGTTTTTCAAGTTCCTAAGAAATTTGGCTTTTGGTCATCGTTGTTAAGTTTCCGGTT
TTCTGAATTTACATCCGAAATATTTTTTGGTTAAATAAAATTCTGGCTTGAGAGTAATTGGCTG
GTGAATGGTAACAAGTCACTTGGTCAAAGCTGTTGTAATTCTTTTCTAATTTCTAGTATCTTCT
GTAGATTGGACACAATGACATGGATTGTTGATCTTTAAAGTCGT

SEQ ID N° : 24, NP_849998.1| ligação de ATP / quinase/ proteína quinase/ proteína serina/treonina quinase/ proteína-tirosinaquinase [Arabidopsis thaliana], AT2G20300
MRNFAMLLLLILLLHSLASFPICFARLFPMSLPFTRSKAHQMHFFHPYLNPSVAPTPSPAFSPNPSRIP
PLRHKGHHRHRRWHLRRNATAVSPSSHDCQQTCVEPLTSTPFGSPCGCVFPMKVQLLLSVAPFSIF
PVTNELEIEVAAGTYLEQSQVKIMGASADSENQGKTVVDINLVPLGEKFDNTTATLIYQRFRHKKV
PLNETVFGDYEVTHISYPGIPSSSPNGDVTGDAPGGLPIPINATTFANKSQGIGFRTIAIIALSGFVLIL
VLVGAISIIVKWKKIGKSSNAVGPALAPSINKRPGAGSMFSSSARSSGSDSLMSSMATCALSVKTFT
LSELEKATDRFSAKRVLGEGGFGRVYQGSMEDGTEVAVKLLTRDNQNRDREFIAEVEMLSRLHH
RNLVKLIGICIEGRTRCLIYELVHNGSVESHLHEGTLDWDARLKIALGAARGLAYLHEDSNPRVIH
RDFKASNVLLEDDFTPKVSDFGLAREATEGSQHISTRVMGTFGYVAPEYAMTGHLLVKSDVYSYG
VVLLELLTGRRPVDMSQPSGEENLVTWARPLLANREGLEQLVDPALAGTYNFDDMAKVAAIASM
CVHQEVSHRPFMGEVVQALKLIYNDADETCGDYCSQKDSSVPDSADFKGDLAPSDSSWWNLTPR
LRYGQASSFITMDYSSGPLEDMENRPHSASSIPRVGGLILPNRSGPLRPMRSRRNFFRLRGSMSEHG
GPSSSRHLWSGNGDWL

SEQ ID N° : 25, Osi093820.1 Oryza sativa, ortólogo At1g54820 putativo
ATGCCAACTAGTAGAGTCGACCTGCTGAGCCGGATGCACTCGCCGTACCTGGTGGGGTTGCTG
GGCTACTGCGCCGACCAGAGCCACCGTCTGCTGGTGTTCGAGTTCATGCCCAACGGCAGCCTC
AAGAGCCACCTCCACCGGCGCGCGCTGGCGCCGGCGGAGCAGCCGCCGCCGCTGGACTGGCA
GACGCGGCTGGGCATCGCGCTGGACTGCGCCCGCGCGCTGGAGTTCCTCCACGAGCACAGCTC
CCCCGCCGTGATCCACCGCGACTTCAAGTGCAGCAACATCCTCCTCGACCACAACTACCGCGC
CCGCGTCTCCGACTTCGGCATGGGCAAGCTCGGCTCCAACAAGGCCAATGGCCAG

SEQ ID N° : 26, Osi093820.1 seqüência de proteína deduzida
MPTSRVDLLSRMHSPYLVGLLGYCADQSHRLLVFEFMPNGSLKSHLHRRALAPAEQPPPLDWQTR
LGIALDCARALEFLHEHSSPAVIHRDFKCSNILLDHNYRARVSDFGMGKLGSNKANGQ

SEQ ID N° : 27, Osi015947.1 Oryza sativa, ortólogo de At1g54820 putativo
ATGGCGGATGCTTGGGCGGCGTCCCCTCCATCTCCTACGGCTCCTTCGCCTTTCGCCGCGGACG
ACCTCGTCGACGCGCGGCTCGCGCCGTGGCCGCCGTTCGCCCCGTGGCCGGCGCCCGGGCAGC
TCCACCACAACCGCCGCGGCGGCGGGCACCCGAACCCGCTCTTCACCATCCTGCCGGCCTCGG
CGCTGGCAATAGGGGTCGTGCTGCTTGTCGCCGTGGTCGTCATCCTGGTGATGACGCGTCGGT
GGAAGCCCGGGGCGGTGGACGGCGCCGGCGGCGCCAGCTGCAACGGCGACAAGCCTGGCGG
CGCCCCCGCGTCCAGCTGCGGCAGCAGCGTCCGCGGCTACAACAACTCCAGGTACTACGCCGC
CGCCGCCGCCGGGTGCATATATGGCGGAAGGCTGGGTTTCTCGGTGCAGCCGCGGAACCGCG
GCGCGCAAGTGTTCACGTACCGGGAGCTGGAGAGCGCGACGGACGGGTTCAGCGAGTGCAAC
GTGGTGGGCCGCGGCGCGTACGGCGTGGTCTTCCGCGGCCGGCTCGGCGACGGCACCACGGC
CGCCATCAAGCGGCTCAAGATGGACGGCCGGCGCGAGGGCGAGCGCGAGTTCCGCATCGAGA
TGGGCGTTGCTATTACTGCTCAGGTCGACCTGCTGAGCCGGATGCACTCGCCGTACCTGGTGG
GGTTGCTGGGCTACTGCGCCGACCAGAGCCACCGTCTGCTGGTTTTCGAGTTCATGCCCAACG
GCAGCCTCAAGAGCCACCTCCACCGGCGC

FIGURA 27 (continuação)

GCGCTGGCGCCGGCGGAGCAGCCGCCGCCGCTGGACTGGCAGACGCGGCTGGGCATCGCGCT
GGACTGCGCCCGCGCGCTGGAGTTCCTCCACGAGCACAGCTCCCCCGCCGTGATCCACCGCGA
CTTCAAGTGCAGCAACATCCTCCTCGACCACAACTACCGCGCCCGCGTCTCCGACTTCGGCAT
GGCCAAGCTCGGCTCCAACAAGGCCAATGGCCAGGTGGCCGCCATCACGGCGATGTGCATAC
AGACGAAGGCGGACTACCGGCCGCTGATGACGGACGTGGTGCAGTCGCTGATCCCCATCGTG
AAGAGCCCACTCATGTCCTGCACCTCCACGCCGCTGAGGCCCGCGCACGGGCACCACCACGTC
GTCTACATGAGCCCGAGCAGGGGCTCTTCCAACGGCGGTGCCCTGGAAACGCGATGCGTCAT
GCACGGCTTGGATTGA

SEQ ID N° : 28, Osi015947.1 seqüência de proteína deduzida, AK106842
MADAWAASPPSPTAPSPFAADDLVDARLAPWPPFAPWPAPGQLHHNRRGGGHPNPLFTILPASAL
AIGVVLLVAVVVILVMTRRWKPGAVDGAGGASCNGDKPGGAPASSCGSSVRGYNNSRYYAAAA
AGCIYGGRLGFSVQPRNRGAQVFTYRELESATDGFSECNVVGRGAYGVVFRGRLGDGTTAAIKRL
KMDGRREGEREFRIEMGVAITAQVDLLSRMHSPYLVGLLGYCADQSHRLLVFEFMPNGSLKSHLH
RRALAPAEQPPPLDWQTRLGIALDCARALEFLHEHSSPAVIHRDFKCSNILLDHNYRARVSDFGMA
KLGSNKANGQVAAITAMCIQTKADYRPLMTDVVQSLIPIVKSPLMSCTSTPLRPAHGHHHVVYMS
PSRGSSNGGALETRCVMHGLD

SEQ ID N° : 29, Osi003977.2 Oryza sativa, ortólogo de At3g58690 putativo
ATGTCGAGCGGCGGCGGCGGCGGCGACGACTACAAGCGGGAGGAGTCGGTGGCACTCATGGT
CATCGTCTCCCTCGCGGCGCTCTCCCTCCTCTCCCTCGTCGCCGCCTTCGCCTACTACTGCTAC
ATCACCCGCAAGGTCTCCCGCCGCCTCCACTCGCTCGACCTCCCCAAGCACCACCGCCGCTCC
TCCTCCTCCTCGCCTCCTCCCATGCCCCCGCCGCTGCCTCCCCCTCCTCCTTCCGCGAATGCCCC
GACGCTCGGGAAGGAGAGCCCGTCCAGCAACTCCGCCAGCGACGGCGCGGCGGCGGCGGTCG
TGGTCGGCGGGGAGAGGGGGGCCGTCCAGGTGTTCAGCTACCGCCAGCTGCACGCCGCCACG
GGCGGGTTCGGGCGGGCGCACGTGGTGGGGCAGGGGAGCTTCGGGGCCGTCTACCGCGGGGT
GCTCCCCGATGGGAGGAAGGTCGCGGTCAAGCTCATGGATCGCCCCGGCAAGCAGGGGGAAG
AGGAGTTCGAGATGGAGGTGGAGCTGCTGAGTCGGCTTCGATCACCTTATCTTTTGGGCTTGA
TTGGCCATTGTTCAGAGGGTGGACACCGGCTGCTGGTCTATGAATTCATGGCCAATGGGGGTC
TCCAGGAGCATCTCTACCCAAATGGAGCTTTTGAAAAGATTGAAACATTTTCGATCTACTTGG
TGAAACAACGCCCGATTTTTGACAAAAATATCGGGCACCCGATTTGTAGGCGCGCCCATTTTT
CTCGTCAACTGATATCCCACAAAGCTGACATTGTAGGTTCCTGTGGTGGTATCTCAAAATTGG
ACTGGCCTACTAGAATGAGAATAGCTCTTGAAGCAGCAAAAGGTCTGGAATATCTTCACGAG
CGTGTTAATCCACCTGTTATACACAGAGACTTCAAGAGCAGCAACATTCTACTGGACAAGGAC
TTCCGTGCAAGAGTGTCAGATTTTGGTTTGGCTAAGCTTGGATCTGATAGAGCTGGTGGTCAT
GTATCAACTCGAGTTCTAGGCACACAAGGTTATGTTGCACCAGATTACGGTGTTGTACTTCTG
GAGTTGCTTACTGGCAGGGTGCCAGTTGATATGAAGAGGCCTCCAGGCGAAGGTGTTCTAGTT
AACTGGGCATTGCCTATGCTCACAGACCGAGAGAAGGTTGTGCAGATCTTGGATCCTGCACTG
GAGGGTCAATATTCATTGAAAGATGCTGTCCAAGTGGCTGCCATTGCTGCCATGTGTGTTCAA
CAAGAGGCTGACTACAGGCCACTAATGGCTGATGTGGTTCAATCGTTGGTTCCGCTTGTCAAG
AATCGTTCTACTCCAAAGACGTGCAACCCTAGTGTCCAAGCGTAG

SEQ ID N° : 30, Osi003977.2 seqüência de proteína deduzida
MSSGGGGGDDYKREESVALMVIVSLAALSLLSLVAAFAYYCYITRKVSRRLHSLDLPKHHRRSSSS
SPPPMPPPLPPPPPSANAPTLGKESPSSNSASDGAAAAVVVGGERGAVQVFSYRQLHAATGGFGRA
HVVGQGSFGAVYRGVLPDGRKVAVKLMDRPGKQGEEEFEMEVELLSRLRSPYLLGLIGHCSEGG
HRLLVYEFMANGGLQEHLYPNGAFEKIETFSIYLVKQRPIFDKNIGHPICRRAHFSRQLISHKADIVG
SCGGISKLDWPTRMRIALEAAKGLEYLHERVNPPVIHRDFKSSNILLDKDFRARVSDFGLA

FIGURA 27 (continuação)

KLGSDRAGGHVSTRVLGTQGYVAPDYGVVLLELLTGRVPVDMKRPPGEGVLVNWALPMLTDRE
KVVQILDPALEGQYSLKDAVQVAAIAAMCVQQEADYRPLMADVVQSLVPLVKNRSTPKTCNPSV
QA

SEQ ID N° : 31, AK102665.1 Oryza sativa (grupo de cultivar japônico) seqüência parcial de cDNA, variante de união de Osi003977

GGCCCCCAAACCCCTCAAAACCCTCGTCGTCTTCTCGCGATCGCTATTCCCTCCTCCATTCCTC
CTCCTCCACCTCCGAGACCTAGGGTTCCAATGTCGAGCGGCGGCGGCGGCGGCGACGACTAC
AAGCGGGAGGAGTCGGTGGCACTCATGGTCATCGTCTCCCTCGCGGCGCTCTCCCTCCTCTCC
CTCGTCGCCGCCTTCGCCTACTACTGCTACATCACCCGCAAGGTCTCCCGCCGCCTCCACTCGC
TCGACCTCCCCAAGCACCACCGCCGCTCCTCCTCCTCCTCGCCTCCTCCCATGCCCCCGCCGCT
GCCTCCCCCTCCTCCTTCCGCGAATGCCCCGACGCTCGGGAAGGAGAGCCCGTCTAGCAACTC
CGCCAGCGACGGCGCGGCGGCGGCGGTCGTGGTCGGCGGGGAGAGGGGGGCCGTCCAGGTGT
TCAGCTACCGCCAGCTGCACGCCGCCACGGGCGGGTTCGGGCGGGCGCACGTGGTGGGGCAG
GGGAGCTTCGGGGCCGTCTACCGCGGGGTGCTCCCCGATGGGAGGAAGGTCGCGGTCAAGCT
CATGGATCGCCCCGGCAAGCAGGGGGAAGAGGAGTTCGAGATGGAGGTGGAGCTGCTGAGTC
GGCTTCGATCACCTTATCTTTTGGGCTTGATTGGCCATTGTTCAGAGGGTGGACACCGGCTGCT
GGTCTATGAATTCATGGCCAATGGGGGTCTCCAGGAGCATCTCTACCCAAATGGAGGTTCCTG
TGGTGGTATCTCAAAATTGGACTGGCCTACTAGAATGAGAATAGCTCTTGAAGCAGCAAAAG
GTCTGGAATATCTTCACGAGCGTGTTAATCCACCTGTTATACACAGAGACTTCAAGAGCAGCA
ACATTCTACTGGACAAGGACTTCCGTGCAAGAGTGTCAGATTTTGGTTTGGCTAAGCTTGGAT
CTGATAGAGCTGGTGGTCATGTATCAACTCGAGTTCTAGGCACACAAGGTTATGTTGCACCAG
AATATGCTTTGACCGGGCATTTGACGACCAAATCCGATGTCTATAGTTACGGTGTTGTACTTCT
GGAGTTGCTTACTGGCAGGGTGCCAGTTGATATGAAGAGGCCTCCAGGCGAAGGTGTTCTAGT
TAACTGGGCATTGCCTATGCTCACAGACCGAGAGAAGGTTGTGCAGATCTTGGATCCTGCACT
GGAGGGTCAATATTCATTGAAAGATGCTGTCCAAGTGGCTGCCATTGCTGCCATGTGTGTTTA
ACAAGAGGCTGACTACAGGCCACTAATGGCTGATGTGGTTCAATCGTTGGTTCCGCTTGTCAA
GAATCGTTCTACTCCAAAGACGTGCAACCCTAGTGTCCAAGCGTAGAAGCCACTTGAAGGGG
AAGTTGAATGCTCAGTTTCCAGGTTGGTGCTTGACTTTTCTGGAAATTTTCATCTACATCTAAC
AACTACTGCTGAGTTTATGAGATCAGTATGCTCCATAACTTAATTCTCTTCCTCTGCCAGTGAA
TTCCTGATTTTGCGAGGTGAATCGTAAGCAGTTAGATTTAGAAATCAAGCATTAATTTAGTTC
GATGGACCAGAAAGCTAGTTTCAGTTTGAAGGGAAGTCCTCTAGTCATGTGCAATTAGTGATC
ATGTTAGGTGGTTTGGCTTAGTACTTATGGTTGTGCCCTGTGGGAATGATCAGAGAGTGTTGTT
TCCAGCTGGCAGAATTGCAATGTACTGCTGAATTTTCTTTTGGCTTGTCAATTATGTTAGGTGT
CTTTGGAGATCCATGTTCGTTGTCTTAACTCTTGTGCTAATACGTTGGTCTCATCAGCTTGGCC
TATAAACATCGCCGATGTATCTTTTTGTATATGTATATAGTAGGATTCTGGAACTTATAAAAAC
AAACATTTGCCAGTTGAGCATTTTCATGC

SEQ ID N° : 32, AK102665, seqüência de proteína deduzida (parcial) Oryza sativa, variante de união de Osi003977

GPQTPQNPRRLLAIAIPSSIPPPPPPRPRVPMSSGGGGGDDYKREESVALMVIVSLAALSLLSLVAAF
AYYCYITRKVSRRLHSLDLPKHHRRSSSSSPPPMPPPLPPPPPSANAPTLGKESPSSNSASDGAAAAV
VVGGERGAVQVFSYRQLHAATGGFGRAHVVGQGSFGAVYRGVLPDGRKVAVKLMDRPGKQGE
EEFEMEVELLSRLRSPYLLGLIGHCSEGGHRLLVYEFMANGGLQEHLYPNGGSCGGISKLDWPTR
MRIALEAAKGLEYLHERVNPPVIHRDFKSSNILLDKDFRARVSDFGLAKLGSDRAGGHVSTRVLGT
QGYVAPEYALTGHLTTKSDVYSYGVVLLELLTGRVPVDMKRPPGEGVLVNWALPMLTDREKVV
QILDPALEGQYSLKDAVQVAAIAAMCV

FIGURA 27 (continuação)

SEQ ID N° : 33, Osi000040.5 Oryza sativa, ortólogo de At4g02010 putativo

ATGGCGCGGATACGGCGCGCGAGCTCCGGACGAGACATGTCAGATAATTTGGTGCAGCATAA
CAGACGCTCAGAAGCAGAAATTTCTACTCATGTAGATGCTGCGCCTCCTGATGCAGCCTCAAA
TACTAGTGCAGCTCCTTCTGGATTAGTACAACCTCCAGTATCTCCTCACAATGCATGTTGTTCA
CATAACATGGTACAAAAGAGAGGGAGCCAAGATTGCCATTGTGTTTACCCAGTAAGAGTTGA
GCTTTTTCTCCGAAATGTTTCACTGACGTCAAATTGGAGCGATGAATTTCTGGGGGAACTTGCT
TCTCAACTCAGTCTGAGGGTTACTCAGTTTGAGATTGTAAATTTCTATGTTGTTGGGGCTTCTG
GGCTAAACATCACAATGTATATAGCTCCCCATACAGGAATTAGTTTCTCAGCCGACCAAGTTA
CCGCAATGAATTATTCACTTAGTCAGCATACAGTTCAGATCAACCCTGTACTAGTTGGTGACT
ACAATCTTCTTAATTTAACCTGGTTCAGGCCACTGGTTCTAGCTCCTGTTTATTTTAGTGTTAA
GGTAGAGAACACTAAAAGGATGTTTGGTGATATCAATATGATACTTCCTATGAATGATATAAA
CATTATTTCTTTATATTTACAGCCCTCTCCAACATTCACAATATCACCTAAACCCTCTCCATCTC
AAGCATCTACAGTACCAAGACACAGTGCGGATACCAGCAATGAAAAACACATGAGCTTGATT
ACTATCATTTGCATATTTATTGGTGCTCTAATTGCTGTCTTGGTGATTGCCATGTTTATTTGCTT
CTGCAAATTAAGGAAAGGGAAAAGGAAGGTTCCTCCAGTTGAGACACCCAAGCAAAGAACAC
CAGATGCGGTTTCTGCAGTGGACTCACTTCCTCGCCCAACAAGTACAAGGTTCCTTGCATATG
ATGAACTGAAAGAAGCAACCAACAACTTTGATCCGTCAAGTATGCTTGGAGAAGGAGGTTTT
GGCCGTGTCTTTAAAGGGGTACTAACTGATGGCACTGCCGTTGCTATAAAGAAGCTTACTAGT
GGAGGGCACCAAGGAGATAAGGAATTTCTAGTTGAAGTGGAGATGCTGAGCAGGTTGCACCA
CCGAAACCTTGTGAAACTCATTGGTTACTATAGTAACCGTACATTGGGAGCTAGCCGCCCCTT
GGACTGGGACACTAGAATGAGGATAGCACTAGATGCTGCCAGGGGATTGGCATACCTTCATG
AGGATTCTCAGCCTTGCGTAATCCACAGGGATTTCAAGGCTTCTAATATATTGCTTGAGGATG
ACTTTCATGCTAAAGTGTCTGATTTTGGTTTGGCCAAACAGGCACCGGAAGGTCGCACAAATT
ATCTCTCAACACGTGTTATGGGAACATTCGGGTATGTAGCACCAGAGTATGCAATGACAGGAC
ACTTGCTTGTAAAAAGCGATGTATATAGCTATGGAGTTGTTCTACTTGAACTACTTACTGGGA
GGAGGCCTGTGGATATGTCACAACCTTCTGGCCAGGAAAATCTAGTGACATGGCTTACACAAA
TGTTTCCTGATCTTTCCACTAAAAGCCTTGTGTCACATCAACCTTTGCTGGCCAAGTTGTCAGG
TGTAGAGATACTTATTTGCTCAATTTTGACTACGCAGGCACGACCAATTTTACGAGATAAAGA
TACGCTAGAAGAACTAGCTGATCCCAAGCTTGGTGGCCAGTATCCAAAAGATGATTTTGTGCG
AGTATGCACAATTGCAGCCGCCTGTGTTTCACCGGAGGCAAGCCAAAGGCCAACAATGGGCG
AGGTGGTGCAGTCACTGAAGATGGTCCAACGCTCTGAGTTCCAGGAATCCATACCAACACCTC
CAGCACGACCCAATGCCTCAGCCCCTTTGAGACTGAGAACATATCGAGAACGGCTTTCTCTGA
AGATCTTCACGAAGGGCGGTAACAGTGGTGAACAAGACCCGGCCACCATGGTCGGCTTCGCC
GACCAGACCGCCGACGTCTCCGCCGCCGCGTTCCAGACCATGTACCGCCTCAACGTCGGCGGC
GCCTACATCCCGCCGTCCAACGACTCCGGCCTCACGCGGCCGTGGTACGACGACACGCCGTTC
GTCCAGGGCCCCCTGCGCGGCCTCGTCTACAACGCCGGCCCGCACTTCCACATCAAGTACCCC
AGCGACGCCGCCGAGTACGCCGCGCCGCCGGAGGTGTACCTCGGCGGCCGCTCCATGGGCAG
GGACCAGCGCCTCAACCAGAACTCCAACCTCACCTGGAGCCTCCACGTGGAGTGCAACTTCAC
CTACGTCGTGCGCCTCCATTTCTGCGAGCTCCAGCTCATCCACGGCAACCAGCGGGTGTTCGA
CATCTACATCAACAACCGGACGGCGCAGACGGACGTGGACGTGCTGGAGATGGCGACGGAGC
GCGGCGTGCCGGTGTACAAGGACTACGCGGTGAGGCTGAGCAACGACACCGCCGACGAGCAT
CTGTGGGTGGCGGTGCACCCCTCGGTGATGCTCCGCCCGCAGTTCTACGACGCCATCCTCAAC
GGCCTCGAGGTGTTCAAGGTGAACAACACCGGCGGCAGCCTGGCGTCGCCGGACCCCGTCCC
GTACAAGCTGCTCGCGGAGAAGGAGCTCGGCTGGGGCGGGCCGCCGGAGTTCTCCACCGACA
ACCCGGCCAACATGGCGTCGGTGATGGGCGGCACGGCGGGCGGCGCGGCGGCGGCCGGGATC
GTGGCGGCCATCTGCGTGGTGGTGTACAGCAACAAGAGGAGCAAGAAGCTGGGCGGCGGCGG
CGCCGACTCGCACACCTCCGCGTGGCTGCCGCTGTACCACTCCCACACCAGCGGCAAGTCGTC
GGGGCACATCACGGCGAACATCGCCGGCATGTGCCGCCACTTCTCGTTCGCGGAGATCAAGG
CGGCGACCAAGAACTTCTCCAACGACCTGGCCATCGG

FIGURA 27 (continuação)

CGTGGGCGGGTTCGGCGTGGTGTACCGCGGCGTGGTGGACGGCGACGTGAAGGTGGCGGTGA
AGCGGTCCAACCCGTCGTCGGAGCAAGGCATAACCGAGTTCCAGACGGAGGTGGAGATGCTC
TCCAAGCTCCGCCACCGCCACCTCGTCTCCCTCATCGGCTTCTGCGAGGAGGACGGCGAGATG
GTGCTCGTCTACGACTACATGGAGCACGGCACCCTGCGCGAGCACCTCTACCACAACGGCGGC
AAGCCCACGCTGTCGTGGCGCCACCGCCTCGACATCTGCATCGGCGCCGCCCGCGGCCTCCAC
TACCTCCACACCGGCGAATCGCACGTCAGCACCGTCGTCAAGGGCAGCTTCGGCTACCTCGAC
CCGGAGTACTACCGCCGGCAGCAGCTCACCGACAAGTCCGACGTCTACTCCTTCGGCGTCGTG
CTGTTCGAGGTGCTCATGGCGCGCCCGGCGCTCGACCCGGCGCTGCCGCGCGACCAGGTCAGC
CTCGCCGACTACGCGCTCGCCTGCAAGCGCGGCGGCGCGCTGCCGGACGTCGTCGACCCGGC
GATCAGGGACCAGATCGCGCCCGAGTGCCTCGCCAAGTTCGCCGACACGGCGGAGAAGTGCC
TCTCCGAGAACGGCACCGAGCGCCCCACCATGGGCGACGTGCTCTGGAACCTCGAGAGCGCG
ATGCACTTCCAGGACGCGTTCGACGCCGCCGCCGGCCGCCCCGTCCCCGCGCTCGACGCCGCC
GCCGGCAGCAGCAGCCACCTCGACGACGGGAGCACGGCCAGCATCAACACGCTGGCCACGTC
GTCCACGTCGCACCCGCACGAGCCGTGCGTCGACGTTGTCTTGGAGCCCGACGACGTCGTCGC
GGAGCGCGCCACCTTCTCGCAGCTCGTCCAGCCCACCGGCCGGTGA

SEQ ID N° : 34, Osi000040.5, seqüência de proteína deduzida
MARIRRASSGRDMSDNLVQHNRRSEAEISTHVDAAPPDAASNTSAAPSGLVQPPVSPHNACCSHN
MVQKRGSQDCHCVYPVRVELFLRNVSLTSNWSDEFLGELASQLSLRVTQFEIVNFYVVGASGLNI
TMYIAPHTGISFSADQVTAMNYSLSQHTVQINPVLVGDYNLLNLTWFRPLVLAPVYFSVKVENTK
RMFGDINMILPMNDINIISLYLQPSPTFTISPKPSPSQASTVPRHSADTSNEKHMSLITIICIFIGALIAV
LVIAMFICFCKLRKGKRKVPPVETPKQRTPDAVSAVDSLPRPTSTRFLAYDELKEATNNFDPSSML
GEGGFGRVFKGVLTDGTAVAIKKLTSGGHQGDKEFLVEVEMLSRLHHRNLVKLIGYYSNRTLGAS
RPLDWDTRMRIALDAARGLAYLHEDSQPCVIHRDFKASNILLEDDFHAKVSDFGLAKQAPEGRTN
YLSTRVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLELLTGRRPVDMSQPSGQENLVTWLTQ
MFPDLSTKSLVSHQPLLAKLSGVEILICSILTTQARPILRDKDTLEELADPKLGGQYPKDDFVRVCTI
AAACVSPEASQRPTMGEVVQSLKMVQRSEFQESIPTPPARPNASAPLRLRTYRERLSLKIFTKGGNS
GEQDPATMVGFADQTADVSAAAFQTMYRLNVGGAYIPPSNDSGLTRPWYDDTPFVQGPLRGLVY
NAGPHFHIKYPSDAAEYAAPPEVYLGGRSMGRDQRLNQNSNLTWSLHVECNFTYVVRLHFCELQ
LIHGNQRVFDIYINNRTAQTDVDVLEMATERGVPVYKDYAVRLSNDTADEHLWVAVHPSVMLRP
QFYDAILNGLEVFKVNNTGGSLASPDPVPYKLLAEKELGWGGPPEFSTDNPANMASVMGGTAGG
AAAAGIVAAICVVVYSNKRSKKLGGGGADSHTSAWLPLYHSHTSGKSSGHITANIAGMCRHFSFA
EIKAATKNFSNDLAIGVGGFGVVYRGVVDGDVKVAVKRSNPSSEQGITEFQTEVEMLSKLRHRHL
VSLIGFCEEDGEMVLVYDYMEHGTLREHLYHNGGKPTLSWRHRLDICIGAARGLHYLHTGESHVS
TVVKGSFGYLDPEYYRRQQLTDKSDVYSFGVVLFEVLMARPALDPALPRDQVSLADYALACKRG
GALPDVVDPAIRDQIAPECLAKFADTAEKCLSENGTERPTMGDVLWNLESAMHFQDAFDAAAGR
PVPALDAAAGSSSHLDDGSTASINTLATSSTSHPHEPCVDVVLEPDDVVAERATFSQLVQPTGR

SEQ ID N° : 35, NM_187872.1 Oryza sativa (grupo de cultivar japônico), Ozsa8316 mRNA predito
ATGGCGCGGATACGGCGCGCGAGCTCCGGACGAGCTCAACTCTCAGGCCGTGAGAATTTGGG
TTTGGTGTCGCGGGAATGGTGGTCATACTACTATGTAGGATTCAGTGCTCTCTGCTGCCGTGGC
GATACTGGATATCTGGAATCTGATCGAAGAACCTCTGATCTAAGTTTGAGTCACTACAAATGG
ATTTGTGATATTGATTATCAAATTCAGGCCAATGGATCCTCCTTTACAAGAAAATGGTCATTAC
TGAACTCTTGTTACTGGAATGATGAGAATTCCTGCATTGCTGGGTTTGCGGAAAGGCTTTTGG
AAGATGCTCGAAGGAAAACCTTCTTAAATTTTTTGGCTGTGGTATTTACTCTGGAATTGATTAC
CAGAATCAATTTGATTGCCCAAACATCAGCTCTAAATGTGGTTGCCCCTGCTATATCTCCATCT
CAAAGTTGGAG

FIGURA 27 (continuação)

ACCAGTTCGCTCCATGTTGTCAAAAGCTAAAGTTGACATTAGTATTTCAGTCAGCGAACAAAG
ACGAAAGAAGCTATATTCATCACCAGCTACTCTATCTGTGCACCCTCCAATGTCAGCGCCAAG
CTATAGCTCCATTTCTGACATGTCAGATAATTTGGTGCAGCATAACAGACGCTCAGAAGCAGA
AATTTCTACTCATGTAGATGCTGCGCCTCCTGATGCAGCCTCAAATACTAGTGCAGCTCCTTCT
GGATTAGTACAACCTCCAGTATCTCCTCACAATGCATGTTGTTCACATAACATGGTACAAAAG
AGAGGGAGCCAAGATTGCCATTGTGTTTACCCAGTAAGAGTTGAGCTTTTTCTCCGAAATGTT
TCACTGACGTCAAATTGGAGCGATGAATTTCTGGGGGAACTTGCTTCTCAACTCAGTCTGAGG
GTTACTCAGTTTGAGATTGTAAATTTCTATGTTGTTGGGGCTTCTGGGCTAAACATCACAATGT
ATATAGCTCCCCATACAGGAATTAGTTTCTCAGCCGACCAAGTTACCGCAATGAATTATTCAC
TTAGTCAGCATACAGTTCAGATCAACCCTGTACTAGTTGGTGACTACAATCTTCTTAATTTAAC
CTGGTTCAGGCCACTGGTTCTAGCTCCTGCTCCAACATTCACAATATCACCTAAACCCTCTCCA
TCTCAAGCATCTACAGTACCAAGACACAGTGCGGATACCAGCAATGAAAAACACATGAGCTT
GATTACTATCATTTGCATATTTATTGGTGCTCTAATTGCTGTCTTGGTGATTGCCATGTTTATTT
GCTTCTGCAAATTAAGGAAAGGGAAAAGGAAGGTTCCTCCAGTTGAGACACCCAAGCAAAGA
ACACCAGATGCGGTTTCTGCAGTGGACTCACTTCCTCGCCCAACAAGTACAAGGTTCCTTGCA
TATGATGAACTGAAAGAAGCAACCAACAACTTTGATCCGTCAAGTATGCTTGGAGAAGGAGG
TTTTGGCCGTGTCTTTAAAGGGGTACTAACTGATGGCACTGCCGTTGCTATAAAGAAGCTTAC
TAGTGGAGGGCACCAAGGAGATAAGGAATTTCTAGTTGAAGTGGAGATGCTGAGCAGGTTGC
ACCACCGAAACCTTGTGAAACTCATTGGTTACTATAGTAACCGTGAGTCATCACAGAACCTAC
TGTGCTATGAGCTTGTTCCTAATGGAAGCCTAGAGGCTTGGCTTCATGGTACATTGGGAGCTA
GCCGCCCCTTGGACTGGGACACTAGAATGAGGATAGCACTAGATGCTGCCAGGGGATTGGCA
TACCTTCATGAGGATTCTCAGCCTTGCGTAATCCACAGGGATTTCAAGGCTTCTAATATATTGC
TTGAGGATGACTTTCATGCTAAAGTGTCTGATTTTGGTTTGGCCAAACAGGCACCGGAAGGTT
GCACAAATTATCTCTCAACACGTGTTATGGGAACATTCGGGTATGTAGCACCAGAGTATGCAA
TGACAGGACACTTGCTTGTAAAAAGCGATGTATATAGCTATGGAGTTGTTCTACTTGAACTAC
TTACTGGGAGGAGGCCTGTGGATATGTCACAACCTTCTGGCCAGGAAAATCTAGTGACATGGG
CACGACCAATTTTACGAGATAAAGATACGCTAGAAGAACTAGCTGATCCCAAGCTTGGTGGC
CAGTATCCAAAAGATGATTTTGTGCGAGTATGCACAATTGCAGCCGCCTGTGTTTCACCGGAG
GCAAGCCAAAGGCCAACAATGGGCGAGGTGGTGCAGTCACTGAAGATGGTCCAACGCTCTGA
GTTCCAGGAATCCATACCAACACCTCCAGCACGACCCAATGTTAGGCAGTCCTCAACCACCTA
TGAATCTGATGGCACTTCATCCATGTTCTCTTCTGGACCCTTTTCAGGCCTCAGCCCCTTTGAG
ACTGAGAACATATCGAGAACGGCTTTCTCTGAAGATCTTCACGAAGGGCGGTAA

SEQ ID N° : 36, NP_912761.1, produto de proteína não denominado [Oryza sativa (grupo de cultivar japônico)]

MARIRRASSGRAQLSGRENLGLVSREWWSYYYVGFSALCCRGDTGYLESDRRTSDLSLSHYKWIC
DIDYQIQANGSSFTRKWSLLNSCYWNDENSCIAGFAERLLEDARRKTFLNFLAVVFTLELITRINLI
AQTSALNVVAPAISPSQSWRPVRSMLSKAKVDISISVSEQRRKKLYSSPATLSVHPPMSAPSYSSISD
MSDNLVQHNRRSEAEISTHVDAAPPDAASNTSAAPSGLVQPPVSPHNACCSHNMVQKRGSQDCH
CVYPVRVELFLRNVSLTSNWSDEFLGELASQLSLRVTQFEIVNFYVVGASGLNITMYIAPHTGISFS
ADQVTAMNYSLSQHTVQINPVLVGDYNLLNLTWFRPLVLAPAPTFTISPKPSPSQASTVPRHSADT
SNEKHMSLITIICIFIGALIAVLVIAMFICFCKLRKGKRKVPPVETPKQRTPDAVSAVDSLPRPTSTRF
LAYDELKEATNNFDPSSMLGEGGFGRVFKGVLTDGTAVAIKKLTSGGHQGDKEFLVEVEMLSRL
HHRNLVKLIGYYSNRESSQNLLCYELVPNGSLEAWLHGTLGASRPLDWDTRMRIALDAARGLAY
LHEDSQPCVIHRDFKASNILLEDDFHAKVSDFGLAKQAPEGCTNYLSTRVMGTFGYVAPEYAMTG
HLLVKSDVYSYGVVLLELLTGRRPVDMSQPSGQENLVTWARPILRDKDTLEELADPKLGGQYPK
DDFVRVCTIAAACVSPEASQRPTMGEVVQSLKMVQRSEFQESIPTPPARPNVRQSSTTYESDGTSS
MFSSGPFSGLSPFETENISRTAFSEDLHEGR

FIGURA 27 (continuação)

SEQ ID N° : 37, Osi001397.3 Oryza sativa, ortólogo At5g56890 putativo

ATGGGGCGGCGTGGAGGAGGAGGACGAGCGGGAGGCGCGTGGATTGGTGGCTGTGTACTTGC
GCTCGCCGCCACCTTGGTCGTCTGCGTGCTCGTGATCCGTGGAGCTGGAGGACTAAATGTAAA
ACCACCTGCAGCAACTAGCAGACGTGTTAACATACAGCAGGAACTATATGCACCAAGCCCAA
CGATCAGCCACAGAGGTCTTGCTATTCCTCCACTTCCAACAACTTCATCCCCAGTTTTCCCACC
TCCCATCAGATCTTATCCTCCACTTCCTGCAAATAAGCCATACCCCAATAGTCCAGTTGTGGCA
CCTCCAAAGGGAAGGCATCACCATTCTTTGCCTGTAAATAATACCCGTGTAAAAGGACCTGCC
TATTCTCCTTCAAATTCTCCCAGTATCCATAGAAAACATGGCATTCCAGTTGCTGCACCTCCAA
AGCAACATTCCAGCAATTTACCACCTTCACATCATAGGCCCCACAAAGGATCCTTTCCTGTCA
TAAGCCCTACTCCACATAAAGCTGATAATGCTTCTGCGACGAAACATGGACGTTCTGGCTTAC
ACCATAGTCCTGCACCTGCACCTGTAGGCTTGCCTCCATCTGAAGGAAATGCACGAGGAAATC
CTGCGTATGCACCACGTCATCCCCATGAATATCACTCGCCTTCAAATTCTCCAGAACCTGGTCT
ACCACCTGTGAATCCGCCTGACAGTCATGCATTTAAAAAGCCCAAGAGTTTGGCACCAGCACC
CCAATCATTTCCGCCACCGCCTCTGAGTTCATATTGCATGGCGTTAAATTGTCAAGATCCTCTG
ACCAATAGTCTCCCTGGAACTACATGTCTTTGTGTGTGGCCAATAAAAGTTGAGCTTCGTTTAG
GTATAGCTTTGTACACATTCTTTGCATTGGTATCAGAACTTGCACAAGATATTGCATCTGGAGT
GCTCATGAAACAAAGTCAAGTTCGTGTAATGGGAGCAAATGCTGCAACTGAGGACCCAGAGA
AAACAGTTGTCCTTATTGATCTTGTGCCACTGGGAGAAAAATTTGACAAGGCAACAGCACTTT
TAGTATTTGAAAGGTTTTGGCACAAGCAGGTCAACATAAACTCTATGCATTTTGGAAACTATG
ATGTGTTATATGTTACCTACCAAGGTCTTCCTCCGTCGCCACCAACAGCTCCCGGGATGAACA
ATGGTCTTAGCAATGTTAATGATCCAAGGTTGCATCCACTTGCTGTTGATGTGGGAAACCACA
GAGAAACAAAAAGCAGGGGCATAATTGTTATAATTGTTCTTTCAAGTGTCTTTGCTTTTATTTT
ATGTTCTGGAGCTGCGTTGGTAATTTGTTTCAAGATTAGAAACCGCAACCATTTAACTGAAGA
GTCACCTATGCCACCGAAGCCTGCAGGTCCTGGGTCTGCAGTTGTTGGGAGCAGGCTAGGAAG
CAGACCTATTTCAGCATCTCCATCTTTCAGCTCAAGCATAGTGACATATAAAGGGACAGCGAA
AACATTTAGCTTGATTGAGATGGAAAGAGCTACACAAAGATTTGACAACTCCAGAATTATTGG
CGAGGGTGGTTTTGGACGTGTTTATGAAGGTATTCTTGAGGATGGAGAACGGGTTGCTGTCAA
AATTCTTAAGAGGGACGACCAGCAAGGTACACGGGAATTTTTAGCTGAGGTTGAGATGCTTA
GCCGATTGCATCACAGAAACCTGGTTAAGTTGATAGGTATATGCACAGAAGAACATATCCGGT
GTCTTGTGTATGAGCTTGTTCCAAATGGTAGTGTGGAATCTCACTTGCACGGATCAGATAAGG
GAACTGCTCCACTTTATTGGGATGCTAGGCTTAAAATTGCACTTGGTGCTGCACGTGCTCTTGC
TTATTTGCATGAAGATTCTAGTCCCCGTGTCATACACCGTGACTTCAAGTCAAGTAACATCTTA
TTGGAACATGACTTCACCCCCAAGGTGTCAGACTTTGGCCTAGCAAGAACAGCCATAGGTGAG
GGGAATGAGCATATTTCAACTCGTGTTATGGGAACTTTCGGGTATGTTGCTCCTGAATACGCA
ATGACTGGGCATCTTCTAGTAAAGAGTGATGTCTACAGCTATGGTGTCGTCCTCCTTGAGCTTC
TTACCGGAAGGAAACCGGTGGATATTTTGAGACCTCCAGGGCAAGAAAATTTAGTTGCATGG
GCTTGTCCTTTTCTTACTAGCAGAGATGGCTTGGAAACAATAATTGATCCATCACTTGGAAAT
AGCATCCTATTTGACAGCATTGCAAAAGTAGCAGCTATTGCTTCTATGTGCGTGCAGCCTGAG
GTGGATCAGCGCCCATTTATGGGTGAAGTTGTTCAAGCTTTGAAGTTGGTATGCGATGAAGGC
AGCGAGTTCAATGAATCAAGAAGTTTCAGCCAAGATTTACATATTCAGGATTCTGGGATTATA
AGTAGAGCAAGCCTGGATGTGGACGTAGAACCTGTTGTATCTGCTGAGCTGTTCAATGCATCA
GCACATTATGACACTCTTGATGCATCTGGTTCTTTTCGACGATATTCAAGTTCAGGTCCTCTTA
GGGTAGGTAGAACCGGGCACAATAGGGGTAGCTCAAGCGAGCACTGTGGTACACAAAGATTC
AGGATAGATTCAGAGTAG

SEQ ID N° : 38, Osi001397.3, seqüência de proteína deduzida

MGRRGGGGRAGGAWIGGCVLALAATLVVCVLVIRGAGGLNVKPPAATSRRVNIQQELYAPSPTIS
HRGLAIPPLPTTSSPVFPPPIRSYPPLPANKPYPNSPVVAPPKGRHHHSLPVNNTRVKGPAYSPS

FIGURA 27 (continuação)

NSPSIHRKHGIPVAAPPKQHSSNLPPSHHRPHKGSFPVISPTPHKADNASATKHGRSGLHHSPAPAP
VGLPPSEGNARGNPAYAPRHPHEYHSPSNSPEPGLPPVNPPDSHAFKKPKSLAPAPQSFPPPPLSSYC
MALNCQDPLTNSLPGTTCLCVWPIKVELRLGIALYTFFALVSELAQDIASGVLMKQSQVRVMGAN
AATEDPEKTVVLIDLVPLGEKFDKATALLVFERFWHKQVNINSMHFGNYDVLYVTYQGLPPSPPT
APGMNNGLSNVNDPRLHPLAVDVGNHRETKSRGIIVIIVLSSVFAFILCSGAALVICFKIRNRNHLTE
ESPMPPKPAGPGSAVVGSRLGSRPISASPSFSSSIVTYKGTAKTFSLIEMERATQRFDNSRIIGEGGFG
RVYEGILEDGERVAVKILKRDDQQGTREFLAEVEMLSRLHHRNLVKLIGICTEEHIRCLVYELVPN
GSVESHLHGSDKGTAPLYWDARLKIALGAARALAYLHEDSSPRVIHRDFKSSNILLEHDFTPKVSD
FGLARTAIGEGNEHISTRVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLELLTGRKPVDILRPP
GQENLVAWACPFLTSRDGLETIIDPSLGNSILFDSIAKVAAIASMCVQPEVDQRPFMGEVVQALKL
VCDEGSEFNESRSFSQDLHIQDSGIISRASLDVDVEPVVSAELFNASAHYDTLDASGSFRRYSSSGPL
RVGRTGHNRGSSSEHCGTQRFRIDSE

SEQ ID N° : 39, Osi000142.1 Oryza sativa, ortólogo At5g56890 putativo
ATGCGGTCGCTCGCTTGTTGCGCGGTGCTTCTTCTCGCTTCGGTTCTTGGATCTACAGGTACTG
ATCTGGGTCCTTCTCCTGTGGTCGCTAACTCGCCAGATGCTCAAGATCAAACTTCTAGCCCTCC
TGAACCTACAATCGCCCTCGGTCCAGTAACTCTTCCAACAGGTTGCTCTGAATTTCTAATGTGG
TGTGGCATTATGGCTCTCAAGATCATAACCAGCATTGACAATTGTCCTTACAATTGCTTTGTCT
CAGCACCCTCTGCTCCATCAGCAAGCCCTCCTGTGGCGAAGGGCGCTGTCAGCCCAGCTGTTC
CCACTCGACCACAAAATGCGCCTACTCCAGTAACACCCCCTAAAGAATATAATGCGCCTCCTC
CAGTTGAGCTTACGCCTCCTGCTCCCACTCATGTGGCGCCAGTAGCTCCCCCGCAAGCTGCAG
TTGAAAATCCTGCACCAGTGCTTCCTGGGACACCTGCTTTGTTGCCTTCTGTTCAGGCTCCTGC
TCCATCTGTGGCCAGGAATCCTAATCTACCGATAGTGCAACCTCCTTCTGTGAACAATCCACC
AAGCGGACCAATTGGATCAGGGAATGGCGTCCCTCCGTATCCACCACCTCAAAGAAGTTTACC
TGCCATTCCTCCTTCCACTTCAGGAGTTCCGCGAGAAAGTGTAAAACCTCCAGTTGCACCACC
TATTATTGCACAAGCACCACGTCAGCAAGCACTGGCACCAAGTAGTGACCATAGCAATGGAA
ACTCAGTGCCCCCAGCAAATACATCACCTCCCCATAAGAATAGTCATATTCCACGTGCACTGC
CACCAAAGGAATCCAGTAGTCAAACTGGCACAGCTCACAAACCGCCAATTAGAGGGCCTCAT
ATTTCTCCAACCATGCCACCAATCCCTCCCCAACCAGGGCCGAAAGCACCATCAGCCCATCCT
ATTTGGGCATTACCTCCACCACCGCCTAATCTAGATTGCAATTCATTAGCATGCCCAGAGCCTC
TAACAGATCCACCTGCGGGAGCCCCATGTGTTTGTGTTCTACCAATCAAAGTTGGAGTCCGCT
TAAGTGTTGATCTGTATTCATTCTTTCCATTGGTATCTGATTTTGCGGAAGAAGTGTCATCTGG
GGTAAATATGGCACAGAGACAGGTTCGTGTTATGGGTGCAAATGTAGCTGGTGATCAGCCTG
ACAAGACAGTAGTTCTTGTTGACCTAGTACCAATGCAAGTGAAGTTTGACAATGCTACTGCTT
TTTTAACGTTTGAAAACTTATGGAGCAAAAAAATTTCCCTAAAACCATCAGTTTTTGGGGACT
ACGAGATTCTGTATGTTGTATATCCAGGGCTTCCTCCTTCTCCACCTTCAGCACCGGAAAGTGT
TGGTGACGGGGCATTTGGAAACAACAGAAATGCAAGGGCAATGAAGCCTCTCGGGGTTGATG
TAGGCAGGCCCAAAAAAAAGAGTCAACGGCAGCCTAATTGCTATTGCCGTCTTGTCTACTGTTA
TAGCATTGATTATTTGCACTCTAGCTGCATGGTTGCTGATAATCAGATTCAGGGGTTCGGATG
GCTTGGCTCAAAGATTTCCACATAGCGCACTTCCAAAATTTTCCAGATCATCTGGTACAGGTC
AAACACTGTTAGCTGGTCGCTATAGTTCACCATCAGGTCCATCAGGTTCATTGGGCTCAAGTA
TCGCAACATATGCAGGACAGGCAAAGACATTCAAATTTGCTGAGATTGAGAAGGCCACAAAC
AGCTTTGATGATTCAACAGTACTTGGAGAGGGTGGCTTCGGATGTGTCTACCAGGGCACGCTT
GAAGATGGAACCAGGGTTGCAGTAAAGGTTCTGAAGAGGTATGATGGCCAAGGCGAGAGAG
AGTTCTTGGCTGAGGTTGAGATGCTGGGACGCTTGCATCACCGAAATTTGGTCAAGTTGTTAG
GCATATGTGTAGAGGAGAACGCAAGATGTCTGGTTTATGAACTTATTCCGAATGGCAGTGTAG
AATCCCATTTGCATGGAGTGGATCTTGAGACAGCACCACTCGATTGGAATGCCCGCATGAAGA
TAGCCTTGGGGGCTGCACGAGCTCTTGCATATTTACACGAAGATTCAAGCCCTTGTGTGATTC
AT

FIGURA 27 (continuação)

CGTGATTTTAAGTCAAGCAACATCCTATTGGAGCATGATTTCACACCGAAAGTGTCTGATTTC
GGACTGGCCAGGACTGCAAGGGGGGAGGGGAACCAGCACATCTCCACTCGTGTCATGGGAAC
ATTTGGATATGTTGCACCGGAGTATGCCATGACAGGACATCTCCTTGTCAAGAGCGATGTGTA
CAGCTATGGAGTAGTGTTGCTTGAGCTCCTCACCGGTAGAAAGCCAGTGGACATGTCTAGGCC
CGGGGGGCAAGAAAACCTAGTTTCATGGGCTCGCCCGCTTCTGACCAATGTGGTAAGCCTACG
TCAAGCTGTTGATCCACTTCTTGGCCCTAACGTACCCCTGGACAATGTCGCAAAAGCAGCTGC
CATCGCATCAATGTGTGTACAACCTGAGGTTGCGCATCGCCCGAGCATGGGTGAAGTAGTGCA
GGCCTTAAAACTTGTTTGCAGTGACGGTGATGAGGGCCTCGGATCAGGAAGTTTCAGCCAGGA
ATTGGCGGCACAAGCTGCAGCAATCTATGATGTAACTGGCATGGAAGCTGAGAGGGTGCTGT
TATCTGAGATGTTTGGCTCAACCCCTGTCTTCACTCCCGCTGCGGATTCAGGTTCTTTCCGAAA
GCAGTCCAGCTCAGGTCCACTGATGACAGGCAAGAACAGAAAGTTCTGGCAGAGATTGCGAA
GCCTATCAAGAGGGAGTATGAGTGAGCATGGTGCCTCACCAGATTTTGAGACGCGCTCGCAGT
GTAGTAATAGGTGA

SEQ ID N° : 40, Osi000142.1, seqüência de proteína deduzida
MRSLACCAVLLLASVLGSTGTDLGPSPVVANSPDAQDQTSSPPEPTIALGPVTLPTGCSEFLMWCGI
MALKIITSIDNCPYNCFVSAPSAPSASPPVAKGAVSPAVPTRPQNAPTPVTPPKEYNAPPPVELTPPA
PTHVAPVAPPQAAVENPAPVLPGTPALLPSVQAPAPSVARNPNLPIVQPPSVNNPPSGPIGSGNGVP
PYPPPQRSLPAIPPSTSGVPRESVKPPVAPPIIAQAPRQQALAPSSDHSNGNSVPPANTSPPHKNSHIP
RALPPKESSSQTGTAHKPPIRGPHISPTMPPIPPQPGPKAPSAHPIWALPPPPPNLDCNSLACPEPLTD
PPAGAPCVCVLPIKVGVRLSVDLYSFFPLVSDFAEEVSSGVNMAQRQVRVMGANVAGDQPDKTV
VLVDLVPMQVKFDNATAFLTFENLWSKKISLKPSVFGDYEILYVVYPGLPPSPPSAPESVGDGAFG
NNRNARAMKPLGVDVGRPKKRVNGSLIAIAVLSTVIALIICTLAAWLLIIRFRGSDGLAQRFPHSAL
PKFSRSSGTGQTLLAGRYSSPSGPSGSLGSSIATYAGQAKTFKFAEIEKATNSFDDSTVLGEGGFGC
VYQGTLEDGTRVAVKVLKRYDGQGEREFLAEVEMLGRLHHRNLVKLLGICVEENARCLVYELIP
NGSVESHLHGVDLETAPLDWNARMKIALGAARALAYLHEDSSPCVIHRDFKSSNILLEHDFTPKVS
DFGLARTARGEGNQHISTRVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLELLTGRKPVDMSR
PGGQENLVSWARPLLTNVVSLRQAVDPLLGPNVPLDNVAKAAAIASMCVQPEVAHRPSMGEVVQ
ALKLVCSDGDEGLGSGSFSQELAAQAAAIYDVTGMEAERVLLSEMFGSTPVFTPAADSGSFRKQS
SSGPLMTGKNRKFWQRLRSLSRGSMSEHGASPDFETRSQCSNR

SEQ ID N° : 41, Osi009054.1 Oryza sativa, ortólogo At2g20300 putativo
GGCAGATCAACTTTGTCTGTTAGCAGGGTGAGCTCTGCATCGGCGTCAATGCTCTCAACAGTG
GCAACTTGCACAACATCAGTAAAGACATTTTCACTTTCCCAGCTTGAAAAGGCCACTGATGGT
TTTGATTCAAAAAGAGTGTTGGGTCAAGGTGGGTTTGGACGCGTCTACCATGGGACCATGGAT
GGCGGAGACGAGATTGCTGTTAAATTGCTCACAAGAGAAGACAGGAGTGGCGATCGAGAGTT
CATTGCGGAGGTTGAAATGCTCAGTCGACTACATCACCGTAATCTTGTCAAATTGATTGGCAT
TTGCATTGAGCACAACAAAAGGTGTCTTGTTTATGAGCTTATACGCAATGGAAGTGTTGAATC
TCACCTTCATGGTGCTGATAAGGCTAAGGGCATGCTGAACTGGGATGTTAGGATGAAAATTGC
TCTGGGTGCAGCTAGAGGCTTAGCATATCTACATGAAGACTCAAACCCTCATGTTATACATCG
TGACTTCAAGGGCAGCAATATATTGTTGGAGGAAGATTTCACTCCTAAGGTTACTGATTTTGG
TTTAGCAAGGGAGGCAACTAATGGAATTCAACCCATTTCTACTAGGGTAATGGGTACTTTTGG
GTACGTTGCTCCAGAATATGCCATGACCGGGCATCTTCTTGTGAAGAGTGATGTCTACAGTTA
CGGTGTTGTCTTGCTGGAGCTTTTATCAGGAAGGAAGCCCGTATGCATGTCTGATACCAATGG
TCCTCAGAACCTTGTGACTTGGGCTCGTCCTCTGCTATGCCATAAGGAGGGTTTGGAGAGGTT
GATCGACCCTTCTCTGAATGGCAATTTCAACTTTGATGATGTAGCTAAAGTAGCATCCATCGCT
TCCATGTGTGTTCACAATGATCCGTCACAGAGGCCATTCATGGGTGAAGTAGTGCAGGCACTG
AAGCTTATTTACAATGATGCGGAGGCGGCTTGCGATGATTCTTACAGCCACAGGGACTCATCG
TGTGACCAGTACGATGACTACCATGGGGCCCT

FIGURA 27 (continuação)

GGCCCTTGACAGCGGTAGTGGTAGCTGGTGGAATAGAAGCTCAAATCCTTCTGGGTTTTTTGA
CAACCGGAACCCCCTACCAGTTATTACCATGGAATACAGCTCTGGCAGGATCGAAGGAGCGC
GTGACCCCCGCTTTGCATTGTCAACAGGTGGTCATGCACAATCACCAGCCCTGCAGAATAGAT
CAGGACCCATCCGGATGAAGAAAAAGCTTGCCTCATTTTACCGGTCTAGAGGTAGCTTCAGCG
AGCATGGCCAGCTGCCTCGCCATTGA

SEQ ID N° : 42, Osi009054.1, seqüência de proteína deduzida
GRSTLSVSRVSSASASMLSTVATCTTSVKTFSLSQLEKATDGFDSKRVLGQGGFGRVYHGTMDGG
DEIAVKLLTREDRSGDREFIAEVEMLSRLHHRNLVKLIGICIEHNKRCLVYELIRNGSVESHLHGAD
KAKGMLNWDVRMKIALGAARGLAYLHEDSNPHVIHRDFKGSNILLEEDFTPKVTDFGLAREATN
GIQPISTRVMGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLELLSGRKPVCMSDTNGPQNLVTWA
RPLLCHKEGLERLIDPSLNGNFNFDDVAKVASIASMCVHNDPSQRPFMGEVVQALKLIYNDAEAA
CDDSYSHRDSSCDQYDDYHGALALDSGSGSWWNRSSNPSGFFDNRNPLPVITMEYSSGRIEGARD
PRFALSTGGHAQSPALQNRSGPIRMKKKLASFYRSRGSFSEHGQLPRH

SEQ ID N° : 43, Osi001078.4 Oryza sativa, ortólogo At2g20300 putativo
ATGCCGCCGCGGCGAGCGGCGCTGCTCCTCCTCGCTCTCGCCGTGTATGTGCCACTGGGAACG
GCAAGCTCAACAACTATCGCGTCGTACTTACTTGGGTTATGGAGTAGAGCACATCGGCATTCT
CTTCCTGCCCCTGCTCCAGCTCCAGCTCCAGCTCCAGCCCCTGAAACTCACCGACCTGGTATTA
GACACCCAGTGCCTAGGCATCACAGGAAAAGGCCTCATGTTGCCCCACCACTACCACCACCAT
CATCATCAGAAAGACAAGTAGCACCATATCAGTTGTTTCCAAGAATTGATGAGTTAGAAATTG
AGATTGCAGCAGGAACATTTCTGAAACAGAGTCAAGTTCGGATAATGGGTGCTGGGAGTAGT
CTTCAAGATCCTGAAAAAACTACTGTCACCGTTGATTTGGTGCCACTAGGTCAGAAGTTTGAT
AGGACTTCAGCCTTACTGACTAGCAACAGATTTTTGCAAAAGAAGGTGCCAATAAACTCGTCC
ATATTTGGTGATTATAACGTAATATATGTTCATTATCCTGGCCTGCCTTCTTTGGTCCCTAGTGT
TCCAGGATCCTTGGGCCCAATAAGCAGTAGCCAATACCCCTTTAGTGCAAATGTACACAATAG
AAGACATCAGAAGATTAACTCTAAAAGTGTTGCCATAATTGCATTATCAGCTGTTGTGCTTGT
ATTAATGAGCTTTGGCATTTGCATCATATGGAAATATAAAGGATTTGAAAAGTCTCGTGGCAC
TGGTCGTGTTTCGAATTCGTCAGCCACAAGAAAAACAGGTATGAGGTCATCATTCTCCAGTAT
GACCAGCTCGACGGCATCTTTTGTTTCAACAATTGCAACATGCCCACCTACAGTTAAGACATT
CTCTATTTCTGAGCTTGAAAAGGCCACAGAAAACTTCAGCTTCAACAAAATAATAGGTGAAGG
AGGGTATGGCCGTGTTTATCGAGGTACAATTGATGATGAAGTTGATGTTGCAGTCAAATTGCT
TACAAGAAAACATCAAAACAGAGATCGTGAGTTTATTGCCGAGGTTGAGATGCTAAGTCGTTT
GCATCATCGTAATCTTGTCAAGCTGATCGGTATATGCATTGAACGGAGCACAAGGTGCTTGGT
GTTTGAGCTTGTTCCAAACGGAAGTGTGGAGTCTCATCTGCACGGTTCTGATAAAATATACGG
CCCACTTGATTTTGATACTAGAATGAAAATAGCCCTTGGTGCAGCAAGGGGTCTGGCCTACCT
ACATGAGGATGCCAATCCCCATGTTATACACCGTGATTTCAAGGCCAGCAACGTTCTTTTGGA
AAATGATTTCACTCCCAAGGTTGCGGATTTCGGGTTGGCAAAGGAAGCCTCAGAAGGAATGG
ACCATATTTCTACTCAGGTCATGGGGACTTTTGGGTACGTTGCCCCGGAGTATGCGATGACCG
GGCATCTCCTGGTGAAAAGCGACGTGTACAGCTACGGTGTCGTGCTGCTGGAGCTCCTCTCGG
GGAGGAAGCCGGTGGACATGACGCAGCCGCCGGGGTCGGAGAACCTCGTCACCTGGGCGCGC
CCCCTCCTCACCGACCGGGACGGCCTCCAGCAGCTGGTCGACCCGTCGATGCCGGCGGCGAGC
TACGGCTTCGAGAAGCTGGCCAAGGCGGCGGCCATCGCGTCCATGTGCGTGCACGTCGAGGC
GTCGCACCGGCCGTTCATGGGCGAGGTCGTGCAGGCCCTGAAGCTCATCTACAACGGCAACA
ACGACGACACCTGCACCAGCGGCTCGTTCGGCGGCGGCGGCGGCGAGGAGTACGAGGACGAG
GAGGCGTCGTCGCCGTGGAACAACCGCTCGTGGAGCCACGACTTCGCCGCGACGCCGCCGCC
GGCGTCGCGCAGGCTGGCGTTCCCGCGGGCTCCGGCCCGCCCGACGACGATGGACTACAGCT
CG

FIGURA 27 (continuação)

GACCCGGCCGACGGGGCGGCGGGGACGTCGTCGGCGTCGGCGCGGCGGCAGCGGTCGACGTC
GAGCCTGGTGCTGGACAAGATCGAGTCGCTGGCGGCGTACGACTGGTCCGGCCCGCTGAGGG
CCAGCAGGGGGAGGAACTTCTACAGGCTGAGGGGAAGCATGAGCGAGCATGGCGGCCATCCT
TCCGAAGATTGCTCCATGGAGGGCTACTGGATGTAG

SEQ ID N° : 44, Osi001078.4, seqüência de proteína deduzida
MPPRRAALLLLALAVYVPLGTASSTTIASYLLGLWSRAHRHSLPAPAPAPAPAPAPETHRPGIRHPV
PRHHRKRPHVAPPLPPPSSSERQVAPYQLFPRIDELEIEIAAGTFLKQSQVRIMGAGSSLQDPEKTTV
TVDLVPLGQKFDRTSALLTSNRFLQKKVPINSSIFGDYNVIYVHYPGLPSLVPSVPGSLGPISSSQYP
FSANVHNRRHQKINSKSVAIIALSAVVLVLMSFGICIIWKYKGFEKSRGTGRVSNSSATRKTGMRSS
FSSMTSSTASFVSTIATCPPTVKTFSISELEKATENFSFNKIIGEGGYGRVYRGTIDDEVDVAVKLLT
RKHQNRDREFIAEVEMLSRLHHRNLVKLIGICIERSTRCLVFELVPNGSVESHLHGSDKIYGPLDFD
TRMKIALGAARGLAYLHEDANPHVIHRDFKASNVLLENDFTPKVADFGLAKEASEGMDHISTQV
MGTFGYVAPEYAMTGHLLVKSDVYSYGVVLLELLSGRKPVDMTQPPGSENLVTWARPLLTDRD
GLQQLVDPSMPAASYGFEKLAKAAAIASMCVHVEASHRPFMGEVVQALKLIYNGNNDDTCTSGS
FGGGGGEEYEDEEASSPWNNRSWSHDFAATPPPASRRLAFPRAPARPTTMDYSSDPADGAAGTSS
ASARRQRSTSSLVLDKIESLAAYDWSGPLRASRGRNFYRLRGSMSEHGGHPSEDCSMEGYWM

SEQ ID N° : 45, TC15181, Saccharum officinarum, homólogo a Osi003977.2
ATGACGGACGGCGGCGACGAGTACAAGCGCGAGGAGTCGGTGACCCTGCTGGTCATCGTCTC
CCTCGCCGCGCTCTCGCTCCTCTCCCTCATCGCGGCCTTCGCCTACTACTGCTACATTACGCGC
AAGGTCTCCCGCCGCCTCCACTCGCTCCATCTCCCCAAGCGTTCCAGCTCCCCTCCTCCGCCTC
CTCCCCGGGCCCCGCCGCGGCAGCAGCAGGGGAAGGACAGCCCGTCCAGCAACTCCGCCAGC
GACGGGGGTGGGGCGGCGGCGGCGATGTCGGTGGTGGTTGCTGGGGAGAGGGGCGTGCAGGT
GTTCAGCTACCGGCAGCTCCACGCGGCCACGGGCGGGTTCGGTCGGGCGCACATGGTTGGGC
AGGGCAGCTTCGGCGCCGTGTACCGCGGGGTGCTCCCCGACGGGAGGAAGGTCGCGGTCAAG
CTCATGGACCGCCCAGGGAAGCAGGGCGAAGAGGAGTTCGAGATGGAGGTGGAGTTGCTGAG
TAGGCTCCGATCGCCGTATTTATTGGGCCTGATTGGCCATTGTTCGGAAGGTGGACACCGTCT
ACTGGTATATGAGTTCATGGCCAATGGGGGTCTCCAGGAGCATCTCTATCCAAACAGAGGTTC
CTGTGGTGGAATCTCAAAATTGGACTGGGACACAAGAATGAGAATTGCACTTGAAGCAGCCA
AAGGTTTGGAATATCTTCATGAGCGTGTTAATCCACCTGTTATACATAGAGACTTCAAGAGTA
GCAATATTCTCCTGGACAAGGACTTCCATGCAAGAGTTTCAGACTTTGGTCTAACCAAACTTG
GATCTGATAGAGCTGGTGGTCATGTCTCAACTCGAGTTTTGGGCACACAAGGCTATGTTGCAC
CAGAATACGCATTGACTGGGCATTTGACTACCAAATCAGATGTGTACAGTTACGGTGTTGTGC
TTTTGGAACTGCTTACTGGCAGAGTACCAGTTGATATGAAGAGGCCTCCTGGAGAAGGTGTTT
TAGTTAACTGGGCATTGCCTATGCTCACTGACCGAGAAAAAGTCGTGCGGATCTTGGATCCAG
CATTGGAAGGTCAATATTCCTTGAAAGATGCTGTCCAAGTGGCTGCTATTGCTGCCATGTGTG
TTCAACCAGAGGCTGACTATAGGCCATTAATGGCTGATGTTGTCCAATCGCTGGTTCCACTTGT
CAAGAACCGTTCTAATCAGAAAGCTTGCAACCCCAATGTGCAATCATCGAAGCCACTCGACTA
G

SEQ ID N° : 46, TC15181, Saccharum officinarum, seqüência de proteína deduzida
MTDGGDEYKREESVTLLVIVSLAALSLLSLIAAFAYYCYITRKVSRRLHSLHLPKRSSSPPPPPPRAP
PRQQQGKDSPSSNSASDGGGAAAAMSVVVAGERGVQVFSYRQLHAATGGFGRAHMVGQGSFGA
VYRGVLPDGRKVAVKLMDRPGKQGEEEFEMEVELLSRLRSPYLLGLIGHCSEGGHRLLVYEFMA

FIGURA 27 (continuação)

NGGLQEHLYPNRGSCGGISKLDWDTRMRIALEAAKGLEYLHERVNPPVIHRDFKSSNILLDKDFH
ARVSDFGLTKLGSDRAGGHVSTRVLGTQGYVAPEYALTGHLTTKSDVYSYGVVLLELLTGRVPV
DMKRPPGEGVLVNWALPMLTDREKVVRILDPALEGQYSLKDAVQVAAIAAMCVQPEADYRPLM
ADVVQSLVPLVKNRSNQKACNPNVQSSKPLD

SEQ ID N° : 47, Glycine max TC196912, seqüência parcial, homólogo de At3g58690
AGCAAGTCAAATGTCATTGGCCATGGTGGATTTGGACTTGTTTACCGGGGAGTGCTCAACGAT
GGCAGGAAAGTTGCAATCAAATTCATGGATCAGGCAGGAAAGCAAGGAGAAGAAGAATTTA
AAGTAGAGGTGGAACTGCTAAGTCGGTTGCATTCTCCATATCTGCTGGCATTGCTTGGGTACT
GCTCTGATAGTAATCATAAATTGTTGGTGTATGAATTTATGGCAAATGGTGGACTGCAGGAAC
ATCTCTATCCTGTCAGCAATTCTATTATTACGCCTGTAAAATTGGATTGGGAAACACGACTAA
GAATAGCACTTGAAGCTGCTAAGGGTTTGGAATATCTTCATGAGCATGTCAGTCCCCCAGTGA
TTCACAGAGATTTTAAGAGTAGCAACATCCTCTTGGACAAGAAATTTCATGCCAAGGTTTCTG
ATTTTGGATTGGCGAAGCTTGGACCTGATAGAGCTGGTGGACATGTTTCAACTCGGGTTTTGG
GCACCCAGGGATATGTTGCCCCTGAATATGCACTAACAGGGCATTTAACAACAAAATCAGAT
GTGTACAGTTATGGTGTTGTACTTTTGGAGCTGCTCACTGGCCGGGTGCCTGTTGATATGAAG
AGACCTCCAGGGGAAGGTGTTCTTGTTTCTTGGGCTCTGCCCCTTTTGACTGATAGAGAAAAG
GTTGTAAAGATTATGGATCCTTCATTGGAGGGACAATACTCTATGAAAGAGGTTGTCCAGGTG
GCTGCAATTGCTGCAATGTGTGTGCAACCAGAGGCAGATTATCGGCCTCTCATGGCTGATGTT
GTGCAGTCATTGGTTCCACTGGTGAAGACTCAAAGGTCCCCTTCAAAGGTAGGAAGCTCCTCT
AGTTTCAATTCCCCTAAGTTGTCCCCTGGCCCAACATTTTAA

SEQ ID N° : 48, Glycine max TC196912, seqüência parcial de proteína deduzida
SKSNVIGHGGFGLVYRGVLNDGRKVAIKFMDQAGKQGEEEFKVEVELLSRLHSPYLLALLGYCSD
SNHKLLVYEFMANGGLQEHLYPVSNSIITPVKLDWETRLRIALEAAKGLEYLHEHVSPPVIHRDFK
SSNILLDKKFHAKVSDFGLAKLGPDRAGGHVSTRVLGTQGYVAPEYALTGHLTTKSDVYSYGVV
LLELLTGRVPVDMKRPPGEGVLVSWALPLLTDREKVVKIMDPSLEGQYSMKEVVQVAAIAAMCV
QPEADYRPLMADVVQSLVPLVKTQRSPSKVGSSSSFNSPKLSPGPTF

SEQ ID N° : 49, TC81885 (Solanum tuberosum) seqüência parcial, homólogo de At3g58690
ACAGGGTGAGGAAGAATTCAAAGTGGAGGTGGAGTTGCTATGCCGCTTGCGCTCACCGTATTT
GCTGTCATTGATTGGCTATTGTTCAGAAAGCAGCCATAAATTGCTTGTTTATGAGTTCATGGCA
AATGGTGGTTTACAGGAGCACTTGTATCCAATTAAAGGTTCCAATAATTGTTGTCCGAAGTTG
GACTGGAAGACCCGGTTAAGAATAGCTTTGGAGGCTGCTAAAGGTTTGGAATATCTACATGA
GCATGTCAATCCCCCAATTATACATAGAGACCTCAAAAGTAGCAATATTCTTTTGGACAAAAA
CTTCCATGCCAAAGTTTCTGACTTTGGATTGGCCAAGCTTGGATCTGATAAAGCTGGTGGCCA
TGTCTCTACCCGCGTTTTGGGGACACAGGGATATGTTGCTCCAGAATATGCATTAACAGGACA
CTTGACCACCAAATCAGATGTTTACAGTTATGGGGTTGTCCTCCTAGAGTTGTTGACTGGCAG
AGTTCCAGTTGATATGAAGAGATCACCTGGCGAAGGTGTTCTTGTTTCTTGGGCATTGCCCCGT
CTCACTGATAGGGAAAAAGTCGTAGAGATAATGGATCCAGCTCTGGAGGGTCAGTATTCAAT
GAAAGAAGTTATTCAAGTAGCTGCTATTGCTGCAATGTGTGTACAGCCCGAGGCTGATTACAG
GCCACTGATGGCAGACGTTGTGCAGTCGTTGGTTCCACTGGTGAAACAACCGCGACCAACTGT
GAAGCCTGGTAGCTCGTCTAGCTTTCACGCCACACAGTCCCCCTCCCCCCATGTTACACAATCT
CCTAAGGCTTAGACTTTTTTCAAGCGAAATGGGCATATCATATCTCTTGATCCTTAATTCTAGT
CCAACTTCTATAACTAGTGGCCATTT

FIGURA 27 (continuação)

AGTTTGTGTTTGGACTATCTAGCTTATGGAACCCCCTTTCATTTAGTAACTTTAATCTAACAAA
TTTTGATGGAGCAAGCCAGCTGATGTAATATTTTCTCATGCTAAACTAGGTGGAGCAAAAACA
GTTTCTCTGTATGTAACATGTCAAAGCTCCCATCTAATATGGGGATTGTATTTGAGTATTCTTG
GTTAGAAGCAGATTCAAGATTTGAAGTTACTATATGATA

SEQ ID N° : 50, TC81885 (Solanum tuberosum) seqüência de proteína deduzida (parcial)
QGEEEFKVEVELLCRLRSPYLLSLIGYCSESSHKLLVYEFMANGGLQEHLYPIKGSNNCCPKLDWK
TRLRIALEAAKGLEYLHEHVNPPIIHRDLKSSNILLDKNFHAKVSDFGLAKLGSDKAGGHVSTRVL
GTQGYVAPEYALTGHLTTKSDVYSYGVVLLELLTGRVPVDMKRSPGEGVLVSWALPRLTDREKV
VEIMDPALEGQYSMKEVIQVAAIAAMCVQPEADYRPLMADVVQSLVPLVKQPRPTVKPGSSSSFH
ATQSPSPHVTQSPKA

SEQ ID N° : 51, DQ202472.1, Nicotiana tabacum CMV 1a proteína de interação 2 (coi2) mRNA, cds completos, homólogo a At2g20300
TAAGATACCACAGCTATGCTGACTTATGAAAGATTTTGGAAAAAGAAGGTGCCTCTCAATAGA
ACGATGTTTGGGGATTATGAAGTGATGCACATTATCTATCCAGGGCTGCCTTCTTCTCCTCCAT
CTGGCATTGATTCTGGTAATGGTCCAACTGGAAGTGCTGTCAATCAACAATTCCCTATTACTGC
TGATTTTGTAAACAAGAGTCAGAGAATGAGTCCCCGAGTCATTTTTCTCATTGCTTCATCAGCA
TTAGTACTGTTGGTGGTATGCTGTGGTGCACTAGTTGTCCTCTTAAAATGCAGGAGGACTGGC
CGACCGTCAAATGCTGTTGGTCCAGTGTTTACACCATCTATGCACAAGAGATCTGGTAAGGGA
ATTGGGTCGACAATATCAAGTAGCCCGACTAGTTCAACATCGATTTCCCTCATTTCTGCTATGC
CTGCTTCTATCCTTTCTGTCAAAACATTTACGCTTGCGGAGCTTGAGAGGGCAACTGACAAGTT
TAGTTTAAAAAGGGTTTTAGGTGAAGGAGGATTTGGACGTGTTTATCATGGTATCTTAGAAGA
CAGAACAGAAGTTGCCGTGAAAGTACTGACTAGGGATAACCAAAATGGAGATCGTGAATTCA
TTGCTGAAGTTGAGATGCTAAGCCGATTGCATCACCGTAACCTGGTGAAATTAATTGGTATAT
GCAGTGAAGAGCGGACTCGCAGCTTGGTATATGAACTTGTTCGGAACGGTAGTGTGGAGTCTC
ATTTGCATGGAAGAGACGGGAGAAAAGAGCCACTTGATTGGGATGTGAGGTTGAAAATTGCT
CTTGGTGCTGCGAGAGGACTAGCTTACCTTCATGAAGATTCTAATCCTCGTGTAATTCATCGTG
ATTTTAAAGCCAGCAATGTTTTGTTAGAAGATGACTTCACGCCCAAGGTTGCAGATTTTGGGT
TAGCAAGGGAAGCAACCGAAGGAAGTCACCACATATCTACAAGAGTCATGGGAACTTTTGGG
TATGTTGCTCCTGAATATGCAATGACGGGACACCTACTTGTTAAAAGTGATGTGTATAGTTAT
GGAGTTGTATTATTGGAGCTTCTCTCCGGAAGAAAACCTGTGGACATGTCTCAACCTCCTGGA
GAAGAAAACCTGGTAACTTGGGCGCGACCTCTTCTGACCACTAGAGAAGGTTTGGAGCAACTT
GTGGATCCTTCTTTGGCTGGAAGCTATGACTTTGATGATATGGCAAAGGTGGCTGCCATTGCTT
CAATGTGCGTTCACCCGGAGGTGACTCAAAGGCCATTTATGGGAGAAGTGGTGCAAGCTCTCA
AACTAATTTATAATGACAATGATGAAACTTGTGCTGATGGATGTAGCCAGAAGGAGTCTTCTC
TGCCAGATTCAGATTTCAAAGGTGTCCCTTCCGATAGCAGTTGGTGGAATGCTGGTGGGGTTA
CACCAAGATTAACATATGGACAAGCCTCCACTTTCATGACCATGGATTACAGTTCTGGTCCGC
TTGAAGAGTTCGAAAACAGACCGTTTTCAGCTTCAAGTTTTAATCTTGGTGGAGGGGCGGGTT
TAACAATAAGCCACGGTAACAGATCAGGTCCTTTGAGGACTGTAAGGAGTAAACCTGCTCTCT
ATAGGTTAAGGGGCAGTATGAGTGAACACGGTGCACTTCTTCCAAGACATGATTGGAGGGAT
GGCACCAATTATGATGCTTCTTTTTAGAGTGATTTGTCAAATGTACAGTGCCGAAGGCCGTTTC
TATTTGGGAAAAAAAGATGGTTCTCCGGTGTAGATTAGGAGTTTGAAATTGCCGCTGTATCCCT
GAGAGAGTGGACTAAGCCTTCTAATTTTGGTAATCTTACATTCATTTTAATAGACAGGAAAGA
TAAACAGGACACTCCTTGTTGTATATGTTGTCAAATTAACTTTTTCTTTAGTCCAATATTGGAT
TGGACTACTAGTCTTTCTAAAAAAAAAAAAAAAAAAA

FIGURA 27 (continuação)

SEQ ID N° : 52, ABB36644.1, CMV 1a proteína de interação 2 [Nicotiana tabacum]
MLTYERFWKKKVPLNRTMFGDYEVMHIIYPGLPSSPPSGIDSGNGPTGSAVNQQFPITADFVNKSQ
RMSPRVIFLIASSALVLLVVCCGALVVLLKCRRTGRPSNAVGPVFTPSMHKRSGKGIGSTISSSPTSS
TSISLISAMPASILSVKTFTLAELERATDKFSLKRVLGEGGFGRVYHGILEDRTEVAVKVLTRDNQN
GDREFIAEVEMLSRLHHRNLVKLIGICSEERTRSLVYELVRNGSVESHLHGRDGRKEPLDWDVRLK
IALGAARGLAYLHEDSNPRVIHRDFKASNVLLEDDFTPKVADFGLAREATEGSHHISTRVMGTFGY
VAPEYAMTGHLLVKSDVYSYGVVLLELLSGRKPVDMSQPPGEENLVTWARPLLTTREGLEQLVD
PSLAGSYDFDDMAKVAAIASMCVHPEVTQRPFMGEVVQALKLIYNDNDETCADGCSQKESSLPDS
DFKGVPSDSSWWNAGGVTPRLTYGQASTFMTMDYSSGPLEEFENRPFSASSFNLGGGAGLTISHG
NRSGPLRTVRSKPALYRLRGSMSEHGALLPRHDWRDGTNYDASF

SEQ ID N° : 53, seqüência codificadora para ABE82646.1, derivado de AC133341, [Medicago truncatula], proteína quinase homóloga a At2g20300
ATGTTAAGTCGTTTGCATCATCGCAATCTAGTGAAACTTATCGGTATATGCATTGAAGGGCGC
AGGCGTTGTCTGGTATACGAGCTTGTTCCTAATGGCAGTGTTGAGTCCCATCTGCATGGTGAT
GACAAGAATAGGGGACCTCTAGATTGGGAAGCACGTATGAAAATTGCCCTTGGAGCTGCGAG
AGGATTGGCGTATCTTCACGAGGATTCCAATCCCCGTGTAATTCATCGAGACTTCAAAGCTAG
CAATGTGTTATTAGAAGACGACTTTACCCCTAAGGTTTCTGATTTCGGTTTAGCAAGAGAAGC
AACTGAGGGGAGTAATCATATTTCTACACGGGTGATGGGGACTTTTGGGTACGTTGCCCCAGA
ATATGCAATGACGGGCCATTTACTGGTTAAAAGTGATGTTTATAGTTACGGTGTTGTGCTTCTT
GAACTTCTCACAGGCAGAAAACCAGTGGATATGTCTCAACCTCAGGGACAGGAGAATCTTGTT
ACCTGGGCACGGGCACTGTTGACTAGTAGAGAAGGTTTAGAGCAGCTAGTGGATCCATCTTTG
GCTGGAGGTTACAACTTTGATGACATGGCAAAGGTAGCAGCTATTGCGTCAATGTGTGTTCAC
TCCGAGGTGACACAGAGACCTTTTATGGGAGAAGTTGTGCAGGCTCTTAAATTAATATACAAT
GACACAGACGAGACTGGTGGAGATTATTGTAGTCAGAAGGACTCCTCTGCTCAGGAGTCTGAT
TTCAGAGGTGAGCTTGCCCCTTCTGATAGCAGTTGGTGGAATGGTGGAGGTTTAACTCCTCGA
TTAACTTATGGACAAGCATCTTCCTTTATCACAATGGAGTACAGTTCTGGTCCACTTGAAGATA
TGGAAAACAGACCGTTTTCAACTTCAAGCTTTAATGGAGATGAGCTATCTTTACCAATTAGGC
ATGGGAATAGGTCAGGTCCCTTAAGAACGACCCGAAGCAAGTTATCCTTATATAGATTTTCAG
GAAGTCGGAGTGAGCATGGGGAAGTTTCGTCTAAGCGAAATTGGATTTGA

SEQ ID N° : 54, ABE82646.1, Proteína quinase [Medicago truncatula], homólogo a At2g20300
MLSRLHHRNLVKLIGICIEGRRRCLVYELVPNGSVESHLHGDDKNRGPLDWEARMKIALGAARGL
AYLHEDSNPRVIHRDFKASNVLLEDDFTPKVSDFGLAREATEGSNHISTRVMGTFGYVAPEYAMT
GHLLVKSDVYSYGVVLLELLTGRKPVDMSQPQGQENLVTWARALLTSREGLEQLVDPSLAGGYN
FDDMAKVAAIASMCVHSEVTQRPFMGEVVQALKLIYNDTDETGGDYCSQKDSSAQESDFRGELA
PSDSSWWNGGGLTPRLTYGQASSFITMEYSSGPLEDMENRPFSTSSFNGDELSLPIRHGNRSGPLRT
TRSKLSLYRFSGSRSEHGEVSSKRNWI

SEQ ID N° : 55, Proteína quinase de Populus sp.
ATGGGCCATAAGCCTCCAGACAAATACAAAGGAGTCCAGGTTTTCACATACAAGGAGCTTGA
AATCGCCACAAACAAATTCAGTGAAGCAAATGTGACATGGAATGAAGGGTATGGAGTGGTGT
ATGGAGGTACCCTAAGTGATGGAACTGTGGCAGCAATTAAGATGCTCCATAGAGCAGGGAAG
CAAGGAGAGCGTGCATTTAGGATAGAGGTGGATTTGCTAAGCAGGTTGCACTCACCATACTTG
GTGGAGCTACTTGGTTATTGTGCTGACCAGAACCACAGGCTCCTAGTATTTGAATTTATGCCTA
ATGGGACTCTTC

FIGURA 27 (continuação)

AACACCATCTCCATCACAAGCAATATCGACCGTTGGATTGGGGGACACGATTGAGAATTGCCC
TTGATTGTGCTAGGGCCCTGGAGTTCCTTCATGAGCTCACAATCCCAGCAGTAATCCATCGTG
ACTTCAAGTGTAGTAACATCTTACTAGACCAAAATTTTCGGGCTAAGGTTTCAGATTTCGGAT
CGGCTAAGATGGGCTCGGAAAGGATCAATGCTCGAAATTCGATGTGTTTGCCGAGTACCACTG
GTTATCTAGCACCAGAGCATGCTTCAACAGGTAAGCTCACCACAAAATCAGATGTATACAGCT
ATGGAGTTGTTCTTTTACAGCTCTTAACTGGTCGTAAACCTGTTGATACCAAGCAGCCCTCTGG
TGAACATGTCCTTGTCTCCTGGGCTCTTCCAAGGCTAACCAACAGAGATAAGGTAGTGGAAAT
GGTTGATCCAGCCATGCAAGACCAGTATTCAAAGAAGGACTTGATCCAGGTGGCTGCTATTGC
AGCTGTGTGTGTGCAACCAGAAGCAGATTATAGGCCTCTCATGACAGATGTTGTGCAGTCTCT
AATCCCTCTGGTCAAGAACCTCTCTTCTGTATCTTCCTCTTGTTCCTCTAAATTTATGAATCAGA
TGCTGTGCAGTCCAAGGCCTATGTA

SEQ ID N° : 56, proteína quinase de Populus sp (TC25047), homólogo a AT1G54820
MGHKPPDKYKGVQVFTYKELEIATNKFSEANVTWNEGYGVVYGGTLSDGTVAAIKMLHRAGKQ
GERAFRIEVDLLSRLHSPYLVELLGYCADQNHRLLVFEFMPNGTLQHHLHHKQYRPLDWGTRLRI
ALDCARALEFLHELTIPAVIHRDFKCSNILLDQNFRAKVSDFGSAKMGSERINARNSMCLPSTTGYL
APEHASTGKLTTKSDVYSYGVVLLQLLTGRKPVDTKQPSGEHVLVSWALPRLTNRDKVVEMVDP
AMQDQYSKKDLIQVAAIAAVCVQPEADYRPLMTDVVQSLIPLVKNLSSVSSSCSSKFMNQMLCSP
RPM

SEQ ID N° : 57, domínio de quinase de SEQ ID N° : 2 identificado pela ferramenta SMART para a identificação do domínio de proteína (Arabidopsis thaliana)
FSEEKKIGNGDVYKGVLSDGTVAAIKKLHMFNDNASNQKHEERSFRLEVDLLSRLQCPYLVELLG
YCADQNHRIXIYEFMPNGTVEHHLHDHNFKNLKDRPQPLDWGARLRIALDCARALXFLHENTIST
VIHRNFKCTNILLDQNNRAKVSDFGLAKTGSDKLNGEISTRVIGTTGYLAPEYASTGKLTTKSDVY
SYGIVLLQLLTGRTPIDSRRPRGQDVLVSWALPRLTNREKISEMVDPTMKGQYSQKDLIQVAAIAA
VCVQPEASYRPLMTDVVHSL

SEQ ID NO 58 promotor GOS2 de Oryza sativa
AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAA
TATAAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCA
CCTACTTTAGTGGCAATCGGGCTAAATAAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGT
AATTAAGTGGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGT
TAAATTATTCGAGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAAATCTTTCTAGCT
GAACTCAATGGGTAAAGAGAGAGATTTTTTTTAAAAAAAATAGAATGAAGATATTCTGAACGT
ATTGGCAAAGATTTAAACATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACA
TCATTAAGGACATGTCTTACTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATT
TTTATTATTTATCTTTTTTCGATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCA
TGTGTGAGTGCACCTCCTCAATACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTAT
TTAATACATTTAGGTAGCAATATCTGAATTCAAGCACTCCACCATCACCAGACCACTTTTAAT
AATATCTAAAATACAAAAAATAATTTTACAGAATAGCATGAAAAGTATGAAACGAACTATTT
AGGTTTTTCACATACAAAAAAAAAAAGAATTTTGCTCGTGCGCGAGCGCCAATCTCCCATATT
GGGCACACAGGCAACAACAGAGTGGCTGCCCACAGAACAACCCACAAAAAACGATGATCTA
ACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAGGCTTTGCGGCCAGGAGAGAGGAG
GAGAGGCAAAGAAAACCAAGCATCCTCCTTCTCCCATCTATAAATTCCTCCCCCCTTTTCCCCT
CTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAAGGACACGCGACTAGCAG
AAGCCGAGCGACCGCCTTCTCGATCCATATCTT

FIGURA 27 (continuação)

CCGGTCGAGTTCTTGGTCGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCTCCC
TTCGGTTGTTCTTGGATTTATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGG
ATCTGTATCTGTGATGATTCCTGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTT
ATCGGTTCGGTTTGATTAGTAGTATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAAT
GGTTTAGGGATCGGAATCTTGCGATTTTGTGAGTACCTTTTGTTTGAGGTAAAATCAGAGCAC
CGGTGATTTTGCTTGGTGTAATAAAGTACGGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTC
TCGATTTGACGAAGCTATCCTTTGTTTATTCCCTATTGAACAAAAATAATCCAACTTTGAAGAC
GGTCCCGTTGATGAGATTGAATGATTGATTCTTAAGCCTGTCCAAAATTTCGCAGCTGGCTTGT
TTAGATACAGTAGTCCCCATCACGAAATTCATGGAAACAGTTATAATCCTCAGGAACAGGGG
ATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTTTTTTCCCAAATATCTTAAAAAGTCAC
TTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATGCTTTTATAGCGTTATCCTAGCTGT
AGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGAAGAACTTATCCGATTTCTGAT
CTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATTCATTTGGATTATTTTTTTT
ATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAACTGTCCTCAATTTTGTT
TTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGTAGAAGTTTCTTTTTGG
TTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAATCGGGATAGTTATACT
GCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCTTGCCACTTTCACCAGCAAA
GTTC

SEQ ID N° : 59, seqüência de ácido nucleico de FBXW de Arabidopsis thaliana NM_122112.2
ATGGAATTTGAGTGCCAAGAGAGACTTGAAGCTGCGAACGATCAAAGAAGTGAGTCTGGTCT
TTTGAATACCGATTTGAGAATAGGTAATGATGGGTCTGTTGAGATTCCAAATGTGAAGTTGTG
TTGTCAGAAAAAGAAGGGAACTTTAGTGCCCAGTGGCTCGAAACAGTTACTCTCTGACAAGG
ATTTGTCTACTACCATTATTGATTTACCTCAGGCATTGATATCTGAGATTCTTAATTGCCTTGA
CCCAAAAGAGTTGGGTCTAGTGTCTTGTGTTTCAACTTATCTGCATAGGTTAGCATCTGAGCAT
CACGCTTGGAAGGAGTTCTATCGTGAGAGATGGGGACTCCCTGTAGTTTTCGGAGCTGCAAGT
TCAGGGCTTTCTGATGAGAGATCATGGAAAGATCTGTTTGTTGAAAGGGAATTTAGGAGTAGG
ACCTTTTTAGGACGTTATAGTATTGATACTCTGTATGGTCACACTGAAGCAGTCCGGACTGTTT
TTCTCTTAGCTTCTGCCAAGCTCGTTTTCACTTCTGGATATGATTCCATTGTTCGGATGTGGGA
CATGGAAGAAGGCTTGTCTATTGCGGCATCTAAACCTCTCGGTTGTACGATTAGGGCACTTGC
TGCTGATACAAAGCTGTTGGTAGCTGGCGGTACTGATGGATTCATTCATTGCTGGAAATCACT
GGATGGTCTCCGGAACCTGTTTGATCTCACAGGTTTTCAGAAAGAGAAGACTGAGTTTCGGCT
ATGGGGACATGAGGGTCCTATTACATCTCTTGCCCTGGACATGACAAGTATCTTTAGCGGTTC
ATGGGATATGTCTGTTCGTATATGGGATCGATCTTCAATGAAGTGTGTCAAAACCTTGAGGCA
TAGTGATTGGGTTTGGGGACTTGCACCTCATGAAACCACCCTCGCTAGCACGTCAGGTTCTGA
TGTATATATTTGGGACGTTAGCAGTGAGACTCCATTGGCGATCATCCCTGATGCTCATGAGGG
TACCACTTACTCTTTGGCAAGAAGCCATACTGGGGATTTCCTGTTTACTGGTGGAGAAGATGG
AGGGATTAAAATGTTTGAGATCAGAAGATACGGTAGTGAAACAAGCGTTGTACTTATATCTCA
ATGGATGCCTCACACTAGTCCTGTCTACTCTCTTTCTTTTGAGTTTCCTTGGCTTGTCTCAGCAT
CCGGTGACGGTAAACTCGCACTTATCGACGTTAGGAAACTGTTAAAAACTAACCGTTGTGCCT
ATTCCAAAAGGATTTCTTCGAGTACTGTGGAACCGCCACAACGAATGCTGCACGGGTTCGGGT
CAAACTTGTTCTCTGTGGACGTGGGTTATGACCGTATAGTGTGTGGAGGTGAAGAAGGCACAG
TCCGGATATGGAACTTCACACAAGCATTGGAGATAGAGCGAAGGACCCGAGCTCTGAAAGGA
ATGAGGCATGAGAATAGGATGAGACGAAGGAGAATGCAAATGGAGATGAACGCAAAGAATG
GAAGGCCTGACCAATGCTCCATTGCTGCTCATAAGAACCCCATCAATGGAGAAAGGAACCGA
GCCTGGCATAGTAAACGAAGAGCAAGCGGGAAGGCGAAGGCCTAA

FIGURA 27 (continuação)

SEQ ID N° : 60, polipeptídeo deduzido de FBXW de Arabidopsis thaliana
MEFECQERLEAANDQRSESGLLNTDLRIGNDGSVEIPNVKLCCQKKKGTLVPSGSKQLLSDKDLST
TIIDLPQALISEILNCLDPKELGLVSCVSTYLHRLASEHHAWKEFYRERWGLPVVFGAASSGLSDER
SWKDLFVEREFRSRTFLGRYSIDTLYGHTEAVRTVFLLASAKLVFTSGYDSIVRMWDMEEGLSIAA
SKPLGCTIRALAADTKLLVAGGTDGFIHCWKSLDGLRNLFDLTGFQKEKTEFRLWGHEGPITSLAL
DMTSIFSGSWDMSVRIWDRSSMKCVKTLRHSDWVWGLAPHETTLASTSGSDVYIWDVSSETPLAI
IPDAHEGTTYSLARSHTGDFLFTGGEDGGIKMFEIRRYGSETSVVLISQWMPHTSPVYSLSFEFPWL
VSASGDGKLALIDVRKLLKTNRCAYSKRISSSTVEPPQRMLHGFGSNLFSVDVGYDRIVCGGEEGT
VRIWNFTQALEIERRTRALKGMRHENRMRRRRMQMEMNAKNGRPDQCSIAAHKNPINGERNRA
WHSKRRASGKAKA

SEQ ID N° : 61 seqüência de ácido nucleico de FBXW de Oryza sativa AK111585
ATGGACTTTGATTGCAAAACAGCAAGAGGAGACTCTTCATCTGTGAATCGTTCATGCATTGTC
ACTGAAGGCACTGTCGTCCAGGCAAAGCCCGTTTCTCACAACGGGAAAGCTAAACACTGGAA
TAGCCTCAGTACATTGAATAACCAGAAGTGCAGTTATGAATTACTTTCCGACCCAAAGAAAAA
TGTTGAAACAAGTGATGGTGAGACCGCCTCAAAGTGTGACTCATGGTGCTTCACTGATTTGCC
ATCCGCATTGGTCTGTGAAGTGCTTGAACACCTCGATCCAAAGGAACTTGGCATTGTATCTTG
CGTCTCCACCCTACTGCATACCCTAGCCACAGATCATCAGGGGTGGAAGAAGTTCTACTGTGA
AAGGTGGGGGATTCCTACTCCTCCAGTCACCCTCAATGGGCCCTTGGTTCCAGGAGGAACTTC
AGATTGGAAGTCTTGGAAAACATTATTTGTGGAGCGGGAGTTTCGAAGTAAATCATTCATGGG
AAGATTCAGTGTGGATGTTCTCCGTGACCACAGTGAGGATGTACGCACTGTGTTCCTTCTAGC
ATCAGTAAATCTGATATTTACTGGTGGTAATGACTCTGTGATCCGAATGTGGGACTTGGAGGA
AGGGCTTTTGATTGATAAGTCCCGCCCACTTTGTTGCACCATCCGGGCCATTGCAGCTGACACT
AGGCTTTTGGTCACTGCAGGAACCAATGCCTTTATTCATTGTTGGAGGGCTGTTGAAGGTAAT
TCTTACCCTTTCCACATCTCTGGGAATGGCACTGACCAGAGTCCTGAGTTTCGCCTTTGGGGGC
ATGAAGGACCTGTGACTTGTCTTGCCTTGGATTCATTGAGGATTTTCAGTGGTAGCTGGGATAT
GACTGTCCGTGTTTGGGACAGATCCGAAATGAAATGTGTTCAGAAGTTCATGCATGCAGATTG
GGTTTGGAGCGTGGCACCTCATGGAAATACTGTTGCCAGTACAGCTGGTAGGGATGCCTATGT
ATGGGATATCAGGAGTGGTGAGTTGGAAAATGTTATTTCCAATGCCCATTATGGTAATGCATT
TTCTTTAGCTCGAACACACCTAGCTGATGTGCTGTTTACCGGAGGAGAGGATGGGGCAATTCG
CCTGTTCAATGTTTCTGAGGTCTCTGATGATGAAGATATTAAGCCAGCTGCTACTTGGGTTCCA
CATACCGGCCCTGTTCATTCCCTCGCTTTTGAGTATCCATGGCTTGTCTCAGCCTCTAGTGATG
GCAGGGTTGCACTGATTGATTTGAGGAAGCTTCTGACCCCAAGAAAGTCATCAAAACAACCAT
TCAGGGTTAAGAACTTTGATCCAAGTTCCATTGAACCTCCACAGAGAATGCTTCATGGCTTTG
GGTGCGATCTTTTTTCTGTCGCCATTGGTGCAGACAGGATTGTCTGTGGAGGCGAGGATGGCG
CTGTCAAAGTCTGGAACTTCTCAGAAGCACTGGAGATTGAGAAGAGGGCACAAGCTCTAAGA
AGTATGAGGCAGGAGAACCGCATGAGGCGAAAAAAGGCACAAGTAGAGATGAATGCAAATG
GTAGAAGGTCTGACCAATGTGGCTCAATAGCCATGAAAAGAAACCAACTGAAGGGTGATAAG
AGTGTCACTTGGCACAGCAAGCGTGCCATCAATGATAAGGTCAAGTCTTAG

SEQ ID N° : 62 polipeptídeo deduzido de FBXW de Oryza sativa
MDFDCKTARGDSSSVNRSCIVTEGTVVQAKPVSHNGKAKHWNSLSTLNNQKCSYELLSDPKKNV
ETSDGETASKCDSWCFTDLPSALVCEVLEHLDPKELGIVSCVSTLLHTLATDHQGWKKFYCERWG
IPTPPVTLNGPLVPGGTSDWKSWKTLFVEREFRSKSFMGRFSVDVLRDHSEDVRTVFLLASVNLIFT
GGNDSVIRMWDLEEGLLIDKSRPLCCTIRAIAADTRLLVTAGTNAFIHCWRAVEGNSYPFHISGNG
TDQSPEFRLWGHEGPVTCLALDSLRIFSGSWDMTVRVWDRSEMKCVQKFMHADWVWSVAPHG
NTVASTAGRDAYVWDIRSGELENVISNAHYGNAFSLARTHLADVLFTGGEDGAIRLFNVSEVSDD
EDIKPAATWVPHTGPVHSLAFEYPWLVSASSDGRVALIDLRKLLTPRKSSKQPFRVKNFDPSSIEPP
QRMLHGFGCDLFSVAIGADRIVCGGEDGAVKVWNFSEALEIEKRAQALRSMRQENRMRRKKAQ
VEMNANGRRSDQCGSIAMKRNQLKGDKSVTWHSKRAINDKVKS

FIGURA 27 (continuação)

SEQ ID N° : 63, seqüência de ácido nucleico de FBXW de Medicago trunculata unido de CR931734

ATGGCATTTGATTGTCCACAGAGTATTAAGGTTTCTCAAAATTTGGTAGAAAATCCAGGTATA
AGCATAGACATAGTTGAATTGAATCCAAAACCCAATTCCAAAGCAATTTCAATTTCATTCCCT
GATTTTGTTACAAATACCAAATCTGAATCTGGAAAAGGTGAGATTTTTAAGGGGAAATCCAAG
CTGAATTCTAAGAAAGGAATCAAAGCAGCCTCTGCTGGTGCTTTGAATCAATCATCAGTTCCT
GGTGATGTGGTTATTGGTAATTGTAGGTCAATTACCGACCTGCCTCCGGTGTTGATATCTGAG
ATTCTGAATTGTCTTGATCCCAAAGAACTTGGAATTGTTTCATGTGTGTCGTTGATTCTGCGTA
GTCTTGCTTCTGAGCATCATGCTTGGAAGGAATTCTATTGTGAAAGGTGGGGGCTTCCAGCAG
TTCCTGAGGCTGTCCTGGATTCGGATTCTGGGGTTGGGGATTCGGTTAAGGATGAAAAGTCAT
GGAAGGATATTTTCGTGGAAAGAGATTATAGGAGTAAGACTTTTATGGGAAGGTATAGTATG
GATGTGTTGTATGGACATACTGAGGCGGTTCGGACTGTTTTCTTGTTGGCTTCTACAAAGCTTA
TATTTACATCTGGGTATGACACTGTTGTGAGGATGTGGAACATGGAAAGTGGGTTGTCGGTTG
CATCGTCTAAACCACTTGGCTGCACCATTCGTGCGGTTGCAGCTGATACAAGGCTGCTAGTTG
CTGGTGGTACTGATGGGTTCATTCATTGTTGGAGGGCTGTTGAAGGTTTGCCACATCTGTTTGA
GCTTAGAAACTCACAACAAAATAAGAATGAGGTTCGACTTTGGGGTCATGATGGTCCGGTTAC
TTCACTTGCTTTAGATTTGACAAGGATTTATAGCGGCTCTTGGGACACGACTGTTCGTGTTTGG
GACCGTCACTCGATGAAATGCACTGTCGTGTTGAGGCATAGTGACTGGGTTTGGGGACTTGTC
CCTCATGATACTACAGTTGTCAGCACATCAGGTTCAAATGTCTATGTTTGGGATACAAATAGT
GGTAATTTGGCAACTGTTGTTCTAAATGCTCATGTTGGTAATACTTATGCTCTGGCAAGAAGCC
ATACTGGGGATTTCATTTTTACTGGGGGTGAAGATGGTTCAATTCACATGTACGAGATTGTTG
ACGGTAGTTATGTGACCGAAGCTCTGCATGTTGCTACATGGGATCCCCACTCTGGTCCTGTGT
ATTCCCTTGCCTTTGAGTTTCCTTGGCTTGTTTCTGCATCAAGTGACGGCAAATTGGCTCTAAT
TGACGTGAGGAAGCTGCTGCGCAGGAGCAAGCGTGCCATTGGAAAGAGAGCTACGAAGGCA
AAGTATTCGGGCGAAGTTAATGTAGAGCCTCCACAGAGGATGTTGCATGGATTTAAGAGTAAT
CTTTTCTCTGTAGGTATAGGAGCTGATCGAATTGTTTGCGGAGGTGAAGAAGGTGTCGTTAGG
ATTTGGAATTTCACAGAAGCTTTGGAAATTGAAAGCAGAGTTCGTGCATTGAGAGGAATGCG
ATTAGAGAACAGAATGAGACGACGTAAGAACCAAACAGAGCTCAACAGTAAAGGCGGTAAG
AGTGATCAATGTTCAGCTGCGGCCAAGAAGAGTTCAGTGACTTGTATTTGGCCGAGTAAACGT
GGGATGGGTGGAAAGACGAAGGCATAG

SEQ ID N° : 64, polipeptídeo deduzido de FBXW de Medicago trunculata

MAFDCPQSIKVSQNLVENPGISIDIVELNPKPNSKAISISFPDFVTNTKSESGKGEIFKGKSKLNSKKG
IKAASAGALNQSSVPGDVVIGNCRSITDLPPVLISEILNCLDPKELGIVSCVSLILRSLASEHHAWKEF
YCERWGLPAVPEAVLDSDSGVGDSVKDEKSWKDIFVERDYRSKTFMGRYSMDVLYGHTEAVRT
VFLLASTKLIFTSGYDTVVRMWNMESGLSVASSKPLGCTIRAVAADTRLLVAGGTDGFIHCWRAV
EGLPHLFELRNSQQNKNEVRLWGHDGPVTSLALDLTRIYSGSWDTTVRVWDRHSMKCTVVLRHS
DWVWGLVPHDTTVVSTSGSNVYVWDTNSGNLATVVLNAHVGNTYALARSHTGDFIFTGGEDGSI
HMYEIVDGSYVTEALHVATWDPHSGPVYSLAFEFPWLVSASSDGKLALIDVRKLLRRSKRAIGKR
ATKAKYSGEVNVEPPQRMLHGFKSNLFSVGIGADRIVCGGEEGVVRIWNFTEALEIESRVRALRG
MRLENRMRRRKNQTELNSKGGKSDQCSAAAKKSSVTCIWPSKRGMGGKTKA

SEQ ID N° : 65, seqüência de ácido nucleico de FBXW de Triticum aestivum contig de ESTs CJ661176.1, CB307121.1 e CJ553648.1

ATGGACTTTGATTGCAATAAGGCAGGAGAGTCTTCAGCCAAGCATTGTTCTAGCATTTGCAAC
GAGGGCACACTCATCCAGGCAAACACCTTAATTCATTGTGGAAAGGCTAAAAAGTGGAATAG
CCTCAACAAATTCAACAACCCGGAGTCAAGTCATGGATCACTTCCAAGGGTTAATGACCCCAA
GGAAGATGGTGAAACAGGAAATGATGCAACTGCCTCTGAGTGTAGCGTTATGTGCTTCACTGA
TCTGCCGTCTGCATTGGTCTGTGAAGTCCTTGCGCGCCTTGATCCAAAGGGGCTTGGGGTTGTA
TCTTGTGTCT

FIGURA 27 (continuação)

CCACAGTTCTGCAGACCCTAGCCACAGATCATCAGGGATGGAAGAAATTCTACTGTGAGAGG
TGGGGACTTCCAAATGCTCCCATCGGACCCCTAGTTCCAGGTGGAACCCCAGATGGGAGGTCG
TGGAAAGCATTGTTTGTGGACCGGGAGTTTCAAAGCAGATCATTCATGGGAAGATTTAGTGTG
GATGTTCTCCGTGGTCACAATGAGGATGTACGCACTGTGTACCTTCTGGCATCAGCAAATCTG
ATATTCACTGGTGGCCATGATTCTGTGGTTCGGATGTGGAATATGGAGGAAGGGCTACTAATT
GATGAGTCCCGCCCATTTGGTTGCACTATCCGGGCAATTGCAGCTGACAGTAGGCTTTTGGTA
ACTGGAGGATCCAAAGCCTTCATTCAGTGTTGGAGGGCTATTGAGGGGGCCTCGCACCTTTCT
CACATTTCTGGGAATGGTACTAACCAGAATTCTGAATTTCGCCTATGGGGGCATGAAGGGCCT
GTGACTTGTCTTTCCTTGGATTCAACAAGGATTTACAGTGGTTCTTGGGATATGACTGTTCGTG
TTTGGGACAGAGCCGAGATGAAGTGTGTGCAAAAGTTCATGCATGCTGACTGGGTTTTGGCCC
TGGCTCCTCATGGAAATACTGTTGCCAGTACAGCTGGTAGAGACGCATATGTGTGGGATATCG
GAAGTGGCGAGTTAACAACTATAATTTCCAATGCCCATGTTGGTAATGCATATTCTGTAGCTC
GAACACACCTGGCAGGTGTGCTGTTTACTGGAGGAGAGGACGGGGCAATTCGCTTGTTCGATG
TTTCTGAGATATCGGATGATGAGAATATTAAACCAGCTGCCACTTGGTTGCCGCATTCTGGCC
CTGTTCATTCTCTTGCTTTTGAGTACCCATGGCTTGTTTCTGCCTCTAGTGATGGCAGGATTGC
ACTAATTGATTTGAGGAAGATCCTGACCCCCCTAAACTCATCAAAGCGCCGTTTCAGGGTTAA
GCCCTTTGATCCAAGTACCGTAGAGCCTCCACAGCGAATGCTTCATGGCTTTGGATGCTATCTT
TTCTCCGTCGGCATCGGTGCAGATAGAATCATATGTGGAGGCGAGGATGGCTCTGTCAGGGTC
TGGAACTTCTCGGAAGCACTGGAGATTGAGAAGAAGGCACAGGCTTTAAGGAGTCTGAGACA
GGAGAACCGCATGAGGCGGAGGAAGGCGCAAGTGGAGATGAATGCAAATGGTAGAAGGGTT
GACCATTGCTCAGTAGCCATGAAAAGGAACCCATTGAAGGGTGATAAGAGTGTCACTTGGCA
AATCAAGCGTCCCATCAGTGACAAGGTCAAGTCTTAG

SEQ ID N° : 66, polipeptídeo deduzido de FBXW de Triticum aestivum
MDFDCNKAGESSAKHCSSICNEGTLIQANTLIHCGKAKKWNSLNKFNNPESSHGSLPRVNDPKED
GETGNDATASECSVMCFTDLPSALVCEVLARLDPKGLGVVSCVSTVLQTLATDHQGWKKFYCER
WGLPNAPIGPLVPGGTPDGRSWKALFVDREFQSRSFMGRFSVDVLRGHNEDVRTVYLLASANLIF
TGGHDSVVRMWNMEEGLLIDESRPFGCTIRAIAADSRLLVTGGSKAFIQCWRAIEGASHLSHISGN
GTNQNSEFRLWGHEGPVTCLSLDSTRIYSGSWDMTVRVWDRAEMKCVQKFMHADWVLALAPH
GNTVASTAGRDAYVWDIGSGELTTIISNAHVGNAYSVARTHLAGVLFTGGEDGAIRLFDVSEISDD
ENIKPAATWLPHSGPVHSLAFEYPWLVSASSDGRIALIDLRKILTPLNSSKRRFRVKPFDPSTVEPPQ
RMLHGFGCYLFSVGIGADRIICGGEDGSVRVWNFSEALEIEKKAQALRSLRQENRMRRRKAQVEM
NANGRRVDHCSVAMKRNPLKGDKSVTWQIKRPISDKVKS

SEQ ID N° : 67, seqüência de ácido nucleico de FBXW de Populus tremuloides
ATGGCATTTGAATGCCAAAAGAGCACTGAGGTGTCGAAAAGTCTTGCCATTTGTGTTGAAAAA
AGTAATTCCAATACGGAACATCTTGAAATTTCGAAGTCTCCTTTAGACTGTTACCCTAAGAAG
GGGATTTTAAGCAGTTTGAGTTCAAAACAGCTTTCTTGTAATGATGTTTTGCTCAATTGTCGCC
GGTCAATTACCGACCTTCCTCCGGCATTGATTTCTGAGATTCTTAGTTGCCTTGATCCTCAAGA
GCTTGGTATAGTTTCATGTGTATCGAGAGTACTTAATAGTCTTGCAACCGAGAATAGTGTTTG
GAAGGAAGTATATGGTGAGAGATGGGGACTCCCTATAGTTCCTGCACCATTGGGTGCAGGGG
TTTCAAATGAGAAGTCATGGAAGGAACTGTTTGTGGAGAGGGAGTACAGGAGTAAGATTTTTT
TGAGTCGTTATAGCATTGATACTTTGCATGGGCATACTGAAGCAGTTAGAACAGTTTCTCTATT
GGCTTCTGCCAAGCTCATTTTTACCTCTGGGTATGATATGATTGTGCGAATGTGGGACATGGA
AGACGGGTTGTATATTGCATCATCACGGCCTCTTGGTTGCACTATACGAGCAGTTGCAGCGGA
CACGAAGCTGTTGGTTGCTGGTGGTACTGATGGTTTTATTCAAGGCTGGAGGGCAGTTGAGGG
TCTCAAGCACTTGTTTGACCTTAAAGGTTCTGAGGTGCCAAACACTGAATTTAGAATTTGGGA
GCATGAGGGACCTATAACCTCTCTTGCCTTGGACCCCACAAGGATTTATAGTGGGTCATGGGA
CATGACTGTTCGTATATGGGATCGTTCCTCACTTGAATGCATAAAGATCTTGAGGCATGGTGA
CTGGGTGTGGAGCCTTGTTCCGCATG

FIGURA 27 (continuação)

ATACTACAGTTGCTAGCACATCAGGTTCAGACGTGTATGTTTGGGACACTAATAGTGGGACTC
TGCTAACTGTTATTCATAGTGCTCATGTAGGTAACACTTATTCTTTGGCTCGAAGCCATACAGA
GGATTTCATTTTTACAGGAGGAGAAGATGGGGCTATGCACATGTTTGAGATCACTGGTCCTAA
ACCTGAGGCTAATGTTTTTAAGGTTGCAACTTGGATGCCTCACTCAGGGCCTGTTTATTCCCTT
GCATTTGAGTTTCCATGGCTTGTTTCAGCTTCTAGTGATGGGAGGCTGTCACTAGTTGATGTGA
GAAAACTACTAAGGACCAACAGACGTTCTCTACGAAAGAATGTTTCAAGGGTTACCGATAAG
GACCGTAACAATGTTGAACCACCTCAGAGGATGTTACATGGGTTTGGGCCCAATTTGTTTTCA
GTAGATGTTGGTGCTGACCGTATTGTATGTGGAGGAGAGGAAGGTGTTGTTAGGATCTGGAAC
TTCTCAAAGGCGCTGGAAAGGGAGCGGACAGCTCGTGCCATGAGAGGAATACGGTTGGAGAA
CAGGATGAGGCGTCGTAGACTTCAAATAGAGATGAGCAGTAAGGGTGGTAGGACTGATCAAT
GCTCTGTTGCAGCCAAGAAGAACCCAATGCATGTAGACAAGAGTGGTGTCTGGCGCAATAAA
CATGGGATGAGTGGCAAGCTGAAAGCATAG

SEQ ID N° : 68, polipeptídeo deduzido de Populus tremuloides
MAFECQKSTEVSKSLAICVEKSNSNTEHLEISKSPLDCYPKKGILSSLSSKQLSCNDVLLNCRRSITD
LPPALISEILSCLDPQELGIVSCVSRVLNSLATENSVWKEVYGERWGLPIVPAPLGAGVSNEKSWKE
LFVEREYRSKIFLSRYSIDTLHGHTEAVRTVSLLASAKLIFTSGYDMIVRMWDMEDGLYIASSRPLG
CTIRAVAADTKLLVAGGTDGFIQGWRAVEGLKHLFDLKGSEVPNTEFRIWEHEGPITSLALDPTRI
YSGSWDMTVRIWDRSSLECIKILRHGDWVWSLVPHDTTVASTSGSDVYVWDTNSGTLLTVIHSAH
VGNTYSLARSHTEDFIFTGGEDGAMHMFEITGPKPEANVFKVATWMPHSGPVYSLAFEFPWLVSA
SSDGRLSLVDVRKLLRTNRRSLRKNVSRVTDKDRNNVEPPQRMLHGFGPNLFSVDVGADRIVCGG
EEGVVRIWNFSKALERERTARAMRGIRLENRMRRRRLQIEMSSKGGRTDQCSVAAKKNPMHVDK
SGVWRNKHGMSGKLKA

SEQ ID N° : 69, seqüência de ácido nucleico de FBXW de Zea mays unido de AC183938.1
ATGGCCTTTGACTGCAACAAGGAAAGAGGAGTTTCTTCGCCTGACAGTTGCTCCCGCATTTGC
ACCGAGGGCACTCTCGTCCAGGCAAATCCCTTATCTCACTACTGGAAGCCTAAACGTTTGGAG
AACTTCAACAAGTTGGAGAACCAGAAGTCCAGTTATGAATCTGCTCCGAGTGGCTCTGATCCA
AAACAGGATGCTGAAGCGAGTGATGAGGCAACTGCCTCACTGTGTGGCTTCAGGTGCTTCACT
GATCTGCCAGCTTCGTTGGTCTGCGAGGTCCTTGCACGTCTCGATGCGAAGGACCTTGGAATT
GTATCCTGTGTCTCCACCCTCCTGCACGCCTTAGCCACAGATCATCAGGGATGGAAGAAGTTA
TACTGCGAAAGGTGGGGGCTTCCAGACCTCCCCACCACCCTGAATGGGCCCTTGCTGCCGGGT
GTCCCCCTAGATGGGAAGTTTTGGAAAACATTTTTCGTGGAGCGGGAGTTTAGGAGCAAATCC
TTCATGGGAAGATTCAATCTGGATATTCTCCGTGGCCACAATGAGGATGTGCGTGCTGTGTCC
CTTTTGGTATCAGCAAATCTGATATTCACAGGCGGCCGCGATTCTGTGGTTAGGATGTGGAAG
ATGGAGGAAGGGTTATTGATTGATACATCCCGTGCACTTGGTGGCACCATCCGGGCGATTGCA
GCTGACTCGAGGCTTTTAGTGAGTGGAGGAACTAATGCCTATATTCAGTGTTGGAGGGCCGTT
GAAGGCAATGGTCAATTATTTCACATCTCTGGAAGTGGTACCGACCAGAATGCGGAGTTTCGC
CTCTGGGGTCATGAAGGTCCTGTGACTTGTCTTGCCCTGGATTCACTGAGGATTTATAGTGGTT
CTTGGGACATGACTGTTCGTGTTTGGGACAGAGACCAGATGGAGTGTGTGCAAAAGTTCATGC
ATGCAGACTGGGTTTGGGATCTTGCTCTGCGTGGAAATACTGTTGCCAGTACTGCTGGTAGAG
ATGCATATGTATGGGATATCAGGGCTGGCGAGTTAACAAGCTTAATTTCCAATGCCCATGTTG
GTAGCGCATATTCAATTGCTTGGACACACCTGCCAGATGTGCTGTTTACTGGTGGAGAGGATG
GTGCTATTCGATTGTTCAATGTTTTTGATGTCTGTGATGATGAGGGTGACATTAAACCAGTGGC
AACTTGGGTGCCACATTCAGGTCCTGTTCATTCTCTTGCTTTTGAGTATCCATGGCTTGTGTCA
GCCTCGAGTGATGGCAGGATTGCACTTATTGATTTGAGGAAGCTTCTGTCCTCCAAAAGTTCA
TCAAAGGGTCTGCGTGTTAAGAGAGTTGATGGAAGCGCAATAGAGCCTCCACAGAGGATGCT
TCATGGCGTTGGATGTGATCTCTTCTCCATCGCTATTGGTGC

FIGURA 27 (continuação)

AGATAGGATTGTATGTGCCGGTGAGGATGGTTCTGTTAGAGTCTGGAACTTCTCAAAAGCATT
GGAGATTGAGAGGAGGGCACAGGCTCTAAGGAGTTTGAGGCAGGAGAACCGCATCAGGCGG
AGGAAATCACAAGCGGAGATGAATACAAATACTAGAAGGCCTGACCAGTGTTCAATAGCCAT
GAAAAAGAACCAACTGAAGGGTGATAAGAGCGCAACTTGGCACAACAAGCGTGCCATTAATG
AGAAGGCCAAGTCTTAG

SEQ ID N° : 70, polipeptídeo deduzido de FBXW de Zea mays
MAFDCNKERGVSSPDSCSRICTEGTLVQANPLSHYWKPKRLENFNKLENQKSSYESAPSGSDPKQ
DAEASDEATASLCGFRCFTDLPASLVCEVLARLDAKDLGIVSCVSTLLHALATDHQGWKKLYCER
WGLPDLPTTLNGPLLPGVPLDGKFWKTFFVEREFRSKSFMGRFNLDILRGHNEDVRAVSLLVSAN
LIFTGGRDSVVRMWKMEEGLLIDTSRALGGTIRAIAADSRLLVSGGTNAYIQCWRAVEGNGQLFHI
SGSGTDQNAEFRLWGHEGPVTCLALDSLRIYSGSWDMTVRVWDRDQMECVQKFMHADWVWDL
ALRGNTVASTAGRDAYVWDIRAGELTSLISNAHVGSAYSIAWTHLPDVLFTGGEDGAIRLFNVFD
VCDDEGDIKPVATWVPHSGPVHSLAFEYPWLVSASSDGRIALIDLRKLLSSKSSSKGLRVKRVDGS
AIEPPQRMLHGVGCDLFSIAIGADRIVCAGEDGSVRVWNFSKALEIERRAQALRSLRQENRIRRRKS
QAEMNTNTRRPDQCSIAMKKNQLKGDKSATWHNKRAINEKAKS

SEQ ID N° : 71, seqüência de ácido nucleico (terminal 3') parcial FBXW de Vitis vinifera parcial contig de ESTs CF210354, CF413646, CF213082
TGGGGAGCTCCGGTTGTTTCGGGGTTTTCAGATGAGAAATCATGGAGGGAATTGTTTGTGGAG
AGAGAGTTTAGGAGCAAAACCTTTAGGGGACGATTTAGTATTGATGTTCTGTATGGTCGCACT
GAGGCGGTTCGAGCTGTTTTCCTTTTGGCCTCTGCAAAGCTCATTTTCACTTCTGGGTATGATT
CCATTGTTCGAATGTGGGATATGGAAGAAGGGTTGTCAATTGCGTGTTCACGGCCTCTCGGTT
GCACAATAAGAGCAGTGGCTGCGGATAAGAAACTGTTGCTTGCGGGGGGTAGTGATGGGTTT
ATTCATTGTTGGAGAGCTGTAGAGGGGCTCTCTTGCTTGTTTGACCTTGTGGGTTCTCAAAACC
TGAGTACTGAGTTCCGAATTTGGGAACATGAAGGTCCTGTGACTTGTCTTGCTTTGGATATTAA
GAGAATTTATAGTGGCTCATGGGACATGACTGTGCGCATATGGGACCGTTCTTCGTTTAAGGT
TGTAAAGGTCTTGAGGCACACAGACTGGGTTTGGGGCCTTGTTCCTCGTGATACTACAGTTGC
TAGCACTTCTGGCTCAGATGTGTATGTTTGGGACGCTGATAGTGGGACTCTGCTGACGATAAT
CTCTAATGCTCATGTGGGTAATGCTTATGCTTTGGCACGGAGCCACACAGGGGATTTTCTATTC
ACTGGGGGAGAAGATGGGGCAATACATATGTTTGAGGTCGTGAGTGATTGCATGGAAAGGAA
TGTATTGGAGGTTTCAACGTGGATTCCTCATTCTGGTCCTGTTCATTCCCTGGCATTTGAGTTTC
CCTGGCTTGTTTCAGCTTCAGCTGATGGGAGGATGTCACTGATTGATGTGAGAAAGCTTTTGC
AAACTTGCAAGCCTTTCTTAGGGAAGAATGTTTCCAAGGTTAGGCACAGGGACCACAAAAGT
GTGGAGCCCCCTCAGAGGATGTTACATGGGTTTGGCTGCAATTTATTCTCAGTGGACATTGGT
GCAGATCGTATTGTCTGTGGAGGTGAGAAAGGTGTGGTTAGAATTTGGAACTTCTCACAAGCT
TTAGAAGCAGAGCAGAGGGCCCGTGCTTTAAGAGGAATACGTTTAGAAACCAGGATGAGGCG
GCGTAGGCTTCAAATAGAGCTGACTAGTAAGGGTAGTCGAACTGATCAATGTTCAGTTGCAGC
CAGTAAGAATCCTATTAATGGTGATAGGAGTGGTGTGTGGCCCAATAAGCGGGGAATGAGTG
GCCGGATGAAAGCATAG

SEQ ID N° : 72, polipeptídeo deduzido parcial de FBXW de Vitis vinifera
WGAPVVSGFSDEKSWRELFVEREFRSKTFRGRFSIDVLYGRTEAVRAVFLLASAKLIFTSGYDSIVR
MWDMEEGLSIACSRPLGCTIRAVAADKKLLLAGGSDGFIHCWRAVEGLSCLFDLVGSQNLSTEFRI
WEHEGPVTCLALDIKRIYSGSWDMTVRIWDRSSFKVVKVLRHTDWVWGLVPRDTTVASTSGSDV
YVWDADSGTLLTIISNAHVGNAYALARSHTGDFLFTGGEDGAIHMFEVVSDCMERNVLEVSTWIP
HSGPVHSLAFEFPWLVSASADGRMSLIDVRKLLQTCKPFLGKNVSKVRHRDHKSVEPPQRMLHGF
GCNLFSVDIGADRIVCGGEKGVVRIWNFSQALEAEQRARALRGIRLETRMRRRRLQIELTSKGSRT
DQCSVAASKNPINGDRSGVWPNKRGMSGRMKA

FIGURA 27 (continuação)

SEQ ID N° : 73, Seqüência de ácido nucleico (extremidade 5') parcial de Senecio cambrensis EST DY662683.1

ATGGCATTTGAGTTAAACCCGAGCACATCGAATTCTGAAAAAGATAAACCTCAGGATAATTCA
TCTGAAAATAATTTACTGATTGATTCGGAAAGCGGGAAACATTTTGCTGCTAAGTTACTGGTT
GACGAGTGTTGTTTGTTGAAAGGTTATAAAACAATTACCGACCTTCCTCCAGCGATAGTTTCT
GAAATATTTAACAGCCTTGATCCAAAGGAACTCGGGATAGTTTCCTGTGTGAATACATCTTTA
TATCGAATTGCATTTGATCACCACGTTTGGAAGGAGTTTTATTGTGAGAGATGGGGACTTCCT
ATTGGAGTCCCCAATACGTCTTTAGAGAAATCATGGAGAGACCTTTTTGTGGAACGTGAGTTT
CGTAGTAAGACGTTTATGGGACCTTACTCTACTGAAACTCTTTATGGTCATACTGAAGCTGTTC
GTACAGTTTTTCTTTTACTTTCAAGAAAGATTATAATTACTTCGGGATATGATTCAGTTATTCG
AATGTGGGACATGGAAGGTGGTTATTCAATTGCTTCGTCCCGTCCTCTTGGTTGTACCATCCGA
GCCGTTACAGCAGATTCAAAACTGTTAATTGCTGGCGGTTCCGATGGTTTTATTCATGGTTGGA
GAGCCCAAGAAGAAGAGGACATCCTT

SEQ ID N° : 74, polipeptídeo deduzido parcial de FBXW de Senecio cambrensis

MAFELNPSTSNSEKDKPQDNSSENNLLIDSESGKHFAAKLLVDECCLLKGYKTITDLPPAIVSEIFNS
LDPKELGIVSCVNTSLYRIAFDHHVWKEFYCERWGLPIGVPNTSLEKSWRDLFVEREFRSKTFMGP
YSTETLYGHTEAVRTVFLLLSRKIIITSGYDSVIRMWDMEGGYSIASSRPLGCTIRAVTADSKLLIAG
GSDGFIHGWRAQEEEDIL

SEQ ID N° : 75, seqüência de ácido nucleico (médio) parcial BXW de Helianthus annuus EST DY916708.1

GCGGGCGACGGAGTCCCCGTCTCTTTAAATGCTGAGTGTTCTGATGAGAGATCTTGGAAGGCA
TTATTTGTGGAACGGGAATTCCGGAGCAAAACATTTCTGGGCCGGTATACAATCGATACGCTT
TACGGTCATACAGAACCCGTTCGCACTCTTTTTGTTTTGCTCTCGAAAAAGCTTATAATAACTT
CCGGTTATGACTCGATTATTCGAATGTGGGACGTTGAAGATGGTTATTCAATCGCTTCATCTCG
GCCTCTGGGTTGCACCATCCGAGCCGTTGCAGCTGATTCTAAGCTGTTAATTGCTGGCGGGTCT
GATGGGTTTATACATGGCTGGAGAGCTGAAGAAGGACACCCTCACTTGTTTGACATAAAAGG
ACCCCAAAACCAGAACACCGAGTTTCGACTTTGGGAACATCAAGGCCCGATAACTTGTATCGG
GCTAGATTCATCTAGGATTTACAGCGGGTCATGGGACATGACCATACGCGTCTGGGATCGTGT
CTCGCTCAAGTGCTTGACTGTCTTGATGCATAACGACTGGGTGTGGAGCCTGGTCCCACATGA
TGGTACTGTAGTAAGTACTTCTGGTTCAGATGTATATATCTGGGAGACTAATACCTTTTCACTT
ATTGACATTATCAAAGATGCCCATGTGGGTAATGCTTATTCTCTAGCTAGAAGTCACACCGGT
AATTTCATCTTCACTGGTGGTGAAGATGGTGCTATACATATGTTTGAGATTACTAGTAATGGAT
GTGGTTCAGTTCGCCGGCTAGCAACGTGNGGTCCACATACGGGTCCTGTATATTCTCTTTCATT
TGAGTTTCCATGGCTAGTTTCTG

SEQ ID N° : 76, polipeptídeo deduzido parcial de FBXW de Helianthus annuus

AGDGVPVSLNAECSDERSWKALFVEREFRSKTFLGRYTIDTLYGHTEPVRTLFVLLSKKLIITSGYD
SIIRMWDVEDGYSIASSRPLGCTIRAVAADSKLLIAGGSDGFIHGWRAEEGHPHLFDIKGPQNQNTE
FRLWEHQGPITCIGLDSSRIYSGSWDMTIRVWDRVSLKCLTVLMHNDWVWSLVPHDGTVVSTSGS
DVYIWETNTFSLIDIIKDAHVGNAYSLARSHTGNFIFTGGEDGAIHMFEITSNGCGSVRRLATXGPH
TGPVYSLSFEFPWLVS

FIGURA 27 (continuação)

SEQ ID N° : 77, seqüência de ácido nucleico (médio) parcial FBXW de Euphorbia esula EST DV129599

GGATGGAAAGCCGTGGAGGGTCTTCCACATTTGTTCAACCTTAAGGGTAGTTCCGAGAATCTT
CCAAACGATGAATTTCGACTTTGGGAACACGAGGGACCTATAACCTCTCTCGCCTTGGATCGC
ACGAGAATTTACAGTGGTTCTTGGGACATGACTGTTCGTGTATGGGATCGTTCCACATTGAAG
TGCCTTAAAATCTTGAGGCATAGCGATTGGGTATGGGGCCTTGTTCCGCACGATACCACCGTT
GCCACCTCATCGGGTTCCGATGTGTATATTTGGGATACTCAAACCGGAACTCTTATAGCCGAT
ATTAATAACGCCCATGTCGGGAACATTTATTCTCTAGCCCGAAGCCACACGGGAGATTTTCTC
TTCACCGGAGGAGAAGATGGGGCTATACATATGTTTGAGATCATTGGTAATAATGGATCCAAG
CCTAAGGTCTACAAGATCGCAACTTGGATTCCACACTCGGGTCCCGTCTACTCCCTCTCGTTTG
AATTTCCGTGGCTTGTTTCGGCTTCTAGTGATGGGAAACTCTCGCTTATAGATGTTCGAAAGCT
ACTTCGAACAAGTAGGCGCTCTATGGGAAAGAAGAATCTTGGTAGGGTTAGCAAAATGGATA
GTAACAATGTGGAGCCACCTCAACGGATGCTCCACGGGTTTGGACCCAATTTGTTTTCGGTAG
ACATTGGGGCTGACCGGATTGTTTGTGGAGGGGA

SEQ ID N° : 78, polipeptídeo deduzido parcial de FBXW de Euphorbia esula

GWKAVEGLPHLFNLKGSSENLPNDEFRLWEHEGPITSLALDRTRIYSGSWDMTVRVWDRSTLKCL
KILRHSDWVWGLVPHDTTVATSSGSDVYIWDTQTGTLIADINNAHVGNIYSLARSHTGDFLFTGG
EDGAIHMFEIIGNNGSKPKVYKIATWIPHSGPVYSLSFEFPWLVSASSDGKLSLIDVRKLLRTSRRS
MGKKNLGRVSKMDSNNVEPPQRMLHGFGPNLFSVDIGADRIVCG

SEQ ID N° : 79, seqüência de ácido nucleico (médio) parcial FBXW de Lycopersicon esculentum EST BI931509

ATCGTCTTGCTTCGGAGCACCATGTGTGGAAGGAGTTCTATTGTGAAAGGTGGGGGCAGCCAA
TATTGTTGGCACCTTTGGGCGCAGAGCATGCAGATGAGAAGTCGTGGAAGGAGCTTTTCGTGG
AGAGGGAGTTCCGTAGTAAAACATTCATGGGACGCTATAGTATTGATACATTGTATGGTCATA
CCGAGGCTGTTAGGACTGTTTCTGTTTTATCTTCAAAGAAACTTCTGTTTACTTCAGGGTATGA
TCAGATAGTGCGAATGTGGGACTTGGAAGAAGGTTTATCTATTGCATCATCTCGCCCTCTTGG
CTGCACTATTCGTGCAGTTGCGGCTGATTCGAGGCTCTTAGTAGCAGGGGGTACAGATGGATT
CATACATGGTTGGAGAGCAGAGGATGGAAATCCACATCTCTTTGATCTTAGATCACCTCAGAA
CCAGAACATGGAATTCAGAGTTTGGGAACATGAAGGACCAATAACATGTCTGGCATTGGATTT
GAATAAGATGTACAGTGGTTCTTGGGACATGAGTATTCGTGTTTGGGATCGTTCTTCTTTGAAA
TGTCTGAAAGTTCTAATGCACAATGACTGGGTTTGGAGCCTTGCTCCCCATGATACAACAGTA
GCGAGCGCTGCAGGCTCAGATCTTTATGTCTGGGAACACTATATTGGGTCAGAACTTGCCACC
ATCACCAATGGTCATGCT

SEQ ID N° : 80, polipeptídeo deduzido parcial de FBXW de Lycopersicon esculentum

RLASEHHVWKEFYCERWGQPILLAPLGAEHADEKSWKELFVEREFRSKTFMGRYSIDTLYGHTEA
VRTVSVLSSKKLLFTSGYDQIVRMWDLEEGLSIASSRPLGCTIRAVAADSRLLVAGGTDGFIHGWR
AEDGNPHLFDLRSPQNQNMEFRVWEHEGPITCLALDLNKMYSGSWDMSIRVWDRSSLKCLKVLM
HNDWVWSLAPHDTTVASAAGSDLYVWEHYIGSELATITNGH

SEQ ID N° : 81, seqüência de ácido nucleico (terminal 3') parcial de FBXW de Aquilegia formosa x Aquilegia pubescens DT753991.1

GTCCCTCGAGACTCTACTATTGCTAGTACTGCCGGGGTGGATATGTATGTTTGGGACATTGAT
AGTGGGGAATTGCTTGCTGTAATTCAAAAGGCTCATGTTGGCAACACTCTCTCTTTGGCCCGA
AGTCACACAGGGAATTTGATTTTCACTGGTGGGGATGATGGGACTATAAGTATGTTTGAGCTT
TTGGACGATTCTACGCTCTTGCATGTGGCAACTTGGAGCCCACATACTGGTGCTGTGCACTCTC
TTGCATT

FIGURA 27 (continuação)

TGAGTTCCCTTGGCTTGTGTCAGCCTCTAGCGACGGAAAGCTCTCATTGATTGACATTAGACG
GTTGCTGAAATCTAGCCAGCGGTCTGTGTCAAGAGACCTCAAAAGGGTAAATTGTCCAGTTTC
GTTGACTGTGGAGGCTCCACAGAGGATGTTACATGGTTTTGGTAGTAATTTATTCTCAGTGGG
CATTGGTTCTGATCGGATTGTATGTGGTGGTGAGGAAGGTGTAGTTAGAATTTGGAACTTTTC
ACAAGCTATGGAGATTGAGCGTAGGGTTCAGGCCTTAAGGGGCATGAGATTAGAAAACCGAA
TGAGGCGGCGGAAGGCCCAAATAGAGATGAACAGCAAGGGTGGTCGCTCTGATCAGTGTTCA
GTTGCAGCTAAGAAGAACCAGATTAATGGAGACAGGGCCTGGCATGGTAGGCGAGGAGCGA
GTGGAAAGCTGAAGGCCTAG

SEQ ID N° : 82, polipeptídeo deduzido parcial de FBXW de Aquilegia formosa x Aquilegia pubescens
VPRDSTIASTAGVDMYVWDIDSGELLAVIQKAHVGNTLSLARSHTGNLIFTGGDDGTISMFELLDD
STLLHVATWSPHTGAVHSLAFEFPWLVSASSDGKLSLIDIRRLLKSSQRSVSRDLKRVNCPVSLTVE
APQRMLHGFGSNLFSVGIGSDRIVCGGEEGVVRIWNFSQAMEIERRVQALRGMRLENRMRRRKA
QIEMNSKGGRSDQCSVAAKKNQINGDRAWHGRRGASGKLKA

SEQ ID N° : 83, seqüência de ácido nucleico (terminal 3') parcial de FBXW de Gossypium hirsutum EST DT466472
GGCCGCTTCGACCGGAGTTGCACTGGAGACTTCCTTTTTACTGGTGGAGAAGATGGAGCGATA
CACATGTTTGAGATTATAAGTGGGTCTGACGAATCTAGTGTCGTGCAGGTTGCAACATGGATT
CCTCACTCAGGTGCTGTTTACTCACTTGCATTTGAGTTTCCCTGGCTTGTTTCAGCTTCCAGTGA
TGGTAAACTTGCACTAATTGATGTTCGGAAGTTGTTGAGGGCCGGTAAGCGCACCTTGGGGAA
AAGGGTTTCAAGGGATAATGACATTGACCAAAGGAACATAGAGCCACCCCAGAGAATGCTCC
ATGGATTTGGGTGCAATTTGTTCTCAGTCGGTATTGGTGCAGATCGAATTGTCTGTGGGGGTG
AAGAAGGTGTTGTTCGTATCTGGAACTTCTCAGAAGCTCTAGAAATTGAGCAGAGGGCACGG
GCACTAAGAGGAATACGTTTGGAGAATAGGATGAGGCGACGCAGACTTCAAATTGAGATGAA
TAATAAGGGTGGTCGAACTGATCAATGTTCTGTTGCAGCCAAGAAGAAAACAGTCAATGGTG
ATAGGAATAGTGTCTGGCACAGTAAACGTAGCATAAGCTCCAAGGTGAAGACATAA

SEQ ID N° : 84, polipeptídeo deduzido parcial de FBXW de Gossypium hirsutum
GRFDRSCTGDFLFTGGEDGAIHMFEIISGSDESSVVQVATWIPHSGAVYSLAFEFPWLVSASSDGKL
ALIDVRKLLRAGKRTLGKRVSRDNDIDQRNIEPPQRMLHGFGCNLFSVGIGADRIVCGGEEGVVRI
WNFSEALEIEQRARALRGIRLENRMRRRRLQIEMNNKGGRTDQCSVAAKKKTVNGDRNSVWHSK
RSISSKVKT

SEQ ID N° : 85, seqüência de ácido nucleico (médio) parcial FBXW de Sorghum bicolor CF770159
GATACATCCCGTGCACTTGGTGGCACCATCCGGGCGATTGCAGCTGACTCTAGGCTTTTAGTG
AGTGGAGGAACTAATGCCTATATTCAGTGTTGGAGGGCTGTTGAAGGCAATGATCACTTATTT
CACATCTCTGGAAGTGGTACTGACCAGAATGCGGAGTTTCGCCTCTGGGGGCATGAAGGTCCT
GTGACTAGTCTTGCCTTGGATTCACTGAGGATTTATAGTGGTTCTTGGGACATGACTGTTCGTG
TTTGGGACAGAGCCCAGATGGATTGCGTGCAAAAGTTCATGCATGCAGACTGGGTGTGGTCTC
TTGCTCTGCGTGGAAATACTGTTGCCAGTACTGCTGGTAGAGATGCATATGTATGGGATATCA
GGGACGGCGAGTTAACAAGCTTAGTTTCCAATGCCCATGTTGGTAATGCATATTCATTGGCTT
GGACACACCTGGCGGATGTGCTGTTTACTGGTGGAGAGGATGGCGCTATTCGCTTGTTCAATG
TTTCTGATGTCTCTGATGATGAGGAGGACATTAAACCAGTGGCAACTTGGGTGCCACATTCAG
GTCCTGTTCATTCTCTTGCTTTTGAGTATCCATGGCTTGTGTCAGCCTCGAGTGATGGCAGGA

FIGURA 27 (continuação)

SEQ ID N° : 86, polipeptídeo deduzido parcial de FBXW de Sorghum bicolor
DTSRALGGTIRAIAADSRLLVSGGTNAYIQCWRAVEGNDHLFHISGSGTDQNAEFRLWGHEGPVT
SLALDSLRIYSGSWDMTVRVWDRAQMDCVQKFMHADWVWSLALRGNTVASTAGRDAYVWDIR
DGELTSLVSNAHVGNAYSLAWTHLADVLFTGGEDGAIRLFNVSDVSDDEEDIKPVATWVPHSGPV
HSLAFEYPWLVSASSDGR

SEQ ID N° : 87, seqüência de ácido nucleico (terminal 3') parcial de FBXW de Ipomea nil EST BJ574759
GGAGAAGACGGAGCCATACACATGTTTGAGGTTGTAAAAAACGCAAGGTTTCATGTAAAGAA
GGTTGCAACATGGATTCCTCACTCGGGCCCTGTTTATTCTCTTGCATTTGAATTCCCCTGGCTT
GTTTCCGCTTCAAGCGATGGAAAACTGGCACTTATCGATGTGAGGAAGTTATTGAAGACAAGT
AGGTATTCAGCAACAGGAAGCCGTCCTTGTAAGGTTGACAATCTGGACCATACAAGTGTTGAG
CCTCCGCAACGAATGCTTCATGGATTTGGGTCCAATTTATTTGCAGTGGATGTTGGTCTGGATC
GTATTGTGTGTGGAGGTGAAGAAGGCGCTGTTAGGATCTGGAACTTCTCACAAGCGTTGGAGA
TAGAGCAGAGGGTCCGTGCTTTACGAGGTTTAAGGTTAGAAAACAGGATGAGGAGGCGTAAG
CTTCAGTCCAACATGAATAGCAAAGGCACCCGGAGTGATCAGTGCTCGGTTGCAGCAAAGAA
GAACCAAATCAATGGCGATAGGAATAGCTGGCAGAACAAGCGTAGGGTAAGCGGGAAGCTC
AAGGCTTAG

SEQ ID N° : 88, polipeptídeo deduzido parcial de FBXW de Ipomea nil
GEDGAIHMFEVVKNARFHVKKVATWIPHSGPVYSLAFEFPWLVSASSDGKLALIDVRKLLKTSRY
SATGSRPCKVDNLDHTSVEPPQRMLHGFGSNLFAVDVGLDRIVCGGEEGAVRIWNFSQALEIEQR
VRALRGLRLENRMRRRKLQSNMNSKGTRSDQCSVAAKKNQINGDRNSWQNKRRVSGKLKA

SEQ ID N° : 89, seqüência de ácido nucleico (extreidade 3') parcial de Solanum tuberosum EST CX161187
TTTTCCTTGGCCCGAATTCACACAGGGAAGCTCCTTATTCCACTGAAGGGGAAGATGGAGCTA
TACGCATGTTCGAGATCATTAGAAATGATAAAGCTTCATCTCAGGCGTGTGGCGACATGGATT
CCTCACTCGGGTGCTGTTTATTCTCTTGCATTTGAGTTCCCTTGGCTTGTTTCAGCTTCCAGTGA
TGGAAAACTCTCACTTATTGATGTGAGAAAGCTGTTGAGGACGAATCGGAGTTATGTGATGGA
GAAGCCTTCGAAGGTTGACAATAGGGCTAAAAATGTAGAGCCACCTCAAAGAATGTTACATG
GATTTGGGAGCAATCTGTTTGCTGTGGATATTGGTCTTGATCGTATCGTATGTGCTGGAGAAG
AAGGCATTGTTAGGATCTGGAACTTCTCAGAAGCTTTGGAGGTTGAGCAGAGAGTTCGAGCAT
TAAGAGGATTAAGGTTAGAGAACAGAATGAGGAGGCGTAAACTTCAGTCTGAAATGACTGGT
AAAGGCAGTCGTGCTGATCAGTGCTCAGTTGCTGCTAAGAAGAATCAATTGAATGGTGATAG
GAACAGCTGGCGCAACAAGCGTAGGGTGACTGGGAAGCTGAAGGCCTAG

SEQ ID N° : 90, polipeptídeo deduzido parcial de FBXW de Solanum tuberosum
FSLARIHTGKLLIPLKGKMELYACSRSLEMIKLHLRRVATWIPHSGAVYSLAFEFPWLVSASSDGK
LSLIDVRKLLRTNRSYVMEKPSKVDNRAKNVEPPQRMLHGFGSNLFAVDIGLDRIVCAGEEGIVRI
WNFSEALEVEQRVRALRGLRLENRMRRRKLQSEMTGKGSRADQCSVAAKKNQLNGDRNSWRN
KRRVTGKLKA

SEQ ID N° : 91, seqüência de ácido nucleico (médio) parcial FBXW de Zamia fischeri EST DY032229.1
GGCAGACATTCTCAATCGCCTCAATGCAAGAGAACTCAGCATAGTTTCATGTGTATGTACCCT
CTTTCGAAGGATTTCTTCAGATAGCCATGGATGGAAGGAGTTCTACTGTGAGAGATGGGGGCT
TCCTGCGCCTAGCAAAATTTCCGCGTCTTCGGCCCCGGGACCAGAGAAGTCGTGGAGAGAGTT
GTATTTGGAGAGAGATGCCCGAAGCAAGGCCTTCCTGGGCCGATTTAATGTAGACATGCTGCA
TGGCCACA

FIGURA 27 (continuação)

CTGCAGCTCTTCGATCTGTTTGCCTTCTGCCATCTGCAAATCTCATATTTACAGCAGGATATGA
CACAGTACTCAGAATGTGGAACATGGAGGAAGGTCTCTTGATCGCTTGTTCCCGGCCTCTTGG
TTGCACGCTCCGTGCCATAGCGGCAGACATGGATATGTTGGTGGTGGCAGGAACTGATCCTTT
CGTGCGCTGTTGGCAAGCAATTTCTGAGCTTCCTAACTTATTTGACATTTCGGGGGTTTCAGGG
CAGAACACTGAATCTTGCCTTCGTGGACATGTTGGACCCATAACTTGCCTTGGCCTAGATTCGT
TGAAAATTTACAGTGGCTCTTGGGACATGAGCATTCGGGTATGGGATAGGGTTA

SEQ ID N° : 92, polipeptídeo deduzido parcial de FBXW de Zamia fischeri
ADILNRLNARELSIVSCVCTLFRRISSDSHGWKEFYCERWGLPAPSKISASSAPGPEKSWRELYLER
DARSKAFLGRFNVDMLHGHTAALRSVCLLPSANLIFTAGYDTVLRMWNMEEGLLIACSRPLGCTL
RAIAADMDMLVVAGTDPFVRCWQAISELPNLFDISGVSGQNTESCLRGHVGPITCLGLDSLKIYSG
SWDMSIRVWDRV

SEQ ID N° : 93, seqüência de ácido nucleico (extremidade 3') parcial de FBXW de Persea americana EST CK756534.1
GCGACATGGCTTCCTCACACGGGCTCTGTTAATTCTCTAGCATTTGAGTTCCCATGGCTTGTGT
CATCTTCCAGTGATGGACGACTTGCTCTGATTGATGTGAGAAAGCTGTTAAGGTCAAGCCGAC
GTACGTCATCAAGACACCATGTTAAAGCTAAACGGTCTGCCCCCGTAAATGTGGAGCCGCCTC
AAAGGATGTTGCATGGGTATGGGTGCAATTTGTTTTCTGTGGATATTGGCGCGGATAGGATTG
TTTGTGGTGGAGAGGAGGGTGTTGTGCGTATCTGGAACTTCTCTAAGGCACTTGAGATTGAGC
AGAGGATTCGGGCTCTGAAAGGGATACGGTTGGAGAACCGAATGAGGCGGCGGAAGATACA
ATTAGAGATGAGCAGTAAGAATCGACCTGGAGATCAATGCTCTGTTGCAGCCAAAAGGAATC
AGATGCCTGGTGATAGAAACAGGGTTTGGCGTGGAAGGCGCAATATGAGTGAAAAGCCGAAG
GCCTAA

SEQ ID N° : 94, polipeptídeo deduzido parcial de FBXW de Persea mericana
ATWLPHTGSVNSLAFEFPWLVSSSSDGRLALIDVRKLLRSSRRTSSRHHVKAKRSAPVNVEPPQRM
LHGYGCNLFSVDIGADRIVCGGEEGVVRIWNFSKALEIEQRIRALKGIRLENRMRRRKIQLEMSSK
NRPGDQCSVAAKRNQMPGDRNRVWRGRRNMSEKPKA

SEQ ID N° : 95, seqüência de ácido nucleico (terminal 3') parcial de FBXW de Glycine max EST CD418593
GTTGCTGCTTGGATTCCTCACTCTGCCCCTGTGTATTCCCTGGCCTTTGAGTTTCCATGGCTTGT
TTCTGCATCAAGTGATGGCCAGCTAGCACTAATTGATGTGAGAAAGCTGTTGAGGATTAGCNA
GCGAGCTCTAGGGAAAAGAGTTTCAAAGGTAAAGCATTTGGGTGGAGACATAGTGGAGCCTC
CACAGAGGATGTTGCATGGATTTAAGAGCCATCTTTTCTCTGTGGATATAGGAGCTGAACGAA
TTGTCAGTGGAGGTGAAGAAGGTGTTGTCAGGATCTGGAATTTCACGGAAGCTTTGGAAATTG
AACGGAGAGCCCGAGCATTAAGAGGAATACGATTAGAGCACAGAATGAGGCGACGGAAGCT
TCAAACAGAGCTGAGCCATAAAAGTGGTCGGAGTGATCAGTGTTCAGTTGCAGCCCAGAAAA
ATTCTGTCACTTGTATTTGGCCCACTAAACGTGGCATGAGCGGAAAGCTGAAAGCATAG

SEQ ID N° : 96, polipeptídeo deduzido parcial de FBXW de Glycine max
VAAWIPHSAPVYSLAFEFPWLVSASSDGQLALIDVRKLLRISXRALGKRVSKVKHLGGDIVEPPQR
MLHGFKSHLFSVDIGAERIVSGGEEGVVRIWNFTEALEIERRARALRGIRLEHRMRRRKLQTELSH
KSGRSDQCSVAAQKNSVTCIWPTKRGMSGKLKA

SEQ ID N° : 97, Motivo 1
WKX(F/V/L)Y(C/R/G)ERWGXP

Onde X é qualquer aminoácido

FIGURA 27 (continuação)

SEQ ID N° : 98, Motivo 2
SL(A/S)FE(F/Y)PWLVS(A/S)S(S/A/G)DG

SEQ ID N° : 99, Motivo 3
I(Y/F)SGSWD(M/T)(T/S)(V/I)R(V/I)WDR

SEQ ID N° : 100, Motivo 4
FTGG(E/D)DG

SEQ ID N° : 101, Motivo 5
E(P/A)PQRMLHG

SEQ ID N° : 102, Região conservada 1
SQWMPHTSPVYSLSFEFPWLVSASGDGKLALIDVRKLL

SEQ ID N° : 103, Região conservada 2
VEPPQRMLHGFGSNLFSVDVGYDRIVCGGEEGTVRIWNFTQALEIERRTRALKGMRHENRMRRRR

SEQ ID N° : 104 promotor GOS2 de Oryza sativa
Atccaagccaagaagagggagagcaccaaggacacgcgactagcagaagccgagcgaccgccttcttcgatccatatcttccggtcgagttcttggtcg
atctcttccctcctccacctcctcctcacagggtatgtgcccttcggttgttcttggatttattgttctaggttgtgtagtacgggcgttgatgttaggaaaggggat
ctgtatctgtgatgattcctgttcttggatttgggatagaggggttcttgatgttgcatgttatcggttcggtttgattagtagtatggttttcaatcgtctggagagct
ctatggaaatgaaatggtttagggtacggaatcttgcgattttgtgagtacctttgtttgaggtaaaatcagagcaccggtgattttgcttggtgtaataaaagta
cggttgtttggtcctcgattctggtagtgatgcttctcgatttgacgaagctatcctttgtttattccctattgaacaaaaataatccaactttgaagacggtcccgtt
gatgagattgaatgattgattcttaagcctgtccaaaatttcgcagctggcttgtttagatacagtagtccccatcacgaaattcatggaaacagttataatcctca
ggaacaggggattccctgttcttccgatttgctttagtcccagaatttttttttcccaaatatcttaaaaagtcactttctggttcagttcaatgaattgattgctacaaa
taatgcttttatagcgttatcctagctgtagttcagttaataggtaataccccctatagtttagtcaggagaagaacttatccgatttctgatctccatttttaattatatg
aaatgaactgtagcataagcagtattcatttggattatttttttttattagctctcacccccttcattattctgagctgaaagtctggcatgaactgtcctcaattttgttttc
aaattcacatcgattatctatgcattatcctcttgtatctacctgtagaagtttcttttttggttattccttgactgcttgattacagaaagaaatttatgaagctgtaatcg
ggatagttatactgcttgttcttatgattcatttcctttgtgcagttcttggtgtagcttgccactttcaccagcaaagttc

SEQ ID N° : 105, prm6999
GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGAATCGTTTTTCTCGTTT

SEQ ID N° : 106, prm7000
GGGGACCACTTTGTACAAGAAAGCTGGGTATCCAATCTTATCGCTTAGG

SEQ ID N° : 107, seqüência de ácido nucleico de comprimento total de FBXW de Brassica rapa AC189583
ATGGAATTAGAGTGCCAAGAGAGAGCTGAAGTTGCGATTGTTGGCGAAGGTTTATCCAATCA
AGGAGCTGAGTTGAGAATTGGAGATGGGTCTGTCGAGGTTCCATGTGTGAAGCTTTGTTATCA
GCAAA

FIGURA 27 (continuação)

AGAAGTCAACTTTGGTGGTGCCCAGGGACAAGGATTTGTCTACAAATACTCATCGATACACCA
TTATCGATTTGCCTCAGGCTTTGATATCCGAGATTCTTAACTGCCTTGACCCAAAAGAGTTGGG
CCTTGTGTCCTGTGTTTCGACTTGTCTGCATAGGTTAGCTTCGGAGCATCACGCGTGGAAAGGC
TTCTACTGCGAGAGATGGGGGCTCCCCGTCGTCTTAGGGGGTAAAACTTCTGAAGAGAGGTCA
TGGAAAGAGCTGTTTATTGAGAGGGAGTTTAGGAGCAAGACTTTTTTAGGACGTCATAGTATT
GATACCCTTTATGGTCACACTGAAGCAGTTCGCACTGTGTTTCTTTTAGCTTCTGCTAAGCTCA
TTTTCACTTCTGGATATGACTGCGTTGTTCGGATGTGGGGAATGGAAGAAGGCTTGTCTATTGC
AGCGTCTAAACCGCTTGGTTGTACTATTAGAGCAGTTGCGGCTGATACGAAGCTCTTGGTAGC
TGGCGGTACTGATGGTTTTATACATTGCTGGAAAGCAGTGGATGGTCTGAGAGACTTGTTTGA
CCTCACGGGTTTTCAGAAAGAGAAGACGGAGTTTCGCCTTTGGGAGCATGAGGGTCCTATTAC
ATCTCTTGCCCTGGATATGACGAGTATCTTTAGCGGTTCGTGGGACATGTCTGTTCGTATATGG
GATAGATCTTCATTCAAGTGTATCAAAACGTTGAGGCATAGCGATTGGGTTTGGGGACTAGCG
CCTCATGAGACAAGCATCGCGTGTGCTGCTGGTTCGGATGTGTACATTTGGGACATTAGCAGT
GAGACTCCAGAGGCGATTATCCATGATGCTCATGAGGGTAACACTTACTCTCTTGCAAGAAGC
CACAGTGGGGATTTCTTGTTTACTGGTGGGGAAGATGGAGGGATCAAAATGTTTGAGATCAGA
AAACACGGTAATGAAACAAGCGTCCTGCTCATATCTCAATGGATGCCTCACACTGGTCCCGTC
TACTCTCTTTCTTTCGAGTTCCCCTGGCTTGTATCAGCATCTGGTGACGGTAAACTCGCTCTTAT
CGACGTTAGGAAGTTGTTGAAGACTAACCGTCGTGGCTTACCCAAGAGGGTTTCTTCGAGAAT
TGTGGAACCCCCACAGAGAATGCTGCACGGGTTCGGTTCAAACCTGTTCTCTGTGGACGTGGG
CTGTGACCGTATAGTTTGTGGAGGTGAAGAAGGCGTAGTTCGGATATGGAACTTCACACAAGC
GTTGGAGATTGAGAGGAGAGCTCGAGCTCTGAAAGGAATGAGGCTTGAGAACAGGATGAGGC
GAAGGAGGATGCAGATGGAGATGAATGCGAAGAGTGGAAGGCCTGACCAATGCTCGATTGCT
GCTCATAAGAACCCTGTTAATGGAGATAGGAATCGAGCGTGGCATACAAAACGAAGAGCAAG
TGGCAAGGCGAAGGCCTAA

SEQ ID N° : 108, seqüência de polipeptídeo de comprimento total de FBXW de Brassica rapa
MELECQERAEVAIVGEGLSNQGAELRIGDGSVEVPCVKLCYQQKKSTLVVPRDKDLSTNTHRYTII
DLPQALISEILNCLDPKELGLVSCVSTCLHRLASEHHAWKGFYCERWGLPVVLGGKTSEERSWKE
LFIEREFRSKTFLGRHSIDTLYGHTEAVRTVFLLASAKLIFTSGYDCVVRMWGMEEGLSIAASKPLG
CTIRAVAADTKLLVAGGTDGFIHCWKAVDGLRDLFDLTGFQKEKTEFRLWEHEGPITSLALDMTSI
FSGSWDMSVRIWDRSSFKCIKTLRHSDWVWGLAPHETSIACAAGSDVYIWDISSETPEAIIHDAHE
GNTYSLARSHSGDFLFTGGEDGGIKMFEIRKHGNETSVLLISQWMPHTGPVYSLSFEFPWLVSASG
DGKLALIDVRKLLKTNRRGLPKRVSSRIVEPPQRMLHGFGSNLFSVDVGCDRIVCGGEEGVVRIW
NFTQALEIERRARALKGMRLENRMRRRRMQMEMNAKSGRPDQCSIAAHKNPVNGDRNRAWHT
KRRASGKAKA

SEQ ID N° : 109, seqüência de ácido nucleico de comprimento total de FBXW de Sorghum bicolor contig of BI075893 CW484775 CF770159 CF770238
ATGGCCTTTGACTGCAACAAGGAAAGAGGAGATTCTTCGCCTGATAATTGCTCCCGCATTTGC
ACTGAGGGTACTCTCGTCCAGGCAAATCCCTTATCTCACTACTGGAAGCCTAAGGGTTTGAAG
AGCCTCAACAAATTGGAGAACCAGAAGTCCAGTCATGGATCTGTTCCGAGAGACGATAATCC
AGAACAAGAAGCTGAAGCGAGCGATGAGGCATCTGCCTCAACCTGTGGCATCAGGTGCTTCA
CCGATCTGCCGGCTGCTTTGGTCTGCGAGGTCCTTGCACGCCTCGATGCAAAGGACCTTGGAA
TTGTATCCTGTGTCTCCACCCTCCTGCATGCCTTAGCCACAGATCATCAGGGATGGAAGAAGT
TATACTGCGAAAGGTGGGGGCTTCCAGACCTTCCCGACGCCCTGAATGGGCCCTTGTTGCCTG
GTGCCCCCCTAGATGGGAAGTCTTGGAAAACATTTTTCGTGGAGCGGGAGTTCAGGAGTAAAT
CATTCATGGGTAGATTCAATGTGGATATTCTCCGTGGCCACAATGAGGATGTTCGAGCCGTCT
TCCTTTTGGTATCAGCAAATCTGATATTCACTGGCGTCACTGGCGGCCGCGATTCTGTGGTTAG
AATGTGGAAAATGGAGGAAGG

FIGURA 27 (continuação)

GTTATTGATTGATACATCCCGTGCACTTGGTGGCACCATCCGGGCGATTGCAGCTGACTCTAG
GCTTTTAGTGAGTGGAGGAACTAATGCCTATATTCAGTGTTGGAGGGCTGTTGAAGGCAATGA
TCACTTATTTCACATCTCTGGAAGTGGTACTGACCAGAATGCGGAGTTTCGCCTCTGGGGGCA
TGAAGGTCCTGTGACTAGTCTTGCCTTGGATTCACTGAGGATTTATAGTGGTTCTTGGGACATG
ACTGTTCGTGTTTGGGACAGAGCCCAGATGGATTGCGTGCAAAAGTTCATGCATGCAGACTGG
GTGTGGTCTCTTGCTCTGCGTGGAAATACTGTTGCCAGTACTGCTGGTAGAGATGCATATGTAT
GGGATATCAGGGACGGCGAGTTAACAAGCTTAGTTTCCAATGCCCATGTTGGTAATGCATATT
CATTGGCTTGGACACACCTGGCGGATGTGCTGTTTACTGGTGGAGAGGATGGCGCTATTCGCT
TGTTCAATGTTTCTGATGTCTCTGATGATGAGGAGGACATTAAACCAGTGGCAACTTGGGTGC
CACATTCAGGTCCTGTTCATTCTCTTGCTTTTGAGTATCCATGGCTTGTGTCAGCCTCGAGTGA
TGGCAGGAGTGATGGCAGGATTGCACTTATTGATTTGAGGAAGCTTCTGACTCCCAAAAGATC
ATCAAAAGGTCTGCGTGTTAAGAGCTTTGATGCAAGTGCCATAGAGCCTCCACAGAGGATGCT
TCATGGCGTTGGATGTGATCTCTTCTCCATCGCCATTGGTGCAGATAGGATTGTATGTGCAGGT
GAGGATGGTTCTGTTAGAGTCTGGAACTTCTCAGAAGCACTGGAGATTGAGAGGAGGGCACA
GGCTCTAAGGAGTTTGAGGCAGGAGAACCGCATGAGGCGGAGGAAGGCACAAGTAGAGATG
AATGCAAATGGTAGAAGGCCTGACCAGTGCTCAATAGCCATGAAAAAGAACCAACTGAAGGC
TGATAAGAGTGCCACTTGGCACAACAAGCGTGCCGTTAATGAGAAGGCCAAGTCTTAG

SEQ ID N° : 110, seqüência de polipeptídeo de comprimento total de FBXW de Sorghum bicolor
MAFDCNKERGDSSPDNCSRICTEGTLVQANPLSHYWKPKGLKSLNKLENQKSSHGSVPRDDNPEQ
EAEASDEASASTCGIRCFTDLPAALVCEVLARLDAKDLGIVSCVSTLLHALATDHQGWKKLYCER
WGLPDLPDALNGPLLPGAPLDGKSWKTFFVEREFRSKSFMGRFNVDILRGHNEDVRAVFLLVSAN
LIFTGVTGGRDSVVRMWKMEEGLLIDTSRALGGTIRAIAADSRLLVSGGTNAYIQCWRAVEGNDH
LFHISGSGTDQNAEFRLWGHEGPVTSLALDSLRIYSGSWDMTVRVWDRAQMDCVQKFMHADWV
WSLALRGNTVASTAGRDAYVWDIRDGELTSLVSNAHVGNAYSLAWTHLADVLFTGGEDGAIRLF
NVSDVSDDEEDIKPVATWVPHSGPVHSLAFEYPWLVSASSDGRSDGRIALIDLRKLLTPKRSSKGL
RVKSFDASAIEPPQRMLHGVGCDLFSIAIGADRIVCAGEDGSVRVWNFSEALEIERRAQALRSLRQ
ENRMRRRKAQVEMNANGRRPDQCSIAMKKNQLKADKSATWHNKRAVNEKAKS

SEQ ID N° : 111, seqüência de ácido nucleico de comprimento total de FBXW de Vitis vinifera AM440865
ATGGCATTTGAATGCCAAGAGAGTATAGAGGTTTGTTTGGTGAAAAATTCAATAGATTCTGAT
GAACAACAGGSTGGATTGAGTGATTTTAATTCCAATTTTAATCATATCACTTCAATTTTTGATG
CCAAATCTCAATTGGAAACCGAAACTGCACACCGAGAAAGTGGAGTGGTTTGTTTGTTGAAG
AATTCAATTGAAGAACGGGCTGGATTGACTCAGTTTAGTCCCGATTCTGATCACAATACTTTA
ATACRTGACTCGGGAAATTCCCAATCGGAAATTGAAATCGGAATCCAAAAACCTTCAACTTCA
AACCCCAAAGTTAACGACACTTCAGGACATTATCGTAGGTTGATAACTGATCTTCCGTCCGCG
TTGATATCGGAAATTCTTAATTGCCTTGACCCGAAAGAGCTTGGTGTGGTTTCGTGTGTCTCAA
CTGTTCTCAATAGGCTAGCGTCYGAGCACTCCGTTTGGAAAGATTTCTATTGTGAGAGGTGGG
GAGCTCCGGTTGTTTCGGGGTTTTCAGATGAGAAATCATGGAGGGAATTGTTTGTGGAGAGGG
AGTTTAGGAGCAAAACCTTTAGGGGACGATTTAGCATTGATGTTCTGTATGGTCACACTGAGG
CGGTTCGAGCTGTTTTCCTTTTGGCCTCTGCAAAGCTCATTTTCACTTCTGGGTATGATTCCATT
GTTCGAATGTGGGATATGGAAGAAGGGTTGTCAATTGCGTGTTCACGGCCTCTCGGTTGCACA
ATAAGAGCAGTGGCTGCGGATAAGAAACTGTTGCTTGCGGGGGGTAGTGATGGGTTTATTCAT
TGTTGGAGAGCTGTAGAGGGGCTCTCTTGCTTGTTTGACCTTGTGGGTTCTCAAAACCTGAGTA
CTGAGTTCCGAATTTGGGAACATGAAGGTCCTGTGACTTGTCTTGCTTTGGATATTAAGAGAA
TTTATAGTGGCTCATGGGACATGACTGTGCGCATATGGGACCGTTCTTCTTTTAAGGTTGTAAA
GGTCTTGAGGCACACAGACTGGGTTTG

FIGURA 27 (continuação)

GGGCCTTGTTCCTCGTGATACTACAGTTGCTAGCACTTCTGGCTCAGATGTGTATGTTTGGGAC
GCTGATAGTGGGACTCTGCTGACGATAATCTCTAATGCTCATGTGGGTAATGCTTATGCTTTGG
CACGGAGCCACACAGGGGATTTTCTATTCACTGGGGGAGAAGATGGGGCAATACATATGTTT
GAGGTCGTGAGTGATTGCATGGAAAGGAATGTATTGGAGGTTTCAACGTGGATTCCTCATTCT
GGTCCTGTTCATTCCCTGGCATTTGAGTTTCCCTGGCTTGTTTCAGCTTCAGCTGATGGGAGGA
TGTCACTGATTGATGTGAGAAAGCTTTTGCAAACTTGCAAGCCTTCCTTAGGGAAGAATGTTT
CCAAGGTTAGGCACAGGGACCACAAAAGTGTGGAGCCCCCTCAGAGGATGTTACATGGGTTT
GGCTGCAATTTATTCTCAGTGGACATTGGTGCAGATCGTATTGTCTGTGGAGGTGAGGAAGGT
GTGGTTAGAATTTGGAACTTCTCACAAGCTTTAGAAGCAGAGCAGAGGGCCCGTGCTTTAAGA
GGAATACGTCTAGAAAACAGGATGAGGCGGCGTAGGCTTCAAATAGAGCTGACTAGTAAGGG
TAGTCGAACTGATCAATGTTCAGTTGCAGCCAGTAAGAATCCTATTAATGGTGATAGGAGTGG
TGTGTGGCACAATAAGCGGGGAATGAGTGGCAGGATGAAAGCATAG

SEQ ID N° : 112, seqüência de polipeptídeo de comprimento total de FBXW de Vitis vinifera
MAFECQESIEVCLVKNSIDSDEQQXGLSDFNSNFNHITSIFDAKSQLETETAHRESGVVCLLKNSIEE
RAGLTQFSPDSDHNTLIXDSGNSQSEIEIGIQKPSTSNPKVNDTSGHYRRLITDLPSALISEILNCLDP
KELGVVSCVSTVLNRLASEHSVWKDFYCERWGAPVVSGFSDEKSWRELFVEREFRSKTFRGRFSI
DVLYGHTEAVRAVFLLASAKLIFTSGYDSIVRMWDMEEGLSIACSRPLGCTIRAVAADKKLLLAG
GSDGFIHCWRAVEGLSCLFDLVGSQNLSTEFRIWEHEGPVTCLALDIKRIYSGSWDMTVRIWDRSS
FKVVKVLRHTDWVWGLVPRDTTVASTSGSDVYVWDADSGTLLTIISNAHVGNAYALARSHTGDF
LFTGGEDGAIHMFEVVSDCMERNVLEVSTWIPHSGPVHSLAFEFPWLVSASADGRMSLIDVRKLL
QTCKPSLGKNVSKVRHRDHKSVEPPQRMLHGFGCNLFSVDIGADRIVCGGEEGVVRIWNFSQALE
AEQRARALRGIRLENRMRRRRLQIELTSKGSRTDQCSVAASKNPINGDRSGVWHNKRGMSGRMK
A

SEQ ID N° : 113, Zea mays – clone de entrada
ATGGCGGACAAGGAGCCTGTCGTGGAGCGATCCGAGGCGTCTGAGGAGGAGGACGCCTCCGC
CGCGGCCGCCGCGGGGGAGGAAGAGGACACGGGTGCCCAGGTGGCCCCCATCGTGCGGCTTG
AGGAGGTACTCGTCACCACCGGTGAGGAGGACGAGGACGTCCTGCTCGACATGAAGGCGAAG
CTGTACCGGTTTGATAAAGACGGGAATCAATGGAAGGAACGGGGCACGGGCACCGTCAAGCT
TCTCAAGAACAAGGAGACCGGCAAGGTCCGACTGGTCATGCGCCAGGCCAAGACGCTTAAGA
TCTGCGCGAACCACCTCGTGGCCCCCACCACGAGAATGCAGGAACATGCCGGCAGCGACAAA
TCGTGCGTCTGGCACGCTTCTGACTTTGCGGATGGCGAACTCAAGGAGGAGATGTTTGCCATC
AGGTTTGGTTCGGTAGAAAACTGCAAGAAGTTCAAGGAGTTGGTTGAGGAGATTGCTGAGTC
ACTTACAGAGAACGAGGACAAGGAAGGTCAAGATGGCTCTTCCACGGCCGGATTGCTGGAGA
AGCTCACTGTGAGCGAGGGCAAATCTGAGGAAAGTGGCAAGTCGGAGTCCACTGATTCTGGC
AAAGTGACCGAAACCAAAGCCGATGCAGCCCCAGCCGAGTAGTTGGTGTGGTTGCACTACCC
AGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTTGCAACGAAC

SEQ ID N° : 114, Zea mays – tradução 1
MADKEPVVERSEASEEEDASAAAAAGEEEDTGAQVAPIVRLEEVLVTTGEEDEDVLLDMKAKLY
RFDKDGNQWKERGTGTVKLLKNKETGKVRLVMRQAKTLKICANHLVAPTTRMQEHAGSDKSCV
WHASDFADGELKEEMFAIRFGSVENCKKFKELVEEIAESLTENEDKEGQDGSSTAGLLEKLTVSEG
KSEESGKSESTDSGKVTETKADAAPAE

FIGURA 27 (continuação)

SEQ ID N° : 115, Zea mays – tradução 2
MADKEPVVERPEAGEEEEDASAAAAAGEEEDTGAQVAPIVRLEEVAVTTGEEDEDVLLDMKAKL
LPIRQRRMRQAKTLKICANHLVASTTKMQEHAGSDKSCVWHAADFADGELKEEMFAIRFGSVEN
CKKFKELVEEIAESLTKNEGKEGEDGGSTAGLLAKLTVSEGKSEGSGKSESTDSGKVTETKADTAP
AE

SEQ ID N° : 116, Arabidopsis thaliana – clone de entrada
ATGCCAACTTTGTACAAAAAAGCAGGCTTCACAATGGCGAGCATTAGCAACGAGCCCGAGCG
TGAGAACAGAGACGAAGAAGAGACTGGAGCCAACGAAGATGAAGACACCGGTGCTCAGGTT
GCTCCTATCGTCAGGCTTGAAGAAGTCGCCGTCACCACCGGTGAAGAAGACGAAGACACCAT
CCTCGATCTGAAATCGAAGTTGTATCGATTTGATAAAGATGGAAGTCAGTGGAAGGAGAGAG
GTGCTGGTACTGTTAAGTTTTTGAAACATAGAGTTTCTGGGAAGATTCGTCTCGTTATGAGGC
AATCGAAAACTTTGAAGATCTGTGCTAATCATCTTGTTGGATCGGGTATGAGTGTTCAGGAAC
ACGCTGGGAATGATAAGTCTTGTGTATGGCACGCTCGTGATTTCTCCGATGGTGAATTGAAGG
ATGAGCTCTTCTGTATCCGGTTTGCTTCAGTTGAGAATTGCAAAGCATTTATGCAAAAGTTCAA
GGAAGTAGCTGAATCTGAAGAAGAGAAAGAAGAGAGCAAAGATGCCTCTGATACCGCTGGTC
TTCTTGAGAAGTTAACAGTGGAAGAGAAGGAAAGTGAGAAGAAACCAGTGGAGAAGGCAGA
GGAAAACAAAAAGAGTGAAGCTGTTGAAGAAAAGAAAACAGAGGAGTCTGTTCCCTCAGCTT
AAGATGCACCCAGCTTTCTTGTACAAAGTTGGCATTATAA

SEQ ID N° : 117, Arabidopsis thaliana – tradução 1
MPTLYKKAGFTMASISNEPERENRDEEETGANEDEDTGAQVAPIVRLEEVAVTTGEEDEDTILDLK
SKLYRFDKDGSQWKERGAGTVKFLKHRVSGKIRLVMRQSKTLKICANHLVGSGMSVQEHAGND
KSCVWHARDFSDGELKDELFCIRFASVENCKAFMQKFKEVAESEEEKEESKDASDTAGLLEKLTV
EEKESEKKPVEKAEENKKSEAVEEKKTEESVPSA

SEQ ID N° : 118, Arabidopsis thaliana – tradução 2
MASISNEPKRENRDEEETGANEDEDTGAQVAPIVRLEEVAVTTGEEDEDTILDLKSKLYRFDKDGS
QWKERGAGTVKFLKHRVSGKIRLVMRQSKTLKICANHLVGSGMSVQEHAGNDKSCVWHARDFS
DGELKDELFCIRFASVENCKAFMQKFKEVAESEEEKEESKDASDTAGLLEKLTVEEKESEKKPVEK
AEENKKSEAVEEKKTEESVPSA

SEQ ID N° : 119, NM_100588.2| Arabidopsis thaliana SIRANBP; AT1G07140 (SIRANBP) ligando Ran GTPase
GCGAAACCCCCAAATTCTCTTCTCTCTATCTCTCTCGCGTACGCCGCACCGTAGCGACCAATTG
AATCCACGAGGAAATCTCAGTAAATACTATGGCGACCAACGAACCCGAGCACGAGCACAGAG
ACGAGGAAGAGGCCGGAGCTAACGAGGATGAGGACACCGGAGCTCAGGTCGCTCCGATCGTT
AGGCTTGAGGAGGTTGCCGTCACTACCGGCGAGGAAGACGAAGATGCCGTCCTTGATCTGAA
ATCGAAGCTTTATCGATTCGATAAGGATGCGAATCAGTGGAAGGAGAGAGGAGCTGGTACTG
TGAAGTTCTTAAAGCATAAGAACACTGGGAAGATTCGTCTCGTTATGAGGCAATCTAAAACTT
TGAAGATCTGTGCTAATCACTTCGTTAAATCGGGCATGAGTGTTCAGGAACACGTTGGGAATG
AAAAGTCATGTGTGTGGCACGCTCGTGACTTTGCTGATGGTGAACTCAAGGATGAGCTTTTCT
GTATCCGATTTGCTTCTATTGAGAATTGCAAAACATTTATGCAAAAGTTCAAGGAAGTTGCTG
AGTCTGAAGAAGAGAAAGAAGAGAGCAAAGATGCCGCTGACACTGCTGGCCTTCTTGAGAAA
TTGACTGTGGAAGAGACAAAAACGGAGGAGAAAACCGAAGCGAAAGCTGTGGAGACGGCAA
AGACTGAAGTGAAAGCAGAAGAAAAGAAAGAGAGCGAGGCAGAGAAATCTGGTGAAGCAAA
GAAAACAGAAGAAAGTGGTCCCTCAACATAAGAAGCGTCATCATTTAAGTTGCCAAATCCTG
GCGAGGTAATTAAGCCTCGAAATGTTTTGATGCATCATGAGTCTACCATTGTCTTGGCACTATC
TATCTATACATGTTTTGTCGAGTCATATCAAGTGGT

FIGURA 27 (continuação)

TGGGGGATCGCTTGGGGTCGGGTTTTACTGAATGCTGAGTCGTTATGGGTCTAATAGTTTTTGA
GTCTAAAGTGTCGGGTGATATGAGAGATTTGGGTGAAATATTTTTTCATGTTGGTTATCTGAA
AGAGTACATTGGTTTCATTGTTTTAAGTTTTCTCATGGATCCTTTTTGGATGGTCTTATTTTGAG
GATACAAATGTGTTTGTCCATGGACAAGAATCAGAAGTCTCCATTTTTTTTTTTTCAAAATTTA
ATGCATGTGATTTTTTAATCTTAAGGTAATTGCCAATTGTTTGACAAAAAAAAGGTAATTGCC
AATCTACTTTGCTCTCTTGTGTTAGTCGATGCTAGTCTTGACCCTGATAAAGATCCAAAATGTA
TATTAACAGTATTCATAAAATTCTTCAGTTATAACGAACTGTTCACGGAAAAAAAATGACATT
GTACTATACTGTGCTAAAATTACCAAAGCATGATTTATATAAATCATATCCAGTAAGACC

SEQ ID N° : 120, At1g07140|F10K1.15 (ORF 93-779)
MATNEPEHEHRDEEEAGANEDEDTGAQVAPIVRLEEVAVTTGEEDEDAVLDLKSKLYRFDKDAN
QWKERGAGTVKFLKHKNTGKIRLVMRQSKTLKICANHFVKSGMSVQEHVGNEKSCVWHARDFA
DGELKDELFCIRFASIENCKTFMQKFKEVAESEEEKEESKDAADTAGLLEKLTVEETKTEEKTEAK
AVETAKTEVKAEEKKESEAEKSGEAKKTEESGPST

SEQ ID N° : 121, gi|90568019|gb|BT024924.1| Arabidopsis thaliana At5g58590 mRNA, cds completos
ATGGCGAGCACTGAGCCAGAGCGTGAGAACCGTGAAGATGAAACCGAAGTCAACGAAGATG
AAGACACCGGAGCTCAGGTAGCTCCGATCGTTAGGCTTGAAGAGGTTGCAGTCACCACCGGC
GAAGAAGATGAAGACGCCGTCCTCGATCTGAAATCGAAGATGTATCGATTCGATAAAGAAGG
GAACCAATGGAAGGAGAGAGGAGCTGGAACTGTGAAGTTATTGAAGCATAAGGAAACTGGA
AAAGTTCGTCTTGTTATGAGACAATCTAAAACCCTAAAGATCTGTGCCAATCACCTCATTTCGT
CTGGGATGAGTGTTCAGGAACATAGTGGGAATGAAAAGTCTTGTTTATGGCACGCTACTGATT
TCTCCGACGGCGAGTTGAAAGATGAGCTTTTCTGCATTCGATTTGCTTCCATTGAGAATTGCAA
AACATTTATGGAGAAGTTCACTGAAATAGCTGAAAGCCAGCAAGTAGGGAAAGAGAGCACCC
AGGGCGACGAGGCTGCTGGTTTAATAGAGAATCTTTCGGTTGAGGAAAATATAAGTGAGGAG
AAAGCCAAGGAAGCAGAAGAGAAAGAGCCTGCAAAGGAAGATAAAGAAACAAAAAAGGAG
AAGGTAGAAGAAGAGAAGAAAACAGAGGCAAGCACTTAA

SEQ ID N° : 122, At5g58590|MZN1.4
MASTEPERENREDETEVNEDEDTGAQVAPIVRLEEVAVTTGEEDEDAVLDLYVTRSVNILVSLPLV
KMYRFDKEGNQWKERGAGTVKLLKHKETGKVRLVMRQSKTLKICANHLISSGMSVQEHSGNEK
SCLWHATDFSDGELKDELFCIRFASIERKMVWKILEWLTDSVASTDCKTFMEKFTEIAESQQVGKE
STQGDEAAGLIENLSVEENISEEKAKEAEEKEPAKEDKETKKEKVTQSVIDMFGVAAKGTIMWCF

SEQ ID N° : 123, AF370443.1|AF370443 proteína de ligação de Ran-1 (RanBP1) de Lycopersicon esculentum
AAAAAAAGCAAAAAGAAAAGAGGCAAAGTTATCAAATCAAAAACCCTACTTCCTCCCTTGTT
GTTCTTCTTCTTCTTCAAATCCGGGGTAAATCTTTCTAAAACCAAAAATCTAGAGTGATGGCAA
GCACTACTGTTGAACCCAAATTGGAGAAGAAAGAAGAGGAAGTAGAAGCAGAAGTAGAAGA
AAACCCAACTGGCGATGATGAAGACACTGGTGCACAAGTTGCTCCCATTGTTAGGCTACAAG
AAGTTGCTGTCTCCACTGGTGAAGAAAATGAACATGTTCTTCTTGATCTGAAATCAAAGTTAT
ACCGATTTGACAAGGAAGGGAGTCAATGGAAAGAGAGAGGTGTTGGGACTGTGAAGCTTCTC
AAGCACAAGGAAACTGGAAAGGTTAGGCTTGTGATGAGGCAATCCAAGACGCTGAAAATCTG
TGCCAACCACTTGGTCCTTCCTACTATGTCGATCCAAGAACATGCTGGGAATGAGAAGTCTTG
TGTGTGGCATGCTGCTGACTTTGCTGATGGAGAACTGAAAGATGAGACTTTTTGCATCCGCTTT
GCTTCAGTTGAGAATTGCAAGGCTTTCAAAGAGAAGGTTGAAGAAATTGCTGAATCTCAACA
GACAAAGTCTGGACAAAGTGAAGAAGCTGGTGCTGCTGCTACTGAACTTATTGAAAAGCTTA
GTGTTGAAAGCAAAGATAAAAATGACAA

FIGURA 27 (continuação)

ACCTGAAGACAAAGAGGCCCCTGCAGCTACAGAGGAAAAGGAAGATAAGAAGGAAGAGAAA
GCTGAAGAAAAGAATTAAAATGTTTGTATTGTCGGTCAGTTTTTTTTTTTGAAAGGTTAATAGC
ATTCGTTGTTTCTGTCTGATCCAATAGATATGATAGATATAAGCAATGTGTCTCTTGTGGTCTT
ATTTGTTGTTGTTGGGATGGTTTGGATTTTCATTTGGGTCTTGAGTCGAGTGCAGCTTCATGTTT
TATTGTGGTCTTTGGTGTTTCCTTGTACTTATTTATTGTGTGTTTGCAAATTTTGGTTCATGGGG
TACACCACCCAGTATAAGGGATTACGTTTGTTAAAACCTTTAGCACTGTTAGTGAGATCATTTT
TGAGTT

SEQ ID N° : 124, AF370443 - Lycopersicon esculentum - proteína de ligação de Ran-1 (RanBP1)
MASTTVEPKLEKKEEEVEAEVEENPTGDDEDTGAQVAPIVRLQEVAVSTGEENEHVLLDLKSKLY
RFDKEGSQWKERGVGTVKLLKHKETGKVRLVMRQSKTLKICANHLVLPTMSIQEHAGNEKSCV
WHAADFADGELKDETFCIRFASVENCKAFKEKVEEIAESQQTKSGQSEEAGAAATELIEKLSVESK
DKNDKPEDKEAPAATEEKEDKKEEKAEEKN

SEQ ID N° : 125 - AU183975 – Oryza sativa – seqüência parcial
CAACCTCTCCCCCCACGATCGCCCTCCTCCCCGAAACTTCTGGAACCGAGAGCAGCAGACGTT
AGCTCTTCTCCTCCGATGGCGGACAAGGAGCCCGTCGTGGACCGCCCCGAGGACGAGGAGGA
GGCCTCCGCCGCCGCCGCCGCCGCGGGCGGCGAGGAGGAGGACACGGGCGCCCAGGTCGCCC
CCATCGTGCGGCTCGAGGAGGTCGCCGTCACCACCGGCGAGGAGGACGAGGACGTGCTCCTC
GACATGAAGGCGAAGCTTTACCGGTTTGACAAGGAGGGGAACCAGTGGAAGGAGCGGGGGA
CGGGCACCGTCAAGCTCCTCAAGCACAAGGAGAACGGCAAGGTCCGCCTCGTCATGCGCCAG
GCCAAGACGCTCAAGATCTGCGCGAACCACCTAGTTGCTTCGACCACGAAGATGCAGGAGCA
TGCGGGCA

SEQ ID N° : 126, AU183975 – Oryza sativa – seqüência parcial
MADKEPVVDRPEDEEEASAAAAAAGGEEEDTGAQVAPIVRLEEVAVTTGEEDEDVLLDMKAKLY
RFDKEGNQWKERGTGTVKLLKHKENGKVRLVMRQAKTLKICANHLVASTTKMQEHAG

SEQ ID N° : 127, BT017282.1| Zea mays clone EL01N0314E02.c seqüência de mRNA
GCACGCGGTAAACACCCCACGTCTCAAGCTGCCCCTTTCTCACCGCCGCGGCCGCCACGCGCA
GCAGCAACCGAGCAGGAGCGCAGCCAGCAAGCTCAGGCCCGAGCCCTACTGCAACCGCCGCT
GCTACCGCACTGAACACCATGGCCGACCCGGCGGAGCACCGTCCAGCGGAGGAGGACGAGGA
AGCCGCGGCGGCCGGCGAGGATGAGGACACCGGCGCTCAGGTCGCGCCCATCGTGAAGCTGG
AGGAGGTCGCCGTTACCACCGGAGAGGAGGACGAGGACGTACTTGTCGACATGAAGGCGAAG
CTCTACCGGTTCGACAAGGAGGCGAACCAGTGGAAAGAACGTGGCACAGGCACTGTGAAGCT
GCTTAAGCACAAGGAGACCGCCAAGGTCCGCCTCGTCATGCGTCAGGCGAAGACGCTTAAGA
TCTGTGCTAACCATCTTGTGGTGGCGACGACTAAGATGCAGGAGCACGCAGGGAGCGACAAG
TCGTGCGTGTGGCACGCGCTGGACTTCGCCGACGGTGAGCTCAAGGAGGAGATGTTTGCTATC
CGTTTTGGATCTGTCGAGAATTGCAAGAAGTTCAAGGACACTGTTGAAGAGATAGCTGAAGA
GCAAGGAAAGACCGAGGTGAAAGAAAACGAGGAAGTCTCCATTGCCGCCGCAGATTTGGTAC
AGAAGTTAACAGTGACAGAGGCTAGCAACGAGGGAGATGCAGAGGAAGAGGAGGCTCCTGC
ATCTGACCATAAGAAGGATGCTAAAGGGTAAAGATTGAAGGCAAACAAGGCCAAATGATTAT
GATTCCATTCATTTCGTTGTCAAGGTTCATCTTCCCCACAAATTTTCATTACAAAGTTTCATTGC
ATTTTCTCCTGAGATGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTGGGTGAAATCTGA
GTTGTCAGTCATCTACATGTCTTTCTTGGTTGTTAGACTTATTCTCAGTCGCCTGTTTATCTGGG
ATGGCTAGGTACATCATGTTTGTACAATTATTTGGGGTTGATTTAAAAAAAAGAAAAAAAAA

FIGURA 27 (continuação)

SEQ ID N° : 128, BT017282.1 – Zea mays
MADPAEHRPAEEDEEAAAAGEDEDTGAQVAPIVKLEEVAVTTGEEDEDVLVDMKAKLYRFDKE
ANQWKERGTGTVKLLKHKETAKVRLVMRQAKTLKICANHLVVATTKMQEHAGSDKSCVWHAL
DFADGELKEEMFAIRFGSVENCKKFKDTVEEIAEEQGKTEVKENEEVSIAAADLVQKLTVTEASNE
GDAEEEEAPASDHKKDAKG

SEQ ID N° : 129, AK069970.1| Oryza sativa (grupo de cultivar japônico) cDNA clone:J023039I06
GGACGCGCAGCAACCGGCAGCAGCAGCAGAGGAGGAGGCCAGGGCAACCCGCAACCCCAAA
CCCTACTCCCGCCGCGCCACCGCCATGGCCGACCCAGCGGAGCACCGTGAGGAGGAGGAGGA
GGCCGCGGCGGCGGCGGCGGGCGAGGACGAGGACACCGGCGCCCAGGTCGCGCCCATCGTCA
AGCTCGAGGAGGTCGCCGTCACCACCGGCGAGGAGGACGAGGAGGTCCTCCTCGACATGAAG
TCGAAGCTGTACCGCTTCGACAAGGAGGGGAACCAATGGAAGGAGAGAGGCACCGGCACGG
TGAAGCTGCTGAAGCACAAGGAGACCGGCAAGGTCCGTCTCGTCATGCGCCAGGCTAAGACG
CTCAAGATCTGCGCCAATCACCTCGTGGCGACGACCACGAAGATGCAGGAGCACGCCGGCAG
CGACAAGTCGTGCGTGTGGCACGCGCTGGACTTCGCCGACGGCGAGCTCAAGGAGGAAATGT
TCGCCATACGCTTTGGATCCGTCGAGAACTGCAAGAAGTTCAGGGAGATGGTAGAAGAGATT
GCTGAGCAGCAAGGAAAGAACGAGGAGAAAGAAAATGAGGAAGTCTCCTCTACTGCAGGGT
TGGTTGAGAAGCTTTCAGTGACCGAGACAAAAAAGGAGGAAAATGCAGAGAAAGAGGAGAC
TCCTGCAGAAGAGGATAAGAAGGACGCTAAAGAGTGAAAGCACCCGCAACCTCTAGGCAGCT
TTGATATGCCTGAAGCGAGTGAAGTGTTAAATGAGCGTCATTTCATTCATTTTGTGGTCAAAG
TTCTTCCTTTCACAAATTTTCATTGTGTTTTCTTTTGGATGAATGTTTGTGCTAGGCAAAATTTG
AGTTGTATCAGTCGGGTTCTTGGTTACTACACTGTTACTTAAACCTTGAGTTGTTTATCCGAGA
TGGGGTGGCGGAGTGGGGGCACAGCATGTTTGTACAATTATTTGAAGTTGCTTTTGGAATGGG
CACGGTTTGGGTTGTCCAAAATTTGGACGGTGATGCCCTGTCACTCACAATTCTTATCTCTGTC
TTGCAATTATATGCGATGTGAAGCTCTTCGCCTCTTCGGTGACGAGTTGTCG

SEQ ID N° : 130, AK069970.1 – Oryza sativa
MADPAEHREEEEEAAAAAAGEDEDTGAQVAPIVKLEEVAVTTGEEDEEVLLDMKSKLYRFDKEG
NQWKERGTGTVKLLKHKETGKVRLVMRQAKTLKICANHLVATTTKMQEHAGSDKSCVWHALD
FADGELKEEMFAIRFGSVENCKKFREMVEEIAEQQGKNEEKENEEVSSTAGLVEKLSVTETKKEEN
AEKEETPAEEDKKDAKE

SEQ ID N° : 131, scaff_IX.923 [767] – Populus sp.
TCTCCTCTCCAAGCAAGAACATCAAATCTAATGGCAAGCACAGAACCCGAACACGAGCACAG
AGAAGATGAGGAAGCTCCGGCTGGCGAAGACGAAGACACCGGTGCTCAGGTTGCTCCGATCG
TCAAGCTCGAAGAAGTTGCTGTCTCTACCGGTGAAGAAGATGAAGATGCGATCCTCGATCTGA
AATCGAAGCTTTATAGATTTGATAAGGACGGGAATCAGTGGAAAGAGAGAGGTGCTGGCACT
GTCAAGTTATTGAAGCATAAAGAGTCTGGAAAAGTTCGTCTTGTTATGAGACAATCTAAGACT
CTCAAGATCTGCGCTAATCATCTCGTGCTGCCGACTATGTCCGTGCAGGAGCACGCAGGGAAT
GATAAGTCGTGTGTGTGGCACGCTACAGATTTCGCTGATGGTGAATTGAAGGATGAGTTGTTC
TGCATTCGTTTCGCATCTGTTGAAAATTGCAAAACCTTTATGGAAATGTTTCAAGAAGTTGCTG
AATCCCAAGAGAGCAAAGAGGAAAATGAGGATGCAACTGTTGCTGCTGATGCATTGGAGAAG
TTGAGTGTTGAAGGGAAGAAAACTGAGGAGAATGCTGGCGAAGAGGCACCTGCTGCAACCAA
GAATGAGGAAACTAAGACTGATACAGATAAGAAAGGGGAGAAGCCTGCTCCCTCAACTTGAG
GGTTGATTTGTTTGTTTTGCCATTCCCTTGGCAGTATGATGAGTTCAGTTGTCAACCATATTTCT
TGTCCATTTTGTGG

FIGURA 27 (continuação)

SEQ ID N° : 132, Populus sp. - comprimento total de tradução
MASTEPEHEHREDEEAPAGEDEDTGAQVAPIVKLEEVAVSTGEEDEDAILDLKSKLYRFDKDGNQ
WKERGAGTVKLLKHKESGKVRLVMRQSKTLKICANHLVLPTMSVQEHAGNDKSCVWHATDFAD
GELKDELFCIRFASVENCKTFMEMFQEVAESQESKEENEDATVAADALEKLSVEGKKTEENAGEE
APAATKNEETKTDTDKKGEKPAPST

SEQ ID N° : 133, sugarcanelcl|SCBGLR1117A05.g gi|1004092| gb|AAB00072.1| (L41692) ciclofilina específica de retina [Bos taurus]
GGAGAACACCTACCCGTCTCCCCACCCTAGCCCCGCCGTCGCGTCCTCTTCGAATCGAATCGC
GCCGCGATCACCTCATCTCATGGCGGACAAGGAGCCTGTCGTGGAGCGACCCGAGGCGGCTG
AGGAGGAGGACGCCTCAGCCGCCGCCGCTGCCGCGGGGGAGGAGGAGGACACGGGCGCCCA
GGTGGCCCCCATCGTGCGGCTTGAGGAGGTAGACGTCACCACCGGCGAGGAGGACGAGGACG
TCCTGCTCGACATGAAGGCGAAGTTGTACCGATTTGACAAAGACGGGAACCAATGGAAGGAA
CGGGGCACCGGCACCGTCAAGCTTCTCAAGCACAAGGAGACCGGCAAGGTCCGACTCGTCAT
GCGCCAGGCCAAGACGCTTAAGATCTGCGCGAACCACCTCGTGGCCTCGACCACGAAGATGC
AGGAGCATGCCGGCAGCGACAAGTCATGCGTCTGGCACGCTGCTGACTTTGCCGATGGCGAA
CTCAAGGAGGAGATGTTTGCCATCAGGTTTGGTTCCGTAGAAAATTGTAAGAAGTTCAAGGAG
TTGGTTGAGGAGATTTCTGAGTCGCTTACAAAGAACGAGGGCAAGGAAAGTGAAGATGGTTC
TTCCACAGCTGGATTGCTGGAGAAGCTTACTGTGAGTGAGCACAAATCTGAGGAAAGTGACA
AGTCGGAGTCCACTGATTCTGGCAAAGTGACCGAAACCAAAGCCGACACAGCCCCAGCTGAG
TAGTTGGTGGTGGCAGATCCTCCCGGTGCCAAATCAGTTTGTGTCCTTCCAGTGGAAGTTTGGT
GCCCATTTGCCATGAAATATGACTGCTAACATTTGGGGGCGAGTTGGAGCAGTCTCAGATGTT
TGGATTTGGTCTAGTCTGGTTTGGTCCGGGCTTGATTTTGTTAAAAACTTAGTACATTGGGGTT
GGAACGTTTGGTATGCGAGTCGCTAGTATTTCACTGCATGGATTGTTATGGATTGCGGTATAA
GGTGCATCTTATATTTTTTTTATTTCGTTCAATACACATACATGTGTTGTATTAATACTTTATGC
CAATGCCCATGGGGAGAGTTGAATTTGAATGTCGAGAGTGG

SEQ ID N° : 134, comprimento total de tradução de tradução de cana de açúcar 1
MADKEPVVERPEAAEEEDASAAAAAAGEEEDTGAQVAPIVRLEEVDVTTGEEDEDVLLDMKAKL
YRFDKDGNQWKERGTGTVKLLKHKETGKVRLVMRQAKTLKICANHLVASTTKMQEHAGSDKSC
VWHAADFADGELKEEMFAIRFGSVENCKKFKELVEEISESLTKNEGKESEDGSSTAGLLEKLTVSE
HKSEESDKSESTDSGKVTETKADTAPAE

SEQ ID N° : 135, sugarcanelcl|SCCCLR2002H12.g gi|417396| sp|P32499|NUP2_YEAST NUCLEOPORIN NUP2 (NUCLEAR PORE PROTEÍNA NUP2) (P95) emb|CAA49587.1| (X69964) nucleoporina [Saccharomyces cerevisiae]
ACAGCAGCAACCGGTAGCAGGAGCGCAGCCAGCAAGCGCAGGCCCCGAGCCGTACTGCAACC
GCCGCTGCCACCGCGCCGAACGCCATGGCCGACCCGGCGGAGCACCGCCCTGCGGAGGAGGA
GGAGGAGGCCGCGGCGGCCGGCGAGGATGAGGACACCGGCGCCCAGGTCGCGCCCATCGTA
AAGCTGGAGGAGGTCGCCGTTACCACCGGAGAGGAGGACGAGGACGTGCTCCTCGACATGAA
GGCGAAGCTTTACCGGTTCGACAAGGAGGGGAACCAGTGGAAAGAGCGCGGCACTGGCACTG
TGAAGCTGCTCAAGCACAAGGAGACCGCCAAGGTCCGCCTCGTTATGCGTCAGGCGAAGACG
CTTAAGATCTGCGCTAACCATCTCGTGGTGGCGACGACTAAGATGCAGGAGCACGCGGGGAG
CGACAAGTCGTGCGTGTGGCACGCGCTGGACTTCGCCGACGGCGAGCTCAAGGAGGAGATGT
TCGCTATCCGTTTTGGATCTGTCGAGAATTGTAAGAAGTTTAAGGATATTGTTGAAGAGATAG
CTGAACAGCAAGGAAAGACTGAGGAGAAAGAAAACGAGGAAGCCTCCACTGCCGCTGATTTG
GTACAGAAGTTAACAGTGACGGAGGCCAGCAAGGAGGAAACTGCGGAGAAAGAGGAGGCTC
CTGCATCTGGCGATAAGAAGGATGCTAAAGATTGAAGGCGTTTATATACGCAAACAATGGTC
AAATGAGTGTCCTTCCATTCATTTTGTTGTCAAGGTT

FIGURA 27 (continuação)

CATCTACCCTCCACAAATTTTCATCGTGTTTTCTCTGGGATGAATGCTTGTGTTAGGTGAAATC
AGAGTCATCAGTCATCTACATGGTTTTCTTGGTCATAGACTGTTATTTTTAAGTCGCCTGTTTA
TCTGGGATGGCTGGGGTCAGCATGTCTGTACAATTATTTGGAGTTGCTTTTGGAATGGACGCG
GTTTGATGAGTAGCCTAAAGCTTGAGATGGTGGTGATGTCTTTTCGCTCCACTTTTCTTATTC

SEQ ID N° : 136, comprimento total de tradução de cana de açúcar 2
MADPAEHRPAEEEEEAAAAGEDEDTGAQVAPIVKLEEVAVTTGEEDEDVLLDMKAKLYRFDKEG
NQWKERGTGTVKLLKHKETAKVRLVMRQAKTLKICANHLVVATTKMQEHAGSDKSCVWHALD
FADGELKEEMFAIRFGSVENCKKFKDIVEEIAEQQGKTEEKENEEASTAADLVQKLTVTEASKEET
AEKEEAPASGDKKDAKD

SEQ ID N° : 137, Medicago parcial lcl|AC135466_14.2 AC135466.16 75025-70551 E EGN_Mt050401 20050623 RanBP1; EVH1 ORF(1, 663)
ATGTCTACAAGCACCGATCATCACGAAGAGGAAGATGTTCCCGCCGGCGATGAAGAAGACAC
CGGCGCTCAAATCGCTCCGATCATCCAACTTCATGAAGTCGCTGTTTCTACTGGTGAAGAAGA
TGAAGAAGCTATTCTCGACCTAAAAGCGAAACTGTACCGATTTGACAAGGTAGGGAATCAAT
GGAAGGAACGAGGGGCAGGAACTGTTAAGTTTTTGAAGCATAAGGTTACTGGAAAAGTTAGG
CTTGTTATGAGACAATCAAAGACTCTTAAGATCTGTGCTAATCATCTCATTTTGCCTAAAATGA
CTGTGCAAGAGCATGCTGGGAATGAGAAATCTTGTGTTTGGCACGCGAAGGACTTTGCTGATG
GTGAATTGAAAGACGAGTTTTTCTGCATTCGATTTGCGTCAATTGAAAATTGCAGAAAGTTCA
TAGAGACATTTCAAGAAATTGCTGAATCCCTGAACAAAGAGGAAAGCAAGGATGCATCCACT
GCTGCTGATCTCCTTGAGAATTTGAGTGTTGAAGGGAAAGCAGATGCAGAAAAGAAAGATAA
AGAGAAATCTGAGGAGAAAACTAAGGAGAAAGAGACACCAGAAAAAGAAAGCAAGGAAGA
TACAGAAAAGAAAGTTGAAGAGCCTGCTTCATCTGCTTGATTATGATATGTTCATATTTTCGA
CAAGGCAATATGGAAAAGTTTGGTGGTTGACCTTCGTGCCACTCTTATCAATACTCTCTCATA
GTACTAGTCTTCAGGTTGTTTTGTTGCTACGTTATTGAGTGGTTGATATCCTTCGTTGAATTATT
GTCAGCAAGGTTGGTTTTTTGAGTCGGGTTTGAATCATTGAGATTGGCGCTTTTGGTCTATGGG
TCTATGGGAGTTGTTATTTTCGTTCTATTCATTTATTTAGTCGAAATTTGAATGAGTCATGTTGG
TTTGGTGTCGGGGATGGGGAGGGTGCAGTCGGGAAAAATTAGTAGTAAGTACAAACAAAGGT
TTTGAATGCATATTATTGAGTATAACATTTTATA

SEQ ID N° : 138, tradução
MSTSTDHHEEEDVPAGDEEDTGAQIAPIIQLHEVAVSTGEEDEEAILDLKAKLYRFDKVGNQWKE
RGAGTVKFLKHKVTGKVRLVMRQSKTLKICANHLILPKMTVQEHAGNEKSCVWHAKDFADGEL
KDEFFCIRFASIENCRKFIETFQEIAESLNKEESKDASTAADLLENLSVEGKADAEKKDKEKSEEKT
KEKETPEKESKEDTEKKVEEPASSA

SEQ ID N° : 139, motivo 2 – Zea mays

EEEDTGAQVAPIVRLEEVAVTTGEEDEDVL

SEQ ID N° : 140, motivo 3 – Zea mays

RQAKTLKICA

SEQ ID N° : 141, motivo 4 – Zea maysNHLVA

SEQ ID N° : 142, motivo 5 – Zea mays

DKSCVWHAADFADGELK

FIGURA 27 (continuação)

SEQ ID N° : 143, motivo 6 - Zea mays
EIAES

SEQ ID N° : 144, motivo 7 – Zea mays

LAKLTVSEG

SEQ ID N° : 145, motivo 2 – Arabidopsis thaliana

EDEDTGAQVAPIVRLEEVAVTTGEEDEDTI

SEQ ID N° : 146, motivo 3 – Arabidopsis thaliana

RQSKTLKICA

SEQ ID N° : 147, motivo 4 – Arabidopsis thaliana

NHLVG

SEQ ID N° : 148, motivo 5 – Arabidopsis thaliana

DKSCVWHARDFSDGELK

SEQ ID N° : 149, motivo 6 – Arabidopsis thaliana

EVAES

SEQ ID N° : 150, motivo 7 – Arabidopsis thaliana

LEKLTVEEK

SEQ ID N° : 151, iniciador prm06703

GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGGCGGACAAGGAGC

SEQ ID N° : 152, iniciador prm06704

GGGGACCACTTTGTACAAGAAAGCTGGGTAGTGCAACCACACCAACTACT

SEQ ID N° : 153, iniciador prm06491

GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACAATGGCGAGCATTAGCAAC

SEQ ID N° : 154, iniciador prm06492

GGGGACCACTTTGTACAAGAAAGCTGGGTGCATCTTAAGCTGAGGGAAC

SEQ ID N° : 155 Oryza sativa – promotor de prolamina

CTTCTACATCGGCTTAGGTGTAGCAACACGACTTTATTATTATTATTATTATTATTATTATT
TTACAAAAATATAAAATAGATCAGTCCCTCACCACAAGTAGAGCAAGTTGGTGAGTTATTGTA
AAGTTCTACAAAGCTAATTTAAAAGTTATTGCATTAACTTATTTCATATTACAAACAAGAGTG
TCAATGGAACAATGAAAACCATATGACATACTATAATTTTGTTTTTATTATTGAAATTATATAA
TTCAAAGAGAATAAATCCACATAGCCGTAAAGTTCTACATGTGGTGCATTACCAAAATATATA
TAGCTTACAAAACATGACAAGCTTAGTTTGAAAAATTGCAATCCTTATCACATTGACACATAA
AGTGAGTGATGAGTCATAATATTATTTTCTTTGCTACCCATCATGTATATATGATAGCCACAAA
GTTACTTTGATGATGATATCAAAGAACATTTTTAGGTGCACCTAACAGAATATCCAAATAATA
TGACTCACTTAGATCATAATAGAGCATCAAGTAAAACTAACACTCTAAAGCAACCGATGGGA
AAGCATCTATAAATAGACAAGCACAATGAAAATCCTCATCATCCTTCACCACAATTCAAATAT
TATAGTTGAAGCATAGTAGTA

FIGURA 27 (continuação)

SEQ ID N° : 156, OsGLK, seqüência codificadora
ATGCTTGCCGTGTCGCCGGCGATGTGCCCCGACATTGAGGACCGCGCCGCGGTGGCCGGCGAT
GCTGGCATGGAGGTCGTCGGGATGTCGTCGGACGACATGGATCAGTTCGACTTCTCCGTCGAT
GACATAGACTTCGGGGACTTCTTCCTGAGGCTGGAGGACGGTGATGTGCTCCCGGACCTCGAG
GTCGACCCGGCCGAGATCTTCACCGACTTCGAGGCAATTGCGACGAGTGGAGGCGAAGGTGT
GCAGGACCAGGAGGTGCCCACCGTCGAGCTCTTGGCGCCTGCGGACGACGTCGGTGTGCTGG
ATCCGTGCGGCGATGTCGTCGTCGGGGAGGAGAACGCGGCGTTTGCCGGGGCTGGAGAGGAG
AAGGGAGGGTGTAACCAGGACGATGATGCGGGGGAAGCGAATGTCGACGATGGAGCCGCGG
CGGTTGAGGCCAAGTCTTCGTCGCCGTCATCGACGACGTCGTCGTCGCAGGAGGCTGAGAGCC
GGCACAAGTCATCCAGCAAGAGCTCCCATGGGAAGAAGAAAGCGAAGGTGGACTGGACGCCT
GAGCTTCACCGGAGGTTCGTGCAGGCGGTGGAGCAGCTCGGCATCGACAAGGCCGTGCCGTC
GAGGATACTTGAGATCATGGGGATCGACTCTCTCACCCGGCACAACATAGCCAGCCATCTTCA
GAAGTACCGGTCACACAGAAAACACATGATTGCGAGAGAGGCGGAGGCAGCGAGTTGGACC
CAACGGCGGCAGATTTACGCCGCCGGTGGAGGTGCTGTTTGCGAAGAGGCCGGAGTCCAACG
CGTGGACCGTGCCAACCATTGGCTTCCCTCCTCCTCCGCCACCACCACCATCACCGGCTCCGAT
TCAACATTTTGCTCGCCCGTTGCATGTTGGGGCCACCCGACGATGGACCCGTCCCGAGTTCCA
GTGTGGCCACCGCGGCACCTCGTTCCCCGTGGCCCGGCGCCACCATGGGTTCCACCGCCGCCG
CCGTCGGACCCTGCTTTCTGGCACCACCCTTACATGAGGGGGCCAGCACATGTGCCAACTCAA
GGGACACCTTGCATGGCGATGCCCATGCCAGCTGCGAGATTTCCTGCTCCACCGGTGCCAGGA
GTTGTCCCGTGTCCAATGTATAGGCCATTGACTCCACCAGCACTGGCGAGCAAGAATCAGCAG
GACGCACAGCTTCAACTCCAGGTTCAACCATCAAGCGAGAGCATCGACGCAGCTATCGGTGA
TGTTTTATCGAAACCGTGGTTGCCTTTGCCTCTTGGACTGAAGCCACCTTCAGTGGACAGTGTG
ATGGGCGAGCTGCAGAGGCAAGGCGTAGCAAACGTTCCTCCAGCGTGTGGATGATATTACCC
AGCTTTCTTGTACAAAGT

SEQ ID N° : 157, OsGLK, seqüência de proteína deduzida
MLAVSPAMCPDIEDRAAVAGDAGMEVVGMSSDDMDQFDFSVDDIDFGDFFLRLEDGDVLPDLEV
DPAEIFTDFEAIATSGGEGVQDQEVPTVELLAPADDVGVLDPCGDVVVGKENAAFAGAGEEKGGC
NQDDDAGEANADDGAAAVEAKSSSPSSSTSSSQEAESRHKSSSKSSHGKKKAKVDWTPELHRRFV
QAVEQLGIDKAVPSRILEIMGIDSLTRHNIASHLQKYRSHRKHMIAREAEAASWTQRRQIYAAGGG
AVAKRPESNAWTVPTIGFPPPPPPPPSPAPMQHFARPLHVWGHPTMDPSRVPVWPPRHLVPRGPAP
PWVPPPPSDPAFWHHPYMRGPAHVPTQGTPCMAMPMPAARFPAPPVPGVVPCPMYRPLTPPALT
SKNQQDAQLQLQVQPSSESIDAAIGDVLSKPWLPLPLGLKPPSVDSVMGELQRQGVANVPPACG

SEQ ID N° : 158, iniciador avançado prm02251
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACAATGCTTGCCGTGTCGC

SEQ ID N° : 159, iniciador reverso prm02252
GGGGACCACTTTGTACAAGAAAGCTGGGTAATATCATCCACACGCTGGA

SEQ ID N° : 160, seqüência de promotor GOS2 de arroz
AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAA
TATAAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCA
CCTACTTTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGT
AATTAAGTGGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGT
TAAATTATTCGAGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTTCTAGCT
GAACTCAATGGGTAAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTCTGAACGT
ATTGGCAAAGATTTAAACATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACA
TCATTAAGGACATGTCTTACTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATT
TTTATTATTTATC

FIGURA 27 (continuação)

TTTTTTCGATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCA
CCTCCTCAATACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTA
GGTAGCAATATCTGAATTCAAGCACTCCACCATCACCAGACCACTTTTAATAATATCTAAAAT
ACAAAAAATAATTTTACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTCACAT
ACAAAAAAAAAAAAGAATTTTGCTCGTGCGCGAGCGCCAATCTCCCATATTGGGCACACAGGC
AACAACAGAGTGGCTGCCCACAGAACAACCCACAAAAAACGATGATCTAACGGAGGACAGC
AAGTCCGCAACAACCTTTTAACAGCAGGCTTTGCGGCCAGGAGAGAGGAGGAGAGGCAAAGA
AAACCAAGCATCCTCCTCCTCCCATCTATAAATTCCTCCCCCCTTTTCCCCTCTCTATATAGGA
GGCATCCAAGCCAAGAAGAGGGAGAGCACCAAGGACACGCGACTAGCAGAAGCCGAGCGAC
CGCCTTCTTCGATCCATATCTTCCGGTCGAGTTCTTGGTCGATCTCTTCCCTCCTCCACCTCCTC
CTCACAGGGTATGTGCCCTTCGGTTGTTCTTGGATTTATTGTTCTAGGTTGTGTAGTACGGGCG
TTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCCTGTTCTTGGATTTGGGATAGAGGG
GTTCTTGATGTTGCATGTTATCGGTTCGGTTTGATTAGTAGTATGGTTTTCAATCGTCTGGAGA
GCTCTATGGAAATGAAATGGTTTAGGGTACGGAATCTTGCGATTTTGTGAGTACCTTTTGTTTG
AGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAAGTACGGTTGTTTGGTCCTCG
ATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTCCCTATTGAACAAAA
ATAATCCAACTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTTAAGCCTGTCCA
AAATTTCGCAGCTGGCTTGTTTAGATACAGTAGTCCCCATCACGAAATTCATGGAAACAGTTA
TAATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTTTTTTCC
CAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATGCTTT
TATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGAA
GAACTTATCCGATTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTAT
TCATTTGGATTATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATG
AACTGTCCTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACC
TGTAGAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGT
AATCGGGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCTT
GCCACTTTCACCAGCAAAGTTC

SEQ ID N° : 161, GARP seqüência de consenso 1
(K/R)(P/M/V/A)(R/K/M)(V/L)(V/D)W(S/T/I/N)(V/A/P/S/T/C/H/Q/D)(E/Q/T/D/S)L(H/D)(R/K/Q/A/H/D/E/L/I)(K/R/Q/S/C/V/A/H)F(V/L/I)(A/K/Q/E/H/D/N/R/S)(A/V/C)(V/G/L/I)(N/E/A/Q/D/G/T/I/K/H)(Q/E/H/L/I/M/K/R/S)L

SEQ ID N° : 162, GARP seqüência de consenso 2
G(I/V/L/P/S/H/Q/A/G)(D/E/K/H/Q/N/A)

SEQ ID N° : 163, GARP seqüência de consenso 3
(A/T)(V/I/Y/F/T)P(K/S)(K/R/T/Q/L/S/G/A)(I/V/L)(L/M/R/K)(D/E/Q/K/R/S)(L/I/F/H/V/M/T/R/A)(M/I/L)(N/K/G/S/D/Q/E)

SEQ ID N° : 164, GARP seqüência de consenso 4
(V/I/M/T/E/L/S/N)(E/D/G/N/Y/K/H/Q/P)(N/G/K/T/S/C/R/D)(I/L)(T/D/A/S)(R/N/I/L/V)(E/H/D/S/A/Y/F)(N/E/H)

SEQ ID N° : 165, GARP seqüência de consenso 5
(V/I/L)(A/K)SHLQ(K/M/I)(Y/F)(R/V)

FIGURA 27 (continuação)

SEQ ID N° : 166, seqüência de consenso 6
SHR(K/R)H(L/M)(L/A/M/I)ARE(A/G/V)EA(A/G)(S/N/T)W

SEQ ID N° : 167, seqüência de consenso de domínio GCT
(H/Q)(P/L)S(N/K/S)E(S/V)(I/V/L)DAAIG(D/E)(V/A)(I/L)(S/T/A/V)(N/K/R)PW(L/T)P(L/P)PLGL(K/N)PP(S/A)(V/M/L)(D/E/G)(G/S)V(M/I)(T/A/S/G)EL(Q/H/E)(R/K)(Q/H)G(V/I)(S/N/P/A)(N/E/T/K)(V/I)P(P/Q)

SEQ ID N° : 168, NM_188697.1| Oryza sativa (grupo de cultivar japônico), seqüência codificadora
ATGCTTGAGGTGTCCACGCTGCGAAGCCCTAAGGCGGATCAGCGGGCGGGCGTCGGCGGCCA
CCATGTCGTCGGCTTCGTCCCGGCGCCGCCGTCGCCGGCCGACGTCGCCGACGAGGTCGACGC
GTTCATCGTCGACGACAGCTGCCTGCTCGAGTACATCGACTTCAGCTGCTGCGACGTGCCGTT
CTTCCACGCCGACGACGGCGACATCCTCCCGGACCTCGAGGTCGACCCCACGGAGCTCCTCGC
CGAGTTCGCCAGCTCCCCGGACGACGAGCCGCCGCCGACGACGTCGGCTCCGGGCCCCGGCG
AGCCAGCTGCTGCTGCAGGAGCCAAGGAAGACGTGAAGGAAGATGGAGCCGCCGCCGCCGCC
GCCGCCGCCGCCGCTGACTACGACGGGTCGCCGCCGCCACCGCGGGGGAAGAAGAAGAAGG
ACGACGAGGAAAGGTCGTCGTCGTTGCCGGAGGAGAAAGACGCGAAGAACGGCGGCGGCGA
CGAGGTCCTGAGCGCGGTGACGACGGAGGATTCCTCGGCCGGTGCCGCCAAGTCGTGCTCGC
CGTCGGCAGAGGGCCACAGCAAGAGGAAGCCGTCGTCGTCGTCGTCATCGGCGGCGGCCGGC
AAGAACTCTCACGGCAAGCGCAAGGTGAAGGTGGACTGGACGCCGGAGTTGCACCGGCGGTT
CGTGCAGGCGGTGGAGCAGCTCGGGATAGACAAGGCCGTGCCGTCCAGGATCCTGGAGCTCA
TGGGCATCGAGTGCCTCACTCGCCACAACATCGCCAGCCATCTCCAGAAATATCGGTCGCACA
GGAAACATCTGATGGCGAGGGAGGCGGAGGCGGCGAGCTGGACGCAGAAGCGGCAGATGTA
CACCGCCGCCGCCGCCGCCGCCGCGGTGGCAGCCGGCGGCGGGCCAAGGAAGGACGCCGCCG
CCGCCACTGCGGCGGTGGCCCCGTGGGTCATGCCGACCATCGGTTTCCCTCCGCCGCACGCGG
CGGCGATGGTGCCTCCCCCGCCGCACCCTCCACCGTTCTGCCGGCCGCCGCTGCACGTGTGGG
GCCACCCGACCGCCGGCGTCGAGCCGACCACCGCGGCGGCGCCACCACCACCCTCGCCGCAC
GCGCAGCCGCCGTTGCTGCCCGTCTGGCCGCGCCACCTGGCGCCGCCGCCGCCGCCGCTGCCG
GCGGCGTGGGCGCACGGCCACCAGCCGGCGCCGGTGGACCCGGCGGCGTACTGGCAGCAACA
GTATAACTTGCAGAGGTTTCCTGTACCGCCGGTGCCTGGAATGGTGCCGCACCCCATGTACAG
ACCGATACCGCCGCCGTCACCGCCGCAGGGGAATAAACTCGCTGCCTTGCAGCTTCAGCTTGA
TGCCCACCCGTCTAAGGAGAGCATAGACGCAGCCATCGGAGATGTTTTAGTGAAGCCATGGCT
GCCGCTTCCCCTCGGCCTCAAGCCACCGTCGCTGGACAGCGTCATGTCTGAGCTGCACAAGCA
GGGCATCCCCAAGGTGCCACCGGCGGCGAGCGGTGCCGCCGGCTGA

SEQ ID N° : 169, NP_913586.1| OSJNBa0086P08.18 [Oryza sativa (grupo de cultivar japônico)]
MLEVSTLRSPKADQRAGVGGHHVVGFVPAPPSPADVADEVDAFIVDDSCLLEYIDFSCCDVPFFH
ADDGDILPDLEVDPTELLAEFASSPDDEPPPTTSAPGPGEPAAAAGAKEDVKEDGAAAAAAAAAA
DYDGSPPPPRGKKKKDDEERSSSLPEEKDAKNGGGDEVLSAVTTEDSSAGAAKSCSPSAEGHSKR
KPSSSSSSAAAGKNSHGKRKVKVDWTPELHRRFVQAVEQLGIDKAVPSRILELMGIECLTRHNIAS
HLQKYRSHRKHLMAREAEAASWTQKRQMYTAAAAAAAVAAGGGPRKDAAAATAAVAPWVMP
TIGFPPPHAAAMVPPPPHPPPFCRPPLHVWGHPTAGVEPTTAAAPPPPSPHAQPPLLPVWPRHLAPP
PPPLPAAWAHGHQPAPVDPAAYWQQQYNLQRFPVPPVPGMVPHPMYRPIPPPSPPQGNKLAALQL
QLDAHPSKESIDAAIGDVLVKPWLPLPLGLKPPSLDSVMSELHKQGIPKVPPAASGAAG

FIGURA 27 (continuação)

SEQ ID N° : 170, AY026772.1| Arabidopsis thaliana fator de transcrição semelhante a golden2 (GLK1) mRNA, cds completos

ATGTTAGCTCTGTCTCCGGCGACAAGAGATGGTTGCGACGGAGCGTCAGAGTTTCTTGATACG
TCGTGTGGATTCACGATTATAAACCCGGAGGAGGAGGAGGAGTTTCCGGATTTCGCTGACCAC
GGTGATCTTCTTGACATCATTGACTTCGACGATATATTCGGTGTGGCCGGAGATGTGCTTCCTG
ACTTGGAGATCGACCCTGAGATCTTATCCGGGGATTTCTCCAATCACATGAACGCTTCTTCAA
CGATTACTACGACGTCGGATAAGACTGATAGTCAAGGGGAGACTACTAAGGGTAGTTCGGGG
AAAGGTGAAGAAGTCGTAAGCAAAAGAGACGATGTTGCGGCGGAGACGGTGACTTATGACG
GTGACAGTGACCGGAAAAGGAAGTATTCCTCTTCAGCTTCTTCCAAGAACAATCGGATCAGTA
ACAACGAAGGGAAGAGAAAGGTGAAGGTGGATTGGACACCAGAGCTACACAGGAGATTCGT
GGAGGCAGTGGAACAGTTAGGAGTGGACAAAGCTGTTCCTTCTCGAATTCTGGAGCTTATGGG
AGTCCATTGTCTCACTCGTCACAACGTTGCTAGTCACCTCCAAAAATATAGGTCTCATCGGAA
ACATTTGCTAGCTCGTGAGGCCGAAGCGGCTAATTGGACACGCAAAAGGCATATCTATGGAG
TAGACACCGGTGCTAATCTTAATGGTCGGACTAAAAATGGATGGCTTGCACCGGCACCCACTC
TCGGGTTTCCACCACCACCACCCGTGGCTGTTGCACCGCCACCTGTCCACCACCATCATTTTAG
GCCCCTGCATGTGTGGGGACATCCCACGGTTGATCAGTCCATTATGCCGCATGTGTGGCCCAA
ACACTTACCTCCGCCTTCTACCGCCATGCCTAATCCGCCGTTTTGGGTCTCCGATTCTCCCTATT
GGCATCCAATGCATAACGGGACGACTCCGTATTTACCGACCGTAGCTACGAGATTTAGAGCAC
CGCCAGTTGCCGGAATCCCGCATGCTCTGCCGCCGCATCACACGATGTACAAACCAAATCTTG
GATTTGGTGGTGCTCGTCCTCCGGTAGACTTACATCCGTCAAAAGAGAGCGTGGATGCAGCCA
TAGGAGATGTATTGACGAGGCCATGGCTGCCACTTCCGTTGGGATTAAATCCGCCGGCTGTTG
ACGGTGTTATGACAGAGCTTCACCGTCACGGTGTCTCTGAGGTTCCTCCGACCGCGTCTTGTGC
CTGA

SEQ ID N° : 171, AAK16743.1| fator de transcrição semelhante a golden2 [Arabidopsis thaliana]

MLALSPATRDGCDGASEFLDTSCGFTIINPEEEEEFPDFADHGDLLDIIDFDDIFGVAGDVLPDLEID
PEILSGDFSNHMNASSTITTTSDKTDSQGETTKGSSGKGEEVVSKRDDVAAETVTYDGDSDRKRK
YSSSASSKNNRISNNEGKRKVKVDWTPELHRRFVEAVEQLGVDKAVPSRILELMGVHCLTRHNVA
SHLQKYRSHRKHLLAREAEAANWTRKRHIYGVDTGANLNGRTKNGWLAPAPTLGFPPPPPVAVA
PPPVHHHHFRPLHVWGHPTVDQSIMPHVWPKHLPPPSTAMPNPPFWVSDSPYWHPMHNGTTPYL
PTVATRFRAPPVAGIPHALPPHHTMYKPNLGFGGARPPVDLHPSKESVDAAIGDVLTRPWLPLPLG
LNPPAVDGVMTELHRHGVSEVPPTASCA

SEQ ID N° : 172, NM_123786.2| Arabidopsis thaliana GLK2 (SEMELHANTE A GOLDEN2 2); fator de ligação / transcrição de DNA (GLK2) mRNA, cds completos

GTATCATTCTTGTTTTCTCTAAATTTTTCAAAAAACCTTTAGAATTTTCATTTTTTTACGATTCC
GATGTTAACTGTTTCTCCGGCTCCAGTACTCATCGGAAACAACTCAAAGGATACTTACATGGC
GGCAGATTTCGCAGATTTTACGACGGAAGACTTGCCGGACTTTACGACGGTCGGGGATTTTTC
CGATGATCTTCTTGATGGAATCGATTACTACGACGATCTTTTCATTGGTTTCGATGGAGACGAT
GTTTTGCCGGATTTGGAGATAGATTCGGAGATTCTTGGGGAATATTCCGGTAGCGGAAGAGAT
GAGGAACAAGAAATGGAGGGTAACACTTCGACGGCATCGGAGACATCGGAGAGAGACGTTG
GTGTGTGTAAGCAAGAGGGTGGTGGTGGTGGTGACGGTGGTTTTAGGGACAAAACGGTGCGT
CGAGGCAAACGTAAAGGGAAGAAAAGTAAAGATTGTTTATCCGATGAGAACGATATTAAGAA
AAAACCTAAGGTGGATTGGACGCCGGAGTTACACCGGAAATTTGTACAAGCGGTGGAGCAAT
TAGGGGTAGACAAGGCGGTGCCGTCTCGAATCTTGGAAATTATGAACGTTAAATCTCTCACTC
GTCACAACGTTGCTAGCCATCTTCAGAAATATAGGTCACATCGGAAACATCTACTAGCGCGTG
AAGCAGAAGCTGCCAGCTGGAATCTCCGGAGACATGCCACGGTGGCAGTGCCCGGAGTAGGA
GGAGGAGGGAAGAAGCCGTGGACAGCT

FIGURA 27 (continuação)

CCTGCCTTAGGCTATCCTCCACACGTGGCACCAATGCATCATGGTCACTTCAGGCCTTTGCACG
TATGGGGTCATCCTACGTGGCCAAAACACAAGCCTAATACTCCGGCGTCTGCTCATCGGACGT
ATCCAATGCCGGCCATTGCGGCGGCTCCGGCATCTTGGCCAGGTCATCCACCGTACTGGCATC
AGCAACCACTCTATCCACAGGGATATGGTATGGCATCATCGAATCATTCAAGCATCGGTGTTC
CCACAAGACAATTAGGACCCACTAATCCTCCCATCGACATTCATCCCTCGAATGAGAGCATAG
ATGCAGCTATTGGGGACGTTATATCAAAGCCGTGGCTGCCGCTTCCTTTGGGACTGAAACCGC
CGTCGGTTGACGGTGTTATGACGGAGTTACAACGTCAAGGAGTTTCTAATGTTCCTCCTCTTCC
TTGAGAAAGATCTCCAAAATTTGTCGAAATCTCAAACTTTTAACTTCATTTTTTTGGTATCTTC
TATGTATTTTTGCAAGGGAAAGACGATAAATCCTTGTTGCTTGATCATATGTATTTCTCTATAT
GAGTGCATGTATCGAAGTTAAGCAACTTTAATATATCGTTATAATCTTCGAT

SEQ ID N° : 173, NP_199232.1| GLK2 (SEMELHANTE A GOLDEN2 2); fator de ligação / transcrição de DNA [Arabidopsis thaliana]

MLTVSPAPVLIGNNSKDTYMAADFADFTTEDLPDFTTVGDFSDDLLDGIDYYDDLFIGFDGDDVLP
DLEIDSEILGEYSGSGRDEEQEMEGNTSTASETSERDVGVCKQEGGGGGDGGFRDKTVRRGKRKG
KKSKDCLSDENDIKKKPKVDWTPELHRKFVQAVEQLGVDKAVPSRILEIMNVKSLTRHNVASHLQ
KYRSHRKHLLAREAEAASWNLRRHATVAVPGVGGGGKKPWTAPALGYPPHVAPMHHGHFRPLH
VWGHPTWPKHKPNTPASAHRTYPMPAIAAAPASWPGHPPYWHQQPLYPQGYGMASSNHSSIGVP
TRQLGPTNPPIDIHPSNESIDAAIGDVISKPWLPLPLGLKPPSVDGVMTELQRQGVSNVPPLP

SEQ ID N° : 174, AY741684.1| Physcomitrella patens proteína 1 semelhante a golden 2 (Glk1) gene, cds completos

TTCTAGGAGGTCAGCTTGGTCCAAGTCCGAATCGCATCCACGCGATAGGATATGTCGATCCAA
ATCACACATTTTTACACTCCCCAGAAGGAGGAAAAGAAGTTTTCGAAACCAGTAGTGAGGAA
TCATAGTTTCGTTTTCCGAAACCTACAAATTCATCATACTCATTTCGGTAACAAATGCTACTTA
AAAGACATGTCAGGAAAACATATAAAGTGACAGTGATTCTGCATCAACGTGACAGGGATCCA
TCAGTGGAGAATGAATAAAAAAAACAATTAGAGAATGGTGAGTAGAATCCTCAGCAAAATTA
CGTTTTATTTATTAATATTATACTTAAAAGTAAAACAGGATGATTCATGATGTCATCACCGAG
GTTGTGGAGAGATTCAAAAAGAGGATGGAGGCTACTTTTATGGGGAGGGTGGGTTCACTTTAT
GCAGTAAAGTAGTAGATCTCTATTTAAAGGGGGGAAAGAGGATGTCTGCTGAAGGCCAGTGC
AGTGGATGAACGCAGGCGCTCGTCTATGTTTCATCCATCCATCCATTGTGTCGATTTAGAGGC
CACCATGGCTTCTGAGGGTGAGGGGAGAGGATAGATAGCATAGAGGGGGGGGGGTGGAGCA
GGCAAGGGCAGGAGTGAGATTGTGGTGGTGGTGTGATTAGGCGATGGAGATATGGCGATGGA
CATGGCGAGGATAGAAGTCGAGCCTCTTGTCGAAGTCCACAACAACACCAACAACTTGCTTGT
GCATTGCGTGCTCGATGCTTTGCCGGATTCTTCACCTTGCTTGAAATCCAGTGAGACTTCGTTT
GAAGCTGTGGTTGTAAAGGAGGACGAGGAGGGAGGGAGGCCGCTGTTTGGCAAGCCTGAGCT
CCACCCTGCCTCACCCTCGAGTGATACAGCGGCTGCTGCAAATGGTGAATTCGCTAGATGCTT
GAATTTTGTGACGGAGGCTGATTGTGGAGATGTAGGCGTACAATGTTTTGAGGATTTTGGTAC
GTTGCCGGACTGTGGTGAGGCGGGGATAAGCAGGGAAGAGGGTGGAGGTAGAGACGGGGAG
CAGGTGGAGTTGTTAGAGTCTATGAGCCTCGATGGTAGTAGGAATTCGGAGAATTTAGAGCTA
GGGGAGCTCGGCGAATTGTTGCAGGGGTCGGAGACGCTGGATTCGGTTCCTGGGAACGAGGT
TGGGGAGGAGGAGGCGCTGCTGTTGGCGAAAGCGGCAAAGGCGACGGGCGTTGTTGTTTCGG
CCTCCGATAGTGGTGAATGTAGCAGTGTCGATAGGAAGGAAAATCAACAAAGTCCGAAATCA
TGTAAAAGTGCCGCACCGGGGAAGAAGAAGGCGAAGGTCGATTGGACGCCGGAGCTTCATCG
GCGCTTTGTCCACGCGGTGGAACAGCTTGGAGTGGAGAAAGCTTTTCCCTCGCGCATACTAGA
ACTGATGGGAGTACAATGTCTCACCCGGCACAATATCGCCAGTCATTTGCAGAAGTATCGCTC
GCATCGTCGCCATCTTGCAGCCAGGGAAGCCGAGGCAGCATCCTGGACTCATCGTCGAGCGTA
CACCCAGATGCCCTGGTCTCGAAGTTCACGGCGCGATGGCCTTCCTTATCTTGTACCCTTACAC
ACCCCTCACATACAACCTCGCCCATCCATGGTCATGGCAATGCAACCACAGCTTCAGACGCAG

FIGURA 27 (continuação)

CACACCCCGGTGTCGACGCCTCTTAAGGTGTGGGGGTATCCTACAGTAGATCATTCAAGTGTA
CACATGTGGCAGCAACCTGCAGTGGCGACCCCATCGTATTGGCAAGCCCCCGATGGCTCTTAC
TGGCAGCATCCTGCCACCAATTATGATGCGTATTCAGCTCGCGCTTGTTATCCCCATCCCATGC
GAGTTTCGTTAGGCACTACGCATGCTGGCTCTCCAATGATGGCTCCAGGATTTCCTGACGAGA
GCTACTACGGTGAAGATGTTCTTGCAGCTACCATGTATCTTTGTAACCAATCATATGACAGTG
AATTAGGACGAGCTGCGGGCGTCGCTGCGTGCAGTAAACCACCGGAGACGCATTTATCGAAG
GAGGTTCTTGATGCAGCCATCGGAGAAGCGCTTGCTAACCCTTGGACTCCCCCGCCCCTGGGA
CTGAAGCCGCCGTCTATGGAGGGAGTAATCGCAGAGCTTCAGCGGCAGGGAATCAACACTGT
GCCTCCCTCTACCTGTTAGCTTTAGTTGTTTGCTGTAGAGAATATTTCATTCTACGATTCGCATT
CCAATGACCGCTGACCGCTGGTGTCGCTTGGCTGGTGTTCATCTCGAGGAATCAATTTGGTGA
AATTTTGGTTATGTCACGGTAGGGGGGTCTATTTTCTCCAGAGTTCCTCCGGAGAGGTGTATG
AAAGGATCGATTTTCTTTCTTCTTTTTTTTTTTTCTTTTTTTTTCTTTTTTTTTTTCTTTTTTTTTTCCT
TCGAATAAGGCTGCCTTGGTGGTGTTTCTTCTTAGTTTTTTAACGAGGGATACCTTATCATATC
TGTCAAGATATGTCACTGAACGTTGCTGCATGTAGACTTACGTTTATGGTAACACTTTCTCAAT
GCCATTAAAAACCGAAACCAGTTCTTCTGATTGTTTCATTGGAAGTATCCTGACAACCTGTCA
GCATCTGTTGGCATGTGGTTTCCTCACCCACACCCTTCTTTCAGTTGGTTTTGAATCTGCATCTA
CTGACTGTTAAGGTTTTACGTGTATCAAGTGTACGATTTTTCCACAAGATAATTTCTCAACAAC
TCTGCTTTCGAAGCACCTTTCTTCTCCACCAATCAAGAGTGGTGGGAAGATGGACCAAGGT

SEQ ID N° : 175, AAV54520.1| proteína 1 semelhante a golden 2 [Physcomitrella patens]
MAMDMARIEVEPLVEVHNNTNNLLVHCVLDALPDSSPCLKSSETSFEAVVVKEDEEGGRPLFGKP
ELHPASPSSDTAAAANGEFARCLNFVTEADCGDVGVQCFEDFGTLPDCGEAGISREEGGGRDGEQ
VELLESMSLDGSRNSENLELGELGELLQGSETLDSVPGNEVGEEEALLLAKAAKATGVVVSASDS
GECSSVDRKENQQSPKSCKSAAPGKKKAKVDWTPELHRRFVHAVEQLGVEKAFPSRILELMGVQ
CLTRHNIASHLQKYRSHRRHLAAREAEAASWTHRRAYTQMPWSRSSRRDGLPYLVPLHTPHIQPR
PSMVMAMQPQLQTQHTPVSTPLKVWGYPTVDHSSVHMWQQPAVATPSYWQAPDGSYWQHPAT
NYDAYSARACYPHPMRVSLGTTHAGSPMMAPGFPDESYYGEDVLAATMYLCNQSYDSELGRAA
GVAACSKPPETHLSKEVLDAAIGEALANPWTPPPLGLKPPSMEGVIAELQRQGINTVPPSTC

SEQ ID N° : 176, AY741685.1| Physcomitrella patens proteína 2 semelhante a golden2 (Glk2) gene, cds completos
GGTGGTGGTGGTAGAAGGCGATGGAGATATGGCGATGGACATGGCGAGGATAGATGAATCAA
CCGCCGTCGAGGTCAACTCGCTTTCGCTTGTGCATTGCGTGTTGGATGGGTTGCCCGATTCGCC
TTGCTTGAAATCCAGCCCGACGTCATTCGAGGAGGCTGTGGCGGAAGGGAGGTCGGTGTTCG
GGGACGAGGAGGACATCATCAACAACAGCAACGACCAGGACAACTCGTCCTCGTGCGGTGCA
GTGGTCACCACCCACGAAGATTTCGCCGAGTGCTTGAATTTTGTGACGGAGGCGGAGTGTGGA
GATGTGGGTGTGCGGTGTTTCGAGGATTTTGACAAGCTGCCGGACTGCGGCGACGAGGGGGA
GACTAGCAAAGCGGAGGAGGAGGGGTGTGTACGAATAGGCGGAGGAGGGGAGCAGGGCGAG
CTGTTAGAGTCTGTGAGCCTTGACTGTAGTAGGAATTCGGAGAATTTAGAGCTTCGGGATCTC
GGGGAATTGTGGGAAGGGTCGGAACGGCCTGACTCAGTGCCAGGGAACGAGGTGGGTGAGG
AGGAGGCGTTGCTGTTGGCGGAGGCGGCCAAGGCGACGGGCGATGTCGTGTCGGCCTCGGAT
AGTGGAGAATGTAGCAGTGTCGATAGGAAGGACAATCAAGCCAGTCCGAAATCCAGTAAAAA
TGCAGCGCCGGGGAAGAAGAAGGCCAAGGTGGACTGGACGCCTGAGCTTCACCGGCGCTTCG
TCCATGCAGTGGAGCAGCTGGGTGTGGAGAAAGCCTATCCATCGCGTATCCTAGAACTGATGG
GCGTGCAATGCTTGACTCGGCACAACATCGCCAGTCACTTGCAGAAGTACCGGTCCCACCGTC
GCCACCTCGCAGCTCGAGAAGCAGAGGCCGCGTCCTGGACGCATCGTCGCACATACACTCAG
GCTCCCTGGCCTCGTAGCTCACGGCGCGATGGCCTCCCTTATCTTGTACCTATACACACCCCTC
ACATACAACCTCGGCCTTCCATGGCCATGGCAATGCAACCGC

FIGURA 27 (continuação)

AGCTTCAGACGCCGCATCATCCGATATCCACTCCTCTCAAGGTCTGGGGCTACCCCACAGTAG
ATCATTCAAATGTACATATGTGGCAGCAACCTGCTGTGGCGACCCCATCTTACTGGCAAGCTG
CCGATGGCTCATACTGGCAACATCCCGCGACCGGTTACGACGCTTTCTCAGCTCGTGCCTGCT
ACTCGCATCCCATGCAGCGAGTTCCTGTAACGACCACGCATGCGGGTTTACCAATTGTGGCGC
CAGGATTTCCTGACGAGAGCTGCTACTACGGCGACGACATGCTTGCAGGCTCCATGTATCTAT
GTAACCAATCATATGATAGTGAAATAGGACGAGCTGCGGGTGTTGCTGCGTGCAGCAAGCCG
ATAGAGACGCATTTGTCCAAAGAGGTGTTGGATGCGGCCATTGGCGAAGCTCTCGCCAATCCC
TGGACTCCCCCACCTCTGGGTCTGAAGCCACCATCCATGGAGGGCGTCATTGCAGAGCTTCAG
CGGCAGGGGATCAACACTGTGCCTCCTTCAACTTGTTAGCTCTCGACGATGTTTTTTTCTTCCT
CTTCCTCTTCCTCTTCTTCTAGAGAGCAATTTTTCTTTCTACGACTCATATTCCAATGACCGCTG
ACCGCTGGTGTCGCTGGTGTCGCTGGTGTTCATCCCGAGGAACTAATTTGGTGAAAATTCGGT
TAAGTTGCGATAGAGGTCTGATCTCCGGAAGGTTTATTTTTTGTCTTCTGGAGAGGTTGTATAA
GAGGTTCCCCTGTTTTGCAATAAGGCTTCCTTGATGAGATTTTCCTTTATTTTTCCCCTGATTCT
TTCTCCTTCCATTTTAGTTTCACAAGAAGGCATCTTCTGGCCATGGCGGTCACGATGTGTCACT
TTGATGCTGCATGTGGACCTTTTTTCTTATATCTGTAGTAGCACTCCTCCATTTCATGAGGAAT
AAAGATCAAACATCACACATCATCACTCGAATTCTTCATCTCCTCCTCCTCCCTTTTCCACTAG
GATGCCTTCTATTGCATTGGTTGTCATGTGACTCAGTCTTTCTCTTACACGCACTTTAACTCGCT
GGATGTCTCACTTTCTGACTGTTCAAGGTGAGGTCTGATAGGTTCGAGACGGGATTGCTTGCG
ATGACTCACCATCTCCTAATTTGGAACCAACATTCCTCCTCAACCTATCAACTCTCACTCAAAC
TTGCATTGTGTGAGCATTGGTC

SEQ ID N° : 177, AAV54521.1| proteína 2 semelhante a golden2 [Physcomitrella patens]
MAMDMARIDESTAVEVNSLSLVHCVLDGLPDSPCLKSSPTSFEEAVAEGRSVFGDEEDIINNSNDQ
DNSSSCGAVVTTHEDFAECLNFVTEAECGDVGVRCFEDFDKLPDCGDEGETSKAEEEGCVRIGGG
GEQGELLESVSLDCSRNSENLELRDLGELWEGSERPDSVPGNEVGEEEALLLAEAAKATGDVVSA
SDSGECSSVDRKDNQASPKSSKNAAPGKKKAKVDWTPELHRRFVHAVEQLGVEKAYPSRILELM
GVQCLTRHNIASHLQKYRSHRRHLAAREAEAASWTHRRTYTQAPWPRSSRRDGLPYLVPIHTPHI
QPRPSMAMAMQPQLQTPHHPISTPLKVWGYPTVDHSNVHMWQQPAVATPSYWQAADGSYWQH
PATGYDAFSARACYSHPMQRVPVTTTHAGLPIVAPGFPDESCYYGDDMLAGSMYLCNQSYDSEIG
RAAGVAACSKPIETHLSKEVLDAAIGEALANPWTPPPLGLKPPSMEGVIAELQRQGINTVPPSTC

SEQ ID N° : 178, AF318580.1|AF318580 Zea mays fator de transcrição putativo ZmGLK1 (Glk1) mRNA, cds completos
AAACAGCGAAACAGTGCAGAGACAAGCTACAACAACCTACCCTAACAGGCCATTCCTTCCAC
TGCACTTCATTCGATCCTGAGCTATAGCTGGCTGCCTGAGCTCGCTCCAAGAAGTGTATCTATA
GTATAGACACTAGGCTTCGTGTGGCGTGCTCGAGATATGCTTGCAGTGTCGCCGTCGCCGGTG
CGGTGTGCCGATGCGGAGGAGTGCGGCGGAGGAGGCGCCAGCAAGGAAATGGAGGAGACCG
CCGTCGGGCCTGTGTCCGACTCGGACCTGGATTTCGACTTCACGGTCGACGACATAGACTTCG
GGGACTTCTTCCTCAGGCTAGACGACGGGGATGACGCGCTGCCGGGCCTCGAGGTCGACCCTG
CCGAGATCGTCTTCGCTGACTTCGAGGCAATCGCCACCGCCGGCGGCGATGGCGGCGTCACGG
ACCAGGAGGTGCCCAGTGTCCTGCCCTTTGCGGACGCGGCGCACATAGGCGCCGTGGATCCGT
GTTGTGGTGTCCTTGGCGAGGACAACGACGCAGCGTGCGCAGACGTGGAAGAAGGGAAAGGG
GAGTGCGACCATGCCGACGAGGTAGCAGCCGCCGGTAATAATAATAGCGACTCCGGTGAGGC
CGGCTGTGGAGGAGCCTTTGCCGGCGAAAAATCACCGTCGTCGACGGCATCGTCGTCGCAGG
AGGCTGAGAGCCGGCGCAAGGTGTCCAAGAAGCACTCCCAAGGGAAGAAGAAAGCAAAGGT
GGATTGGACGCCGGAGCTTCACCGGAGATTCGTTCAGGCGGTGGAGGAGCTGGGCATCGACA
AGGCGGTGCCGTCCAGGATCCTCGAGATCATGGGGATCGACTCCCTCACGCGGCATAACATA
GCCAGCCATCTGCAGAAGTACCGTTCCCACAGGAAGCACATGCTTGCGAGGGAGGTGGAGGC
AGCGACGTGGACGACGCACCGGCGGCCGATGTACGCTGCC

FIGURA 27 (continuação)

CCCAGCGGCGCCGTGAAGAGGCCCGACTCTAACGCGTGGACCGTGCCGACCATCGGTTTCCCT
CCGCCGGCGGGGACCCCTCCTCGTCCGGTGCAGCACTTCGGGAGGCCACTGCACGTCTGGGGC
CATCCGAGTCCGACGCCAGCGGTGGAGTCACCCCGGGTGCCAATGTGGCCTCGGCATCTCGCC
CCCCGCGCCCCGCCGCCGCCGCCGTGGGCTCCGCCACCGCCAGCTGACCCGGCGTCGTTCTGG
CACCATGCTTACATGAGGGGGCCTGCTGCCCATATGCCAGACCAGGTGGCGGTGACTCCATGC
GTGGCAGTGCCAATGGCAGCAGCGCGTTTCCCTGCTCCACACGTGAGGGGTTCTTTGCCATGG
CCACCTCCGATGTACAGACCTCTCGTTCCTCCAGCACTCGCAGGCAAGAGCCAGCAAGACGCG
CTGTTTCAGCTACAGATACAGCCATCAAGCGAGAGCATAGATGCAGCAATAGGTGATGTCTTA
ACGAAGCCGTGGCTGCCGCTGCCCCTCGGACTGAAGCCCCCTTCGGTAGACAGTGTCATGGGC
GAGCTGCAGAGGCAAGGCGTAGCGAATGTGCCGCAAGCTTGTGGATGATCCCCGGGGCTGCA
TAGCTGCTAGCTCGTCCCTGCACGCAGTGCAACAAAAGGAATTTGTCGACATGCCTTTTATGT
TTTGAATTCCCGTGAACCGTTTATGGAGCACTCTCAACTCGTTCAGGGTCAGATGAACAGAAC
ATTATAGGAGTACATTCTTTGAGGTTTCAGGTTTGAGGGAAATCAACCTGAGTGCATCCAAGT
ACAGTGTACTGCCACAGTGCCACGTCGGAAAATGAAGTCTTGACCCGAAT

SEQ ID N° : 179, AAK50392.1|AF318580_1 fator de transcrição putativo ZmGLK1 [Zea mays]

MLAVSPSPVRCADAEECGGGGASKEMEETAVGPVSDSDLDFDFTVDDIDFGDFFLRLDDGDDALP
GLEVDPAEIVFADFEAIATAGGDGGVTDQEVPSVLPFADAAHIGAVDPCCGVLGEDNDAACADVE
EGKGECDHADEVAAAGNNNSDSGEAGCGGAFAGEKSPSSTASSSQEAESRRKVSKKHSQGKKKA
KVDWTPELHRRFVQAVEELGIDKAVPSRILEIMGIDSLTRHNIASHLQKYRSHRKHMLAREVEAAT
WTTHRRPMYAAPSGAVKRPDSNAWTVPTIGFPPPAGTPPRPVQHFGRPLHVWGHPSPTPAVESPR
VPMWPRHLAPRAPPPPPWAPPPPADPASFWHHAYMRGPAAHMPDQVAVTPCVAVPMAAARFPA
PHVRGSLPWPPPMYRPLVPPALAGKSQQDALFQLQIQPSSESIDAAIGDVLTKPWLPLPLGLKPPSV
DSVMGELQRQGVANVPQACG

SEQ ID N° : 180, AF318579.1|AF318579 Zea mays fator de transcrição putativo GOLDEN 2 mRNA, cds completos

CTCATCTTCTCTTCTTCTTCCCCCAGTCCTCCCCCCCTTTCCCCTCTTCGGCCTCTCTCGCTCAGC
TCACTCTTCATTAAGCGAGCTCACGTCGTTCCTCCCTTCTTCTTCTTCTTATCCGTTGTTCAATT
CGTTCAGCTAGCCGGCCAGAGATCGAGCATCATCTCCATCATCGATTCATCCATCTCATCTTTC
TCTTCTTCTATTCTATGTCGTCGTCGTCCCAGATTAGATCGAAGCTGCTAGCAGTCTATCCAGT
CTCTAGCTAGCTAGCTAGATCAAGCCCGCAGACTATACAATATAATACAAGCTAGCTACGTGC
TTATTATTGCTTTATTAGTCTAGAATAATCTTGATATATACGATCGAACATATATCTCGATCTC
CACCGATACACACCCGGCCCTATCCATGCTTGAGGTGTCGACGCTGCGCGGCCCTACTAGCAG
CGGCAGCAAGGCGGAGCAGCACTGCGGCGGCGGCGGCGGCTTCGTCGGCGACCACCATGTGG
TGTTCCCGACGTCCGGCGACTGCTTCGCCATGGTGGACGACAACCTCCTGGACTACATCGACT
TCAGCTGCGACGTGCCCTTCTTCGACGCTGACGGGGACATCCTCCCCGACCTGGAGGTAGACA
CCACGGAGCTCCTCGCCGAGTTCTCGTCCACCCCTCCTGCGGACGACCTGCTGGCAGTGGCAG
TATTCGGCGCCGACGACCAGCCGGCGGCGGCAGTAGCACAAGAGAAGCCGTCGTCGTCGTTG
GAGCAAACATGTGGTGACGACAAAGGTGTAGCAGTAGCCGCCGCCAGAAGAAAGCTGCAGA
CGACGACGACGACGACGACGACGGAGGAGGAGGATTCTTCTCCTGCCGGGTCCGGGGCCAAC
AAGTCGTCGGCGTCGGCAGAGGGCCACAGCAGCAAGAAGAAGTCGGCGGGCAAGAACTCCA
ACGGCGGCAAGCGCAAGGTGAAGGTGGACTGGACGCCGGAGCTGCACCGGCGGTTCGTGCAG
GCGGTGGAGCAGCTGGGCATCGACAAGGCCGTGCCGTCCAGGATCCTGGAGATCATGGGCAC
GGACTGCCTCACAAGGCACAACATTGCCAGCCACCTCCAGAAGTACCGGTCGCACAGAAAGC
ACCTGATGGCGCGGGAGGCGGAGGCCGCCACCTGGGCGCAGAAGCGCCACATGTACGCGCCG
CCAGCTCCAAGGACGACGACGACGACGGACGCCGCCAGGCCGCCGTGGGTGGTGCCGACGAC
CATCGGGTTCCCGCCGCCGCGCTTCTGCCGCCCGCTGCACGTGT

FIGURA 27 (continuação)

GGGGCCACCCGCCGCCGCACGCCGCCGCGGCTGAAGCAGCAGCGGCGACTCCCATGCTGCCC
GTGTGGCCGCGTCACCTGGCGCCGCCCCGGCACCTGGCGCCGTGGGCGCACCCGACGCCGGT
GGACCCGGCGTTCTGGCACCAGCAGTACAGCGCTGCCAGGAAATGGGGCCCACAGGCAGCCG
CCGTGACGCAAGGGACGCCATGCGTGCCGCTGCCGAGGTTTCCGGTGCCTCACCCCATCTACA
GCAGACCGGCGATGGTACCTCCGCCGCCAAGCACCACCAAGCTAGCTCAACTGCATCTGGAG
CTCCAAGCGCACCCGTCCAAGGAGAGCATCGACGCAGCCATCGGAGATGTTTTAGTGAAGCC
ATGGCTGCCGCTTCCACTGGGGCTCAAGCCGCCGTCGCTCGACAGCGTCATGTCGGAGCTGCA
CAAGCAAGGCGTACCAAAAATCCCACCGGCGGCTGCCACCACCACCGGCGCCACCGGATGAC
ATACTATCTCGTCGACAATACATGCATGTACAATAGACGGATGTCTAGTAGTATAGTAGATTG
CTGCTAGCTAGCTAGCCGCTAGAGTGTAGTAGCATATGCATGCCTTTTTTTTCTTCTTCTTTTTT
TGCCCTTATTATTAAGCTGGCTAATAGCGATTGAGATGGAGCTTGACACAGATCTGCTAGCTA
GCTATTTGAGGGTTTCTCTTGTATGCTACCTATTGCTGCTGCTTGCTGCTGAGACAAGTAATGT
ACGTACGCCTGTGCAAACGACACATCGGATTGTATTGTATTACTAGTTATATGAACAACAACA
ATAATAATAGGCACATGCATGCCTGCCTAGTATGTGAGCTGCTAGCTAGCTACATGCATGTTG
CGCATTCCTTTCC

SEQ ID N° : 181, AAK50391.1|AF318579_1 fator de transcrição putativo GOLDEN 2 [Zea mays]

MLEVSTLRGPTSSGSKAEQHCGGGGGFVGDHHVVFPTSGDCFAMVDDNLLDYIDFSCDVPFFDAD
GDILPDLEVDTTELLAEFSSTPPADDLLAVAVFGADDQPAAAVAQEKPSSSLEQTCGDDKGVAVA
AARRKLQTTTTTTTTEEEDSSPAGSGANKSSASAEGHSSKKKSAGKNSNGGKRKVKVDWTPELHR
RFVQAVEQLGIDKAVPSRILEIMGTDCLTRHNIASHLQKYRSHRKHLMAREAEAATWAQKRHMY
APPAPRTTTTTDAARPPWVVPTTIGFPPPRFCRPLHVWGHPPPHAAAAEAAAATPMLPVWPRHLA
PPRHLAPWAHPTPVDPAFWHQQYSAARKWGPQAAAVTQGTPCVPLPRFPVPHPIYSRPAMVPPPP
STTKLAQLHLELQAHPSKESIDAAIGDVLVKPWLPLPLGLKPPSLDSVMSELHKQGVPKIPPAAATT
TGATG

SEQ ID N° : 182, GLK1 TC269481 seqüência parcial de trigo

GAATTGGGCACGAGGACCGCCCCATTGTGCCGGACGCATAATCGCCGTCGTCGACAACGTCGT
CGTCCACGGAGGCCGAGAGCCGCCACAAGTCATCAAGCAAGAACTCCCACGGAAAGAAGAA
AGCCAAGGTGGACTGGACGCCGGAGCTGCACCGGAGGTTCGTGCAGGCCGTAGAGCAGCTCG
GCATCGACAAGGCAGTGCCGTCTAGGATTTTGGAGATCATGGGGATCAACTCACTCACTCGGC
ACAACATAGCAAGCCATTTGCAGAAGTACCGGTCTCACCGGAAGCACATGATTGCGCGGGAG
GCGGAGGCGGCGAGCTGGACCCAACGACGACAGATGTACGCCGCCGGCGGGCCGGCCGCGG
CCGTGAAGAGGCAGGACTCGAACATGTGGACCGTGCCAACCATCGGCTTCGCGCCGGCGCAC
CCTCCTCCTCCTCCTCCCCCTCCTTCACCGGCAGCCATGCAGCACTACGCTCGGCCGTTGCACG
TCTGGGGTCACCCTACGATGGACTCGCCCCGGATGCCGATGTGGCCGAGGCACACAATATCCC
GCGCCCCGATGCCGGCGTGGGCTCCCCCGCCGCCGCCGCCGTCTGACCCGGCTTTCTGGCACC
ACCCTTACATGAGGGGGCCCGCAGCGTATATGCCGACCCATGGGACTCCTTGCATGGCAATGC
CGATGACACCGAAATTTCCTGCTGCACCCGTGCCCGTGGCCATGCCGTGCCCAGTCTATGCGT
CCCCCTCCCCGGCACCAGCGCTGGCGAGCAAGAGCCAACAAGATTCGCAGCTCCAGCTCCAA
GCACAACCATCAAATGAGAGCATAGACGCGGCCATTGGTGACGTTTTATCCAAACCGTGGCTG
CCGCTGCCGCTGGGGCTGAAGCCCCCTTCGTTGGGCAGCGTCATGGGCGAGCTTGAAAGGCA
AGGCGTGGCCAATGTGCCCCAAGCCTGCGGGTGAGCGTTGGAGGGTGCATCCGCGTGCACTC
CATGCATGACTGCTGCTCCGTTCGATGGAGCAGCTAGCAGCAACAACAACATCGTCGACATGT
CTTTGTTCCTTTGAACGTTTTGTGGATCTCTCTCTCTCTCTTGGTCAACTTGTGCATAATTTC
GATCAAGAGAAACATCGTTATTTTGAGTTCACTCCCGAGACACATTTTTGTTACACCTAAACTT
TAAGTTTGCGGTAACAGCAGCATCAACAAC

FIGURA 27 (continuação)

SEQ ID N° : 183, TaGLK1|TC269481 (Triticum aestivum) seqüência parcial
MGINSLTRHNIASHLQKYRSHRKHMIAREAEAASWTQRRQMYAAGGPAAAVKRQDSNMWTVPT
IGFAPAHPPPPPPPPSPAAMQHYARPLHVWGHPTMDSPRMPMWPRHTISRAPMPAWAPPPPPSDP
AFWHHPYMRGPAAYMPTHGTPCMAMPMTPKFPAAPVPVAMPCPVYASPSPAPALASKSQQDSQ
LQLQAQPSNESIDAAIGDVLSKPWLPLPLGLKPPSLGSVMGELERQGVANVPQACG

SEQ ID N° : 184, GLK1 TC2836 seqüência parcial de cebola
TGCAAATATTGTTCAACGTATATAAAGACAATTTAGTATGCTAACAATCTCTCCCCTTGAAAG
TTCTAATCCAGATGCAAAAGACGAGGGAGATTTTGTACTAGAAGACATCAGCTTTGATGACTT
GTTCGTAGGATTCGATGAACTTCCCGACCTTGAAATTGACCCTTGTGATATATTTGCTGATTTT
TCTTTCAATAGTAGCGAGGAGGGAGATGGAAATAACAAGGGTTCAACATCGGAAGAAAATGA
TGTACAGCATAAAAGAGACGGATACTATAAGGAGCACTCAGGTAAAGAAGAGACGGAGGTT
GTGAGTGCAAGAACAAGGGAGGATGAAAAGAAGCCGCCTTCATCTAAATCAGAGAATGTTAA
AAGCCTAAAGTCATCAGGCAAAAAGTCATCTCAGTCTAAAAAGAAAGCTAAGGTGGACTGGA
CTCCGGAGCTTCATCGGAGATTTGTTCAGGCAGTGGAGCAACTTGGGGTAGACAAAGCAGTA
CCATCCAGGATTTTAGAGCTCATGGGAATTGATTGTCTCACAAGACATAATATTGCAAGCCAT
TTACAGAAATATCGGTCGCACAGAAAGCATTTGCTAGCAAGAGAAGCAGAAGCAGCAAGCTG
GAGTCACAGACGTCAGTTGTACGTGAGCAGCACTAACAATGGAAAGAGGAGGAGGAAAAAC
AATGAACATGAAACCTAGTTGGGCCCCACAGTCACAGCCCATTGGTGGATTTCCTCCCAACAA
TGCATCATCCCTCAACATCAGAACTTCAGAACCTTTGCACGTCTGGGGCCCATCCACAGGCAG
TGGAGCATCCTCAGCTAGTTCCAGTCTG

SEQ ID N° : 185, AcGLK1|TC2836 (Allium cepa) seqüência parcial
MLTISPLESSNPDAKDEGDFVLEDISFDDLFVGFDELPDLEIDPCDIFADFSFNSSEEGDGNNKGSTS
EENDVQHKRDGYYKEHSGKEETEVVSARTREDEKKPPSSKSENVKSLKSSGKKSSQSKKKAKVD
WTPELHRRFVQAVEQLGVDKAVPSRILELMGIDCLTRHNIASHLQKYRSHRKHLLAREAEAASWS
HRRQLYVSSTNNGKRRRKNNEHET

SEQ ID N° : 186, TC136874 seqüência parcial semelhante a GLK2 de cevada
CATCGTGGGCGATGAGGTTTGCAGCGCGGTGACGACGGACGATTCTTCCGCTGCGGTCGGGTC
CGAGAACAGCAAGTCGTCGGCGTCGGCAGAGGGCCACAGCAAGAGGACGTCGGCAGCCGCA
GCGACCAAGAGCTCCCACGGCAGGCGGAAGGTGAAGGTGGACTGGACGCCGGAGTTGCACCG
GCGGTTCGTGCAGGCGGGGGAGCAGCTGGGGCTGGACAAGGCGGTGCCGTCGCGGATCCTGG
AGCTCATGGGCAACGAGTACCGTCTCACGCGCCACAACATCGCCAGCCATCTCCAGAAGTACC
GGTCACACAGGAAGCACCTGATGGCGAGGGAGGGCGAGGCGGGTAGCTGGACGCCGCAGCC
GCAAATGTGCTGGGGTGCGAAGGTGGTGAGGGTGGGGCGGGCTCTGACGTTGTAGGGGTTAT
ATAGCGGGGTGGCGGCGGGG

SEQ ID N° : 187, semelhante a HvGLK2, seqüência parcial (Hordeum vulgare)
SLSIVGDEVCSAVTTDDSSAAVGSENSKSSASAEGHSKRTSAAAATKSSHGRRKVKVDWTPELHR
RFVQAGEQLGLDKAVPSRILELMGNEYRLTRHNIASHLQKYRSHRKHLMAREGEAGSWTPQPQM
CWGAKVVRVGRALTL

SEQ ID N° : 188, TC96378 seqüência parcial semelhante a GLK1 de sorgo
GCACGAGGCTTGAAATTCCCTCCGCCGCCGCCGCCGCCGCCGCCTCCTCCTCCTCCTTCTCCTC
ATCCGATGCAGCACTTCGGAAGGCCGCTGCATGCCTGGGGCCATCCGACGCCAACGGTGGAG
TCACCCCGGGTGCCAATGTGGCCTCGGCATCTCGTCCCCCGCACCCCGCCGCCGCCGTGGGCT
CCTCCGCCGCCATCTGACCCGGCGTTCTGGCACCATGCTTACATGAGGGGGCCTGCACACATG
CCGGGCCA

FIGURA 27 (continuação)

GGTGACTCCCTGCGTGGCAGTGCCAATGCCAGCTGCGCGTTTCCCTGCTCCACCCGTGAGGGG
TGCTTTGCCATGTCCACCTCCAATGTACAGACCTCTCGTTCCTCCAACACTCGATACGCAGTTT
CAGCTACAGACACAGCCATCAAGTGAGAGCATAGATGCAGCAATAGGTGATGTTTTAACGAA
ACCGTGGCTGCCACTGCCCCTGGGACTGAAGCCCCCTTCAGTAGACAGTGTCATGGGCGAGCT
GCAGAGGCAAGGTGTGGCAAATGTGCCACCAGCTTGTGGATGATCCCCGGGGCTCCATAGCT
AGCTGCTTGTCCCTGCAGTCCAACAAAAAGAATTTGTCGACATGCCTTTTCTGTTTCAAATTTT
CATGAACCATTTATGGACCACTCTCTCTTAGTTAACTTGTTCAGGGTCAGAGAGCAAAACAGT
ACATCCTTTCAGGTTTGAGGGAAATCAACCTGAGTACATCCAAGTACAATGTGCCATGACGGA
ATAATGAAGTCTTGGCCTTGGCCCGAATAATCAGTAGCTGCCATCTC

SEQ ID N° : 189, semelhante a SbGLK1 (Sorghum bicolor), seqüência parcial
ARGLKFPPPPPPPPPPPPPSPHPMQHFGRPLHAWGHPTPTVESPRVPMWPRHLVPRTPPPPWAPPPP
SDPAFWHHAYMRGPAHMPGQVTPCVAVPMPAARFPAPPVRGALPCPPPMYRPLVPPTLDTQFQL
QTQPSSESIDAAIGDVLTKPWLPLPLGLKPPSVDSVMGELQRQGVANVPPACG

SEQ ID N° : 190, seqüência parcial SCSBSB1052F04.g CA171658 SCSBSB1052F04 semelhante a GLK2 de cana de açúcar
CATCGACAAGGCCGTGCCGTCCCGGATCCTCGAGATAATGGGCATGGATTGCCTCACAAGAC
ACAACATTGCCAGCCACCTCCAGAAGTACCGGTCGCACAGAAAGCACCTGATGGCGCGGGAG
GCGGAGGCCGCCACCTGGGCGCAGAAGCGGCACATGTACGCGGCGGCCGGCGGAGTAGCCCC
GAGGACGGACGCACCACACGGCAGCCGGCCGTGGGTGGTGCCGACCATCGGGTTCCCGCCGC
CGGCGCCCCCGCCCTTCTGCCGGCCGCTGCACGTGTGGGGCCACCCGCCCCACGCCGCCGCCG
CTGAAGCCGCTGCTGCGGTGGCGGCGACTCCGACTCCGATGCTGCCCGTGTGGCCGCGGCACC
TGGCGCCGCCCCGGCCGCTGCCGCCGTGGGCGCACCCGCACCCGCCGGCCGTGGACCCGGCG
TTCTGGCACCAGCAGTACAACGCGGCCAGGAAATGGGGTCCACAGGCAACCGCAGTGACGCA
AGGGACGCCGGCGTGCGTGCCGNCGNCCGCCGCCGNCATGCTGCCGAGGTTTTCCCGGGTTG
GCCCCGGGGGGGGGGGGGGAAAAAAAATTTTTT

SEQ ID N° : 191, semelhante a SoGLK2 (Saccharum officinarum), seqüência parcial
IDKAVPSRILEIMGMDCLTRHNIASHLQKYRSHRKHLMAREAEAATWAQKRHMYAAAGGVAPRT
DAPHGSRPWVVPTIGFPPPAPPPFCRPLHVWGHPPHAAAAEAAAAVAATPTPMLPVWPRHLAPPR
PLPPWAHPHPPAVDPAFWHQQYNAARKWGPQATAVTQGTPACVPXXAAXMLPRFSRVGPGGGG
GKKIF

SEQ ID N° : 192, AK122161.1| Oryza sativa (grupo de cultivar japônico) cDNA clone:J033147G01, seqüência de insersão total
GGCAACTAACAATACCTCGCACCTAGTCACCACCGCACCAACCGCGCGCGACCGGCCCCCCGT
CAGTCGAAAGCTGAGCCTGAGCGAGCAACTGATGCCGCTAGCTATAGCTGCTGCCTGCAGCA
CGCAGCGCCCGTAACCTAACGTATCATTTACTTCCATTGCACTGCACATCCTGTGAGTTCGTTG
TCTGAGTGAGAGCTGGACGTGCTGCACCGAGCTCCTGGCCGAGATGCTTGCCGTGTCGCCGGC
GATGTGCCCCGACATTGAGGACCGCGCCGCGGTGGCCGGCGATGCTGGCATGGAGGTCGTCG
GGATGTCGTCGGACGACATGGATCAGTTCGACTTCTCCGTCGATGACATAGACTTCGGGGACT
TCTTCCTGAGGCTGGAGGACGGTGATGTGCTCCCGGACCTCGAGGTCGACCCGGCCGAGATCT
TCACCGACTTCGAGGCAATCGCGACGAGTGGAGGCGAAGGTGTGCAGGACCAGGAGGTGCCC
ACCGTCGAGCTCTTGGCGCCTGCGGACGACGTCGGTGTGCTGGATCCGTGCGGCGATGTCGTC
GTCGGGGAGGAGAACGCGGCGTTTGCCGGGGCTGGAGAGGAGAAGGTAGGGTGTAACCAGG
ACGATGATGCGGGGGAAGCGAATGTCGACGATGGAGCCGCGGCGGTTGAGGCCAAGTCTTCG
TCGCCGTCATCGACGACGTCGTCGTCG

FIGURA 27 (continuação)

CAGGAGGCTGAGAGCCGGCACAAGTCATCCAGCAAGAGCTCCCATGGGAAGAAGAAAGCGA
AGGTGGACTGGACGCCTGAGCTTCACCGGAGGTTCGTGCAGGCGGTGGAGCAGCTCGGCATC
GACAAGGCCGTGCCGTCGAGGATACTTGAGATCATGGGGATCGACTCTCTCACCCGGCACAA
CATAGCCAGCCATCTTCAGAAGTACCGGTCACACAGAAAACACATGATTGCGAGAGAGGCGG
AGGCAGCGAGTTGGACCCAACGGCGGCAGATTTACGCCGCCGGTGGAGGTGCTGTTGCGAAG
AGGCCGGAGTCCAACGCGTGGACCGTGCCAACCATTGGCTTCCCTCCTCCTCCGCCACCACCA
CCATCACCGGCTCCGATTCAACATTTTGCTCGCCCGTTGCATGTTTGGGGCCACCCGACGATG
GACCCGTCCCGAGTTCCAGTGTGGCCACCGCGGCACCTCGTTCCCCGTGGCCCGGCGCCACCA
TGGGTTCCACCGCCGCCGCCGTCGGACCCTGCTTTCTGGCACCACCCTTACATGAGGGGCCA
GCACATGTGCCAACTCAAGGGACACCTTGCATGGCGATGCCCATGCCAGCTGCGAGATTTCCT
GCTCCACCGGTGCCAGGAGTTGTCCCGTGTCCAATGTATAGGCCATTGACTCCACCAGCACTG
GCGAGCAAGAATCAGCAGGACGCACAGCTTCAACTCCAGGTTCAACCATCAAGCGAGAGCAT
CGACGCAGCTATCGGTGATGTTTTATCGAAACCGTGGTTGCCTTTGCCTCTTGGACTGAAGCC
ACCTTCAGTGGACAGTGTGATGGGCGAGCTGCAGAGGCAAGGCGTAGCAAACGTTCCTCCAG
CGTGTGGATGATATTTGCATGATAGATTGATGGTTCCGTTACTGACTGATTGCTTATCTGGTCA
GTTCACCAAAAAATGGGAAAATTCGCCGACATGTCCTTTTATTTTTCTTTTGGCCTAGCGTTTC
TTGGACATCTCGTTTAATTTTTAACTTGTGCAGAGTTCGAAAGACATCTTTGTTTCGACTCCGG
GAGAAATTAACCTGCGTATTGTAAATGTAAAATGT

SEQ ID N° : 193, AK122161.1| Oryza sativa OsGLK1 seqüência de proteína deduzida (cds 233-1600)
MLAVSPAMCPDIEDRAAVAGDAGMEVVGMSSDDMDQFDFSVDDIDFGDFFLRLEDGDVLPDLEV
DPAEIFTDFEAIATSGGEGVQDQEVPTVELLAPADDVGVLDPCGDVVVGEENAAFAGAGEEKVGC
NQDDDAGEANVDDGAAAVEAKSSSPSSTTSSSQEAESRHKSSSKSSHGKKKAKVDWTPELHRRFV
QAVEQLGIDKAVPSRILEIMGIDSLTRHNIASHLQKYRSHRKHMIAREAEAASWTQRRQIYAAGGG
AVAKRPESNAWTVPTIGFPPPPPPPPSPAPIQHFARPLHVWGHPTMDPSRVPVWPPRHLVPRGPAPP
WVPPPPPSDPAFWHHPYMRGPAHVPTQGTPCMAMPMPAARFPAPPVPGVVPCPMYRPLTPPALA
SKNQQDAQLQLQVQPSSESIDAAIGDVLSKPWLPLPLGLKPPSVDSVMGELQRQGVANVPPACG

SEQ ID N° : 194 Oryza sativa Orysa_REV parcial CTR seqüência de ácido nucleico

TGTGCTAAGGCATCCATGCTACTGCAGAGTGTCCCACCTGCAGTTTTGGTTCGATTTTTGAGGG
AACATCGTTCTGAATGGGCGGATTATAACTTCGATGCATATTCAGCTTCATCTCTGAAGACAA
GCTCATGTTCACTTCCTGGGTTGCGGCCTATGAGATTTTCTGGGAGCCAGATCATTATGCCACT
TGCTCACACGGTGGAGAATGAAGAGATTTTAGAAGTTGTCCGTCTTGAAGGACAAGCACTTAC
ACATGATGATGGTCTTATGTCTAGAGATATTCACCTGCTTCAGCTTTGCACTGGAATAGATGA
GAAATCAATGGGATCCTGCTTCCAGCTTGTCTTTGCACCAATCGATGAGCTTTTCCCTGATGAT
GCTCCGTTAATATCTTCAGGCTTTCGTGTTATACCGCTGGACATGAAAACAGATGGTACACCT
GCTGGTAGAACATTAGATTTGGCATCTAGCCTTGAGGTTGGTTCAACTGCACAGCCCACAGGG
GATGCATCTATGGATGACTGTAATCTACGATCAGTGCTGACAATTGCCTTTCAGTTCCCTTATG
AAATGCATCTCCAAGACAGCGTTGCAACTATGGCCCGGCAATATGTCCGCAGTATTGTTTTCT
CTGTTCAGAGAGTATCAATGGCTGTTTCTCCTTCTCGGTCTGGCTTGAATGCTGGGCAGAAGAT
AATTTCAGGCTTCCCTGAAGCCCCAACGCTAGCTCGTTGGATTTGCCAAAGCTACCAGTTCCA
TTTGGGGGTGGAGTTACTTAGGCAGGCAGATGATGCTGGGGAAGCACTATTGAAAATGCTAT
GGGATTACGAAGACGCTATTTTGTGCTGTTCTTTCAAGGAAAAGCCTGTATTTACTTTTGCCAA
CGAGATGGGACTAAACATGCTAGAAACATCTCTCGTCGCTCTCCAAGATCTCTCACTGGACAA
GATATTTGATGAAGCCGGTAGGAAGGCCCTATACAACGAGATCCCGAAATTGATGGAACAGG
GTTACGTGTACCTGCCTGGTGGAGTGTGCTTGTCCGGGATGGGGCGCCATGTTTCTTTCGAGC
AAGCTGTAGCATGGAAGGTGC

FIGURA 27 (continuação)

SEQ ID N° : 195 Oryza sativa Orysa_REV polipeptídeo deduzido de CTR parcial
CAKASMLLQSVPPAVLVRFLREHRSEWADYNFDAYSASSLKTSSCSLPGLRPMRFSGSQIIMPLAH
TVENEEILEVVRLEGQALTHDDGLMSRDIHLLQLCTGIDEKSMGSCFQLVSAPIDELFPDDAPLISS
GFRVIPLDMKTDGTPAGRTLDLASSLEVGSTAQPTGDASMDDCNLRSVLTIAFQFPYEMHLQDSV
ATMARQYVRSIVSSVQRVSMAISPSRSGLNAGQKIISGFPEAPTLARWICQSYQFHLGVELLRQAD
DAGEALLKMLWDYEDAILCCSFKEKPVFTFANEMGLNMLETSLVALQDLSLDKIFDEAGRKALY
NEIPKLMEQGYVYLPGGVCLSGMGRHVSFEQAVAWKV

SEQ ID N0: 196 Oryza sativa Orysa_REV CTR seqüência de ácido nucleico
GCTCAAGAGACAAGCGGGGAGGTTGTATACGCTTTGGGGAGGCAACCTGCTGTTTTGCGGAC
ATTTAGTCAGAGGTTGAGTAGAGGCTTCAATGATGCTATAAGTGGTTTCAATGATGATGGTTG
GTCTGTCATGGGTGGGGATGGCATTGAAGATGTGATCATTGCTTGCAATGCAAAGAAGGTTAG
GAATACTAGCACTTCGGCCAATGCTTTTGTAACTCCAGGAGGTGTTATATGTGCTAAGGCATC
CATGCTACTGCAGAGTGTCCCACCTGCAGTTTTGGTTCGATTTTTGAGGGAACATCGTTCTGAA
TGGGCGGATTATAACTTCGATGCATATTCAGCTTCATCTCTGAAGACAAGCTCATGTTCACTTC
CTGGGTTGCGGCCTATGAGATTTTCTGGGAGCCAGATCATTATGCCACTTGCTCACACGGTGG
AGAATGAGGAGATTTTAGAAGTTGTCCGTCTTGAAGGACAAGCACTTACACATGATGATGGTC
TTATGTCTAGAGATATTCACCTGCTTCAGCTTTGCACTGGAATAGATGAGAAATCAATGGGAT
CCTGCTTCCAGCTTGTCTCTGCACCAATCGATGAGCTTTTCCCTGATGATGCTCCGTTAATATC
TTCAGGCTTTCGTGTTATACCGCTGGACATGAAAACAGATGGTACACCTGCTGGTAGAACATT
AGATTTGGCATCTAGCCTTGAAGTTGGTTCAACTGCACAGCCCACAGGGGATGCATCTATGGA
TGACTGTAATCTACGATCAGTGCTGACAATTGCCTTTCAGTTCCCTTATGAAATGCATCTCCAA
GACAGCGTTGCAACTATGGCCCGGCAATATGTCCGCAGTATTGTTTCCTCTGTTCAGAGAGTA
TCAATGGCTATTTCTCCTTCTCGGTCTGGCTTGAATGCTGGGCAGAAGATAATTTCAGGCTTCC
CTGAAGCCCCAACGCTAGCTCGTTGGATTTGCCAAAGCTACCAGTTCCATTTGGGGGTGGAGT
TACTTAGGCAGGCAGATGATGCTGGGGAAGCACTATTGAAAATGCTATGGGATTACGAAGAC
GCTATTTTGTGCTGTTCTTTCAAGGAAAAGCCTGTGTTTACTTTTGCCAACGAGATGGGACTAA
ACATGCTAGAAACATCTCTCGTCGCTCTCCAAGATCTCTCACTGGACAAGATATTTGATGAAG
CCGGTAGGAAGGCCCTATACAACGAGATCCCGAAATTGATGGAACAGGGTTACGTGTACCTG
CCTGGTGGAGTGTGCTTGTCCGGGATGGGGCGCCATGTTTCTTTCGAGCAAGCTGTAGCATGG
AAGGTGCTCGGAGAAGACAACAATGTGCACTGCCTCGCCTTCTGCTTCGTCAACTGGTCCTTC
GTGTGA

SEQ ID N° : 197 Oryza sativa Orysa_REV polipeptídeo deduzido de CTR
AQETSGEVVYALGRQPAVLRTFSQRLSRGFNDAISGFNDDGWSVMGGDGIEDVIIACNAKKVRNT
STSANAFVTPGGVICAKASMLLQSVPPAVLVRFLREHRSEWADYNFDAYSASSLKTSSCSLPGLRP
MRFSGSQIIMPLAHTVENEEILEVVRLEGQALTHDDGLMSRDIHLLQLCTGIDEKSMGSCFQLVSAP
IDELFPDDAPLISSGFRVIPLDMKTDGTPAGRTLDLASSLEVGSTAQPTGDASMDDCNLRSVLTIAF
QFPYEMHLQDSVATMARQYVRSIVSSVQRVSMAISPSRSGLNAGQKIISGFPEAPTLARWICQSYQ
FHLGVELLRQADDAGEALLKMLWDYEDAILCCSFKEKPVFTFANEMGLNMLETSLVALQDLSLD
KIFDEAGRKALYNEIPKLMEQGYVYLPGGVCLSGMGRHVSFEQAVAWKVLGEDNNVHCLAFCFV
NWSFV

SEQ ID N° : 198 Oryza sativa Orysa_REV seqüência de ácido nucleico Os10g33960 AK102830
GACGTCAGGTTATTTTTTTTCTCCTCGTCTCCTGACGATTTCGCTTTTGCTCGAAATCTCCCAG
GTTTGTTGTTGTTGTGTTGAAGGCGGAGAAGAGGGGAGGCTTTGGTGGTGGTGGAGGAGGAG
GATGGCGGCGGCGGTGGCGATGCGGAGCGGCAGCGGCAGCGACGGCGGCGGCGGCGGGTAC
GACAAGG

FIGURA 27 (continuação)

CCGGGATGGACTCCGGCAAGTACGTGCGGTACACGCCGGAGCAGGTGGAGGCGCTGGAGAGG
GTGTACGCCGAGTGCCCCAAGCCCAGCTCCTCCCGCCGCCAGCAGCTGCTCCGCGACTGTCCC
ATCCTCGCCAACATCGAGCCCAAGCAGATCAAGGTCTGGTTCCAGAACAGAAGGTGCCGAGA
TAAGCAGCGGAAGGAGGCATCAAGGCTTCAGGCCGTGAACCGAAAATTGACGGCGATGAATA
AGCTTCTCATGGAGGAGAATGAGCGTCTTCAGAAGCAGGTCTCCCAGCTGGTCCATGAGAAC
GCGTACATGAAGCAGCAACTTCAGAATCCGTCATTGGGCAATGATACAAGCTGTGAATCAAA
TGTGACCACTCCTCAGAACCCTCTGAGAGATGCAAGTAACCCGTCTGGACTCCTTACAATTGC
GGAGGAGACCCTGACAGAGTTCCTCTCCAAGGCTACAGGGACTGCTGTTGATTGGGTGCCAAT
GCCTGGGATGAAGCCTGGTCCGGATTCGTTTGGTATTGTGGCCGTTTCACATGGTTGCCGTGGT
GTTGCTGCCCGTGCCTGTGGTTTGGTGAATCTAGAACCAACAAAGATCGTGGAGATCTTAAAA
GACCGCCCATCTTGGTTCCGTGATTGTCGAAGTCTTGAAGTCTTCACAATGTTTCCAGCTGGAA
ATGGTGGCACGATCGAACTTGTTTACATGCAGATGTATGCTCCTACTACTTTGGTTCCTGCACG
AGATTTTTGGACACTTAGATACACAACTACAATGGAGGATGGCAGCCTTGTGGTCTGTGAGAG
ATCATTGAGTGGTTCTGGAGGTGGTCCAAGTACAGCCTCCGCACAGCAATTTGTAAGAGCTGA
GATGCTTCCTAGCGGCTATCTAGTGCGCCCATGCGAGGGTGGTGGCTCCATCGTGCATATTGT
GGACCATCTGGATCTTGAGGCTTGGAGTGTTCCAGAAGTGCTTCGGCCACTCTACGAGTCATC
TAGGGTAGTTGCTCAGAAAATGACTACTGCAGCACTACGGCACATCAGACAAATTGCTCAAG
AGACAAGCGGGGAGGTTGTATACGCTTTGGGGAGGCAACCTGCTGTTTTGCGGACATTTAGTC
AGAGGTTGAGTAGAGGCTTCAATGATGCTATAAGTGGTTTCAGTGATGATGGTTGGTCTGTCA
TGGGTGGGGATGGCATTGAAGATGTGATCATTGCTTGCAATGCAAAGAAGGTTAGGAATACT
AGCACTTCGGCCAATGCTTTTGTAACTCCAGGAGGTGTTATATGTGCTAAGGCATCCATGCTA
CTGCAGAGTGTCCCACCTGCAGTTTTGGTTCGATTTTTGAGGGAACATCGTTCTGAATGGGCG
GATTATAACTTCGATGCATATTCAGCTTCATCTCTGAAGACAAGCTCATGTTCACTTCCTGGGT
TGCGGCCTATGAGATTTTCTGGGAGCCAGATCATTATGCCACTTGCTCACACGGTGGAGAATG
AGGAGATTTTAGAAGTTGTCCGTCTTGAAGGACAAGCACTTACACATGATGATGGTCTTATGT
CTAGAGATATTCACCTGCTTCAGCTTTGCACTGGAATAGATGAGAAATCAATGGGATCCTGCT
TCCAGCTTGTCTCTGCACCAATCGATGAGCTTTTCCCTGATGATGCTCCGTTAATATCTTCAGG
CTTTCGTGTTATACCGCTGGACATGAAAACAGATGGTACACCTGCTGGTAGAACATTAGATTT
GGCATCTAGCCTTGAAGTTGGTTCAACTGCACAGCCCACAGGGGATGCATCTATGGATGACTG
TAATCTACGATCAGTGCTGACAATTGCCTTTCAGTTCCCTTATGAAATGCATCTCCAAGACAGC
GTTGCAACTATGGCCCGGCAATATGTCCGCAGTATTGTTTCCTCTGTTCAGAGAGTATCAATG
GCTATTTCTCCTTCTCGGTCTGGCTTGAATGCTGGGCAGAAGATAATTTCAGGCTTCCCTGAAG
CCCCAACGCTAGCTCGTTGGATTTGCCAAAGCTACCAGTTCCATTTGGGGGTGGAGTTACTTA
GGCAGGCAGATGATGCTGGGGAAGCACTATTGAAAATGCTATGGGATTACGAAGACGCTATT
TTGTGCTGTTCTTTCAAGGAAAAGCCTGTGTTTACTTTTGCCAACGAGATGGGACTAAACATG
CTAGAAACATCTCTCGTCGCTCTCCAAGATCTCTCACTGGACAAGATATTTGATGAAGCCGGT
AGGAAGGCCCTATACAACGAGATCCCGAAATTGATGGAACAGGGTTACGTGTACCTGCCTGG
TGGAGTGTGCTTGTCCGGGATGGGGCGCCATGTTTCTTTCGAGCAAGCTGTAGCATGGAAGGT
GCTCGGAGAAGACAACAATGTGCACTGCCTCGCCTTCTGCTTCGTCAACTGGTCCTTCGTGTG
ACCCAAAATCCAATCCACCATGCTCGCAGTCATCGTCGTTGTCGTCAGATCTCCATTTTTTGCA
ATCTCTGTAGAAGAGAGGACACCAAACCAACCACAAACCTGTCAGCTGTTAGCTAATGATTCT
ACTAGCTTTTCTAGATCTTTGAAGGCAAAATGCTGCCGGTGCCTGTAAAGAGTGTGTGCGTGC
GTGTGTAAATTTGGGCGCGTTTTGCACTAATTTGATCTGGGGTAAGGAGGCGCTGGCTGCTGC
TGTGCTGTAGTTTGGTAAAAGTTGAGTACGCTAACCTTGGATGGTCTCGAACTGTATGATGAA
ATTCCTAGCAGTAAAATTTCAACTTGGATCCGCG

FIGURA 27 (continuação)

SEQ ID N° : 199 polipeptídeo deduzido de Oryza sativa Orysa_REV AK102830

MAAAVAMRSGSGSDGGGGGYDKAGMDSGKYVRYTPEQVEALERVYAECPKPSSSRRQQLLRDC
PILANIEPKQIKVWFQNRRCRDKQRKEASRLQAVNRKLTAMNKLLMEENERLQKQVSQLVHENA
YMKQQLQNPSLGNDTSCESNVTTPQNPLRDASNPSGLLTIAEETLTEFLSKATGTAVDWVPMPGM
KPGPDSFGIVAVSHGCRGVAARACGLVNLEPTKIVEILKDRPSWFRDCRSLEVFTMFPAGNGGTIE
LVYMQMYAPTTLVPARDFWTLRYTTTMEDGSLVVCERSLSGSGGGPSTASAQQFVRAEMLPSGY
LVRPCEGGGSIVHIVDHLDLEAWSVPEVLRPLYESSRVVAQKMTTAALRHIRQIAQETSGEVVYAL
GRQPAVLRTFSQRLSRGFNDAISGFNDDGWSVMGGDGIEDVIIACNAKKVRNTSTSANAFVTPGG
VICAKASMLLQSVPPAVLVRFLREHRSEWADYNFDAYSASSLKTSSCSLPGLRPMRFSGSQIIMPLA
HTVENEEILEVVRLEGQALTHDDGLMSRDIHLLQLCTGIDEKSMGSCFQLVSAPIDELFPDDAPLIS
SGFRVIPLDMKTDGTPAGRTLDLASSLEVGSTAQPTGDASMDDCNLRSVLTIAFQFPYEMHLQDSV
ATMARQYVRSIVSSVQRVSMAISPSRSGLNAGQKIISGFPEAPTLARWICQSYQFHLGVELLRQAD
DAGEALLKMLWDYEDAILCCSFKEKPVFTFANEMGLNMLETSLVALQDLSLDKIFDEAGRKALY
NEIPKLMEQGYVYLPGGVCLSGMGRHVSFEQAVAWKVLGEDNNVHCLAFCFVNWSFV

SEQ ID N° : 200 Oryza sativa Orysa_HOX10 seqüência de ácido nucleico AY425991.1

CTCTCTCCCGTGCTGCTCTCCCTTCTCCTCTCCTCAATTACCACGCCGCAGCCGCAGCCCCCCA
ACCCTAGCTAGTACGGCGGGCGACCTGAGGGGGTAGCTGCCTACCTCTATACCTGCTATCTCC
GTTCGATCCGGGTCGGGCTCGTGCTGCTGGGGGCGGGGGCAATCAATCGGCGAGCTCGGTGA
TCCATGGCCGTGGCTGTGGGGTGGCTGTGAGGGTGCCCCCGGCGGCGCTCCCCTCCGCGCCTG
CCGGCGAGGGGGCTCGGACTGAAGGGATCTAGGCGAGCTGAAAATTGAAGTGCAGGCAAGG
AGATAAGAGCAGCGTCCAAATTGTGAGTACTTCATTAGCAGGAGGTAGTGGTTGTGCTTGCTT
GGCTCCTTTGCAATTTGGCTTTGGCGAGGTAGCAATGGCTGCGGCAGTGGCAATGCGAGGGAG
TAGCAGTGATGGAGGTGGCTATGATAAGGTTTCCGGGATGGACTCCGGTAAATATGTGCGCTA
CACGCCTGAGCAGGTGGAGGCGCTTGAGCGGGTGTACGCCGATTGCCCCAAGCCAACCTCCTC
CCGCAGGCAGCAACTGCTGCGTGAGTGCCCCATACTTGCTAACATTGAGCCCAAGCAGATCAA
GGTCTGGTTCCAGAACAGAAGGTGCCGGGATAAGCAGCGGAAGGAGTCTTCACGGCTTCAGG
CTGTCAACAGGAAATTGACGGCAATGAACAAGCTACTTATGGAAGAGAATGAGCGACTCCAG
AAGCAGGTCTCCCAATTGGTTCATGAGAATGCCCACATGCGACAGCAGCTGCAGAATACTCCG
CTGGCAAATGATACAAGCTGTGAATCAAATGTGACTACCCCTCAAAACCCTTTAAGGGATGCA
AGTAACCCCTCTGGGCTCCTTTCAATTGCAGAGGAGACCTTGACAGAGTTCCTCTCAAAGGCT
ACTGGTACAGCTATTGATTGGGTCCAGATGCCTGGGATGAAGCCTGGTCCGGATTCGGTTGGT
ATTGTGGCCATTTCACATGGTTGCCGTGGTGTTGCTGCCCGTGCCTGTGGTTTGGTGAACCTAG
AACCAACAAAAGTGGTAGAGATATTGAAAGATCGTCCATCTTGGTTCCGTGATTGTCGAAACC
TGGAAGTCTTTACAATGATTCCAGCAGGAAATGGAGGAACGGTTGAACTTGTCTACACACAGT
TGTATGCTCCAACAACTTTAGTTCCTGCACGAGATTTTTGGACGTTACGGTACACAACCACAA
TGGAAGATGGCAGTCTTGTGGTCTGTGAGAGATCTTTAAGTGGTTCAGGGGGCGGTCCAAGTG
CTGCCTCTGCTCAGCAATATGTGAGAGCGGAAATGCTTCCAAGTGGATACCTGGTTCGCCCAT
GTGAAGGTGGGGGATCAATTGTGCACATAGTGGACCATCTGGATCTTGAGGCATGGAGTGTTC
CTGAGGTGCTTCGGCCACTCTATGAATCTTCAAGGGTAGTCGCTCAGAAAATGACTACTGCGG
CACTCCGGCACATCAGACAAATTGCTCAAGAAACAAGTGGGGAAGTGGTGTATGCCTTGGGG
AGGCAACCAGCAGTGCTACGGACTTTTAGTCAAAGGCTGAGCAGAGGCTTTAACGATGCCATT
AGTGGTTTCAATGATGATGGGTGGTCTATAATGGGCGGAGACGGTGTTGAAGATGTAGTTATT
GCTTGCAACTCAACTAAGAAAATTAGGAGTAACAGCAATGCTGGCATCGCCTTTGGAGCCCCC
GGAGGTATTATATGTGCTAAGGCATCAATGTTACTGCAGAGTGTTCCTCCAGCAGTACTGGTT
CGATTTCTGAGGGAGCATAGATCTGAATGGGCCGATTACAATATTGATGCATATTTGGCTTCA
ACTCTGAAAACAAGTGCATGTTCACTTACTGGGTTGCGACCCATGAGATTTTCTGGGAGCCAA
A

FIGURA 27 (continuação)

TCATCATTCCACTTGCTCACACAGTTGAGAATGAGGAGATTCTTGAAGTTGTTCGCCTTGAGG
GTCAACCTCTTACTCATGATGAAGCTCTTCTTTCAAGGGATATCCACCTGCTTCAGCTCTGCAC
TGGAATAGACGAGAAGTCTGTGGGATCCTCCTTTCAGCTTGTGTTTGCACCGATTGATGATTTC
CCAGATGAAACTCCATTGATTTCTTCTGGCTTCCGTGTTATACCACTTGATATGAAAACAGATG
GTGCATCCTCTGGTAGGACATTAGATTTGGCATCTAGTCTTGAAGTAGGTTCAGCAACAGCTC
AAGCCTCCGGAGATGCATCTGCAGATGATTGTAACTTGCGATCTGTTCTGACGATCGCTTTTCA
ATTCCCTTACGAGTTGCATCTCCAAGACAGTGTTGCAGCTATGGCTCGCCAATATGTCCGTAG
CATTGTTTCTGCTGTGCAAAGAGTGTCAATGGCTATCTCTCCCTCTCAAACTGGTCTAAATGCC
GGACAGAGGATAATCTCTGGTTTCCCTGAAGCAGCAACCCTTGCTCGATGGGTTTGCCAGAGC
TACCATTACCATCTAGGGGTAGAGTTACTTAGTCAATCAGATGGAGATGCAGAACAATTGTTG
AAGATGCTATGGCATTACCAAGATGCTATTTTGTGCTGCTCATTCAAGGAAAAACCGGTGTTT
ACATTTGCCAACAAAGCAGGACTGGACATGCTAGAAACTTCCCTTGTCGCCTTACAGGACCTC
ACATTGGACAGGATCTTTGATGAGCCTGGAAAAGAAGCATTGTTCTCAAACATTCCCAAATTG
ATGGAGCAGGGCCATGTCTACCTGCCATCAGGCGTGTGCATGTCAGGAATGGGTCGGCATGTT
TCTTTCGATCAGGCCGTGGCTTGGAAAGTGCTTGCCGAGGATAGCAATGTTCACTGCCTGGCC
TTCTGTTTCGTCAACTGGTCCTTTGTGTGACCATCCCACATCCACTGCGAAGTGAAGATTCATT
TTTGCTTGTTCAAGACTGTTTTGCACTAGATCTAGACCTGAGAATCTAGTAGTATTTGCGAAAT
TTCTCGCTTATGTAAATGTAATCTCGCTCCCATTCCATCAAGTACCCAAGTACAAGGCAGTGG
CCAAGTGGTTTTAGTTGTAGGGGGATCGGAGCAATCCTGACGGTTGGGTACTTGGGTTCCTTC
CAGCAATGTAAAATCGGATCCTGTAGGGCGTCGAAAACTGCATTGGCAAGATTCTTCGAAATG
CAGACCAAATTTAGCTAGTACATCGTAACTTTGGC

SEQ ID N° : 201 polipeptídeo deduzido de Oryza sativa Orysa_HOX10 AY425991.1
MAAAVAMRGSSSDGGGYDKVSGMDSGKYVRYTPEQVEALERVYADCPKPTSSRRQQLLRECPIL
ANIEPKQIKVWFQNRRCRDKQRKESSRLQAVNRKLTAMNKLLMEENERLQKQVSQLVHENAHM
RQQLQNTPLANDTSCESNVTTPQNPLRDASNPSGLLSIAEETLTEFLSKATGTAIDWVQMPGMKPG
PDSVGIVAISHGCRGVAARACGLVNLEPTKVVEILKDRPSWFRDCRNLEVFTMIPAGNGGTVELV
YTQLYAPTTLVPARDFWTLRYTTTMEDGSLVVCERSLSGSGGGPSAASAQQYVRAEMLPSGYLV
RPCEGGGSIVHIVDHLDLEAWSVPEVLRPLYESSRVVAQKMTTAALRHIRQIAQETSGEVVYALGR
QPAVLRTFSQRLSRGFNDAISGFNDDGWSIMGGDGVEDVVIACNSTKKIRSNSNAGIAFGAPGGIIC
AKASMLLQSVPPAVLVRFLREHRSEWADYNIDAYLASTLKTSACSLTGLRPMRFSGSQIIIPLAHTV
ENEEILEVVRLEGQPLTHDEALLSRDIHLLQLCTGIDEKSVGSSFQLVFAPIDDFPDETPLISSGFRVIP
LDMKTDGASSGRTLDLASSLEVGSATAQASGDASADDCNLRSVLTIAFQFPYELHLQDSVAAMAR
QYVRSIVSAVQRVSMAISPSQTGLNAGQRIISGFPEAATLARWVCQSYHYHLGVELLSQSDGDAEQ
LLKMLWHYQDAILCCSFKEKPVFTFANKAGLDMLETSLVALQDLTLDRIFDEPGKEALFSNIPKLM
EQGHVYLPSGVCMSGMGRHVSFDQAVAWKVLAEDSNVHCLAFCFVNWSFV

SEQ ID N° : 202, Arabidopsis thaliana Arath_REV seqüência de ácido nucleico AF188994
AGAGTCTTCAAAACTTTTGCAGCTTCAATTGTACCTGGGTTTCTTCTTCATTGTTCCTAAGGTTT
CTGTGTCCTTCAATTCTTCTGATATAATGCTTCTTTAAGAGAGTTGACATCATCACTTTCTTGG
GGTACTCTTCTCTGTTTCTCCCCAGAAAATCCAACTCTGTAATTTTGGGTCTTTATTCTGTTTTT
CTCTTTGAAGAATCTTTAAAATTCTCAGATCTTCTGAATCTCTCTTCTTTAAAACTTTTTTTAAC
TTTATTTTTTGTACTCGCTTCTTTGCCTTCATTTTTCTCGTATCCACATGTCGTTGGTCTTTCGCT
ACAAGCCACGACCGTAGAATCTTCTTTTGTCTGAAAAGAATTACAATTTACGTTTCTCTTACGA
TACGACGGACTTTCCGAAGAAATTAATTTAAAGAGAAAAGAAGAAGAAGCCAAAGAAGAAG
AAGAAGCTAGAAGAAACAGTAAAGTTTGAGACTTTTTTTGAGGGTCGAGCTAAAATGGAGAT
GGCGGTGG

FIGURA 27 (continuação)

CTAACCACCGTGAGAGAAGCAGTGACAGTATGAATAGACATTTAGATAGTAGCGGTAAGTAC
GTTAGGTACACAGCTGAGCAAGTCGAGGCTCTTGAGCGTGTCTACGCTGAGTGTCCTAAGCCT
AGCTCTCTCCGTCGACAACAATTGATCCGTGAATGTTCCATTTTGGCCAATATTGAGCCTAAGC
AGATCAAAGTCTGGTTTCAGAACCGCAGGTGTCGAGATAAGCAGAGGAAAGAGGCGTCGAGG
CTCCAGAGCGTAAACCGGAAGCTCTCTGCGATGAATAAACTGTTGATGGAGGAGAATGATAG
GTTGCAGAAGCAGGTTTCTCAGCTTGTCTGCGAAAATGGATATATGAAACAGCAGCTAACTAC
TGTTGTTAACGATCCAAGCTGTGAATCTGTGGTCACAACTCCTCAGCATTCGCTTAGAGATGC
GAATAGTCCTGCTGGATTGCTCTCAATCGCAGAGGAGACTTTGGCAGAGTTCCTATCCAAGGC
TACAGGAACTGCTGTTGATTGGGTTCAGATGCCTGGGATGAAGCCTGGTCCGGATTCGGTTGG
CATCTTTGCCATTTCGCAAAGATGCAATGGAGTGGCAGCTCGAGCCTGTGGTCTTGTTAGCTT
AGAACCTATGAAGATTGCAGAGATCCTCAAAGATCGGCCATCTTGGTTCCGTGACTGTAGGAG
CCTTGAAGTTTTCACTATGTTCCCGGCTGGTAATGGTGGCACAATCGAGCTTGTTTATATGCAG
ACGTATGCACCAACGACTCTGGCTCCTGCCCGCGATTTCTGGACCCTGAGATACACAACGAGC
CTCGACAATGGGAGTTTTGTGGTTTGTGAGAGGTCGCTATCTGGCTCTGGAGCTGGGCCTAAT
GCTGCTTCAGCTTCTCAGTTTGTGAGAGCAGAAATGCTTTCTAGTGGGTATTTAATAAGGCCTT
GTGATGGTGGTGGTTCTATTATTCACATTGTCGATCACCTTAATCTTGAGGCTTGGAGTGTTCC
GGATGTGCTTCGACCCCTTTATGAGTCATCCAAAGTCGTTGCACAAAAAATGACCATTTCCGC
GTTGCGGTATATCAGGCAATTAGCCCAAGAGTCTAATGGTGAAGTAGTGTATGGATTAGGAA
GGCAGCCTGCTGTTCTTAGAACCTTTAGCCAAAGATTAAGCAGGGGCTTCAATGATGCGGTTA
ATGGGTTTGGTGACGACGGGTGGTCTACGATGCATTGTGATGGAGCGGAAGATATTATCGTTG
CTATTAACTCTACAAAGCATTTGAATAATATTTCTAATTCTCTTTCGTTCCTTGGAGGCGTGCT
CTGTGCCAAGGCTTCAATGCTTCTCCAAAATGTTCCTCCTGCGGTTTTGATCCGGTTCCTTAGA
GAGCATCGATCTGAGTGGGCTGATTTCAATGTTGATGCATATTCCGCTGCTACACTTAAAGCT
GGTAGCTTTGCTTATCCGGGAATGAGACCAACAAGATTCACTGGGAGTCAGATCATAATGCCA
CTAGGACATACAATTGAACACGAAGAAATGCTAGAAGTTGTTAGACTGGAAGGTCATTCTCTT
GCTCAAGAAGATGCATTTATGTCACGGGATGTCCATCTCCTTCAGATTTGTACCGGGATTGAC
GAGAATGCCGTTGGAGCTTGTTCTGAACTGATATTTGCTCCGATTAATGAGATGTTCCCGGAT
GATGCTCCACTTGTTCCCTCTGGATTCCGAGTCATACCCGTTGATGCTAAAACGGGAGATGTA
CAAGATCTGTTAACCGCTAATCACCGTACACTAGACTTAACTTCTAGCCTTGAAGTCGGTCCA
TCACCTGAGAATGCTTCTGGAAACTCTTTTTCTAGCTCAAGCTCGAGATGTATTCTCACTATCG
CGTTTCAATTCCCTTTTGAAAACAACTTGCAAGAAAATGTTGCTGGTATGGCTTGTCAGTATGT
GAGGAGCGTGATCTCATCAGTTCAACGTGTTGCAATGGCGATCTCACCGTCTGGGATAAGCCC
GAGTCTGGGCTCCAAATTGTCCCCAGGATCTCCTGAAGCTGTTACTCTTGCTCAGTGGATCTCT
CAAAGTTACAGTCATCACTTAGGCTCGGAGTTGCTGACGATTGATTCACTTGGAAGCGACGAC
TCGGTACTAAAACTTCTATGGGATCACCAAGATGCCATCCTGTGTTGCTCATTAAAGCCACAG
CCAGTGTTCATGTTTGCGAACCAAGCTGGTCTAGACATGCTAGAGACAACACTTGTAGCCTTA
CAAGATATAACACTCGAAAAGATATTCGATGAATCGGGTCGTAAGGCTATCTGTTCGGACTTC
GCCAAGCTAATGCAACAGGGATTTGCTTGCTTGCCTTCAGGAATCTGTGTGTCAACGATGGGA
AGACATGTGAGTTATGAACAAGCTGTTGCTTGGAAAGTGTTTGCTGCATCTGAAGAAAACAAC
AACAATCTGCATTGTCTTGCCTTCTCCTTTGTAAACTGGTCTTTTGTGTGATTCGATTGACAGA
AAAAGACTAATTTAAATTTACGTTAGAGAACTCAAATTTTTGGTTGTTGTTTAGGTGTCTCTGT
TTTGTTTTTTAAAATTATTTTGATCAAATGTTACTCACTTTCTTCTTTCAAAAAAAAAA

SEQ ID N° : 203, polipeptídeo deduzido de Arabidopsis thaliana Arath_REV deduced polypeptide AF188994

MEMAVANHRERSSDSMNRHLDSSGKYVRYTAEQVEALERVYAECPKPSSLRRQQLIRECSILANI
EPKQIKVWFQNRRCRDKQRKEASRLQSVNRKLSAMNKLLMEENDRLQKQVSQLVCENGYMKQQ
LTTVVNDPSCESVVTTPQHSLRDANSPAGLLSIAEETLAEFLSKATGTAVDWVQMPGMKPGPDSV
GIFAISQRCNGVAARACGLVSLEPMKIAEILKDRPSWFRDCRSLEVFTMFPAGNGGTIELVYMQTY
A

FIGURA 27 (continuação)

PTTLAPARDFWTLRYTTSLDNGSFVVCERSLSGSGAGPNAASASQFVRAEMLSSGYLIRPCDGGGS
IIHIVDHLNLEAWSVPDVLRPLYESSKVVAQKMTISALRYIRQLAQESNGEVVYGLGRQPAVLRTF
SQRLSRGFNDAVNGFGDDGWSTMHCDGAEDIIVAINSTKHLNNISNSLSFLGGVLCAKASMLLQN
VPPAVLIRFLREHRSEWADFNVDAYSAATLKAGSFAYPGMRPTRFTGSQIIMPLGHTIEHEEMLEV
VRLEGHSLAQEDAFMSRDVHLLQICTGIDENAVGACSELIFAPINEMFPDDAPLVPSGFRVIPVDAK
TGDVQDLLTANHRTLDLTSSLEVGPSPENASGNSFSSSSSRCILTIAFQFPFENNLQENVAGMACQY
VRSVISSVQRVAMAISPSGISPSLGSKLSPGSPEAVTLAQWISQSYSHHLGSELLTIDSLGSDDSVLK
LLWDHQDAILCCSLKPQPVFMFANQAGLDMLETTLVALQDITLEKIFDESGRKAICSDFAKLMQQ
GFACLPSGICVSTMGRHVSYEQAVAWKVFAASEENNNNLHCLAFSFVNWSFV

SEQ ID N° : 204, Zea mays Zeama_HDIII RLD1 (folha enrolada 1) seqüência de ácido nucleico AY501430.1

GAGAGCTAAGAGCAACAAGGCGACGAGATTTGTGTGGCTACGATTTTGCGTTGCCGACGGAA
CATGGGAACAATGGTTGCAGCGGTGGCGTTGCGAGGGGGAAGCAGCGATAGCGGAGGATTTG
ATAAGGTTCCCGGGATGGACTCGGGAAAATACGTGCGCTACACCCCGGAGCAAGTTGAGGTG
CTCGAGCGGCTCTACATAGATTGCCCCAAGCCAAGCTCCTCACGGCGGCAGCAACTGCTGCGC
GAGTGTCCTATACTCTCCAACATTGAGCCAAAGCAGATCAAGGTCTGGTTCCAGAACCGGAGG
TGCCGCGATAAGCAGCGGAAGGAGTCTTCGCGGCTTCAGGCTGTGAACAGAAAGTTGACGGC
AATGAACAAGCTTCTAATGGAAGAGAATGAGCGGCTTCAGAAGCAAGTCTCTCAGTTGGTTC
ATGAAAACGCGCACATGCGGCAGCAGCTGCAGAATACTTCATTGGCCAATGACACAAGCTGT
GAATCAAATGTCACTACCCCTCCAAACCCTATAAGGGACGCAAGTAACCCTTCTGGACTCCTT
GCGATTGCAGAGGAGACCTTCACAGAGTTCCTCTCAAAGGCTACTGGGACAGCTATTGATTGG
GTCCAGATGCCTGGGATGAAGCCTGGTCCGGATTCAGTTGGTATCGTGGCCATTTCGCATGGT
TGCCGTGGCGTTGCTGCCCGCGCCTGTGGTTTGGTGAATCTAGAACCAACAAAAGGCATAGAG
ATCTTGAAAGATCGTCCCTCTTGGTTCCGTGATTGCCGAAGTCTTGAAGTGTTTACAAGGTTTC
CAGCTGGAAATGGGGGAACAATTGAACTTATTTACATGCAGATGTATGCCCCAACAACTTTAG
TCCCTGCACGTGATTTTTGGACACTACGATACACGACAACAATGGAAGATGGCAGCCTTGTGG
TCTGTGAGAGATCCTTGAGTGGTTCTGGTGGTGGTCCAAATGCAGCCTCTACACAACAATTTG
TTAGGGCTGAGATGCTTCCAAGTGGGTATTTAGTTCGCCCATGCGAAGGTGGAGGATCAATTG
TGCATATAGTGGACCATCTAGATCTCGAAGCATGGAGTGTTCCTGAAGTGCTTCGACCACTGT
ATGAGTCTTCTAGAGTTGTTGCTCAGAAAATGACTACTGTGGCACTGCGCCACCTTAGACAAA
TTGCTCAAGAAACAAGTGGAGAAGTAGTGTACGCCTTGGGAAGGCAACCTGCAGTCCTACGG
ACCTTTAGTCAAAGACTAAGCAGGGGGTTTAATGATGCCATTAGCGGTTTCAATGATGATGGC
TGGTCTGTAATGGGGGGAGATGGCATTGAAGACGTTGTTATTGCTTGCAACTCAACCAAGAAA
ATTAGGAATACCAGCAATGCTGGGATTACATTTGGAGCCCCAGGAGGCATTATTTGTGCGAAG
GCATCCATGTTACTGCAGAGCGTCCCACCGGCAGTACTGGTACGATTTCTGAGGGAGCATAGA
TCTGAATGGGCTGATTACAATATTGATGCGTATTTGGCTTCATCATTGAAGACCAGTGCGTGCT
CACTTCCTGGGTTGCGACCCATGAGATTTTCTGAGGGGCAGATGATCATGCCACTTGCTCACA
CAGTTGAGAACGAGGAGATTCTTGAAGTTGTCCGCCTTGAAGGTCAGCCTCTTACTCATGATG
AAGCTCTTCTTTCAAGGGACATCCACCTTCTCCAGCTTTGCACTGGAATAGATGAGAAATCTG
TGGGTTCCTCCTTCCAGCTTGTGTTTGCACCAATTGATGAGCATTTCCCAGATGATGCTCCATT
GATTTCTTCTGGCTTTCGTGTCATACCACTTGATGTGAAAACAGATGGTGTATCCTCTGGTAGG
ACGCTAGATTTGGCATCTAGTCTTGATGTGGGCTCTGCTGCACCCCAAGCCTCAGGGGATGCA
TCTCCAGATGACTGCAGTTTGAGATCTGTGCTGACAATCGCCTTTCAATTCCCGTATGAGATGC
ACCTTCAGGACAGCGTTGCAGCAATGGCCCGTCAATATGTTCGTAGTGTCATTTCTGCTGTGC
AAAGAGTGTCGATGGCTATATCTCCCTCCCAATCTGGTCTAAATGCTGGGCATAGGATGCTTT
CTGGCTTCCCTGAAGCTGCCACACTTGCTAGATGGGTTTGCCAGAGTTATCACTACCATCTAG
GTATGGAATTACTTAATCAATCAGATGGAGCTGGTGAAGCATTGTTGAAAATGCTCTGGCATC
ATCCAGATGCTGTTCTGTGCTGCTCCTTTAAGGAGAAACCTATGTTTACGTTTGCAAACAAGG
CAGGGCTGGAC

FIGURA 27 (continuação)

ATGTTAGAAACATCTCTTGTTGCCCTGCAAGACCTGACGCTAGACAAGATCTTCGACGAGTCA
GGAAGGAAAGCACTATTCTCAGACATCTCGAAACTAATGGAACAGGGCTACGCGTACCTGCC
GTCAGGCGTGTGCATGTCAGGAATGGGCCGCCATGTTTCTTTCGACCAAGCTGTAGCGTGGAA
GGTGCTCGGCGAGGATAGCAACATCCACTGCCTGGCGTTCTGCTTCGTCAACTGGTCCTTCGT
GTGACTGACGCCCAACCCGTCGGGTGGTTATGGCGAGCTGACCCTTTTTTGTATGGAAAATCA
TCAGCTGTGACAAAAGGCATATGTTGCATGCACTAGTTGAAGAAAC

SEQ ID N° : 205, Zea mays Zeama_HDIII RLD1 (folha enrolada 1) polipeptídeo deduzido AY501430.1

MGTMVAAVALRGGSSDSGGFDKVPGMDSGKYVRYTPEQVEVLERLYIDCPKPSSSRRQQLLREC
PILSNIEPKQIKVWFQNRRCRDKQRKESSRLQAVNRKLTAMNKLLMEENERLQKQVSQLVHENAH
MRQQLQNTSLANDTSCESNVTTPPNPIRDASNPSGLLAIAEETFTEFLSKATGTAIDWVQMPGMKP
GPDSVGIVAISHGCRGVAARACGLVNLEPTKGIEILKDRPSWFRDCRSLEVFTRFPAGNGGTIELIY
MQMYAPTTLVPARDFWTLRYTTTMEDGSLVVCERSLSGSGGGPNAASTQQFVRAEMLPSGYLVR
PCEGGGSIVHIVDHLDLEAWSVPEVLRPLYESSRVVAQKMTTVALRHLRQIAQETSGEVVYALGR
QPAVLRTFSQRLSRGFNDAISGFNDDGWSVMGGDGIEDVVIACNSTKKIRNTSNAGITFGAPGGIIC
AKASMLLQSVPPAVLVRFLREHRSEWADYNIDAYLASSLKTSACSLPGLRPMRFSEGQMIMPLAH
TVENEEILEVVRLEGQPLTHDEALLSRDIHLLQLCTGIDEKSVGSSFQLVFAPIDEHFPDDAPLISSGF
RVIPLDVKTDGVSSGRTLDLASSLDVGSAAPQASGDASPDDCSLRSVLTIAFQFPYEMHLQDSVAA
MARQYVRSVISAVQRVSMAISPSQSGLNAGHRMLSGFPEAATLARWVCQSYHYHLGMELLNQSD
GAGEALLKMLWHHPDAVLCCSFKEKPMFTFANKAGLDMLETSLVALQDLTLDKIFDESGRKALF
SDISKLMEQGYAYLPSGVCMSGMGRHVSFDQAVAWKVLGEDSNIHCLAFCFVNWSFV

SEQ ID N° : 206, Populus trichocarpa Poptr_HDIII seqüência de ácido nucleico AY919617

CTGTGTTTGATTATTTATTTTGTGGTGGAAATGGCCATGGAAGTGGCCCCTCTGCACCGAGAG
AGCAGCAGTAGTGGGAGCATAAACAAGCACCTTACTGATGATGGGAAGTATGTTCGGTACAC
AGCAGAGCAAGTGGAGGCTCTTGAAAGAGTTTATGCGGAGTGCCCTAAGCCCAGCTCTTTGCG
TAGGCAACAATTGATTAGAGAGTGCCCCATTTTAGCTAATATTGAGCCGAAGCAGATCAAAGT
CTGGTTTCAAAATCGCAGGTGTAGAGAGAAGCAGAGAAAAGAGTCCTCAAGACTCCAGACTG
TGAACAGGAAATTGACAGCAATGAATAAGCTGTTGATGGAGGAGAATGATCGCCTTCAAAAA
CAGGTGTCCCAGCTGGTGTGCGAAAATGGTTTCATGCGGCAACAACTGCAAACTGCACCAAC
AGCAACTGATGCAAGCTGTGACTCTGTGGTTGCCACTCCACAACATTCACTAAGAGATGCCAA
TAACCCTGCTGGACTCCTCTCAATAGCTGAGGAGACCTTGTCAGAGTTCCTTGCAAAGGCCAC
AGGAACTGCTCTTGAGTGGGTCCAGATGCCTGGAATGAAGCCTGGTCCGGATTCGATTGGGAT
TTTTTCCATTTCACAAAGGTGCGGTGGAGTGGCAGCAAGAGCCTGCGGTCTTGTAAGTTTAGA
GCCTAAAAAGATTGCAGAGATCCTTAAAGATCGTTCATCTTGGTTCCGTGATTGTCGGAACCT
TGAAGTTTTCACTATGTTTCCTGCTGGAAATGGTGGAACAATTGAACTAGTGTACAGTCAGAT
ATATGCTCCAACTACTCTTGCTCCAGCACGGGATATGTGGACTCTGAGATACACTACAAGTTT
AGAGAATGGCAGCCTTGTGGTCTGTGAGAGATCACTTTCTGGTTATGGTGCTGGCCCCGATGC
AGCTGCCGCTGCCCAGTTTGTGAGGGCTGAAATGCTTCCTAGTGGCTATTTGATAAGGCCATG
TGAAGGAGGGTCAATTATCCACATCGTGGATCACCTGAATCTTCAGGCATGGAGTGTGCCTGA
GGTGCTTCGACCACTTTATGAATCATCAAAAGCAGTGGCACAGAAAATGACCATTGCGGCACT
GCGTTATGTCAGGCAAGTAGCTCATGAAACCAGTGGTGAGGTGGTGTATGGTTTGGGCAGGC
AACCAGCTGTTCTGAGAACCTTTAACCAAAGATTAAGCAGAGGTTTCAATGATGCCATCAATG
GGTTTAATGATGATGGCTGGTCACTGATGAATGCGGATGGAGCTGAGGATGTAATAATTGCAG
TTAACTCAACCAAGAATTTGATTGGTGCTAATAATTCTGCTCATTCTCTTTCGTTCTTGGGAGG
GATTCTTTGTGCGAAAGCTTCCATGCTACTGCAAAATGTGCATCCTGCTGTGCTGGTCTGTTTC
CTAAGGGAGCATCACGCTGAGTGGGCTGATTTCAGTGTTGATGCCTATT

FIGURA 27 (continuação)

CTGCTGCATTGTGGAAGGCTGGTTCATATGCATATCCAGGAATGAGGCCCATGAGGTTTACTG
GGAGCCAAATCACCATGCCACTGGGCCACACAATTGAACAAGAAGACTTGCTTGAAGTTATTC
GACTTGAAGGCCATTCTTTTGCACAAGAAGATGCTTTTGTTTCACAGGACATCCATCTCCTACA
GATATGTAGTGGAATTGATGAGAATGCTGTTGGAGCCTGTTCTGAACTAGTTTTTGCTCCAATT
GATGAAACGTTTCCAGATGATGCTCCACTGCTACCTTCTGGTTTCCGCATCATTTCTCTGGAAT
CAAAAGCAAAGGATACACAGGAGGTATTGACTACAAATTGTACTCTGGATCTAACATCAAGT
CTTGAAGCGGGGCTGGCAATAAACCACACTGCAGTGGATGGCTCATCGTGTCATAGTTTGCGA
TCAGTGTTGACTATTGCCTTCCAGTTCCCTTTTGAGAGCAATCTACAGGATAATGTTGCCACCA
TGGCACGTCAGTACGTACGTAGTGTGATTTCATCTGTCCAGAGGGTTGCCATGGCCATATCTC
CATCAGGATTGAGTCCAGTTTTGGGACCAAAGCTATCTGCAGGATCTCCTGAAGCTCTAACCT
TGGCTCACTGGATCTGCCAAAGTCACAGGCAAGTTCTGCTTAAGTTTTCATCATGTTACCATTT
AGGAGCAGAATTACTGAGATCTGATTCTGTGGGTGGAGATTCTGTCCTGAAACATCTATGGCA
TCATCCAGATGCAATTTTATGTTGTTCATTGAAGTCGCTGCCAGTTTTCATCTTTGCCAATCAG
GCAGGGCTTGACATGCTGGAGACAACTCTAGTGGCTCTACAAGATATTACACTGGATAAGATA
TTCAATGAATCTGGACGCCAGGCATTGTATACAGAATTTGCAAAGTTAATGCAACAGGGATTT
GCATGCTTGCCTGCTGGAATCTGCATGTCAACAATGGGGCGTAATGTTTCATATGAGCAAGCT
GTTGCTTGGAAAGTGCTATCTGCAGAAGAGAACGCTGTTCATTGCATTGCTTTCTCTTTTGTAA
ACTGGTCTTTTTTGTGAACTCCACAGTTCATTTATCCAACTATGCAGTCGAATACGAGAACAAC
GGTATGGTAGAGATTTTTGAAAATGAAAAGAGA

SEQ ID N° : 207, polipeptídeo deduzido de Populus trichocarpa Poptr_HDIII AY919617
MAMEVAPLHRESSSSGSINKHLTDDGKYVRYTAEQVEALERVYAECPKPSSLRRQQLIRECPILAN
IEPKQIKVWFQNRRCREKQRKESSRLQTVNRKLTAMNKLLMEENDRLQKQVSQLVCENGFMRQQ
LQTAPTATDASCDSVVATPQHSLRDANNPAGLLSIAEETLSEFLAKATGTALEWVQMPGMKPGPD
SIGIFSISQRCGGVAARACGLVSLEPKKIAEILKDRSSWFRDCRNLEVFTMFPAGNGGTIELVYSQIY
APTTLAPARDMWTLRYTTSLENGSLVVCERSLSGYGAGPDAAAAAQFVRAEMLPSGYLIRPCEGG
SIIHIVDHLNLQAWSVPEVLRPLYESSKAVAQKMTIAALRYVRQVAHETSGEVVYGLGRQPAVLR
TFNQRLSRGFNDAINGFNDDGWSLMNADGAEDVIIAVNSTKNLIGANNSAHSLSFLGGILCAKAS
MLLQNVHPAVLVCFLREHHAEWADFSVDAYSAALWKAGSYAYPGMRPMRFTGSQITMPLGHTIE
QEDLLEVIRLEGHSFAQEDAFVSQDIHLLQICSGIDENAVGACSELVFAPIDETFPDDAPLLPSGFRII
SLESKAKDTQEVLTTNCTLDLTSSLEAGLAINHTAVDGSSCHSLRSVLTIAFQFPFESNLQDNVATM
ARQYVRSVISSVQRVAMAISPSGLSPVLGPKLSAGSPEALTLAHWICQSHRQVLLKFSSCYHLGAE
LLRSDSVGGDSVLKHLWHHPDAILCCSLKSLPVFIFANQAGLDMLETTLVALQDITLDKIFNESGR
QALYTEFAKLMQQGFACLPAGICMSTMGRNVSYEQAVAWKVLSAEENAVHCIAFSFVNWSFL

SEQ ID N° : 208, Medicago trunculata Medtr_HDIII seqüência de ácido nucleico unido de AC138171.17
AGGGTGTGATTGTTTAAGTTAATTTGAGATTAGGAAGATGGCTATGGCTGTTGCACAACAACA
AAGAGATAACAGCATTGAGAGACACCTTGATTCGTCTGGCAAATATGTGAGGTACACTGCTG
AACAGATTGAAGCTTTGGAAAAGGTTTATGTGGAATGCCCTAAGCCTAGTTCATTGAGAAGGC
AACAGCTGATTCGGGAGTGCCCGGTTCTGGCCAACATTGAGCCTAAGCAGATCAAGGTTTGGT
TTCAGAATAGGAGGTGTAGGGAGAAGCAGAGAAAAGAGGCTTCTCAGCTTCAGAGTGTGAAC
AGGAAACTTTCTGCGATGAATAAGCTGTTGATGGAGGAAAATGAGAGGCTGCAGAAGCAGGT
TTCACAGCTGGTGAATGAGAATGGATTTATGCGCCAGCAACTACACCCTACCCCAGCAGCTCC
AAATGCTGACGGTAGTGGCGTTGATTCCGCGGCTGCTGCTCCTATGAACTCATTGAGAGATGC
TAATAGCCCTGCTGGATTCCTATCAATTGCGGAGGAGACATTGACAGAGTTCCTTTCAAAGGC
TACAGGAACTGCTGTCGATTG

FIGURA 27 (continuação)

GGTCCAGATGCCTGGGATGAAGCCTGGTCCGGATTCGGTTGGGATATTTGCCATTTCTCAAGG
TGGCAACGGAGTGGCAGCTCGAGCCTGTGGTCTTGTTAGTTTAGAACCTACTAAGATTGTGGA
GATATTAAAAGATCGCCCAACTTGGTACCGTGATTGTCGGAGTTCAGAAGTTTTCACAATGTT
CCCAGCTGGAAATGGAGGAACAATTGAACTTGTTTACACACAGACATATGCTCCAATGACACT
GGCTTCTGCTCGCGACTTCTGGACTCTAAGATACACTACAAATTTGGAAAACGGAAGTGTTGT
GGTTTGTGAAAGGTCACTGTCTGGTACTGGTGCTGGCCCTAATGCTGCAGCCGCCTCACAGTT
TGAGAGGGCTGAAATGCTCCCTAGTGGCTATTTGATTCGACCATGTGAAGGTGGAGGATCGAT
CATCCACATTGTAGACCACCTAAACCTGCAGGCATGGAGTGTGCCAGAAGTGCTGCGGCCGAT
CTATGAATCGTCGCAAATGGTAGCTCAGAGACTGACAATTGCGGCACTTCGCTATATCAGGCA
AGTAGCTCAAGAAACAAGTGGTGACGTGGTGTATAGCATGGGTCGGCAACCTGCAGTTCTTA
GAACTTTTAGCCAACGGTTGAGCAGAGGTTTCAATGACGCTGTCAATGGATTCAATGATAATG
GTTGGTCTGTTCTGAACTGTGATGGTGCTGAGGGTGTTACTATTTCAGTAAATTCAATCAAGA
ATTTGAGTGGCACTTCTAATCCAGCAAGTTCCCTTTCACTCCTTGGAGGAATTGTCTGTGCAAA
AGCTTCTATGTTACTCCAAAACACCACTCCTGCTGTTTTAGTTCGCTTTCTGAGGGAGCATCGC
TCGGAGTGGGCTGATTTTAGTGTTGATGCCTTTTCTGCTGCATCACTTAAAGCTGGCTCCTATG
GCTATCCTGGAATGAGGTCTACAAAGTTCACCGGCAATCAAGCAATCATGCCTCTTGGACATA
CAATTGAACATGAAGAGATGCTAGAAATTATCCGCCTTGAAGGTCTTGCTCAAGATGATTCTT
TTGTTTCTAGGGATGTTCATCTCTTACAGGTGCTTCCTCTGACCCTTGTTATGTTATGTACTGGA
ATTGATGAGAATGCTGTGGGGGCTTGTTCCGAGCTCATATTTGCTCCAATTGATGACATGTTCC
CAGAAGATGCTCCCTTAGTGCCTTCTGGTTTCCGCATTGTCCTGTTGAATTCTCAACCAGGTGA
TACAAAGAACACAACAACAGCAAATCGAACCTTGGATTTGACATCTGGTCTTGAAGTAAGCC
CGGCAACAGCTCATGCTAACGGAGACGCATCGTGTCCTAACAATCGATGTGTGTTGACTGTTG
CCTTTCAGTTTCCTTTTGAGAGCGGTCTGCAGGATAATGTTGCAGCCATGGCACGTCAATATGT
CCGGCGTGTAGTTTCTGCCGTGCAGGCGGTTGCAACGGCTATATCTCCATCCAGTGTTAACAC
TTCTGGTGGAGCAAAGCTCTCCCCTGGCACTCCAGAAGCACTTACACTAGCTCAATGGATCTG
CCAGAGTTATAGTCATCATCTGGGCGCGCAACTGCTGAGATCTGATTCTCTTATTGGTGATATG
CTACTGAAACATTTGTGGCATCATCCAGATGCTATTTTATGCTGCTCTTTGAAGCAAGTGCCCG
TATTCATCTTTGCTAACCAGGCTGGCCTTGACATGTTGGAAACAACTCTAGTGGCTCTACAAG
ATATCACACTGGACAAAATATTTGATGAGTCTGCACGCAAGAATTTGATTGCATATTTTGCGA
AGTTAATGCAGCAGGGGTTTGCTTGTATGCCAGCTGGGATCTGCATGTCAACAATGGGGCGAC
ATGCTTCATATGATCAAGCCGTCGCGTGGAAAGTGCATGCTGAAGACAACAGTGTTCATTGCT
TGGCTTTCTCATTCATTAATTGGTCATTTATATGACCTATTGCTGGATTTAAACCCCCACTTTTT
TTTCCCACTTGTTGAATACAAAA

SEQ ID: 209, polipeptídeo deduzido de Medicago trunculata Medtr_HDIII AC138171.17
MAMAVAQQQRDNSIERHLDSSGKYVRYTAEQIEALEKVYVECPKPSSLRRQQLIRECPVLANIEPK
QIKVWFQNRRCREKQRKEASQLQSVNRKLSAMNKLLMEENERLQKQVSQLVNENGFMRQQLHP
TPAAPNADGSGVDSAAAAPMNSLRDANSPAGFLSIAEETLTEFLSKATGTAVDWVQMPGMKPGP
DSVGIFAISQGGNGVAARACGLVSLEPTKIVEILKDRPTWYRDCRSSEVFTMFPAGNGGTIELVYT
QTYAPMTLASARDFWTLRYTTNLENGSVVVCERSLSGTGAGPNAAAASQFERAEMLPSGYLIRPC
EGGGSIIHIVDHLNLQAWSVPEVLRPIYESSQMVAQRLTIAALRYIRQVAQETSGDVVYSMGRQPA
VLRTFSQRLSRGFNDAVNGFNDNGWSVLNCDGAEGVTISVNSIKNLSGTSNPASSLSLLGGIVCAK
ASMLLQNTTPAVLVRFLREHRSEWADFSVDAFSAASLKAGSYGYPGMRSTKFTGNQAIMPLGHTI
EHEEMLEIIRLEGLAQDDSFVSRDVHLLQVLPLTLVMLCTGIDENAVGACSELIFAPIDDMFPEDAP
LVPSGFRIVLLNSQPGDTKNTTTANRTLDLTSGLEVSPATAHANGDASCPNNRCVLTVAFQFPFES
GLQDNVAAMARQYVRRVVSAVQAVATAISPSSVNTSGGAKLSPGTPEALTLAQWICQSYSHHLG
AQLLRSDSLIGDMLLKHLWHHPDAILCCSLKQVPVFIFANQAGLDMLETTLVALQDITLDKIFDES
ARKNLIAYFAKLMQQGFACMPAGICMSTMGRHASYDQAVAWKVHAEDNSVHCLAFSFINWSFI

FIGURA 27 (continuação)

SEQ ID N° : 210, seqüência de ácido nucleico parcial de Saccharum officinarum Sacof_HDIII contig de CA125167.1 CA217027.1 CA241276.1 CA124509.1

GATTCGGTTGGTATCGTGGCCATTTCGCATGGTTGCCGTGGTGTTGCTGCCCGCGCCTGTGGTT
TGGTGAATCTAGAACCAACAAGAGTCATAGAGATCTTGAAAGATCGTCCCTCTTGGTTCCGTG
ATTGTCGAAGTCTTGAAGTGTTTACAATGTTTCCAGCTGGAAATGGGGGAACAGTTGAACTTA
TCTACATGCAGATGTATGCCCCAACAACTTTAGTCCCTGCACGTGATTTTTGGACGTTACGATA
CACGACAACAATGGAAGATGGTAGCCTTGTGGTCTGTGAGAGATCCTTGAGTGGTTCTGGAGG
CGGTCCAAATGCAGCCTCTGCACAGCAATTTGTTAGGGCTGAAATGCTTCCAAGTGGGTATTT
AGTTCGCCCATGCGAAGGTGGAGGTTCGATTGTGCATATAGTGGACCATCTAGATCTCGAGGC
ATGGAGTGTTCCTGAAGTGCTTCGACCGCTGTATGAGTCTTCCAGAGTAGTTGCTCAGAAAAT
GACTACGGTGGCACTGCGCCACCTTAGACAAATTGCTCAAGAAACAAGTGGAGAAGTAGTGT
ACGCCTTGGGAAGGCAACCTGCAGTACTACGGACCTTTAGTCAAAGACTAAGCAGGGGTTTTA
ATGATGCCATTAGTGGTTTCAATGATGATGGCTGGTCTGTAATGGGGGGAGATGGCATTGAAG
ACGTTGCTGTTGCTTGCACCTCAACCAAGAAAATTAGGAATAACAGCAATGCGGGGATTACAT
TTGGAGCCCCTGGAGGCATTATTTGTGCGAAGGCATCCATGTTACTGCAGAGTGTCCCACCGG
CAGTACTGGTCCGATTTCTGAGGGAGCATAGATCTGAATGGGCTGATTACAATATTGATGCGT
ATTTGGCTTCATCACTGAAGACCAGTGCATGCTCACTTCCGGNGTTGCAACCTATGAGATNTT
CTGGGGGGCAGATGATCATGCCACTTGCTCACACAGTGGAGAATGAGGAGATTCTTGAAGTT
ATCCGCCTTGAAGGACATCCTCTTACTCATGATGAAGCTCTTCTTTCAAGAGACATCCACCTTC
TCCAGCTTTGCACTGGAATAGATGAGAAATCTGTGGGTTCCTCCTTCCAGCTTGTGTTTGCACC
AATTGATGAGCACTTTCCAGATGATGCTCCATTGATTTCTTCTGGCTTTCGTGTCATACCACTT
GATATGAAAACAGATGGTGTATCCTCTGGCAGGACGCTAGATTTGGCATCTAGTCTTGATGTG
GGTTCTGCTGCACCCCAAGCCTCAGGGGATGCATCTCCAGATGACTGTAATTTGAGATCTGTG
CTGACAATCGCCTTTCAATTCCCTTACGAGATGCACCTTCAGGACAGTGTTGCAACTATGGCTC
GTCAATATGTTCGTAGTGTTGTTTCTGCTGTGCAAAGAGTGTCGATGGCTATATCTCCCTCCCA
ATCTGGTCTAAATGCTCAGAGGACACTTTCTGGCTTCCCTGAAACTGCCACACTTGCTAGATG
GGTTTGCCAGAGTTATCACTACCATCTAGGCGTGGAATTACTTAATCAATCAGATGAAGCTGG
CGAAGCATTGTTGAAAATGCTCTGGCATCATCCAGACGCTGTTCTGTGCTGCTCTTTTAAGGA
GAAACCTATGTTTACGTTTGCAAACAAGGCAGGGCTTGACATGTTAGAAACATCTCTTATTGC
CCTGCAAGACCTGACACTAGACAAGATTTTCGACGAGTCAGGAAGGAAAGCAATATTCTCAG
ATATCTCAAAACTAATGGAACAGGGCTACGCATACCTGCCGTCAGGTGTGTGCATGTCAGGAA
TGGGTCGCCATGTTTCTTTCGACCAAGCTGTGGCGTGGAAGGTGCTCGGTGAGGACAGCAACG
TGCACTGCCTGGCTTTCTGCTTCGTCAACTGGTCCTTCGTGTAACGCCCACCCATCCGATGGGG
TGGCGAGCCTGTCGGTCTGTGTGATGAGCCAGCCCAATTTTGTATGATGATCCTGATAAGGAA
ATCATTGACCCGGGCAAAATGCTTATGGTGCATGCACTATTTTGAGGCCCTGAAATCCCGGGT
TTTATTAAAAAAACTGGAGAATTTTCCCAGGTTTTTTCCCACTTTTCGTTGGCCCCCCCACCCA
CAGGGTGTTTGTACAATTTCTTTGGCGAGGGGCACCCCACTCCTGCAGTTTTTTTTTCCATAC

SEQ ID N° : 211, polipeptídeo parcial deduzido de Saccharum officinarum Sacof_HDIII

DSVGIVAISHGCRGVAARACGLVNLEPTRVIEILKDRPSWFRDCRSLEVFTMFPAGNGGTVELIYM
QMYAPTTLVPARDFWTLRYTTTMEDGSLVVCERSLSGSGGGPNAASAQQFVRAEMLPSGYLVRP
CEGGGSIVHIVDHLDLEAWSVPEVLRPLYESSRVVAQKMTTVALRHLRQIAQETSGEVVYALGRQ
PAVLRTFSQRLSRGFNDAISGFNDDGWSVMGGDGIEDVAVACTSTKKIRNNSNAGITFGAPGGIIC
AKASMLLQSVPPAVLVRFLREHRSEWADYNIDAYLASSLKTSACSLPXLQPMRXSGGQMIMPLAH
TVENEEILEVIRLEGHPLTHDEALLSRDIHLLQLCTGIDEKSVGSSFQLVFAPIDEHFPDDAPL

FIGURA 27 (continuação)

ISSGFRVIPLDMKTDGVSSGRTLDLASSLDVGSAAPQASGDASPDDCNLRSVLTIAFQFPYEMHLQ
DSVATMARQYVRSVVSAVQRVSMAISPSQSGLNAQRTLSGFPETATLARWVCQSYHYHLGVELL
NQSDEAGEALLKMLWHHPDAVLCCSFKEKPMFTFANKAGLDMLETSLIALQDLTLDKIFDESGRK
AIFSDISKLMEQGYAYLPSGVCMSGMGRHVSFDQAVAWKVLGEDSNVHCLAFCFVNWSFV

SEQ ID N° : 212, seqüência de ácido nucleico parcial de Triticum aestivum Triae_HDIII contig de CD905903 BM135681.1 BQ578798.1 CJ565259.1

AGGGTATTGTTGCCGTTTCACATGGTTGCCGAGGTGTCGCTGCCCGTGCCTGTGGTCTGGTGA
ATCTACAACCAACAAAGATTGTGGAGATCTTGAAAGACCGCCCATCTTGGTTCCGTGATTGTC
GGAGTCTTGAATTTTTTACGATGCTTCCAGCTGGAAACGGCGGGACCATTGAACTCGTTTACA
TGCAGATGTATGCTCCTACCACTTTAGTTCCTGCACGTGATTTTTGGACGCTGAGATACACAAC
TACAATGGAAGATGGAAGTCTCGTGGTTTGTGAGAGATCTTTGAGTGGTTCAGGAGGTGGTCC
AAGTACTGCCTCAGCACAGCAATTTGTAAGGGCTGAGATGCTTCCTAGTGGCTATTTAGTTCG
GCCATGCGACGGTGGGGGTTCAATTGTGCATATAGTGGATCATCTGGACCTTGAGGCTTGGAG
TGTTCCAGAAGTGCTTCGCCCGCTCTATGAGTCTTCTAGGGTAGTTGCTCAGAAAATGACTACT
GCGGCATTACGACACATCAGACAGATTGCTCAAGAAACTAGTGGGGAGGTTGTATATGCTCTA
GGGAGGCAACCTGCTGTTCTGCGGACATTTAGTCAGAGGCTGAGTAGAGGATTTAATGATGCT
ATAAGTGGCTTCAATGATGATGGTTGGTCTGTAATGGCTGGAGACGGCATTGAAGATGTGATT
ATTGCTTGCAACTCAAAAAAGATTAGAAGCAATAACACTGCTCCCAATGCGTTTATAGCTCCT
GGAGGTGTTATATGTGCTAAAGCATCAATGTTACTGCAGAGTGTTCCACCAGCAGTACTGGTT
CGGTTTCTGAGGGAACATCGGTCTGAATGGGCTGATTATAACTTTGATGCATATTCGGCTTCA
GCGCTGAAATCAAGCTCATGCTCACTTCCTGGGTTGCGTCCCATGAGATTTTCTGGGAGCCAG
ATCATCATGCCACTTGCTCACACAGTGGAGAATGAGGAGATTCTAGAAGTTGTTCGTCTTGAA
GGGCAAGCGCTCGATGAGGGTCTTTTATCAAGAGATATCCACCTGCTTCAGTTTTGCACTGGA
CTAGACGAGAAATCAATGGGATCCTGCTTCCAACTTGTCTTTGCACCAATTGATGAGCTTTTCC
CTGATGATGCTCCATTAATATCTTCAGGCTTTCGTGTTATACCACTGGACATGAAAACAGATG
GTGCACCCGCCGGTAGAACACTAGATTTGGCCTCTAGCCTTGAAGCTGGTTCAACTACACTGC
AAGCCTCAGGCGGTGCGGATGATTGTAATCTACGATCAGTGCTGACAATTGCCTTTCAATTCC
CTTACGAGATGCATCTCCAAGATAGTGTTGCAACTATGGCCCGGCAGTATGTCCGCAGCATTG
TCTCCGCCGTTCAGAGAGTGTCGATGGCTATCTCTCCCTCTCGATCTGGCTTGAATGCTGAACA
GAAGATAATTTCTGGCTTCCCTGAAGCTGCAACACTAGCTCGCTGGATATGCCAAAGTTACCG
GTTCCATCTGGGGGTCGAGTTATTTAGACAGGCAGATGAAGCTGGGGAATCTTTATTGAGAAT
GCTCTGGGATCATGAAGATGCTATTTTGTGCTGTTCTTTCAAGGAAAAGCCTGTATTTACGTTT
GCAAACGAGATGGGAATTAACATGTTGGAAACATCTTTCGTTGCCCTTCAAGATCTCTCGTTG
GACAAGATATTTGATGAAGCTGGTAGAAAGGCACTATACTCCGAGATCCCAAAGTTGATGGA
GCAGGGCTTTGTTTACCTGCCAGGGGGTGTGTGCTTGTCTGGGATGGGCCGCCATGTCTCATTT
GAGAGTGCTGTAGCATGGAAAGTAGTTGGTGAGGACAACAATGTGCACTGCCTCGCCTTCTGC
TTCGTCAACTGGTCTTTCGTGTGATCATCCATCACCCTTGCCTGTCCCATGCAGCTCCATTTTTG
CTCTCTCTGTAGGGGAGAACCGCCGCCACTGGAAAGTAGTTTTACGTTATCTTTAACTAGTTTG
TTTCTAGATTTGAAGAGCCAGCATCGGGAAGAATTCTGCCTAGTGAAAATTG

SEQ ID N° : 213, polipeptídeo parcial deduzido de Triticum aestivum Triae_HDIII

GIVAVSHGCRGVAARACGLVNLQPTKIVEILKDRPSWFRDCRSLEFFTMLPAGNGGTIELVYMQM
YAPTTLVPARDFWTLRYTTTMEDGSLVVCERSLSGSGGGPSTASAQQFVRAEMLPSGYLVRPCDG
GGSIVHIVDHLDLEAWSVPEVLRPLYESSRVVAQKMTTAALRHIRQIAQETSGEVVYALGRQPAV
LRTFSQRLSRGFNDAISGFNDDGWSVMAGDGIEDVIIACNSKKIRSNNTAPNAFIAPGGVICAKASM
LLQSVPPAVLVRFLREHRSEWADYNFDAYSASALKSSSCSLPGLRPMRFSGSQIIMPLAHTVENEEI
LEVVRLEGQALDEGLLSRDIHLLQFCTGLDEKSMGSCFQLVFAPIDELFPDDAPLISSGFR

FIGURA 27 (continuação)

VIPLDMKTDGAPAGRTLDLASSLEAGSTTLQASGGADDCNLRSVLTIAFQFPYEMHLQDSVATMA
RQYVRSIVSAVQRVSMAISPSRSGLNAEQKIISGFPEAATLARWICQSYRFHLGVELFRQADEAGES
LLRMLWDHEDAILCCSFKEKPVFTFANEMGINMLETSFVALQDLSLDKIFDEAGRKALYSEIPKLM
EQGFVYLPGGVCLSGMGRHVSFESAVAWKVVGEDNNVHCLAFCFVNWSFV

SEQ ID N° : 214, seqüência de ácido nucleico parcial de Hordeum vulgare Horvu_HDIII compilada de BU996988.1 BJ452342.1 BJ459891.1

GCGCTGGGGAGGCAGCCTGCTGTTCTGCGGACATTTAGTCAGAGGCTGAGTAGAGGATTTAAT
GATGCTATAAGTGGCTTCAATGATGATGGTTGGTCTGTAATGGCCGGAGATGGCATTGAAGAT
GTGATTATCGCTTGCAACTCAAAGAAGATTAGGAGCAATAACACTGCTCCCAATGCTTTTATA
GCTCCTGGAGGTGTTATATGTGCTAAAGCATCAATGTTACTGCAGAGTGTTCCACCAGCAGTA
CTGGTTCGATTTCTAAGGGAACATCGGTCTGAATGGGCTGATTATAACTTTGATGCATATTCG
GCTTCAGCGCTGAAATCAAGTTCATGCTCACTTCCTGGGTTGCGCCCTATGAGATTTTCTGGGA
GCCAGATCATCATGCCACTTGCTCACACGGTGGAGAATGAGGAGATTCTAGAAGTTGTTCGTC
TTGAAGGGCAAGCGCTCGATGAGGGTCTTTTATCAAGAGATATCCACCTGCTTCAGTTTTGCA
CTGGAATAGACGAGAAATCAATGGGGTCCTGCTTCCAACTTGTCTTTGCCCCAATTGATGAGC
TTTTCCCTGATGATGCTCCATTAATATCTTCAGGCTTCCGTGTTATACCACTGGACATGAAAAC
AGATGGTGCACCCACCGGTAGAACACTAGATTTGGCCTCTAGCCTTGAAGCTGGTTCAACTAC
ACTGCAAGCCTCAGGCAATGCGGACGATTGTAATCTACGATCAGTGCTGACAATTGCCTTTCA
ATTCCCTTACGAGATGCATCTCCAAGATAGTGTTGCAACCATGGCCCGGCAGTATGTCCGCAG
CATTGTCTCTGCCGTTCAGAGAGTGTCGATGGCTATCTCTCCCTCTCGATCTGGTTTGAATGCT
GAACAGAAGATAATTTCTGGCTTCCCTGAAGCTGCAACACTTGCTCGCTGGATATGCCAAAGC
TACCGGTTCCATCTGGGGGTCGAGTTATTTAGACAGGCAGATGAAGCTGGGGAATCTTTATTG
AGAATGCTCTGGGATCATGAAGATGCTATTTTGTGCTGTTCTTTCAAGGAAAAGCCTGTATTTA
CGTTTGCAAACGAGATGGGGATTAACATGTTGGAAACATCTTTCGTTGCCCTTCAAGACCTCT
CGTTGGACAAGATATTTGATGAAGCTGGTAGAAAGGCACTATACTCTGAGATCCCAAAGTTGA
TGGAGCAGGGCTTTGTTTACCTGCCAGGTGGTGTGTGCTTGTCTGGGATGGGCCGCCATGTCT
CCTTCGAGAATGCTATAGCATGGAAAGTAGTCGGTGAGGACAACAATGTGCACTGCCTTGCCT
TCTGCTTCGTCAACTGGTCTTTCGTGTGATCATCCTCCCAAGGCAATCCGTGCCCGTCGCACGC
AGCTCCATTTTTGCTCTCTCTCTCTGTAGGAGAGAACCGCTGCCGCTGGAAAGTAGTTTCATGC
TATCTTTAACTAGTTTGTTTGTAGATTTGAAGAGCCAGCATCCGGAAGAATTCTGCCTTGTGTA
AATCTGCTCACATTTCCACTAATTTACCCAGGCAA

SEQ ID N° : 215, polipeptídeo parcial deduzido de Hordeum vulgare Horvu_HDIII

ALGRQPAVLRTFSQRLSRGFNDAISGFNDDGWSVMAGDGIEDVIIACNSKKIRSNNTAPNAFIAPG
GVICAKASMLLQSVPPAVLVRFLREHRSEWADYNFDAYSASALKSSSCSLPGLRPMRFSGSQIIMP
LAHTVENEEILEVVRLEGQALDEGLLSRDIHLLQFCTGIDEKSMGSCFQLVFAPIDELFPDDAPLISS
GFRVIPLDMKTDGAPTGRTLDLASSLEAGSTTLQASGNADDCNLRSVLTIAFQFPYEMHLQDSVAT
MARQYVRSIVSAVQRVSMAISPSRSGLNAEQKIISGFPEAATLARWICQSYRFHLGVELFRQADEA
GESLLRMLWDHEDAILCCSFKEKPVFTFANEMGINMLETSFVALQDLSLDKIFDEAGRKALYSEIP
KLMEQGFVYLPGGVCLSGMGRHVSFENAIAWKVVGEDNNVHCLAFCFVNWSFV

SEQ ID N° : 216, seqüência de ácido nucleico parcial de Phyllostachys praecox Phypr_HDIII DQ013803

CGGCGGGTACGACAAGGCCGGGATAGACTCGGGCAAGTACGTGCGATACACGCCGGAGCAG
GTGGAGGCGCTCGAGCGGATGTATGCTGAGTGCCCCAAGCCCAGCTCCACGCGCAGGCAGCA
GCTGCTGCGCGAGTGCCCGATTCTGGCCAACATCGAGCCCAAGCAGATCAAGGTCTGGTTCCA
GAACCGAAGGTGCCGTGATAAGCAGCGGAAGGAGGCTTCCCGGCTTCAGGCCGTGAACAGAA
AATTGACCGCAA

FIGURA 27 (continuação)

TGAACAAGCTTCTCATGGAAGAGAATGATCGTCTCCAGAAGCAGGTTTCCCAGCTGGTTCATG
AGAATGCATACATGAAGCAGCAGCTGCAGAATCCATCATTGGCCAATGATACAAGCTGTGAA
TCAAATGTGACCACTCAAAACCCTCTGAAGGATGCAAGTAACCCTTCTGGACTCCTTTCAATT
GCGGAGGAGACCTTGACAGAGTTCCTCTCCAAGGCTACAGGGACTGCTGTTGATTGGGTTCAG
ATGCCTGGGATGAAGCCTGGTCCGGATTCGGTTGGTATTGTGGCCATTTCACATGGTTGCCGT
GGTGTTGCTGCCCGTGCCTGTGATTTGGTGAATCTAGAACCAACAAAAGTTGTGGAGATCTTG
AAAGACCGTCCATCTTGGTTCTGTGATCGTCAAAGCCTGGAAGTCTTCACAATGTTTCCAGCT
GGAAATGGAGGAACAATTGAACTTGTCTACACGCAGTTGTATGCTCCAACAACTTTAGTCCCT
GCACGTGATTTTTGGACTCTAAGGTACACAACCACAATGGAAGATGGCAGCCTTGTGGTCTGT
GAGAGATCCTTGAGTGGTTCAGGGGGTGGTCCAAGTGCAGCCTCTGCACAGCAATTTGTAAGA
GCCGAAATGCTTCCAAGTGGATATTTAGTTCGCCCATGCGAGGGTGGGGGCTCAATTGTGCAC
ATAGTGGACCATCTGGACCTTGAGGCTTGGAGTGTTCCTGAAGTGCTTCGGCCGCTCTACGAG
TCGTCTAGGGTAGTTGCCCAGAAAATGACTACAGCGGCACTACGGCACATCAGACAAATTGCT
CAAGAAACTAGTGGGGAGGTTGTATATGCTTTGGGGAGGCAGCCTGCCGTTCTACGGACATTT
AGTCAGAGGTTGAGTAGAGGATTTAACGATGCTATAAGTGGTTTCAATGATGATGGTTGGTCT
GTCATGGGTGGAGATGGCATTGAAGATGTAATCATTGCTTGCAACTCAAAGAAGATTAGGAA
CAATAGCACTGCTGCCAATGCTTTTGGAGCCCCTGGAGGCGTTATATGTGCTAAGGCATCCAT
GTTACTGCAGAGTGTTCCACCAGCAGTACTGGTTCGGTTTCTGAGGGAACATCGGTCTGAATG
GGCTGATTATAACTTTGATGCATATTCGGCTTTAGCACTGAAGACGAGCTCATGTTCACTTCCT
GGGTTGCGGCCTACGAGATTTTCTGGGAGCCAGATCATCATGCCACTTGCTCACACAGTGGAG
AATGAGGAGATTCTAGAAGTCATTCGTCTTGAAGGGCAGGCACTTACACATGATGAAGGTCTT
TTATCAAGAGATATCCACCTGCTTCAGCTTTGCACTGGAATAGATGAGAAATCAATGGGATCC
TGCTTCCAGCTTGTCTTTGCACCCATCGATGAGCTTTTCCCTGATGATGCCCCATTGATATCTTC
AGGCTTTCGTGTTATACCACTGGACATGAAAACAGATGGTGCACCTACTGGTAGAACATTAGA
TTTGGCCTCTAGCCTTGAAGTTGGTTCAACTACACAACAAGCTACAGGGGATGCATCTTTGGA
TGATTGTAGGAATCTGCGATCAGTGCTGACAATTGCCTTTCAATTCCCTTATGAAATTCATCTC
CAGGATAGTGTTGCAACTATGGCCCGGCAATACGTCCGTAGCGTTGTTTCTGCTGTGCAGAGA
GTGTCGATGGCTATTTCTCCCCCTCGATCTGGTGTGAATGCTGGGCAGAAGATATTTTCTGGCT
TCCCTGAAGCCGCAACACTTGCTCGCTGGATTTGCCAAAGCTACCAGTTCCATTTGGGAGTGG
AGTTACTTAGGCAGGCAGATGAAGCCGGGGAATCATTATTGAGAATGCTCTGGGATTACGAA
GATGCTATCTTGTGCTGTTCTTTCAAGGAAAAGCCTGTATTTACTTTTGCCAACGAGATGGGAC
TTAACATGTTAGAAACATCTCTCGTTGCTCTACAAGACCTCTCATTGGACAAGATATTTGATGA
AACTGGTAGAAAGGCACTACATTCGGAGATCCCAAAATTGATGGAACAGGGCTATGTTTACCT
GCCGGCTGGTGTGTGCTTGTCCGGAATGGGCCGCCATGTCTCTTTCGAGCAAGCTGTAGCATG
GAAAGTACTTGGTGAGGACAACAATGTGCACTGCCTGGCCTTCTGCTTCGTCAACTGGTCTTT
CGTGTGATCCATCCGACGCAATCCATGTCCGTTGTGAGCAGCACCATTTTCGCATTCTCTGTAG
AAGAGAACCATCGTCGCTGTTAGTTAATGCTGTCTTTAACTAGTTTCTAGATTTGGAAAAGCC
AGTCTGCAAAAAAAAAAAAAAAAAAA

SEQ ID N° : 217, polipeptídeo parcial deduzido de Phyllostachys praecox Phypr_HDIII
GGYDKAGIDSGKYVRYTPEQVEALERMYAECPKPSSTRRQQLLRECPILANIEPKQIKVWFQNRRC
RDKQRKEASRLQAVNRKLTAMNKLLMEENDRLQKQVSQLVHENAYMKQQLQNPSLANDTSCES
NVTTQNPLKDASNPSGLLSIAEETLTEFLSKATGTAVDWVQMPGMKPGPDSVGIVAISHGCRGVA
ARACDLVNLEPTKVVEILKDRPSWFCDRQSLEVFTMFPAGNGGTIELVYTQLYAPTTLVPARDFW
TLRYTTTMEDGSLVVCERSLSGSGGGPSAASAQQFVRAEMLPSGYLVRPCEGGGSIVHIVDHLDLE
AWSVPEVLRPLYESSRVVAQKMTTAALRHIRQIAQETSGEVVYALGRQPAVLRTFSQRLSRGFND
AISGFNDDGWSVMGGDGIEDVIIACNSKKIRNNSTAANAFGAPGGVICAKASMLLQSVPPAVLVRF
LREHRSEWADYNFDAYSALALKTSSCSLPGLRPTRFSGSQIIMPLAHTVENEEILEVIRLEGQA

FIGURA 27 (continuação)

LTHDEGLLSRDIHLLQLCTGIDEKSMGSCFQLVFAPIDELFPDDAPLISSGFRVIPLDMKTDGAPTGR
TLDLASSLEVGSTTQQATGDASLDDCRNLRSVLTIAFQFPYEIHLQDSVATMARQYVRSVVSAVQ
RVSMAISPPRSGVNAGQKIFSGFPEAATLARWICQSYQFHLGVELLRQADEAGESLLRMLWDYED
AILCCSFKEKPVFTFANEMGLNMLETSLVALQDLSLDKIFDETGRKALHSEIPKLMEQGYVYLPAG
VCLSGMGRHVSFEQAVAWKVLGEDNNVHCLAFCFVNWSFV

SEQ ID N° : 218 promotor GOS2 de Oryza sativa
AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAA
TATAAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCA
CCTACTTTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGT
AATTAAGTGGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGT
TAAATTATTCGAGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTTCTAGCT
GAACTCAATGGGTAAAGAGAGAGATTTTTTTTAAAAAAAATAGAATGAAGATATTCTGAACGT
ATTGGCAAAGATTTAAACATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACA
TCATTAAGGACATGTCTTACTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATT
TTTATTATTTATCTTTTTTCGATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCA
TGTGTGAGTGCACCTCCTCAATACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTAT
TTAATACATTTAGGTAGCAATATCTGAATTCAAGCACTCCACCATCACCAGACCACTTTTAAT
AATATCTAAAATACAAAAAATAATTTTACAGAATAGCATGAAAAGTATGAAACGAACTATTT
AGGTTTTTCACATACAAAAAAAAAAAGAATTTTGCTCGTGCGCGAGCGCCAATCTCCCATATT
GGGCACACAGGCAACAACAGAGTGGCTGCCCACAGAACAACCCACAAAAAACGATGATCTA
ACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAGGCTTTGCGGCCAGGAGAGAGGAG
GAGAGGCAAAGAAAACCAAGCATCCTCCTCCTCCCATCTATAAATTCCTCCCCCCTTTTCCCCT
CTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAAGGACACGCGACTAGCAG
AAGCCGAGCGACCGCCTTCTTCGATCCATATCTTCCGGTCGAGTTCTTGGTCGATCTCTTCCCT
CCTCCACCTCCTCCTCACAGGGTATGTGCCCTTCGGTTGTTCTTGGATTTATTGTTCTAGGTTGT
GTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCCTGTTCTTGGATTT
GGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTCGGTTTGATTAGTAGTATGGTTTTCAA
TCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTAGGGTACGGAATCTTGCGATTTTGTGAGT
ACCTTTTGTTTGAGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAAGTACGGTT
GTTTGGTCCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTCCCT
ATTGAACAAAAATAATCCAACTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTT
AAGCCTGTCCAAAATTTCGCAGCTGGCTTGTTTAGATACAGTAGTCCCCATCACGAAATTCAT
GGAAACAGTTATAATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAG
AATTTTTTTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTAC
AAATAATGCTTTTATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTT
AGTCAGGAGAAGAACTTATCCGATTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAG
CATAAGCAGTATTCATTTGGATTATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGA
AAGTCTGGCATGAACTGTCCTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCC
TCTTGTATCTACCTGTAGAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAA
TTTATGAAGCTGTAATCGGGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTT
CTTGGTGTAGCTTGCCACTTTCACCAGCAAAGTTC

SEQ ID N° : 219, prm03263 (Orysa_REV1 parcial CTR)
GGGGACAAGTTTGTACAAAAAAGCAGGCTTGTGCTAAGGCATCCATGCTAC

SEQ ID N° : 220 prm03264 (Orysa_REV1 parcial CTR)
GGGGACCACTTTGTACAAGAAAGCTGGGTGCACCTTCCATGCTACAGCTTG

FIGURA 27 (continuação)

SEQ ID N° : 221, prm01983 (Orysa_REV1)
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACAATGGCGGCGGCGGTGG

SEQ ID N° : 222, prm01984 (Orysa_REV1)
GGGGACCACTTTGTACAAGAAAGCTGGGTGGATTTTGGGTCACACGAAGGACCA

SEQ ID N° : 223 Oryza sativa – variante de SEQ ID N° : 197
TGTGCTAAGGCATCCATGCTACTGCAGAGTGTCCCTCCTGCAGTTTTGGTTCGATTTTTGAGGG
AACATCGTTCTGAATGGGCGGATTATAACTTCGATGCATATTCAGCTTCATCTCTGAAGACAA
GCTCATGTTCACTTCCTGGGTTGCGGCCTATGAGATTTTCTGGGAGCCAGATCATTATGCCACT
TGCTCACACGGTGGAGAATGAAGAGATTTTAGAAGTTGTCCGTCTTGAAGGACAAGCACTTAC
ACATGATGATGGTCTTATGTCTAGAGATATTCACCTGCTTCAGCTTTGCACTGGAATAGATGA
GAAATCAATGGGATCCTGCTTCCAGCTTGTCTTTGCACCAATCGATGAGCTTTTCCCTGATGAT
GCTCCGTTAATATCTTCAGGCTTTCGTGTTATACCGCTGGACATGAAAACAGATGGTACACCT
GCTGGTAGAACATTAGATTTGGCATCTAGCCTTGAGGTTGGTTCAACTGCACAGCCCACAGGG
GATGCATCTATGGATGACTGTAATCTACGATCAGTGCTGACAATTGCCTTTCAGTTCCCTTATG
AAATGCATCTCCAAGACAGCGTTGCAACTATGGCCCGGCAATATGTCCGCAGTATTGTTTCCT
CTGTTCAGAGAGTATCAATGGCTGTTTCTCCTTCTCGGTCTGGCTTGAATGCTGGGCAGAAGAT
AATTTCAGGCTTCCCTGAAGCCCCAACGCTAGCTCGTTGGATTTGCCAAAGCTACCAGTTCCA
TTTGGGGGTGGAGTTACTTAGGCAGGCAGATGATGCTGGGGAAGCACTATTGAAAATGCTAT
GGGATTACGAAGACGCTATTTTGTGCTGTTCTTTCAAGGAAAAGCCTGTATTTACTTTTGCCAA
CGAGATGGGACTAAACATGCTAGAAACATCTCTCGTCGCTCTCCAAGATCTCTCACTGGACAA
GATATTTGATGAAGCCGGTAGGAAGGCCCTATACAACGAGATCCCGAAATTGATGGAACAGG
GTTACGTGTACCTGCCTGGTGGAGTGTGCTTGTCCGGGATGGGGCGCCATGTTTCTTTCGAGC
AAGCTGTAGCATGGAAGGTGC

SEQ ID N° : 224, seqüência de ácido nucleico de comprimento total Brassica rapa Revoluta unido de AC189324.1
ATGGAGATGGCGGTGGCGAATCACCGAGAGAGAAGCAGTGATAGCATGAACAGACATTTAGA
CAGCAGCGGCAAGTACGTCAGGTACACCGCTGAGCAAGTCGAGGCCCTCGAGCGTGTCTACG
CCGAGTGTCCCAAGCCTAGCTCGCTCCGGAGACAGCAGTTGATCCGTGAATGTTCTATCTTGG
CCAACATCGAGCCTAAGCAGATCAAAGTCTGGTTCCAAAACCGGAGGTGTCGAGATAAGCAG
AGGAAAGAGGCGTCGAGGCTCCAGAGCGTAAACAGGAAGCTTTCGGCTATGAACAAGCTGTT
GATGGAGGAGAACGATAGGTTGCAGAAGCAAGTTTCTCAGCTCGTCTGTGAAAATGGATACA
TGAAGCAGCAGCTTACTACTACTGTTGTGAATGATCCAAGCTGTGATTCTGTGGTGACAACTC
CTCAGCATTCGCTTAGAGACGCTAATAGTCCTGCTGGGCTGCTCTCGATCGCAGAGGAGACAT
TGGCAGAGTTCCTATCCAAGGCTACAGGAACTGCTGTTGATTGGGTTCAGATGCCTGGGATGA
AGCCTGGTCCGGATTCGGTTGGGATCTTTGCTATATCGCAAAGATGCAGTGGGGTGGCAGCTC
GAGCCTGTGGTCTTGTTAGTTTAGAGCCTGTCAAGATTGCAGAGATACTCAAAGATAGGCCAT
CTTGGTTCCGTGACTGTAGGAGCCTTGAAGTCTTCACTATGTTCCCGGCTGGTAACGGCGGCA
CCATTGAGCTCGTCTATATGCAGACATATGCACCAACGACTCTGGCTCCTGCCCGCGATTTCTG
GACCCTGAGATACACAACGAGCCTAGACAATGGCAGTTTTGTGGTTTGTGAGAGGTCACTCTC
TGGTTCTGGTGCTGGTCCTAACGCTGCATCAGCTTCTCAGTTTGTAAGAGCAGAAATGCTTTCT
AGTGGGTATCTAATAAGGCCTTGTGATGGTGGCGGTTCCATTATTCACATTGTCGATCACCTTA
ATCTTGAGGCTTGGAGTGTTCCCGATGTGCTTCGACCCCTTTACGAGTCATCTAAAGTCGTAGC
ACAGAAAATGACCATTTCAGCATTGAGGTATATCAGGCAACTAGCTCAAGAGTCTAATGGTG
AATTAGTGTATGGATTAGGAAGACAGCCTGCTGTTTTAAGAACATTTAGCCAAAGATTAAGCA
GGGGTTTTAATGATGCGGTTAATGGGTTTGGTGACGACGGATGGTCTACAATGCATTGTGATG
GGGCTGAAGACATAATCGTTGCTATTAACTCTACAAAGCATTTGAATAATATGTCTAATTCTCT
TTCCTTCCTTGGAGGTGTCCTTTGTGCCAAGGCTTCTATGCTT

FIGURA 27 (continuação)

CTCCAAAATGTTCCTCCTGCGGTTTTGATCCGGTTTCTAAGAGAGCATCGGTCCGAGTGGGCT
GACTTCAATGTTGATGCATATTCTGCTGCAACGCTTAAAGCCGGTTCTTTTGCTTATCCGGGTA
TGAGGCCTACAAGGTTCACCGGGAGCCAAATCATAATGCCTCTAGGACATACCATCGAACAC
GAAGAAATGCTTGAAGTTGTTAGACTAGAAGGTCATTCTCTTGCACAAGAAGATGCGTTCATG
TCCCGTGATGTCCATCTTCTTCAGGTATGTACTGGGATTGATGAGAACGCTGTTGGAGCATGTT
CAGAACTAATATTTGCTCCCATCAATGAGATGTTCCCGGATGATGCTCCACTTGTTCCTTCTGG
ATTCAGAGTCATACCCGTTGATGCTAAAACGGGAGATGCGCAAGATCTGTTGACCGCTAACCA
CCGGACGCTAGACTTGACTTCTAGCCTTGAGGTTGGTCCAACACCAGAGAACGCTTCTGGAAA
CTCTTCTTCTAGCTCAAGCTCAAGATGTATCCTCACCATCGCGTTTCAGTTCCCTTTTGAAAAC
AACTTGCAAGACAATGTTGCTGGTATGGCTTGTCAGTACGTGCGGAGCGTGATCTCATCGGTT
CAACGTGTTGCAATGGCCATCTCACCCTCTGGGATAAGCCCAAGTCTGGGATCCAAACTGTCC
CCGGGGTCTCCTGAAGCTGTTACACTTGCCCAATGGATCTCACAAAGTTACAGTCACCACTTA
GGCTCGGAGTTGCTGACGGTTGACTCGCTTGGAAGCAACGACTCGGTATTGAAACTTCTATGG
GATCACCAAGATGCCATTTTGTGTTGCTCTTTAAAGCCTCAGCCAGTGTTTATGTTCGCGAACC
AAGCTGGTTTAGACATGTTAGAGACGACGCTTGTAGCCTTACAAGACATTACACTCGAAAAGA
TCTTTGATGAATCTGGTCGAAAGGCTCTCTGCTCCGATTTCGCCAAGCTAATGCAACAGGGAT
ATGCATGCTTGCCTTCAGGAATATGTTTGTCTACGATGGGAAGACATGTGACTTATGAACAAG
CTGTTGCTTGGAAAGTGTTTGCTCCTGCTGCTGCATCATCCGAAAACAACGATGGCAATGATA
ATACTTTGCACTGTCTTGCTTTCTCCTTTGTAAACTGGTCGTTTGTGTAA

SEQ ID N° : 225, seqüência de polipeptídeo de comprimento total de Brassica rapa Revoluta
MEMAVANHRERSSDSMNRHLDSSGKYVRYTAEQVEALERVYAECPKPSSLRRQQLIRECSILANI
EPKQIKVWFQNRRCRDKQRKEASRLQSVNRKLSAMNKLLMEENDRLQKQVSQLVCENGYMKQQ
LTTTVVNDPSCDSVVTTPQHSLRDANSPAGLLSIAEETLAEFLSKATGTAVDWVQMPGMKPGPDS
VGIFAISQRCSGVAARACGLVSLEPVKIAEILKDRPSWFRDCRSLEVFTMFPAGNGGTIELVYMQT
YAPTTLAPARDFWTLRYTTSLDNGSFVVCERSLSGSGAGPNAASASQFVRAEMLSSGYLIRPCDGG
GSIIHIVDHLNLEAWSVPDVLRPLYESSKVVAQKMTISALRYIRQLAQESNGELVYGLGRQPAVLR
TFSQRLSRGFNDAVNGFGDDGWSTMHCDGAEDIIVAINSTKHLNNMSNSLSFLGGVLCAKASMLL
QNVPPAVLIRFLREHRSEWADFNVDAYSAATLKAGSFAYPGMRPTRFTGSQIIMPLGHTIEHEEML
EVVRLEGHSLAQEDAFMSRDVHLLQVCTGIDENAVGACSELIFAPINEMFPDDAPLVPSGFRVIPV
DAKTGDAQDLLTANHRTLDLTSSLEVGPTPENASGNSSSSSSSRCILTIAFQFPFENNLQDNVAGMA
CQYVRSVISSVQRVAMAISPSGISPSLGSKLSPGSPEAVTLAQWISQSYSHHLGSELLTVDSLGSNDS
VLKLLWDHQDAILCCSLKPQPVFMFANQAGLDMLETTLVALQDITLEKIFDESGRKALCSDFAKL
MQQGYACLPSGICLSTMGRHVTYEQAVAWKVFAPAAASSENNDGNDNTLHCLAFSFVNWSFV

SEQ ID N° : 226, Ginkgo biloba Revoluta seqüência de ácido nucleico de comprimento total DQ385525
ATGGCAGTAATAGCACAGAAAGACGTTAAAGGTGCCATGGATACGAGCAAGTATGTTCGCTA
TACTTCTGAACAAGTCGAAGCTCTTGAACGCGTTTACAGCGAGTGTCCGAAGCCTAGTTCTCT
TCGTCGGCAGCAGCTTATTAGAGAGTGTCCAATACTTTCTAATATTGAGCCCAAGCAAATCAA
AGTTTGGTTCCAAAATCGCAGGTGTCGAGAGAAACAGAGGAAGGAGGCATCACGCCTTCAAA
CTGTTAACAGGAAGCTTACGGCAATGAACAAATTGCTGATGGAGGAAAATGATCGCCTTCAA
AAGCAGGTTTCACAGTTGGTGTACGAGAACGGTTATATGAGACAGCAGCTACAGAATGCATC
TGTAGCAACAACAGACACAAGTTGTGAGTCTGTTGTGACTAGTGGTCAGCACCAGCATAATCC
AACACCTCAGCATCCTCCAAGGGATGCTAGCCCCGCTGGACTCCTGTCTATCGCAGAGGAGAC
CTTGGCAGAGTTCCTTTCAAAGGCTACAGGAACTGCTGTTGACTGGGTCCAGATGCCTGGGAT
GAAGCCTGGTCCGGATTCGATT

FIGURA 27 (continuação)

GGTATTGTGGCTATTTCACACAGTTGTAGTGGAGTAGCTGCACGAGCTTGCGGTCTTGTGGGT
TTAGAACCTACAAAGATTGCAGAAATTCTCAAAGATCGCCCATCTTGGCTTCGTGATTGCCGT
TGCCTTGATGTGTTGACTCCCTTTCCTACTGGAAATGGAGGGACAATTGAGCTTTTATACATGC
AGACATATGCTCCCACAACTTTAGCTTCTGCTAGAGATTTTTGGACTCTGAGATACACCACAG
TATTGGAAGATGGTAGTCTTGTGGTTTGTGAAAGGTCCTTAAGTGGTGCCCAGGGTGGTCCAA
GCATAGCTCCTGCACAGCACTTTGTAAGAGCAGAAATGCTTCCCAGTGGGTATTTGATAAGAC
CTTGTGAAGGTGGAGGTTCTATTATCCATATTGTTGACCATATGGATTTGGAGCCATGGAGTG
TGCCTGAGGTGTTACGTCCACTATATGAGTCATCTACTGTACTCGCCCAAAAGATGACTATTG
CAGCCTTGCGCCGCATACGCCAAATAGCACAGGAGGTTACAGGTGAAGTGGTCTTTGGTTGGG
GTAGGCAGCCAGCTGTTCTGCGGACATTTAGCCAGAGACTGAGCAGGGGTTTCAATGAGGCT
GTGAATGGCTTCACAGATGATGGATGGTCTTTGATGGGTAATGATGGAATGGAGGACGTGACT
ATTGCAATAAATTCATCTCCAAGCAAACTCCTAGGTTCTCAGGTTAATTCTTCTAATGGGCTTA
CAGCTGTTGGTGGGGGTATACTGTGTGCAAAGGCATCCATGCTTTTACAGAATGTGCCTCCAG
CATTACTTGTTCGCTTCTTGCGGGAGCATCGATCGGAGTGGGCAGATTGTAACATTGACGCCT
ATTCTGCAGCTGCATTAAAGGCAAGTCCTTATAGTGTCCCAGGTTCACGAGCAGGAGGCTTTT
CAGGAAGTCAAGTTATCCTCCCCCTGGCACATACCGTGGAACATGAAGAGTTCTTGGAGGTCA
TTAAGCTGGAGGGCCATGGCCTCACACAGGAAGAAGCTGTGCTATCTAGGGATATGTTTCTGT
TGCAGCTTTGCAGTGGAATTGATGAAAATGCAGCTGGTGCTTGTGCCCAACTTGTGTTTGCAC
CAATTGATGAATCATTTGCTGATGATGCTCCTTTGTTGCCGTCTGGTTTCCGAGTTATTCCTCTG
GACTCTAGAACAGATGGTACTAGTGGTCCCAATCGGACATTGGATCTGGCTTCAGCTCTTGAG
GTTGGATCAGCTGGAACTAGAACTTCTGGCGATTCTGGAGCCAACTCGTTCAATCTAAGATCT
GTGTTAACTATTGCATTCCAATTCACTTATGAGAATCATTTGCGAGAAAATGTGGCATCCATG
GCCCGTCAATATGTACGCAGTGTTGTAGCATCTGTGCAGAGGGTTGCCATGGCATTAGCCCCC
TCTCGACTGAGTTCACATGTGGGGCCAAGGCTACCACCTGGCACTCCAGAAGCACTTACTCTT
GCTCGGTGGATTTGTCAAAGCTACAGATTCCACCTAGGTGTAGAGCTGCTTCGCGCTGATTGT
GAGGCCAGTGAATCCGTGCTGAAACTACTCTGGCACCATTCAGATGCAATTGTGTGCTGTTCT
TTGAAGTCGCTACCAGTTTTCACATTTGCAAATCAAGCAGGCCTTGACATGCTGGAGACAACC
CTGGTTGCCTTGCAAGATATATCATTGGACAAAATACTCGATGAAAATGGACGCAAAAGTTTA
TGCTCAGATTTTGCACAAATTATGCAACAGGGCTATGCTTATCTGCCTGCGGGGATATGTGTCT
CTAGTATGGGCAGGCCTGTTTCATATGACCGGGCTGTTGCTTGGAAGGTCCTAAATGATGCAG
ACAGCACCCACTGCATGGTTTTCATGTTTATGAATTGGTCTTTCATGTGA

SEQ ID N° : 227, seqüência de polipeptídeo de comprimento total Ginkgo biloba Revoluta
MAVIAQKDVKGAMDTSKYVRYTSEQVEALERVYSECPKPSSLRRQQLIRECPILSNIEPKQIKVWF
QNRRCREKQRKEASRLQTVNRKLTAMNKLLMEENDRLQKQVSQLVYENGYMRQQLQNASVAT
TDTSCESVVTSGQHQHNPTPQHPPRDASPAGLLSIAEETLAEFLSKATGTAVDWVQMPGMKPGPD
SIGIVAISHSCSGVAARACGLVGLEPTKIAEILKDRPSWLRDCRCLDVLTPFPTGNGGTIELLYMQT
YAPTTLASARDFWTLRYTTVLEDGSLVVCERSLSGAQGGPSIAPAQHFVRAEMLPSGYLIRPCEGG
GSIIHIVDHMDLEPWSVPEVLRPLYESSTVLAQKMTIAALRRIRQIAQEVTGEVVFGWGRQPAVLR
TFSQRLSRGFNEAVNGFTDDGWSLMGNDGMEDVTIAINSSPSKLLGSQVNSSNGLTAVGGGILCA
KASMLLQNVPPALLVRFLREHRSEWADCNIDAYSAAALKASPYSVPGSRAGGFSGSQVILPLAHT
VEHEEFLEVIKLEGHGLTQEEAVLSRDMFLLQLCSGIDENAAGACAQLVFAPIDESFADDAPLLPSG
FRVIPLDSRTDGTSGPNRTLDLASALEVGSAGTRTSGDSGANSFNLRSVLTIAFQFTYENHLRENVA
SMARQYVRSVVASVQRVAMALAPSRLSSHVGPRLPPGTPEALTLARWICQSYRFHLGVELLRADC
EASESVLKLLWHHSDAIVCCSLKSLPVFTFANQAGLDMLETTLVALQDISLDKILDENGRKSLCSD
FAQIMQQGYAYLPAGICVSSMGRPVSYDRAVAWKVLNDADSTHCMVFMFMNWSFM

FIGURA 27 (continuação)

SEQ ID N° : 228, seqüência de ácido nucleico de ciomprimento total Gossypium barbadense Revoluta AY966446.1

ATGGCTATGGCGATGACACATCATCACAACAGGGAAAGCAGCATAGACAAGCATTTAGATAC
AGGCAAGTATGTTAGGTACACAGCTGAACAAGTTGAAGCTTTGGAACGTGTTTATGCTGAATG
TCCAAAACCAAGTTCTTTGCGTAGGCAACAGTTAATAAGGGAATGTCCTATTCTTTCTAACATT
GAGCCTAAACAGTTCAAAGCTTTGTTCCAAAATCGCAGGTGTAGAGAGAAACAAAGGAAAGA
AGCTTCAAGGCTTCAAACAGTAAATAGAAAATTAACAGCAATGAATAAGCTGTTAATGGAAG
AGAATGATAGGTTGCAAAAACAGGTTTCTCAGTTGGTTTGTGAGAATGGGTACATGAGACAG
CAATTGCATACTGTAAATGCATCGGCGACTGATGCGAGCTGTGACTCTGCGGTAACCACTCCT
CAGCATTCACTGAGAAATGCTAATAACCCTGCTGGACTCCTCTCTATCGCGGAGGAGACCTTG
GCAGAGTTCCTTTCTAAGGCTACGGGAACTGCTGTCAATTGGGTCCAGATGCCTGGGATGAAG
CCTGGTCCAGATTCAGTTGGGATATTTGCCACTTCCCAAAGTTGTAGTGGAATGGCAGCTCGA
GCTTGTGGTCTTGTAAGTTTAGAACCTACAAAGATTGCAGAGATTCTTAAAGATCGTCCATCTT
GGTTCCGGGACTGTCGGAAGCTTGAAGTTTTCACCATGTTTCCAGCTGGCAATGGTGGTACGA
TTGAGCTTGTATACACACAGATGTTTGCTCCTACTACATTGGCTCCTGCAAGGGACTTTTGGAC
TCTAAGGTACACTACAACTTTAGAGAACGGCAGTCTCGTGGTGTGTGAGAGGTCTCTTTCGGG
TTCCGGTGCTGGCCCGAGTGTGGCTTCCGCAGCTCAATTTGTGAGAGCTGAAGTGCTCCCTAG
TGGCTATTTGATTAGGCCTTGTGAGGGTGGAGGGTCGATTATCCATATTGTTGACCACTTGAAT
CTTGAGGCATGGAGTGTGCCGGAGGTCTTGCGCCCCCTTTATGAATCATCCAGAGTAATTGCT
CAGAAAATGACTATTCCGGCGCTGCGCTATGTTAGGCAGATTGCTCAAGAGACAAGCGGCGA
GGTGGTTTACAGTTTGGGCAGGCAGCCTGCTGTGCTTAGAACATTTAGCCAAAGATTAAGCAG
GGGCTTCAATGAGGCAATAAATGGATTCAACGAAGATGGCTGGTCGATAATGAATTGTGATG
GTACAGAGGATGTGATAATTGCTATTAATTCCGGCAAGAGCTTGAGCAACAGTTCGAATTTGA
CCACCGGTCTTTCATTCCTCGGGGGCGTTCTATGTGCAAAGGCATCCATGCTGCTTCAAAATGT
TCCTCCTGCTGTTCTTGTCCGATTCTTGAGGGAACATCGTTTGGAATGGGCTGATTTCAATGTC
GATGCTTATTCAGCTGCATCATTGAAGGCCGGAACATACACTTATCCTGGAATGAGACCTACA
AGTTTTACTGGGAGTCAGATCATCATGCCACTCGGCCAGACGGTCGAACATGAAGAGCTGCTC
GAAGTTATAAGACTCGAAGGCCAGTCTCTTACGCAAGAAGACGCGCTTCTATCAAGGGACATT
CATCTATTACAGATATGTAGTGGAATTGATGACAATGCGGTTGGGGCCTGTTCGGAGCTTGTG
TTTGCACCAATAGATGAGATGTTTCCGGATGATGCTGCCTTACTACCTTCAGGATTCCGCATTA
TCCCGTTGGAGTCAAAACCAGATTCGTTGGCTACAAATCGGACTTTGGATCTTACTTCGAGTCT
CGAAGTGGGGCCTGCAACAAGCCAAGCTGCCGGAGATTCTCCGAGTCAGAATGCACGATCGG
TGTTGACAATTGCCTTCCAGTTTCCCTTCGATACAAATCTTCGGGATAATGTTGCGACCATGGC
ACGCCAGTATGTCCGTAGCGTCATTTCTTCTGTCCAGAGGRTTGCTATGGCCATATCTCCATGY
GGATCGAGCCCAACTATAGGACCAAAGCCATCTCCAGGTTCTCCCGAAGCGCTTACGTTGGCT
CATTGGATCTGCCAGAGCTACAGTTTCCATTTGGGGGAAGAGTTGTTGAAATCCGAATCACTT
GGTGGCGACTCAGTATTGAAGAATCTTTGGCAACATCAGGATGCAATATTGTGTTGTTCGTTG
AAGTCTGTACCGGTTTTCATCTTCGCAAATCAGGCCGGTCTAGACATGCTTGAGACAACTCTA
GTGGATCTACCAGACATCACACTGGACAAAATATTCGATGAGTCGGGACGGAAGGCATTGTG
CTCTGATTTCACCAAGTTGATGCAGCAGGGATTCACTCACTTGCTGGCCGGAGTTTGCATGTC
GACAATGGGCCGCCACGTCTCGTATGAACAAGCTGTTGCTTGGAAAGTACTTGCAGCTGATGC
AAACACTGTCCACTGCTTGGCATTCTCTTTTATAAACTGGTCTTTTGTGTGA

SEQ ID N° : 229, seqüência de polipeptídeo de comprimento total de Gossypium barbadense Revoluta

MAMAMTHHHNRESSIDKHLDTGKYVRYTAEQVEALERVYAECPKPSSLRRQQLIRECPILSNIEPK
QFKALFQNRRCREKQRKEASRLQTVNRKLTAMNKLLMEENDRLQKQVSQLVCENGYMRQQLHT
VNASATDASCDSAVTTPQHSLRNANNPAGLLSIAEETLAEFLSKATGTAVNWVQMPGMKPGPDS
VGIFATSQSCSGMAARACGLVSLEPTKIAEILKDRPSWFRDCRKLEVFTMFPAGNGGTIELVYTQM
F

FIGURA 27 (continuação)

APTTLAPARDFWTLRYTTTLENGSLVVCERSLSGSGAGPSVASAAQFVRAEVLPSGYLIRPCEGGG
SIIHIVDHLNLEAWSVPEVLRPLYESSRVIAQKMTIPALRYVRQIAQETSGEVVYSLGRQPAVLRTFS
QRLSRGFNEAINGFNEDGWSIMNCDGTEDVIIAINSGKSLSNSSNLTTGLSFLGGVLCAKASMLLQ
NVPPAVLVRFLREHRLEWADFNVDAYSAASLKAGTYTYPGMRPTSFTGSQIIMPLGQTVEHEELL
EVIRLEGQSLTQEDALLSRDIHLLQICSGIDDNAVGACSELVFAPIDEMFPDDAALLPSGFRIIPLESK
PDSLATNRTLDLTSSLEVGPATSQAAGDSPSQNARSVLTIAFQFPFDTNLRDNVATMARQYVRSVI
SSVQRXAMAISPCGSSPTIGPKPSPGSPEALTLAHWICQSYSFHLGEELLKSESLGGDSVLKNLWQH
QDAILCCSLKSVPVFIFANQAGLDMLETTLVDLPDITLDKIFDESGRKALCSDFTKLMQQGFTHLLA
GVCMSTMGRHVSYEQAVAWKVLAADANTVHCLAFSFINWSFV

SEQ ID N° : 230, seqüência de ácido nucleico de ciomprimento total Lycopersicon esculentum Revoluta BT013577

ATGGCTATGGTGGCTCAACAGCATAGGGAGAGTAGTAGTGGTAGTATTACTAAACATCTTGAT
AGTAGTGGAAAGTATGTTCGATATACGGCTGAGCAAGTTGAGGCTTTGGAGAGGGTTTATGCA
GAGTGTCCTAAGCCTAGCTCGTTGCGTCGACAGCAATTGATCCGTGAATGTCATATTCTGTCG
AATATCGAGCCTAAGCAGATCAAAGTTTGGTTTCAGAACAGAAGGTGTCGAGAGAAGCAAAG
GAAAGAGTCTTCTCGATTGCAGACTGTGAACAGAAAGTTGTCTGCGATGAATAAACTGTTGAT
GGAGGAGAATGACCGCTTGCAGAAACAAGTCTCGCAGCTTGTATGTGAAAATGGCTATATGC
GGCAACAATTGCAAAGTGTATCGGCGGCCACTACTGATGTAAGTTGTGAATCAGTGGTAACCA
CTCCTCAGCATTCCCTTAGAGATGCTAACAACCCTGCTGGACTACTACCAATTGCAGAAGAAA
CCTTGGCAGAGTTCCTTTCTAAGGCTACAGGAACTGCTGTCGATTGGGTCCCGATGCCTGGGA
TGAAGCCTGGTCCGGATTCAGTTGGGATTTTTGCCATCTCACACAGTTGCAGTGGAGTGGCAG
CCCGAGCATGTGGTCTTGTTAGTTTAGAGCCAACAAAGATTGCTGATATCCTCAAAGATCGAC
CTTCTTGGTTCCGCGACTGCCGGAATGTTGAAGTTATCACAATGTTTCCTGCTGGAAATGGTGG
TACAGTTGAGCTTTTGTATACCCAGATATATGCTCCCACAACTCTGGCTCCCGCGCGTGATTTT
TGGACGCTGAGATACACAACAACCCTAGACAATGGTAGTCTCGTGGTTTGTGAAAGATCCCTA
TCTGGTAATGGGCCTGGCCCAAATCCTACTGCTGCTTCCCAGTTTGTAAGAGCTCAAATGCTTC
CATCTGGATATCTGATCCGACCGTGTGATGGTGGAGGATCAATCATACATATTGTTGATCACC
TGAATCTTGAGGCATGGAGTGCCCCTGAGATTTTGCGTCCACTCTATGAATCGTCGAAAGTTG
TGGCACAGAAAATGACTATTGCAGCACTGCGATATGCAAGGCAACTAGCTCAAGAGACTAGC
GACGAGGTAGTATATGGTCTAGGAAGGCAACCTGCTGTTCTTCGAACATTTAGCCAGAGATTA
TGCAGAGGGTTCAATGATGCCATCAATGGATTCGGTGACGATGGCTGGTCAATGTTAAGTTCA
GATGGTGCTGAAGATGTCATAGTTGCTGTCAATTCAAGGAAGAACCTCGCAACCACCTCCATT
CCTCTTTCCCCGCTTGGTGGCGTCCTTTGTGCCAAAGCATCAATGCTACTCCAGAATGTCCCCC
CTGCCGTACTGGTTCGGTTTCTGAGGGAGCACCGTTCAGAGTGGGCAGACTTTAATGTTGATG
CTTTTGTAGCTTCTGCATTGAAGTCTTGTCCGTATACATATCCTGGGATGAGGCCTACCAGATT
TACCGGTAGCCAGATAATAATGCCACTTGGCCACACAATTGAGCATGAAGAAATGCTTGAAG
TTATTAGACTTGAAGGGCACTCTATTGGACAGGAAGATGCTTTTATGCCAAGAGATATTCACC
TTTTACAGATGTGTAGTGGCACTGATGAGAATGCAGTTGGAGCCTGTTCTGAACTAGTTTTTGC
CCCTATTGATGAGATGTTTCCAGATGATGCACCTTTGCTTCCCTCCGGATTTCGCGTTATTCCT
CTCGAATCAAAATCAGGTGATGCCCAGGATACGTTGAACGCACATAGAACATTAGATCTAGC
ATCAAGTCTTGAAGTGGGCCCAGCAACAAACTCGACTACTGGAGATGCAGCTTCTTGTTACAG
TGCACGATCTGTACTGACAATCGCTTTCCAATTTCCATTTGAGGACAATCTTCAGGACAATGTG
GCTACCATGGCCCGACAGTATGTTCGCAGTGTGGTTTCGTCTGTCCAACGAGTTGCCATGGCA
ATATCTCCTACGGGAATGAATCCTACACTGGGGGCCAAGCTTTCTCCAGGCTCACCAGAAGCT
GTAACATTGTCGCATTGGATCTGCCAGAGCTATAGTTATCACATGGGAACAGAGTTACTTCGA
GCTGATTCTAGTGGCGATGAATCAGTACTAAAGAATCTCTGGCAACACCAGGATGCAATTTTG
TGCTGCTCATTGAAGTCGCTGCCAGTTTTCATTTTTGCTAATAAGGCCGGACTCGACATGCTGG
AGA

FIGURA 27 (continuação)

CAACATTAGTTGCATTACAGGACATTTCTCTAGATAAGATATTTGATGAATCCGGAAGGAAAG
TGTTGCTCTCAGAATTTGCCAAGATTATGGAACAGGGTTTCGCGTGTTTGCCTGGTGGTATCTG
CATGTCAACAATGGGACGACATATTTCATATGAACAAGCTATCGCGTGGAAAGTGTTTGCGTC
GTCTGAAGAAAATGCTGTCCACTGTTTGGCCTTTTCGTTTATCAACTGGTCATTCGTGTAA

SEQ ID N° : 231, seqüência de polipeptídeo de comprimento total de Lycopersicon esculentum Revoluta
MAMVAQQHRESSSGSITKHLDSSGKYVRYTAEQVEALERVYAECPKPSSLRRQQLIRECHILSNIEP
KQIKVWFQNRRCREKQRKESSRLQTVNRKLSAMNKLLMEENDRLQKQVSQLVCENGYMRQQLQ
SVSAATTDVSCESVVTTPQHSLRDANNPAGLLPIAEETLAEFLSKATGTAVDWVPMPGMKPGPDS
VGIFAISHSCSGVAARACGLVSLEPTKIADILKDRPSWFRDCRNVEVITMFPAGNGGTVELLYTQIY
APTTLAPARDFWTLRYTTTLDNGSLVVCERSLSGNGPGPNPTAASQFVRAQMLPSGYLIRPCDGG
GSIIHIVDHLNLEAWSAPEILRPLYESSKVVAQKMTIAALRYARQLAQETSDEVVYGLGRQPAVLR
TFSQRLCRGFNDAINGFGDDGWSMLSSDGAEDVIVAVNSRKNLATTSIPLSPLGGVLCAKASMLL
QNVPPAVLVRFLREHRSEWADFNVDAFVASALKSCPYTYPGMRPTRFTGSQIIMPLGHTIEHEEML
EVIRLEGHSIGQEDAFMPRDIHLLQMCSGTDENAVGACSELVFAPIDEMFPDDAPLLPSGFRVIPLE
SKSGDAQDTLNAHRTLDLASSLEVGPATNSTTGDAASCYSARSVLTIAFQFPFEDNLQDNVATMA
RQYVRSVVSSVQRVAMAISPTGMNPTLGAKLSPGSPEAVTLSHWICQSYSYHMGTELLRADSSGD
ESVLKNLWQHQDAILCCSLKSLPVFIFANKAGLDMLETTLVALQDISLDKIFDESGRKVLLSEFAKI
MEQGFACLPGGICMSTMGRHISYEQAIAWKVFASSEENAVHCLAFSFINWSFV

SEQ ID N° : 232, seqüência codificadora para polipeptídeo CLV3 de Saccharum officinarum
ATGAGGATGTTCTTCCGGCACTTGCTATCTTGCATCCTCCTGCTCTTGCTAATGTCACACTTGC
CAAGCCCGATCCTCGGCTTGAGAACATTGAGAGGAGAGGAGGCGAAGAGCGACTTGAGGCGC
CATGAGCATGAGCTTCCGCCAGCGGTATCTCCTTCAAAAGAGGTGGACAACGATGACGCTGCC
GCCTCCAGTAAGTTCACAGTTTCAAGAAGAATGGTTCCTCAAGGTCCAAACCCTCTCCACAAC
AGATGA

SEQ ID N° : 233, polipeptídeo CLV3 de Saccharum officinarum
MRMFFRHLLSCILLLLLMSHLPSPILGLRTLRGEEAKSDLRRHEHELPPAVSPSKEVDNDDAAASSK
FTVSRRMVPQGPNPLHNR

SEQ ID N° : 234, Prm05843 iniciador avançado
GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGAGGATGTTCTTCCGG

SEQ ID N° : 235, Prm05844 iniciador reverso
GGGGACCACTTTGTACAAGAAAGCTGGGTTCCTCTCATCTGTTGTGGAG

SEQ ID N° : 236 seqüência promotora de Oryza sativa Prolamina RP6
CCTTCTACATCGGCTTAGGTGTAGCAACACGACTTTATTATTATTATTATTATTATTATTATTAT
TTTACAAAAATATAAAATAGATCAGTCCCTCACCACAAGTAGAGCAAGTTGGTGAGTTATTGT
AAAGTTCTACAAAGCTAATTTAAAAGTTATTGCATTAACTTATTTCATATTACAAACAAGAGT
GTCAATGGAACAATGAAAACCATATGACATACTATAATTTTGTTTTTATTATTGAAATTATATA
ATTCAAAGAGAATAAATCCACATAGCCGTAAAGTTCTACATGTGGTGCATTACCAAAATATAT
ATAGCTTACAAAACATGACAAGCTTAGTTTGAAAAATTGCAATCCTTATCACATTGACACATA
AAGTGAGTGATGAGTCATAATATTATTTTTCTTGCTACCCATCATGTATATATGATAGCCACAA
AGTTACTTTGATGATGATATCAAAGAACATTTTTAGGTGCACCTAACAGAATATCCAAATAAT
ATGACTCACTTAGATCATAATAGAGCATCAAGTAAAACTAACACTCTAAAGCAACCGATGGG
AAAGCATCTATAAATAGACAAGCACAATGAAAATCCTCATCATCCTTCACCACAATTCAAATA
TTATAGTTGAAGCATAGTAGTAGAATCCAACAACA

FIGURA 27 (continuação)

SEQ ID N° : 237, domínio CLE, seqüência de consenso
(S/E/R/M/P/L)(K/E/D/R/S)R(K/I/L/R/F/Q/V/M)(V/I/S/L)(P/L/R)(R/T/Q/C/N/G/K/S)(N/G)(S/P)(D/N/Y)P(I/L/Q/R/Y/H)(H/L/I)(H/N)

SEQ ID N° : 238, domínio CLE, seqüência de consenso
(S/P)(R/E/K)R(M/L/I)(V/S/I)P(Q/G/C/T/S)GP(N/D)P(L/Q/H)H(H/N)

SEQ ID N° : 239, CV276412 Populus trichocarpa x Populus deltoides cDNA clone WS0179_L07 3', seqüência de mRNA
GCAAATTCCCCCTGCTCTAAAATCCATTTTCCTATCTATCAACCTCTCTGGTCTCTATATCCCCA
CCATTTTCATCATGAGGAATAAAAATCCTCAGCTGTTCCTTATTTTCCTGATAGTGTTCCTGGT
ACTCGTTCATGGCACCACTTGCCGCGATGCCAAAAGATCTACTAGCAATGGAGAAACAGTAC
AAGGGTCGAAAACCAAACATTCTTCAATGTTTCTACAAGCGCTTTCTTCCATTTTTAAGGCTTC
AGAATCCAGCACTAACAACATCAAGGCACTTCACACCGTCTCGCGCAGGCTTGTTCCTTGTGG
ACCAAACCCACTCCACAACTGATCATCGAATCAACAAAATTCCAATCTGTTTCCATTTGTTGA
AATATATATAGTGTTTTTAAAATATTTTTGTTGAATAATCTATACATATTTTCCCTATACATCCA
GTTTTGAAAAAAGAAAAAAAAAAAAAAAAA

SEQ ID N° : 240, CV276412 Populus trichocarpa x Populus deltoides, seqüência de proteína deduzida
MRNKNPQLFLIFLIVFLVLVHGTTCRDAKRSTSNGETVQGSKTKHSSMFLQALSSIFKASESSTNNI
KALHTVSRRLVPCGPNPLHN

SEQ ID N° : 241, CI670037 Oryza sativa (grupo de cultivar japônico) cDNA clone J06B5214G11T3 5', seqüência de mRNA
TCTCTCTCTCTCACACACACACACACACAAACTAGAGACTTATGCCATGAGGAGGTTCTCCAA
GCAGCACCTGGTCCCTTTCATTCTGCTCCTGCTACTTGTCATGTCTCACTTGCCAATCTCAAGC
CTTGGTTCAAGGAGAGCATTCAGAGAAGAAGCTGTCAGTGGTTTCAGAAGCCATGAGCTTGCT
CCAACCATGGCTCCTTCTCAAGAGAAAGAAGCAGGCGTCGTCGCCGGCGCCGGTAGCATCTGT
GGCCAGAAGTATGCTGTCTCAAGAAGAATGGTTCCTCAGGGACCAAACCCTCTCCACAACTGA
TGAACTCATGCACTGCAATCTGCCGTGATCATCAGCTTCTCCAATACAGGTGAATGGTGCAGC
CATCATTGGTTACCTTTAAGTCCGTCTGCTTAATTAACTAGATATTTCATAGTATTTTATCACA

SEQ ID N° : 242, CI670037 Oryza sativa (grupo de cultivar japônico) cDNA clone J06B5214G11T3 5', seqüência de proteína deduzida
MRRFSKQHLVPFILLLLLVMSHLPISSLGSRRAFREEAVSGFRSHELAPTMAPSQEKEAGVVAGAG
SICGQKYAVSRRMVPQGPNPLHN

SEQ ID N° : 243, CA259979 Saccharum officinarum cDNA clone SCPRRT3029C01 5', seqüência de mRNA
TTTGAAACTGTTACAACTGGTTCGCCTGCACCTGCCTGCCCGGAACGACCCACGCGTCCGCGC
TCTCTCTCTCCCTCTCCCTCTACCTCTCTCTAACATATGCCATGAGGATGTTCTTCCGGCACTAG
CTATCTTGCATCCTACTGCTCTAGGTAATGTCACACTTGCCAAGCTCGATCCTCGGCTTGAGAA
CATCGAGAGGAGAGGAGGCGAAGAGCGACTTGAGGCGCCATGAGCATGAGCTTCCGCCAGCC
GTATCTCCTTCAAAAGAGGTGGACAACGACGACGCTGCCGCCGCCAGTAAGTTAACAGTTTCA
AGAAGAATGGTTCCTCAAGGTCCAAACCCTCTCCACAACAGATGAGAGGATCGGAGCAGCAT
GCTGCTGCTGCTTCTGGTATGGTCTTCAGCCAAGGTAGCAGCTAGCTAGCTCTTCTCAGTCTCA
GCTAATGGTGTA

FIGURA 27 (continuação)

GTGAGTGGCACGCACTCAACAAAGTACCCTTCCTCTCTTTCCTTTTTTTTTTAACCCGTCATATG
AATGATACAAATGGATGCATATATAAGTTGAATACTACTTCCCATGTAATGTGTTTTTGTTGCA
TTGATTTCATTTATAGGCAGCTTTGTACATAGATTTTTATGATATTAAGGTGTAAAAGTTGTTG
TGCTAGATGGATCTAGATTTTGAATGTTTG

SEQ ID N° : 244, CA259979 clone de cDNA clone de Saccharum officinarum SCPRRT3029C01 5', seqüência de proteína deduzida (possivelmente o termina N é perdido)
MSHLPSSILGLRTSRGEEAKSDLRRHEHELPPAVSPSKEVDNDDAAAASKLTVSRRMVPQGPNPLH
NR

SEQ ID N° : 245, NM_117965.3| Arabidopsis thaliana CLE2 (CLAVATA3/ESR-RELATED); ligação do receptor (CLE2) mRNA, cds completos
GCAACGAGAAAAACCGTCCTTGGTGACAACTCATGCTTGCACTCAAGCCTTAAGCTAGCTAAA
CCTATCTCGCGCACTACTAGAATTCAAATAAAACTCTATAAATAGAAACCCTCATGAGATCTC
TTCTTTCCTCATATACACTCATACACACCACGTGAACAATCTATCTCTCTTTCTATTGCTTTTCT
ATATATACAGAAACTAATTAATTGTATCTGTAATGGCTAAGTTAAGCTTCACTTTCTGCTTCTT
GTTGTTTCTTCTGTTATCCTCAATCGCCGCTGGAAGCCGCCCTCTTGAGGGGGCTCGGGTCGGG
GTGAAGGTGAGAGGCCTAAGCCCTTCTATCGAGGCTACGAGTCCGACTGTAGAGGATGATCA
AGCTGCGGGTAGCCATGGGAAATCTCCAGAGCGGTTAAGCCCAGGAGGACCCGACCCACAAC
ATCACTAGTTATTTTGTGTTTTTCAATTTCTTCGACATGTTTATTACTTATCAATAATTTGGTTG
CAACGAAGCTGTTTTTCTTTTTTGT

SEQ ID N° : 246, NP_193586.1 Arabidopsis thaliana CLE2 (CLAVATA3/ESR-RELATED); ligação do receptor
MAKLSFTFCFLLFLLLSSIAAGSRPLEGARVGVKVRGLSPSIEATSPTVEDDQAAGSHGKSPERLSP
GGPDPQHH

SEQ ID N° : 247, AF343656.1, polipeptídeo semelhante a CLE de Brassica napus, mRNA, cds completos
CTTTCATATTGTATCTCCCGCAGCCAAACAAAAACTTTTTTTTTGACAAAGATAGAACATGAA
GATCAAGAGTTTGATATTGGCTTCCTCTTTTCTGATTCTTGCCTTCATTCATCACTCAGAATCA
GCTTCATTTCGGAGTTTGCTGATGAAGAATGGATTGTACGAAGAAGAAGAAGCAAAAATTCT
ATTGGGCGACTCCAAAGAAACGATAACTAATTCTACAGCTTTGGAGTCTAAACGGATAATTCC
GACGGGTCCAAATCCACTTCACAACAGGTAACTTTGATCATTTAAGAACAAGATATGTTGTGA
GTATGTCTCTTCCTCTGTTTTTGTTACAAGTATCTATCTTACTGTGTAATGTAATGGTGAATGTA
TAATACATTTTCTAATTAAATTACCTTATTTAAAAAA

SEQ ID N° : 248, AAL57176.1 Brassica napus semelhante a CLE
MKIKSLILASSFLILAFIHHSESASFRSLLMKNGLYEEEEAKILLGDSKETITNSTALESKRIIPTGPNP
LHNR

SEQ ID N° : 249, AF126009.1, Arabidopsis thaliana CLAVATA3 (CLV3) gene, cds completos
AACAGTTTCTATATTTCTCTCTGTATCTCTCTCACTCAGTCACTTTCTCTCTAAAAAATGGATTC
TAAAAGCTTTGTGCTACTACTACTACTCTTCTGCTTCTTGTTCCTTCATGATGCTTCTGGTCTCA
CTCTCTCTTTCACTCCTCTATATGTCTATAACCCATGTAAATGGATTCTTATAAGTCTCAACAT
A

FIGURA 27 (continuação)

ACTTTTTTTTGTTATTTACTATTTCACCAGATCTCACTCAAGCTCATGCTCACGTTCAAGGACTT
TCCAACCGCAAGGTTATCTTCAACTTGTACTCATTAAGGCCTCTCAGCATTCATGTGTTATGTT
CATGTAGATGTCCGGTCCAGTTCAACAACTGTTTCATTGCTTTAGTTGTCACGAGAAATATTTG
TATATATTATTATGGTGTGCAAAACATAGTAAAATGTTGTTCAATTGGCAGATGATGATGATG
AAAATGGAAAGTGAATGGGTTGGAGCAAATGGAGAAGCAGAGAAGGCAAAGACGAAGGGTT
TAGGACTACATGAAGAGTTAAGGACTGTTCCTTCGGGACCTGACCCGTTGCACCATCATGTGA
ACCCACCAAGACAGCCAAGAAACAACTTTCAGCTCCCTTGACCTAATCTCTTGTTGCTTTAAA
TTATTTCATATTGTAAATTACTTTCTGCTTTATCGGTTTTACCATTTCGGGAGTCTTTTTTGTGT
GCAATCTGTTTCGTTTGGTGTAGTTTCATGAAAGTGAATGTAAGATATGCATTACGTTTGTTGC
TGAAGTG

SEQ ID N° : 250, AF126009_1 CLAVATA3 [Arabidopsis thaliana]
MDSKSFVLLLLLFCFLFLHDASDLTQAHAHVQGLSNRKMMMMMKMESEWVGANGEAEKAKTKG
LGLHEELRTVPSGPDPLHHHVNPPRQPRNNFQLP

SEQ ID N° : 251, OsSYR seqüência codificadora
ATGGAAGGTGTAGGTGCTAGGCAGAGGAGGAACCCTCTGATACCCAGACCAAACGGTTCAAA
GAGGCATCTGCAGCATCAGCATCAGCCAAATGCTGCCGAGAAGAAGACCGCCGCGACATCGA
ATTACTTCAGTATCGAGGCGTTCCTCGTGCTCGTCTTCCTCACCATGTCATTGCTCATACTTCC
ATTGGTGCTTCCCCCATTGCCTCCGCCGCCATCGCTGCTGCTGCTGCTGCCAGTCTGCCTGCTC
ATCCTGCTGGTTGTGCTGGCCTTCATGCCAACGGATGTGCGGAGCATGGCTTCCTCTTACTTGT
AAATACATCTCCTAGGGGAATTTATTTTTGTTTTTGA

SEQ ID N° : 252, OsSYR seqüência de proteína deduzida
MEGVGARQRRNPLIPRPNGSKRHLQHQHQPNAAEKKTAATSNYFSIEAFLVLVFLTMSLLILPLVL
PPLPPPPSLLLLLPVCLLILLVVLAFMPTDVRSMASSYL

SEQ ID N° : 253, prm08170, códon de partida em negrito, locais AttB1 em itálicos
GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGGAAGGTGTAGGTGCTAGG

SEQ ID N° : 254, prm08171, complementaridade reversa, local AttB2 em itálico
GGGGACCACTTTGTACAAGAAAGCTGGGTCAAAAACAAAAATAAATTCCCC

SEQ ID N° : 255, seqüência de promotor GOS2 de arroz
AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAAA
TATAAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATCCA
CCTACTTTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCTTAGT
AATTAAGTGGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCATGAAGT
TAAATTATTCGAGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTTCTAGCT
GAACTCAATGGGTAAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTCTGAACGT
ATTGGCAAAGATTTAAACATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACA
TCATTAAGGACATGTCTTACTCCATCCCAATTTTTATTTAGTAATTAAAGACAATTGACTTATT
TTTATTATTTATCTTTTTTCGATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCA
TGTGTGAGTGCACCTCCTCAATACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTAT
TTAATACATTTAGGTAGCAATATCTGAATTCAAGCACTCCACCATCACCAGACCACTTTTAAT
AATATCTAAAATACAAAAAATAATTTTACAGAATAGCATGAAAAGTATGAAACGAACTATTT
AGGTTTTTCACATACAAAAAAAAAAAGAATTTTGCTCGTGCGCGAGCGCCAATCTCCCATATT
GGGCACACAGGCAACAACAGA

FIGURA 27 (continuação)

GTGGCTGCCCACAGAACAACCCACAAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAA
CAACCTTTTAACAGCAGGCTTTGCGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGC
ATCCTCCTCCTCCCATCTATAAATTCCTCCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAA
GCCAAGAAGAGGGAGAGCACCAAGGACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTTC
GATCCATATCTTCCGGTCGAGTTCTTGGTCGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGG
TATGTGCCCTTCGGTTGTTCTTGGATTTATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTA
GGAAAGGGGATCTGTATCTGTGATGATTCCTGTTCTTGGATTTGGGATAGAGGGGTTCTTGAT
GTTGCATGTTATCGGTTCGGTTTGATTAGTAGTATGGTTTTCAATCGTCTGGAGAGCTCTATGG
AAATGAAATGGTTTAGGGTACGGAATCTTGCGATTTTGTGAGTACCTTTTGTTTGAGGTAAAA
TCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAAGTACGGTTGTTTGGTCCTCGATTCTGGTA
GTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTCCCTATTGAACAAAAATAATCCAA
CTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTTAAGCCTGTCCAAAATTTCGC
AGCTGGCTTGTTTAGATACAGTAGTCCCCATCACGAAATTCATGGAAACAGTTATAATCCTCA
GGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTTTTTTCCCAAATATCT
TAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATGCTTTTATAGCGTT
ATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGAAGAACTTATC
CGATTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATTCATTTGGA
TTATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAACTGTCCT
CAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGTAGAAGT
TTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAATCGGGAT
AGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCTTGCCACTTTC
ACCAGCAAAGTTC

SEQ ID N° : 256, motivo conservado 1a
YFS

SEQ ID N° : 257, motivo conservado 1b
YFT

SEQ ID N° : 258, motivo conservado 1c
YFG

SEQ ID N° : 259, motivo conservado 1d
YLG

SEQ ID N° : 260, motivo conservado 2
(V/A/I)LAFMP(T/S)

SEQ ID N° : 261, motivo conservado 3
(S/P)YL

SEQ ID N° : 262, homólogo SYR de arroz 1 (XP_472637), codificado por SEQ ID N° : 277
MYLLSPRNGDEEDEQEEIQELISDDEPPNLKLASCATAASSSSSSGSDMEKGRGKACGGGSTAPPPP
PPSSSGKSGGGGGSNIREAAASGGGGGVWGKYFSVESLLLLVCVTASLVILPLVLPPLPPPPSMLML
VPVAMLVLLLALAFMPTTTSSSSSAGGGGGGGRNGATTGHAPYL

FIGURA 27 (continuação)

SEQ ID N° : 263, homólogo SYR de arroz 2, seqüência de proteína deduzida (AP008218)

MLLEHLMITMEEQMFREQQMQRGGRHHQHHTTREQEQQQKQQQRRRLMNNATNGGGGDGGSR
CYFSTEAILVLACVTVSLLVLPLILPPLPPPPTLLLLLPVCLLALLVVLAFMPTDMRTMASSYL

SEQ ID N° : 264, homólogo SYR de milho (AY110705), codificado por SEQ ID N° : 278

MASRSSAMEGGAAIQRRNAVKRHLQQRQQEADFLDKKVIASTYFSIGAFLVLACLTVSLLILPLXX
XXXXXXXXXLLWLPVCLLVLLVVLAFMPTDVRSMASSYL

SEQ ID N° : 265, homólogo SYR de trigo, seqüência de proteína deduzida (CK211328)

MDSQFGALERGGSRQRRSPVLARPNTTKRHIQQQRANAADKKVVMPNYFSIEAFFVLACLTVSLLI
LPLVLPPLPPPPSLLLFVPVCLLILLMVLAFMPTDMRSMATSYL

SEQ ID N° : 266, homólogo SYR cevada (CB871444), codificado por SEQ ID N° : 286

MDSQFGAMDRGGSRQRSSPVLARPNTAKRQMQQQRANAADKKVVIPNYFGVEAFFVLACLTVSL
LILPLVLPPLPPPPSLLLLLPVCLLILLMVLAFMPTDVRSMATSYL

SEQ ID N° : 267, homólogo SYR 1 de cana de açúcar codificado por SEQ ID N° : 287 (CA165713)

MEGGGQIQRRNNAVKRHLQQRQQEADFLDKKVIASTYFSIEAFLVLACLTVSLLILPLVLPXLPAP
ASLLLWLPVWLLELLIVLAFMPTDVRSMASSYL

SEQ ID N° : 268, homólogo SYR 2 de cana de açúcar codificado por SEQ ID N° : 288(CA242805)

MEGGGQIQRRNNAVKRHLQQRQQEADFLDKKVIASTYFSIEAFLVLACLTVSLLILPLVLPPLPPPP
SLLLWLPVCLLILLIVLAFMPTDVRSMASSYL

SEQ ID N° : 269, homólogo SYR de sorgo, codificado por SEQ ID N° : 289 (CX611532)

MASRSSALEGGGAAIQRRNNAVKRHLQQRQQEADFHDKKVIASTYFSIGAFLVLACLTFSLLILPL
VLPPLPPPPSLLLWLPVCLLVLLVVLAFMPTDVRSVAASYL

SEQ ID N° : 270, homólogo SYR 1 de Arabidopsis thaliana, codificado por SEQ ID N° : 290 (NM_115853)

MIREISNLQKDIINIQDSYSNNRVMDVGRNNRKNMSFRSSPEKSKQELRRSFSAQKRMMIPANYFS
LESLFLLVGLTASLLILPLVLPPLPPPPFMLLLVPIGIMVLLVVLAFMPSSHSNANTDVTCNFM

SEQ ID N° : 271, homólogo SYR 2 de Arabidopsis thaliana, codificado por SEQ ID N° : 291 (NM_180078)

MIREFSSLQNDIINIQEHYSLNNNMDVRGDHNRKNTSFRGSAPAPIMGKQELFRTLSSQNSPRRLIS
ASYFSLESMVVLVGLTASLLILPLILPPLPPPPFMLLLIPIGIMVLLMVLAFMPSSNSKHVSSSSTFM

FIGURA 27 (continuação)

SEQ ID N° : 272, homólogo SYR de videira (CF404276), codificado por SEQ ID N° : 279

MSIEQPEADSRLSEGPLINLQDRYLSGIMEARGRRNSAPLQVERKNPTPPMAEGKKMEYNRTPLSR
ENSRRLIPASYFSLESLLLLICLTASLLILPLILPPLPPPPFMLLLLPIGILAVLMILAFMPSNVRDLTYT
YV

SEQ ID N° : 273, homólogo SYR de cítrico (CF830612), codificado por SEQ ID N° : 280

MNSDNSESRQRLSKGIINLQDRYPTSIMDRGVRKIATPPVEKRKVEYHRSYSQGASRKLFSASYFTL
ESLLLLVCLTASLLILPLVLPPLPPPPFLLLLVPIXILAVLLVLAFMPSNVRDITSTYV

SEQ ID N° : 274, homólogo SYR 1 de tomate (AI774560), codificado por SEQ ID N° : 282

MNMDMESSEAKLRSSKGFINLEEHQQYFNNIMEGNKMEHKRSFTQGHGKKMLSMNYFSLESIILL
LGLTASLLLLPLMLPPLPPPPFMLLLVPIFILVVLMILAFMPSNVRNVTCSYL

SEQ ID N° : 275, homólogo SYR 2 de tomate, seqüência parcial (BG125370), codificado por SEQ ID N° : 281

MEGNKMEHKRSFTQGHGKKMLSMNYFSLESIILLLGLTASLLLLPLMLPPLPPPPFMLLLVPIFILV
VLMILAFMPSNVRNVTCSYL

SEQ ID N° : 276, proteína ARGOS de Arabidopsis thaliana (AY305869), codificado por SEQ ID N° : 292

MDVGRNNRKNMSFRSSPEKSKQELRRSFSAQKRMMIPANYFSLESLFLLVGLTASLLILPLVLPPLP
PPPFMLLLVPIGIMVLLVVLAFMPSSHSNANTDVTCNFM

SEQ ID N° : 277 homólogo SYR de Oryza sativa, mRNA, (XM_472637), que codifica a proteína de SEQ ID N° : 262

ATGTACTTGTTGAGCCCAAGAAATGGCGACGAGGAGGACGAACAGGAGGAAATCCAGGAGCT
GATCAGCGACGACGAGCCGCCCAATCTCAAGTTGGCATCCTGCGCCACTGCAGCCAGCAGCA
GCAGCAGCAGCGGCAGCGACATGGAGAAGGGAAGAGGTAAAGCCTGCGGCGGCGGGAGTAC
GGCGCCGCCGCCGCCGCCGCCGTCGTCGTCAGGTAAATCCGGCGGCGGCGGCGGCAGCAATA
TCAGGGAGGCGGCGGCTAGCGGCGGCGGCGGCGGCGTGTGGGGCAAGTACTTCTCGGTGGAG
TCGCTGCTCCTGCTGGTGTGCGTGACGGCGTCGCTGGTGATCCTCCCGCTCGTGCTGCCGCCGC
TGCCCCCGCCGCCGTCGATGCTGATGCTGGTGCCGGTGGCGATGCTGGTGCTGCTGCTGGCGC
TGGCGTTCATGCCGACGACGACGTCGTCGTCGTCCGCCGGCGGCGGCGGCGGCGGCGGC
CGCAATGGGGCGACGACGGGACATGCTCCCTACTTGTAG

SEQ ID N° : 278, homólogo de milho SYR, mRNA (AY110705), que codifica a proteína da SEQ ID N° : 264

TTCACATTACACTCATGACTCGTCTTAGGACGAATCCTGCAGCTGCAAAACACAGAAAATCTG
GCCAAGACCTATTTATCTATTTACAGGTAAGAGGAGGCCATGCTGCGCACATCTGTCGGCATG
AAGGCCAGTACAACCAGCAAGACGAGCAGGCAGACCGGCAGCCACAGCAGNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNCCAGCGGCAGTATCAGCAGCGAGACGGTGAGGCAGGCGAG
CACGAGGAATGCCCCGATGCTGAAGTAGGTGGACGCGATGACCTTCTTGTCGAGGAAATCCG
CCTCCTGCTGACGCTGCTGCAGATGCCGCTTCACGGCATTCCTCCTTTGTATTGCCGCCCCTCC
TTCCATCGCGCTAGATCGGCTTGCCATGTGTTCTTTGGTGTAGGCGGAGGTGGAGACCAGTCG
CAGTGAGTGGTTGCCAAAGTAAGCAAGAAGTGAAAGGCGGTGTAGAACCGCCTTGNGTTTTC
TGAACGTTTTGTAATCAGATCTGAGCTCGGGTGGTCGAA

FIGURA 27 (continuação)

SEQ ID N° : 279, clone de cDNA de Vitis vinifera cDNA CSECS024D03 3' seqüência de mRNA (CF404276, complemento reverso), codifica SEQ ID N° : 272

TACTTTGAACCGGTTGGAACTTGGTGACCGGTAAAGGAGAAATTTAAATGAGTATTGAGCAGC
CTGAAGCAGATTCAAGACTGTCTGAAGGACCTTTGATCAATCTGCAAGATCGATATTTGAGTG
GCATCATGGAAGCGAGAGGAAGGAGGAACTCTGCTCCTCTGCAAGTTGAGAGGAAAAACCCT
ACTCCTCCTATGGCCGAAGGAAAGAAGATGGAATATAATAGAACTCCCCTCTCACGAGAGAA
CAGCAGGAGACTGATCCCAGCAAGCTATTTCAGCTTGGAGTCATTGCTTTTGCTCATCTGTCTC
ACGGCTTCATTGCTGATCCTTCCCCTGATACTACCACCATTGCCACCCCCTCCTTTCATGCTGC
TTCTGCTCCCCATTGGCATTCTAGCAGTGCTTATGATCTTGGCTTTCATGCCTTCTAATGTCAG
AGATTTGACTTATACATATGTGTAAATGGTGGTGTTCAAAAGTGCACCTCTTCTCATCATACAA
TTTTTGTTATTTGTCTTGATATTCAGATGAAAATCAGTTATTTTATTTCTTGATTAAAAAAAAA
AAAAAAAA

SEQ ID N° : 280, clone de cDNA de Citrus reticulata CR_CEa05B13, seqüência de mRNA (CF830612), que codifica SEQ ID N° : 273 (nt 313-693)

AGGGTTTCTTCAAAGATAGGTAGCCATTTGCACATTTGAATCTGCTTGTTGGATATTGTCAAGG
AGGCTGTTGGAATTAGGCCACATTTCAGAATCTGGTTTCATCCTGGATCGCTGGGCATTTGAA
GGCATTTTGTGATCATCGCTGTTTAAAATTTGGCCGCATATTAGAATCTGGGTTCTCATCGGTT
TTCCGTACATTTACCTTGACCACATTTTGATATCTGGGTTGAGCTGCATTTTAGCCTTCGTATTT
AAAAGGACTTGATCTAATCTGGGGTCTTGGTGAGCCGGGGCAACTGATCATAGTAAATGAATT
CTGATAATTCTGAGTCGAGACAGAGACTATCAAAGGGCATTATAAACTTGCAAGATCGATATC
CGACCAGCATTATGGATCGTGGTGTAAGAAAAATTGCAACTCCTCCGGTCGAGAAGAGGAAA
GTTGAGTATCACCGAAGTTACTCGCAAGGGGCATCCAGAAAACTGTTTTCGGCAAGCTATTTC
ACCCTGGAATCATTGCTTTTGCTCGTATGTCTGACGGCCTCATTGCTGATCCTGCCATTGGTGC
TTCCGCCCTTGCCGCCCCCGCCATTCCTGCTGCTTCTGGTTCCTATANGTATTCTAGCCGTGCTT
TTGGTCTTGGCATTCATGCCTTCTAATGTAAGAGATATAACTTCCACGTACGTGTAAATGGTGT
TGCT

SEQ ID N° : 281, seqüência 5' de clone de cDNA de Lycopersicon esculentum cTOF8M10, seqüência de mRNA parcial (BG125370) que codifica SEQ ID N° : 275

GAAAAAAATGTATTTAATCATTATGTAAAAAACAAGTGAATCTACTTTGATATTTTCTTCTAA
ATTAAACCACACAATTAAAGATATGAGCAAGTCACATTCCTAACATTAGAAGGCATAAAAGC
TAAGATCATAAGAACAACAAGAATGAAAATTGGGACTAACAACAACATAAAAGGTGGTGGTG
GCAATGGTGGAAGCATCAATGGCAAAAGTAACAAAGATGCTGTAAGACCAAGTAACAAAATA
ATTGACTCTAAGCTAAAATAATTCATTGACAACATTTTCTTGCCATGTCCTTGTGTAAATGATC
TCTTATGCTCCATCTTATTGCCTTCCATAATGTTGTTGAAATATTGTTGATGTTCCTCCAAATTA
ATAA

SEQ ID N° : 282, seqüência 5' clone de cDNA de Lycopersicon esculentum cDNA clone cTOF8M10, seqüência de mRNA (AI774560), que codifica SEQ ID N° : 274

TTTGTTAAAGATTGGCACATTTTCAAGTTCAGTATTCATTCGATTTTTGATATCTACATAAAAA
AAAAGTGTCCTGGTACTACTCAATATTCCTCAGAACGACTTCATATTCAGGTCTCGAATTCAA
AACCTCACATCAAGATTCTTAGGAAATTTCAAGATTGGTTGAAAAACTCATATCCTTCTCTAA
GTTTCAAGATTGGTTCCAAATTAAAACTCGAGACTTCTGAGTAAGAGCGTACGACTAGTAATG
AACATGGACATGGAATCATCAGAGGCAAAATTGAGATCATCAAAAGGGTTTATTAATTTGGA
GGAACATCAACAATATTTCAACAACATTATGGAAGGCAATAAGATGGAGCATAAGAGATCAT
TTACACAAGGACATGGCAAGAAAATGTTGTCAATGAATTATTTTAGCTTAGAGTCAATTATTT
TGTTACTTGGTCTTACAGCATCTTTGTTACTTTTGCCATTGATGCTTCCACCATTGCCACCACCA
CCTTTTATGTTGTTGTTAGTCCCAATTTTCATTCTTGTTGTTCTTATGATCTTAGCTTTTATGCCT
TCTAATGTTAGGAATGTGACTTGCTCATATCTTTAATTGTGTGGTTTAATTTA

FIGURA 27 (continuação)

SEQ ID N° : 283, promotor de proteína do grupo de mobilidade alta
CATGCGGCTAATGTAGATGCTCACTGCGCTAGTAGTAAGGTACTCCAGTACATTATGGAATAT
ACAAAGCTGTAATACTCGTATCAGCAAGAGAGAGGCACACAAGTTGTAGCAGTAGCACAGGA
TTAGAAAAACGGGACGACAAATAGTAATGGAAAAACAAAAAAAAACAAGGAAACACATGGC
AATATAAATGGAGAAATCACAAGAGGAACAGAATCCGGGCAATACGCTGCGAAAGTACTCGT
ACGTAAAAAAAAAGAGGCGCATTCATGTGTGGACAGCGTGCAGCAGAAGCAGGGATTTGAAAC
CACTCAAATCCACCACTGCAAACCTTCAAACGAGGCCATGGTTTGAAGCATAGAAAGCACAG
GTAAGAAGCACAACGCCCTCGCTCTCCACCCTCCCACCCAATCGCGACGCACCTCGCGGATCG
GTGACGTGGCCTCGCCCCCCAAAAATATCCCGCGGCGTGAAGCTGACACCCCGGGCCCACCC
ACCTGTCACGTTGGCACATGTTGGTTATGGTTCCCGGCCGCACCAAAATATCAACGCGGCGCG
GCCCAAAATTTCCAAAATCCCGCCCAAGCCCCTGGCGCGTGCCGCTCTTCCACCCAGGTCCCT
CTCGTAATCCATAATGGCGTGTGTACCCTCGGCTGGTTGTACGTGGGCGGGTTACCCTGGGGG
TGTGGGTGGATGACGGGTGGGCCCGGAGGAGGTCCGGCCCCGCGCGTCATCGCGGGGCGGGG
TGTAGCGGGTGCGAAAAGGAGGCGATCGGTACGAAAATTCAAATTAGGAGGTGGGGGGCGG
GGCCCTTGGAGAATAAGCGGAATCGCAGATATGCCCCTGACTTGGCTTGGCTCCTCTTCTTCTT
ATCCCTTGTCCTCGCAACCCCGCTTCCTTCTCTCCTCTCCTCTTCTCTTCTCTTCTCTGGTGGTGT
GGGTGTGTCCCTGTCTCCCCTCTCCTTCCTCCTCTCCTTTCCCCTCCTCTCTTCCCCCCTCTCACA
AGAGAGAGAGCGCCAGACTCTCCCCAGGTGAGGTGAGACCAGTCTTTTTGCTCGATTCGACGC
GCCTTTCACGCCGCCTCGCGCGGATCTGACCGCTTCCCTCGCCCTTCTCGCAGGATTCAGCC

SEQ ID N° : 284, Medicago sativa
MVRCFSLGSVLILIALAASMVVLPLMLPPLPPPPLALLFFPVGIMAALVVLAFSPSENVKNVVVYSS
SSSGIANSKR

SEQ ID N° : 285, Medicago sativa
MIMVASKEKTNSGGCMFRYSVLILSLLALSILVLPLVMPPLPPPPLLLLLVPVFIMLLLFFIAFSPSKK
VPNKASFVS

SEQ ID N° : 286, CB871444 HC03O01y CH clon de cDNA de Hordeum vulgare HC03O01 3-PRIME, seqüência de complemento reverso de mRNA que codifica SEQ ID N° : 266

CTTCCGAGAAAGATGCTTCATATTGCACTCATCTTATGCCAACACACAATCAGAAACAAAAAC
CACGCAGCAAGCCTATTTACAAGTAAGAGGTGGCCATGCTCCGCACATCAGTCGGCATGAAG
GCCAGCACCATCAACAGGATGAGCAAGCAGACCGGCAAGAGCAGCAGTAGCGATGGCGGTG
GGGGCAACGGAGGCAGCACCAATGGCAGTATGAGCAATGAAACGGTGAGGCAGGCGAGCAC
GAAGAACGCTTCGACGCCGAAGTAGTTCGGTATGACAACCTTCTTGTCAGCAGCATTTGCCCT
CTGCTGCTGCATTTGCCTCTTTGCAGTATTTGGTCTGGCTAACACAGGGCTGCTCCTCTGCCTA
CTACCGCCTCTGTCCATTGCACCGAACTGGCTATCCATCTATTCTTTGGTGAGTGCAGATCAGC
GGCTATCCAGGAACTGAGATGAATAAGAACTTGTGGAATCTGAAGGTTTTTAACAGATCTGAA
GTCTTTGTGAGTCCGTGACTGAACTGCAGATCATCTGCTGTGGAAGAACTGCACCGCAAGTGG
CAATACCTTCGATCCTGAAACTTGCATTTTGGTGATGCCCGTACCACGC

SEQ ID N° : 287, CA165713 SCUTRZ3071B09.g RZ3 clone de cDNA de Saccharum officinarum SCUTRZ3071B09 5', seqüência de mRNA, que codifica SEQ ID N° : 267
GGAGAACGTGCTTGTTCAGGTGGTTCAACAGTGCGGACCAGAGAACAATGCGAGGTTCAGGA
TTAAAGGTATTCTGGCTTCAGAGGCAGTTCTTCCAAAGCAAGTGACAGGCGATTCCATCACCG
GAGCT

FIGURA 27 (continuação)

AAGATCTTATTACAACATTCAGAAACACAGGAGTCCTCCCACCGCCTTTCACTTCTTGCTTACT
TCTGCAACCACTCGCCGTGACTGATCTCCACGCACACCAAAGAACACAACACGTGGCAAGCC
GATCTAGCGCGATGGAAGGAGGGGGGCAAATACAGAGGAGGAATAATGCCGTGAAGCGGCA
CCTGCAGCAGCGGCAGCAGGAGGCGGATTTCCTCGACAAGAAGGTCATCGCGTCCACCTATTT
CAGCATCGAGGCGTTCCTCGTGCTCGCCTGCCTCACCGTCTCGCTGCTGATACTGCCGCTTGTG
CTGCCGNCGCTGCCGGCGCCGGCGTCGCTGCTGCTGTGGCTGCCGGTCTGGCTGCTCGAACTG
CTGATTGTGCTTGCCTTCATGCCGACAGATGTGCGCAGCATGGCCTCCTCTTACTTGTAAAAAA
ATAGATAAATAGGCCACCTTTGGCAATTTTCTGGGGTTTGGAGGTG

SEQ ID N° : 288, complemento reverso de CA242805 SCSFFL3090G07.b Saccharum officinarum FL3 Saccharum officinarum cDNA clone SCSFFL3090G07 3', seqüência de mRNA, que codifica SEQ ID N° : 268

GATTTTTTCCAAGCCAGTGACCGGCGATTCATCAACCGGAGCTGAGATCTTATACAACATTCA
GAAACACAGGAGTCCTCGCACCGCTTTCCACTCTTGGCTAATTTTGCAACCACTCGCCGTGAC
TGATCTCCACGCACACCAAAGAACACAACACGTGGCAACCCGATCTAGCGCGATGGAAGGAG
GGGGGCAAATACAGAGGAGGAATAATGCCGTGAAGCGGCACCTGCAGCAGCGGCAGCAGGA
GGCGGATTTCCTCGACAAGAAGGTCATCGCGTCCACCTATTTCAGCATCGAGGCGTTCCTCGT
GCTCGCCTGCCTCACCGTCTCGCTGCTGATACTGCCGCTGGTGCTGCCGCCGCTGCCGCCGCCG
CCGTCGCTGCTGCTGTGGCTGCCGGTCTGCCTGCTCATCCTGCTGATTGTGCTGGCCTTCATGC
CGACAGATGTGCGCAGCATGGCCTCCTCTTACTTGTAATAAATAGATAAATAGGCCACCTTGG
TCAATATTCTGTGATTTGGAGGTGATTCGTCCTGAGATGAGTCTCGATTGATGTCAGCTACTCC
CAAGGGGAAATGCATGTAACACTTGGTGGCCGACGGTGGCAAGATAATCATGCTACCATGTC
AGTTAAACC

SEQ ID N° : 289, CX611532 ANR1_25_E08.b1_A002 clone de cDNA de Sorghum bicolor de raízes anaeróbicas ANR1_25_E08_A002 3', seqüência de mRNA que codifica SEQ ID N° : 269

GCTGTTGGTGTTGTTCTTGAGATCTCTTTCTTGATCTTGTGTGGGATTAAAGAGGGATTCTTGC
CTTCCTACGGGAGAAAGAGAAAAGGGGAAGAACGTGCTTGTTCCGGTGGTTCAACAGTGCGG
AGACCCGGAGAACAATGCGAGGTTCAGGATTAAAGGTGTTCTGGCTTCAGGTGCAGTTCTTCC
AAAGCAGGTGACAGGCGATTCGATCACCGGAGCTCAGATCTGACGAAAACAAAACACAGTCC
TCCTCCCACCGCCTTTCAGTTCTTGCTTACTTCTGCAACCACTCACTGCGACCGTACACCAAAG
AACACTGCACATGGCAAGCCGATCTAGCGCGCTGGAAGGAGGGGGGGCAGCAATACAGCGG
AGGAATAATGCCGTGAAGCGGCACCTGCAGCAGCGGCAGCAGGAGGCGGATTTCCACGACAA
GAAGGTCATCGCGTCCACCTACTTCAGCATCGGCGCGTTCCTGGTGCTCGCCTGCCTCACCTTC
TCGCTGCTCATCCTGCCGCTGGTGCTGCCGCCGCTGCCGCCGCCGCCGTCGCTGCTGCTGTGGC
TGCCGGTCTGCCTGCTCGTCCTGCTGGTTGTGCTGGCCTTCATGCCGACAGATGTGCGCAGCGT
GGCGGCCTCTTACTTGTAAATAGCCAGATAAATAGGCCACCACCTTTGGCCAGTTTTCTGTGTT
TTCGGGGGTGATTCGTCCTAAGATGAGTCATGATCGAGTGTAATGTGAAGCATCTTTTCCAGG
GGTAGTAGATTTCAATGAAGT

SEQ ID N° : 290, NM_115853.3| proteína desconhecida de Arabidopsis thaliana (AT3G59900) mRNA, cds completos, que codifica SEQ ID N° : 270

TTGTCTTCCTCATTTCCCTACTAGTACTTGTTTCACACAGTTTCTTGATCCAACCAAAACCAAT
ACACAAAGCTTCTCAAACTCCTTCACCTCAAAGCTTCTTCCTTTACATCTGAATCGTTGAGTTA
ACTCGGATTTGTTCTGCATCCTCTGTTTCTGAATCGTGGGCCATCCTTATTTTGTCTCGAATTCT
TCACCAATTGCTTCGATCAAGCTGCATTGGTTAACCAGTTGCCCTAAAGATCAGATCTTTGAG
CAAAATTTTGTCACTGATCTTCTAAATCCAAACCAGACACAGCAAAACAACCTCTGTAGATGA
TTCGAGAAATCTCAAACTTACAAAAAGATATTATAAACATTCAAGACAGTTATTCGAACAACC
GAGTCATGGACGTCGGAAGAAACAACCGGAAAAACATGAGCTTTCGAAGTTCGCCGGAGAAA
AGCAAGCAAGA

FIGURA 27 (continuação)

GTTACGGCGGAGTTTCTCGGCGCAGAAAAGGATGATGATCCCGGCGAATTATTTCAGTTTAGA
GTCTCTGTTCCTATTGGTTGGTCTAACGGCATCTCTGTTAATACTTCCGTTAGTTTTGCCGCCGT
TACCTCCGCCTCCGTTTATGCTGCTATTGGTTCCCATTGGGATTATGGTTTTACTCGTCGTTCTT
GCCTTCATGCCTTCTTCTCATTCTAATGCTAATACAGATGTAACTTGCAATTTCATGTAAATCT
GAAATTTATTATATGATGAT

SEQ ID N° : 291, NM_180078.2| proteína desconhecida de Arabidopsis thaliana (AT2G44080) mRNA, cds completos, que codifica SEQ ID N° : 271

CCCACTTTATTTCTTCTTCTCTGGTTTTCTTACCAAAAGAAACTTTCTTCGTCTTCCTCTGTATTT
AAGCTTTAACACCCTGTTTTTGGTTTCCAACGTTCAATCTTCATCTTCTTCTCGCTGAAGGTGTG
TTTGGCTCTAACGGTTTAAAGCTTTCTTGAAACACCAATTGAATCTTTTCTCTCTACCGGCAAA
AAAAAAAGATTAGTCCTTTTAGGTCTGGAAACGCCAAGATCACTCGTTCTAAACCTTAGATTT
TGTCTGCATTTCGGGATAATCATTTCATCGTCAGGGTTCTTCAACCAAACTACATTTACAGAAG
AAGAAGAAGAAGAAAAGTTCGTTACTTTTTATGCGTTTGGATAAACAAACTCAAGTTTCTTCT
TCATACATCGATCTGATTTTCCAGATCAAACTTCGAAAAGAGAAAAAGCCTTCTTTAAATGAT
TCGTGAGTTCTCCAGTCTACAAAACGACATCATAAACATTCAAGAACATTATTCTCTCAACAA
CAACATGGACGTGAGAGGAGATCATAACCGGAAAAACACGAGTTTTCGTGGTTCAGCTCCAG
CTCCGATTATGGGGAAGCAAGAATTGTTTCGGACATTGTCGTCGCAGAACAGTCCAAGGAGG
CTAATATCAGCGAGTTACTTCAGTTTAGAATCAATGGTTGTGCTTGTTGGTCTCACAGCATCTC
TCTTGATCTTACCGTTGATTCTTCCACCATTGCCTCCTCCTCCTTTTATGCTGCTTTTGATTCCTA
TTGGGATTATGGTTTTGCTTATGGTTCTTGCTTTCATGCCTTCTTCTAATTCCAAACATGTTTCT
TCTTCTTCCACTTTTATGTAATAAACGTTTCTTTAATTGAAGAAAGAAATCCTTAAACAAA

SEQ ID N° : 292, AY305869.1| mRNA de proteína indutível por auxina de Arabidopsis thaliana (ARGOS) mRNA, cds completos, que codifica SEQ ID N° : 276

TTGTCTTCCTCATTTCCCTACTAGTACTTGTTTCACACAGTTTCTTGATCCAACCAAAACCAAT
ACACAAAGCTTCTCAAACTCCTTCACCTCAAAGCTTCTTCCTTTACATCTGAATCGTTGAGTTA
ACTCGGATTTGTTCTGCATCCTCTGTTTCTGAATCGTGGGCCATCCTTATTTTGTCTCGAATTCT
TCACCAATTGCTTCGATCAAGCTGCATTGGTTAACCAGTTGCCCTAAAGATCAGATCTTTGAG
CAAAATTTTGTCACTGATCTTCTAAATCCAAACCAGACACAGCAAAACAACCTCTGTAGATGA
TTCGAGAAATCTCAAACTTACAAAAAGATATTATAAACATTCAAGACAGTTATTCGAACAACC
GAGTCATGGACGTCGGAAGAAACAACCGGAAAAACATGAGCTTTCGAAGTTCGCCGGAGAAA
AGCAAGCAAGAGTTACGGCGGAGTTTCTCGGCGCAGAAAAGGATGATGATCCCGGCGAATTA
TTTCAGTTTAGAGTCTCTGTTCCTATTGGTTGGTCTAACGGCATCTCTGTTAATACTTCCGTTAG
TTTTGCCGCCGTTACCTCCGCCTCCGTTTATGCTGCTATTGGTTCCCATTGGGATTATGGTTTTA
CTCGTCGTTCTTGCCTTCATGCCTTCTTCTCATTCTAATGCTAATACAGATGTAACTTGCAATTT
CATGTAAATCTGAAATTTATTATATGATGAT

SEQ ID N° : 293, promotor de HMG, variante

CATGCGGCTAATGTAGATGCTCACTGCGCTAGTAGTAAGGTACTCCAGTACATTATGGAATAT
ACAAAGCTGTAATACTCGTATCAGCAAGAGAGAGGCACACAAGTTGTAGCAGTAGCACAGGA
TTAGAAAAACGGGACGACAAATAGTAATGGAAAAACAAAAAAAAAACAAGGAAACACATGGC
AATATAAATGGAGAAATCACAAGAGGAACAGAATCCGGGCAATACGCTGCGAAAGTACTCGT
ACGTAAAAAAAAGAGGCGCATTCATGTGTGGACAGCGTGCAGCAGAAGCAGGGATTTGAAAC
CACTCAAATCCACCACTGCAAACCTTCAAACGAGGCCATGGTTTGAAGCATAGAAAGCACAG
GTAAGAAGCACAACGCCCTCGCTCTCCACCCTCCCACCCAATCGCGACGCACCTCGCGGATCG
GTGACGTGGCCTCGCCCCCCAAAAATATCCCGCGGCGTGAAGCTGACACCCCGGGCCCACCC
ACCTGTCACGTTGGCACATGTTG

FIGURA 27 (continuação)

GTTATGGTTCCCGGCCGCACCAAAATATCAACGCGGCGCGGCCCAAAATTTCCAAAATCCCGC
CCAAGCCCCTGGCGCGTGCCGCTCTTCCACCCAGGTCCCTCTCGTAATCCATAATGGCGTGTGT
ACCCTCGGCTGGTTGTACGTGGGCGGGTTACCCTGGGGGTGTGGGTGGATGACGGGTGGGCCC
GGAGGAGGTCCGGCCCCGCGCGTCATCGCGGGGCGGGGTGTAGCGGGTGCGAAAAGGAGGC
GATCGGTACGAAAATTCAAATTAGGAGGTGGGGGGCGGGGCCCTTGGAGAATAAGCGGAATC
GCAGATATGCCCCTGACTTGGCTTGGCTCCTCTTCTTCTTATCCCTTGTCCTCGCAACCCCGCTT
CCTTCTCTCCTCTCCTCTTCTCTTCTCTTCTCTGGTGGTGTGGGTGTGTCCCTGTCTCCCCTCTCC
TTCCTCCTCTCCTTTCCCCTCCTCTCTTCCCCCCTCTCACAAGAGAGAGAGCGCCAGACTCTCC
CCAGGTGAGGTGAGACCAGTCTTTTTGCTCGATTCGACGCGCCTTTCACGCCGCCTCGCGCGG
ATCTGACCGCTTCCCTCGGCCTTCTCGCAGGATTCAGCC

FIGURA 27 (continuação)

RESUMO

"MÉTODO PARA AUMENTAR A RESISTÊNCIA À TENSÃO ABIÓTICA EM PLANTAS COM RELAÇÃO ÀS PLANTAS DE CONTROLE, E, USO DE UMA CONSTRUÇÃO"

A presente invenção diz respeito no geral ao campo da biologia molecular e diz respeito a um método para melhorar várias características do crescimento vegetal pela modulação da expressão em uma planta de um ácido nucleico que codifica um GRP (Proteína Relacionada com o Crescimento). A presente invenção também diz respeito a plantas tendo expressão modulada de um ácido nucleico que codifica um GRP, cujas plantas tem características de crescimento melhoradas com relação às plantas do tipo selvagem correspondentes ou outras plantas de controle. A invenção também fornece construções úteis nos métodos da invenção. A GRP pode ser uma das seguintes: Quinase semelhante ao receptor de extensina (ERLK), polipeptídeo F-Box WD40 (FBXW), Proteína de ligação de RAN (RANBP), Fator de Transcrição semelhante a Golden2 (GLK), Polipeptídeo de domínio zíper de leucina de homeodomínio REV delta, proteína CLE e proteína reguladora do crescimento de semente (SYR).

A requerente apresenta novas vias das reivindicações para melhor esclarecer e definir o presente pedido.

REIVINDICAÇÕES

1. Método para aumentar a resistência à tensão abiótica em plantas com relação às plantas de controle, caracterizado pelo fato de que compreende modular a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo SYR, cujo polipeptídeo SYR compreende um domínio rico em leucina, precedido pelo motivo 1 de tripeptídeo conservado (um de SEQ ID N° : 256, 257, 258 ou 259)) e seguido pelo motivo 2 conservado (SEQ ID N° : 260), em que a dita resistência à tensão abiótica aumentada é a eficiência de captação de nutriente aumentada, com relação às plantas de controle.

2. Método de acordo com a reivindicação 1, caracterizado pelo fato de que o dito polipeptídeo SYR tem, em ordem crescente de preferência, pelo menos 27%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ou mais identidade de seqüência para o polipeptídeo SYR representado por SEQ ID N° : 252.

3. Método de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que o dito ácido nucleico que codifica um polipeptídeo SYR é representado por qualquer um do ácido nucleico SEQ ID NOs dadas na Tabela II ou uma porção destes ou uma seqüência capaz de hibridizar com qualquer um dos ácidos nucleicos das SEQ ID NOs dadas na Tabela II.

4. Método de acordo com qualquer uma das reivindicações de 1 a 3, caracterizado pelo fato de que a dita seqüência de ácido nucleico codifica um ortólogo ou parálogo de qualquer uma das SEQ ID NOs dadas na Tabela II.

5. Método de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que a dita proteína SYR além disso compreende o motivo 3 conservado (SEQ ID N° : 261).

6. Método de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que a dita eficiência de captação de

nutriente resulta no rendimento aumentado de semente e/ou biomassa aumentada.

7. Método de acordo com a reivindicação 6, caracterizado pelo fato de que o dito rendimento aumentado de semente compreende pelo menos o peso total aumentado de sementes, peso de mil núcleos e/ou número aumentado de sementes enchidas.

8. Método de acordo com a reivindicação 6, caracterizado pelo fato de que a dita biomassa aumentada é biomassa de broto aumentada e/ou biomassa de raiz aumentada.

9. Método de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que a dita eficiência de captação de nutriente aumentada ocorre sob condições de seca brandas.

10. Método de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que a dita expressão modulada é realizada por introdução e expressão em uma planta de um ácido nucleico que codifica um polipeptídeo SYR.

11. Método de acordo com a reivindicação 10, caracterizado pelo fato de que o dito ácido nucleico está operacionalmente ligado a um promotor constituído, preferivelmente a um promotor GOS2.

12. Método de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que o dito ácido nucleico que codifica um polipeptídeo SYR é de origem vegetal, preferivelmente de uma planta monocotiledônea, ainda preferivelmente da família Poaceae, mais preferivelmente do gênero Oryza, mais preferivelmente de *Oryza sativa*.

13. Uso de uma construção, caracterizado pelo fato de ser em um método para a fabricação de plantas tendo resistência à tensão abiótica aumentada, dita construção compreendendo

(a) ácido nucleico que codifica um polipeptídeo SYR como definido em qualquer um das reivindicações de 1 a 5;

(b) uma ou mais seqüências de controle capazes de conduzir a expressão da seqüência de ácido nucleico de (a); e opcionalmente

(c) uma seqüência de terminação de transcrição,

e em que um das ditas seqüências de controle é um promotor constituído, preferivelmente um promotor GOS2 e em que a dita resistência à tensão abiótica aumentada é a eficiência de captação de nutriente aumentada, com relação às plantas de controle.