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

(21) **PI0619242-4 A2**

BRPI0619242A2

(22) Data de Depósito: 29/11/2006
(43) Data da Publicação: 08/11/2011
(RPI 2131)

(51) Int.Cl.:
C12N 5/10
C12N 15/82
C07K 14/415

(54) Título: PROTEÍNA, MÉTODOS PARA AUMENTAR PRODUÇÃO DE SEMENTE E/OU AUMENTAR TAXA DE CRESCIMENTO DE PLANTAS E TAXA DE CRESCIMENTO, CÉLULA VEGETAL, CONSTRUÇÃO, E, MÉTODOS PARA PRODUZIR UMA PLANTA TRANSGÊNICA, PARA MELHORAR CARACTERÍSTICAS DE CRESCIMENTO DE PLANTAS E CARACTERÍSTICAS DE CRESCIMENTO EM RELAÇÃO A CORRESPONDENTES PLANTAS DO TIPO SELVAGEM, PARA AUMENTAR PRODUÇÃO VEGETAL, NÚMERO DE SEMENTES EM PLANTAS E PRODUÇÃO DE SEMENTE EM PLANTAS

(30) Prioridade Unionista: 01/12/2005 EP 05111597.0, 05/12/2005 EP 05111691.1, 07/12/2005 EP 05111786.9, 12/12/2005 EP 05111996.4, 21/12/2005 EP 05112562.3, 30/12/2005 EP 05113110.0, 30/12/2005 EP 05113111.8, 05/12/2005 US 60/742352, 08/12/2005 US 60/748903, 09/12/2005 US 60/749219, 14/12/2005 US 60/750143, 23/12/2005 US 60/753650, 04/01/2006 US 60/756042, 04/01/2006 US 60/756086, 05/12/2005 US 60/742352, 04/01/2006 US 60/756042, 04/01/2006 US 60/756086, 12/12/2005 EP 05111996.4, 21/12/2005 EP 05112562.3, 30/12/2005 EP 05113110.0, 30/12/2005 EP 05113111.8, 05/12/2005 US 60/742352, 08/12/2005 US 60/748903, 09/12/2005 US 60/749219, 14/12/2005 US 60/750143, 23/12/2005 US 60/753650, 05/12/2005 EP 05111691.1, 07/12/2005 EP 05111786.9, 05/12/2005 US 60/742352, 04/01/2006 US 60/756042, 04/01/2006 US 60/756086

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

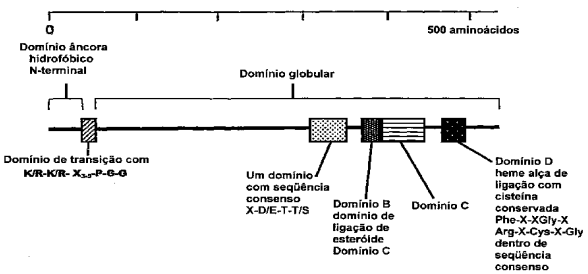
(72) Inventor(es): Ana Isabel Sanz Molinero, Christian Dammann, Christophe Reuzeau, Valerie Frankard

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

(86) Pedido Internacional: PCT US2006045721 de 29/11/2006

(87) Publicação Internacional: WO 2007/064724de 07/06/2007

(57) Resumo: PROTEÍNA, MÉTODOS PARA AUMENTAR PRODUÇÃO DE SEMENTE E/OU AUMENTAR TAXA DE CRESCIMENTO DE PLANTAS E TAXA DE CRESCIMENTO, CÉLULA VEGETAL, CONSTRUÇÃO, E, MÉTODOS PARA PRODUZIR UMA PLANTA TRANSGÊNICA, PARA MELHORAR CARACTERÍSTICAS DE CRESCIMENTO DE PLANTAS E CARACTERÍSTICAS DE CRESCIMENTO EM RELAÇÃO A CORRESPONDENTES PLANTAS DO TIPO SELVAGEM, PARA AUMENTAR PRODUÇÃO VEGETAL, NÚMERO DE SEMENTES EM PLANTAS E PRODUÇÃO DE SEMENTE EM PLANTAS. A presente invenção geralmente se refere ao campo de biologia molecular e se relaciona a um método para melhorar várias características de crescimento vegetal modulando expressão em uma planta de um ácido nucleico codificando uma GRP (Proteína Relacionada a Crescimento). A presente invenção também se relaciona a plantas tendo expressão modulada de um ácido nucleico codificando uma GRP, cujas plantas tem características de crescimento melhoradas em relação a correspondentes plantas do tipo selvagem ou outras plantas de controle. A invenção também fornece construções úteis nos métodos da invenção. A GRP pode ser uma das seguintes: Regulador de Produção de Sementes (SYR), FG- GAP, CYP90B, CDC27, fatores de transcrição AT-hook, fatores de transcrição DOF e Inibidores de Quinase Dependentes de Ciclina (CKIs).



 PI0619242-4

"PROTEÍNA, MÉTODOS PARA AUMENTAR PRODUÇÃO DE SEMENTE E/OU AUMENTAR TAXA DE CRESCIMENTO DE PLANTAS E TAXA DE CRESCIMENTO, CÉLULA VEGETAL, CONSTRUÇÃO, E, MÉTODOS PARA PRODUZIR UMA PLANTA TRANSGÊNICA, PARA MELHORAR CARACTERÍSTICAS DE CRESCIMENTO DE PLANTAS E CARACTERÍSTICAS DE CRESCIMENTO EM RELAÇÃO A CORRESPONDENTES PLANTAS DO TIPO SELVAGEM, PARA AUMENTAR PRODUÇÃO VEGETAL, NÚMERO DE SEMENTES EM PLANTAS E PRODUÇÃO DE SEMENTE EM PLANTAS"

A presente invenção geralmente se refere ao campo de biologia molecular e se relaciona a um método para melhorar várias características de crescimento vegetal modulando expressão em uma planta de um ácido nucleico codificando uma GRP (Proteína Relacionada a Crescimento). A presente invenção também se relaciona a plantas tendo expressão modulada de um ácido nucleico codificando uma GRP, cujas plantas têm características de crescimento melhoradas em relação a plantas tipo selvagem correspondentes ou outras plantas de controle. A invenção também fornece construções úteis nos métodos da invenção.

Dada a sempre crescente população mundial, e a área diminuída de terra disponível para agricultura, permanece como um principal

Segue-se a página 2"

objetivo de pesquisa melhorar a eficiência de agricultura e melhorar a diversidade de plantas em horticultura. Meios convencionais para melhoramentos culturais e horticulturais utilizam técnicas de melhoramento seletivo para identificar plantas tendo características desejáveis. Entretanto, tais técnicas de melhoramento seletivo têm diversas desvantagens, em outras palavras que estas técnicas são tipicamente de trabalho intensivo e resultam em plantas que freqüentemente contêm complementos genéticos heterogêneos que podem nem sempre resultar na característica desejável sendo passada adiante a partir de plantas parentais. Avanços em biologia molecular permitiram ao ser humano manipular o germoplasma de animais e plantas. Engenharia genética de plantas exige o isolamento e manipulação de material genético (tipicamente na forma de DNA ou RNA) e a subseqüente introdução de tal material genético em uma planta. Tal tecnologia levou ao desenvolvimento de plantas tendo várias características econômicas, agronômicas ou horticulturais melhoradas. Característica de particular interesse econômico são características de crescimento tal como alta produção. Produção normalmente é definida como a produção mensurável de valor econômico de uma safra. Isto pode ser definido em termos de quantidade e/ou qualidade. Produção é diretamente dependente de diversos fatores, por exemplo, o número e tamanho dos órgãos, arquitetura vegetal (por exemplo, o número de galhos), produção de sementes e mais. Desenvolvimento de raiz, absorção de nutrientes e tolerância a estresse também podem ser fatores importantes para determinar produção.

Produção de sementes é uma característica particularmente importante, uma vez que as sementes de muitas plantas são importantes para nutrição humana e animal. Safras tal como, milho, arroz, trigo, canola e soja respondem por mais da metade do consumo calórico humano total, quer através de consumo das próprias sementes ou através de consumo de produtos de carne criados com sementes processadas. Elas também são uma fonte de

açúcares, óleos e muitos tipos de metabólitos usados em processos industriais. Sementes contêm um embrião (a fonte de novos brotos e raízes) e um endosperma (a fonte de nutrientes para crescimento de embrião durante germinação e durante crescimento precoce de plântulas). O desenvolvimento de uma semente envolve muitos genes, e requer a transferência de metabólitos das raízes, folhas e caules para a semente em desenvolvimento. O endosperma, em particular, assimila os precursores metabólicos de carboidratos, óleos e proteínas e sintetiza estes em macromoléculas de armazenamento para preencher o grão.

Outra importante característica para muitas safras é vigor precoce. Melhorar vigor precoce é um importante objetivo de programas modernos de melhoramento de arroz tanto em cultivares de arroz temperados como tropicais. Raízes longas são importantes para ancoramento apropriado ao solo em arroz semeado em água. Onde arroz é semeado diretamente em campos alagados, e onde plantas devem emergir rapidamente através de água, brotos maiores estão associados com vigor. Onde sementeira em sulcos é praticada, mesocótilos e coleóptilos maiores são importantes para boa emergência de plântula. Vigor precoce também pode resultar de aptidão vegetal aumentado devido a, por exemplo, as plantas serem melhor adaptadas ao seu ambiente (i.e. sendo mais capazes de lidar com vários fatores de estresse abiótico ou biótico). Plantas tendo vigor precoce também mostram melhor estabelecimento da safra (com a safra crescendo de uma maneira mais uniforme, i.e. com a maioria das plantas atingindo os vários estágios de desenvolvimento substancialmente ao mesmo tempo), e mostram melhor crescimento e freqüentemente melhor produção.

Uma característica importante adicional é aquela de tolerância melhorada a estresse abiótico. Estresse abiótico é uma causa primária de perda de safra ao redor do mundo, reduzindo produções médias para a maioria das principais cultivares em mais do que 50% (Wang *et al.*, Planta (2003)

218: 1-14). Estresses abióticos podem ser causados por seca, salinidade, extremos de temperatura, toxicidade química e estresse oxidativo. A capacidade de melhorar tolerância vegetal a estresse abiótico seria de grande vantagem econômica para fazendeiros ao redor do mundo e iria permitir o cultivo de safras durante condições adversas e em territórios onde cultivo de safras pode não ser possível de outra maneira.

Produção de safra portanto pode ser aumentada otimizando um dos fatores mencionados acima.

Dependendo da utilização, a modificação de certas características de produção pode ser favorecida sobre outras. Por exemplo para aplicações tal como produção de forragem ou madeira, ou recurso de biocombustível, um aumento nas partes folhosas de uma planta pode ser desejável, e para aplicações tal como produção de trigo, amido ou óleo, um aumento em parâmetros de semente pode ser particularmente desejável. Mesmo entre os parâmetros de semente, alguns podem ser favorecidos sobre outros, dependendo da aplicação. Vários mecanismos podem contribuir para aumentar produção de sementes, quer isto seja na forma de tamanho de semente aumentado ou número de sementes aumentado.

Uma abordagem para aumentar produção (de sementes) em plantas pode ser através de modificação dos mecanismos de crescimento inerentes de uma planta. Um tal mecanismo é o ciclo celular.

Foi verificado agora que várias características de crescimento podem ser melhoradas em plantas modulando expressão em uma planta de um ácido nucleico codificando uma GRP (Proteína Relacionada a Crescimento) em uma planta. A GRP pode ser uma das seguintes: Regulador de Produção de Sementes (SYR), FG-GAP, CYP90B, CDC27, fatores de transcrição AT-hook, fatores de transcrição DOF e Inibidores de Quinase Dependentes de Ciclina (CKIs).

FUNDAMENTOS

Regulador de Produção de Sementes (SYR)

Existe uma necessidade contínua de encontrar novos genes para estimulação de produção de sementes e diversas abordagens foram usadas até agora, por exemplo através de manipulação de níveis de hormônios vegetais (WO 03/050287), através de manipulação do ciclo celular (WO 2005/061702), através de manipulação de genes envolvidos em resposta a estresse salino (WO 2004/058980) entre outras estratégias.

SYR é uma proteína nova que até agora não foi caracterizada. SYR mostra alguma homologia (cerca de 48% de identidade de seqüência no nível de DNA, cerca de 45% no nível de proteína) com uma proteína de *Arabidopsis* chamada ARGOS (Hu et al., Plant Cell 15, 1951-1961, 2003; Patente Norte-Americana 2005/0108793). Hu et al. postularam que ARGOS é uma proteína de função exclusiva e é codificada por um único gene. Os principais fenótipos de superexpressão de *ARGOS* em *Arabidopsis* são biomassa de folhas aumentada e florescimento atrasado.

FG-GAP

Proteínas FG-GAP são proteínas transmembrana putativas. Elas são caracterizadas pela presença de um ou mais domínios FG-GAP (Pfam número de acesso PF01839) e pela presença de um peptídeo sinal N-terminal e um domínio transmembrana na metade C-terminal da proteína.

Uma tal proteína, DEX1, foi isolada de *Arabidopsis* e foi relatada desempenhando um papel durante desenvolvimento de pólen (Paxson-Sowders et al. Plant physiol. 127, 1739-1749, 2001). Plantas *Dex1* mutantes mostram ser defeituosas em formação padrão de parede de pólen. O gene *DEX1* codifica uma proteína de 896 aminoácidos que é predita localizar a membrana plasmática, com resíduos 1 até 860 sendo localizados fora da célula, resíduos 880 até 895 no lado citoplasmático da membrana, e aminoácidos 861 até 879 representando um domínio transmembrana potencial. Doze sítios potenciais de N-glicosilação estão presentes em DEX1.

Portanto, a proteína tem o potencial de ser fortemente modificada e interagir com vários componentes da parede celular. DEX1 mostra a maior similaridade de seqüência com uma proteína tipo hemolisina de *V. cholerae*, ao passo que um segmento de aproximadamente 200 aminoácidos de DEX1 (aminoácidos 439-643) também mostra similaridade limitada com o domínio de ligação de cálcio de alfa-integrinas. Nesta região estão pelo menos dois conjuntos de ligantes putativos de ligação de cálcio que também estão presentes em uma proteína calmodulina de *Arabidopsis* prevista (AC009853). Portanto, parece que DEX1 pode ser uma proteína de ligação de cálcio. DEX1 parece ser uma proteína vegetal exclusiva; homólogos não estão presentes em bactérias, fungos, ou animais.

As alterações observadas em plantas *dex1*, assim como a estrutura prevista de DEX1, geram diversas possibilidades para o papel da proteína em formação de parede de pólen (Paxson-Sowders et al., 2001):

(a) DEX1 pode ser uma proteína ligante. Ela pode estar associada com a membrana de micrósporo e participar em acoplamento ou da primexina ou esporopolenina à membrana plasmática. Ausência da proteína da superfície de micrósporo pode resultar em alterações estruturais na primexina. S numerosos sítios potenciais de N-glicosilação são consistentes com acoplamento de DEX1 à parede de calose, à intina, ou ambas.

(b) DEX1 pode ser um componente da matriz de primexina e desempenha um papel na polimerização inicial da primexina. Mudanças em concentrações de íon Ca+2 parecem ser importantes para síntese de parede de pólen; beta-glucana sintase é ativada por concentrações micromolares de Ca+2 durante formação de parede de calose.

(c) DEX1 pode ser parte do ER bruto e pode estar envolvido em processamento e/ou transporte de precursores de primexina para a membrana. O aparecimento atrasado e alterações gerais na primexina são consistentes com uma ausência geral de precursores de primexina. A matriz

de primexina é inicialmente composta de polissacarídeos, proteínas, e celulose, seguido pela incorporação de materiais mais resistentes. Portanto, DEX1 pode participar na formação ou transporte de qualquer número de componentes diferentes.

CYP90B

Brassinoesteróides (BRs) são uma classe de hormônios vegetais que são importantes para promover crescimento, divisão e desenvolvimento vegetal. O termo BR coletivamente se refere a mais do que quarenta derivados de esterol poli-hidroxilado naturalmente ocorrentes, com similaridade estrutural com hormônios esteróides animais. Dentre estes, brassinolide mostrou ser o mais ativo biologicamente (para revisão, Clouse (2002) Brassinosteroids. The Arabidopsis Book: 1–23).

A via biossintética de BR foi elucidada usando análises bioquímicas e mutacionais. BRs são sintetizados através de pelo menos duas vias bioquímicas ramificadas começando a partir do mesmo precursor inicial, campesterol (Fujioka *et al.* (1997) physiol Plant 100:710-715). Os genes verificados de biossíntese de BR foram encontrados codificando principalmente monooxigenases de citocromo P450 (CYP) (Bishop and Yokota (2001) Plant Cell physiol 42:114-120). Superfamília CYP de enzimas catalisa oxidação de muitos produtos químicos, e no presente caso mais especificamente catalisa reações oxidativas essenciais na biossíntese de BRs. Uma das etapas importantes identificadas consiste na hidroxilação da cadeia lateral esteróide de intermediários de BR campestanol e 6-oxocampestanol para formar 6-deoxocatasterona e catasterona respectivamente. Estas duas etapas oxidativas paralelas também são coletivamente chamadas da etapa inicial de alfa-hidroxilação de esteróide C-22 (Choe *et al.* (1998) Plant Cell 10: 231-243). Em *Arabidopsis*, uma enzima CYP específica, CYP90B1 ou DWF4, realiza esta etapa (para referência geral em nomenclatura de CYP vegetal, Nelson *et al.* (2004) Plant pH ys 135: 756-772).

Plantas mutantes de Arabidopsis carecendo de atividade de alfa hidroxilase de esteróide 22 devido à inserção de um T-DNA no locus de DWF4 apresentaram um fenótipo definhado devido à carência de prolongamento celular (Choe *et al.* (1998) Plant Cell 10: 231-243). Estudos bioquímicos de alimentação com intermediários de biossíntese de BR mostraram que todos os compostos anteriores resgataram o fenótipo, ao passo que os precursores conhecidos fracassaram ao fazer isto.

Foram geradas plantas transgênicas de Arabidopsis e tabaco, ambas dicotiledôneas, que superexpressaram ectopicamente um fragmento genômico DWF4 de Arabidopsis, usando o promotor 35S de vírus do mosaico da couve-flor (Choe *et al.* (2001) Plant J 26(6): 573-582). Caracterização fenotípica das plantas mostrou que o comprimento de hipocótilo, altura de planta em maturidade, número total de galhos e número total de sementes foram aumentados nas transgênicas comparado com plantas de controle. Choe *et al.* encontrou que a produção aumentada de sementes foi devido a um maior número de sementes por planta, aumento de tamanho de semente estando dentro do intervalo de desvio padrão. Estes experimentos são adicionalmente descritos em WO00/47715.

Patente Norte-Americana 6.545.200 se refere a fragmentos de ácidos nucleicos isolados codificando genes biossintéticos de esterol, e mais especificamente reivindica uma seqüência de nucleotídeos codificando um polipeptídeo tendo atividade de C-8,7 esterol isomerase. Seqüências de nucleotídeos parciais codificando DWF4 são descritas.

Patente Norte-Americana 2004/0060079 se refere a um método de produzir uma planta monocotiledônea modificada tendo uma característica desejada. É fornecido um exemplo no qual a seqüência de nucleotídeos codificando DWF4 de arroz (referida ou como OsDWF4 ou CYP90B2) é colocada sob o controle de um promotor constitutivo, o promotor de actina de arroz. Quatorze das trinta e seis plantas transgênicas de

arroz expressando a construção quimérica mostram um número aumentado de sementes por estaca se comparadas com plantas de controle não transformadas. De acordo com os requerentes, o aumento de produção nas transgênicas comparadas com os tipos selvagens é devido a um aumento em número total de sementes, pois nenhuma diferença significante é encontrada no "peso de 10 grãos".

CDC27

Dependendo da utilização, a modificação de certas características de produção pode ser favorecida sobre outras. Por exemplo para aplicações tal como produção de forragem ou madeira, ou recurso de biocombustível, um aumento nas partes folhosas de uma planta pode ser desejável, e para aplicações tal como produção de trigo, amido ou óleo, um aumento em parâmetros de semente pode ser particularmente desejável. Mesmo entre os parâmetros de semente, alguns podem ser favorecidos sobre outros, dependendo da aplicação. Vários mecanismos podem contribuir para aumentar produção de sementes, quer isto seja na forma de tamanho de semente aumentado ou número de sementes aumentado. Um tal mecanismo é o ciclo celular.

Progressão através do ciclo celular é fundamental para o crescimento e desenvolvimento de todos os organismos multicelulares e é crucial para proliferação celular. Os principais componentes do ciclo celular são altamente conservados em levedura, mamíferos, e plantas. O ciclo celular é tipicamente dividido nas seguintes fases seqüenciais: G0 – G1 – S – G2 – M. Replicação ou síntese de DNA geralmente ocorre durante a fase S ("S" é para síntese de DNA) e segregação mitótica dos cromossomos ocorre durante a fase M (o "M" é para mitose), com fases de intervalos intervenientes, G1 (durante a qual células crescem antes de replicação de DNA) e G2 (um período depois de replicação de DNA durante o qual a célula se prepara para divisão). Divisão celular é completada depois de citocinese, a última etapa da

fase M. Células que saíram do ciclo celular e se tornaram quiescentes são ditas estarem na fase G0. Células nesta fase podem ser estimuladas para manter o ciclo celular na fase G1. O "G" em G1, G2 e G0 representa "intervalo". Conclusão do processo de ciclo celular permite que cada célula filha durante divisão celular receba uma cópia completa do genoma parental.

Divisão celular é controlada por dois eventos principais de ciclo celular, em outras palavras início de síntese de DNA e início de mitose. Cada transição para cada um destes eventos chave é controlada por um ponto de verificação representado por complexos protéicos específicos (envolvidos em replicação e divisão de DNA). A expressão de genes necessários para síntese de DNA no limite de G1/S é regulada pela família E2F de fatores de transcrição em mamíferos e células vegetais (La Thangue, 1994; Muller *et al.*, 2001; De Veylder *et al.*, 2002). Entrada no ciclo celular é regulada/desencadeada por um complexo E2F/Rb que integra sinais e permite ativação de transcrição de genes de ciclo celular. A transição entre as diferentes fases do ciclo celular, e portanto progressão através do ciclo celular, é dirigida pela formação e ativação de diferentes proteínas quinases heterodiméricas de serina/treonina, geralmente referidas como quinases dependentes de ciclina (CDKs). Um pré-requisito para atividade destas quinases é a associação física com uma ciclina específica, o momento de ativação sendo amplamente dependente de expressão de ciclina. Ligação de ciclina induz mudanças conformacionais no lobo N-terminal do CDK associante e contribui para a localização e especificidade de substrato do complexo. CDKs monoméricos são ativados quando eles estão associados com ciclinas e assim têm uma atividade de quinase. Níveis protéicos de ciclina flutuam no ciclo celular e portanto representam um fator principal para determinar momento de ativação de CDK. A ativação periódica destes complexos contendo ciclinas e CDK durante ciclo celular media a regulação temporal de transições de ciclo celular (momentos de verificação).

Mecanismos existem para assegurar que replicação de DNA ocorra apenas uma vez durante o ciclo celular. Por exemplo, proteínas CDC16, CDC23 e CDC27 são parte de um complexo de alto peso molecular conhecido como o complexo promotor de anáfase (APC) ou ciclossomo, (veja Romanowski and Madine, Trends in Cell Biology 6, 184-188, 1996, e Wuarin and Nurse, Cell 85, 785-787 (1996). O complexo em levedura é composto de pelo menos oito proteínas, das proteínas contendo TPR (repetição de tetratricopeptídeo) CDC16, CDC23 e CDC27, e cinco outras subunidades chamadas APC1, APC2, APC4, APC5 e APC7 (Peters et al. 1996, Science 274, 1199-1201). O APC direciona seus substratos para degradação proteolítica catalisando a ligação de moléculas de ubiquitina a estes substratos. Proteólise dependente de APC é requerida para a separação das cromátides-irmãs em transição de metáfase para anáfase e para a saída final de mitose Entre os substratos de APC estão a proteína inibidora de anáfase Pds1p e ciclinas mitóticas tal como ciclina B, respectivamente (Ciosk et al. 1998, Cell 93, 1067-1076; Cohen-Fix et al. 1996, Genes Dev 10, 3081-3093; Sudakin et al. 1995, Mol Biol Cell 6, 185-198; Jorgensen et al. 1998, Mol Cell Biol 18, 468-476; Townsley and Ruderman 1998, Trends Cell Biol 8, 238-244). Para se tornarem ativos como uma ligase de ubiquitina, pelo menos CDC16, CDC23 e CDC27 necessitam ser fosforilados na fase M (Ollendorf and Donoghue 1997, J Biol Chem 272, 32011-32018). APC ativado persiste ao longo de G1 do ciclo celular subseqüente para prevenir aparecimento prematuro de ciclinas tipo B, o que iria resultar em uma entrada incontrolada na fase S (Irniger and Nasmyth 1997, J Cell Sci 110, 1523-1531). Foi demonstrado em levedura que mutações em qualquer de pelo menos dois dos componentes de APC, CDC16 e CDC27, pode resultar em super-replicação de DNA sem passagens intervenientes através de fases M (Heichman and Roberts 1996, Cell 85, 39-48). Este processo de replicação de DNA nuclear sem subseqüente mitose e divisão celular é chamado endoreduplicação de

DNA, e leva a tamanho celular aumentado.

CDC16, CDC23 e CDC27 são todas proteínas contendo repetições de tetratricopeptídeo (TPR; 34 aminoácidos de comprimento). Uma seqüência consenso mínima sugerida do motivo TPR é como segue: X_3-W-X_2-L-G-X_2-Y-X_8-A-X_3-F-X_2-A-X_4-P-X_2, onde X é qualquer aminoácido (Lamb et al. 1994, EMBO J 13, 4321-4328). Os resíduos consenso podem exibir degenerescência significante e pouca ou nenhuma homologia está presente em resíduos não consenso. É a hidrofobicidade e o tamanho dos resíduos consenso, ao invés de sua identidade, que parece ser de importância. Motivos TPR estão presentes em uma ampla variedade de proteínas funcionais em levedura e eucariotos superiores em mitose (incluindo os componentes de proteínas APC CDC16, CDC23 e CDC27), transcrição, processamento, importação de proteínas e neurogênese (Goebl and Yanagida 1991, Trends Biochem Sci 16, 173-177). O TPR forma uma estrutura α–helicoidal; repetições em tandem se organizam em uma estrutura super-helicoidais idealmente adequada como interfaces para reconhecimento de proteínas (Groves and Barford 1999, Curr Opin Struct Biol 9, 383-389). Dentro da α–hélice, dois domínios anfipáticos estão normalmente presentes, um na região NH_2 terminal e o outro perto da região COOH terminal (Sikorski et al. 1990, Cell 60, 307-317).

CDC27 (também conhecido como Hobbit; outros nomes incluem CDC27, BimA, Nuc2 ou makos) foi isolado de vários organismos, incluindo *Aspergillus nidulans*, levedura, drosófila, humano e várias plantas (tal como *Arabidopsis thaliana* e *Oryza sativa*). O gene codificando CDC27 está presente como uma cópia única na maioria de genomas, mas duas cópias podem ser excepcionalmente encontradas dentro do mesmo genoma, por exemplo em *Arabidopsis thaliana*. Os dois genes codificando proteínas CDC27 foram nomeados CDC27A e CDC27B (referências MIPS At3g16320 e At2g20000 respectivamente).

Pedido de patente Internacional publicado, WO01/02430 descreve seqüências de CDC27A (CDC27A1 e CDC27A2) e CDC27B. Também descrita neste documento é uma seqüência de aminoácidos truncada de CDC27B na qual 161 aminoácidos estão ausentes da região NH_2 terminal. Referência é feita neste documento a número de acesso de GenBank AC006081 para o gene CDC27B codificando um polipeptídeo CDC27B truncado na região NH_2 terminal. O documento relata a região NH_2 terminal como sendo conservada em homólogos de CDC27 de diferente origem. As seqüências de CDC27 mencionadas em WO01/02430 são descritas como sendo úteis para modificar endoreduplicação.

Endoreduplicação de DNA ocorre naturalmente em plantas florescendo, por exemplo durante desenvolvimento de sementes. Endoreduplicação leva a núcleos aumentados com elevado teor de DNA. Foi sugerido que o aumentado teor de DNA durante endoreduplicação pode sustentar expressão gênica aumentada durante desenvolvimento de endosperma e enchimento de sementes, uma vez que ele coincide com atividade enzimática aumentada e acumulação protéica neste momento (Kowles et al., (1992) Genet. Eng. 14:65–88). Em espécies de cereal, o endosperma celular armazena as reservas da semente durante a fase marcada por endoreduplicação. A magnitude de endoreduplicação de DNA é altamente correlacionada com peso fresco de endosperma, o que implica um importante papel de endoreduplicação de DNA na determinação de massa de endosperma (Engelen-Eigles *et al.* (2000) Plant Cell Environ. 23:657–663). Em milho por exemplo, o endosperma constitui 70 a 90% de massa de semente; assim, fatores que mediam desenvolvimento de endosperma em um grande grau também determinam produção de grãos de milho, através de peso de semente individual. Endoreduplicação aumentada é portanto tipicamente indicativa de biomassa de semente aumentada mas não é de maneira nenhuma relacionada a número de sementes aumentado.

Fator de transcrição AT-hook

Um domínio AT-hook é encontrado em polipeptídeos pertencendo a uma família de fatores de transcrição associados com remodelagem de Cromatina. O motivo AT-hook é constituído de 13 ou mais (algumas vezes cerca de 9) aminoácidos que participam em ligação de DNA e que têm uma preferência por regiões ricas em A/T. Em *Arabidopsis* existem pelo menos 34 proteínas contendo domínios AT-hook. Estas proteínas compartilham homologia ao longo da maioria da seqüência, com o domínio AT-hook sendo uma região particularmente altamente conservada.

Pedido de Patente Internacional WO 2005/030966 descreve diversos fatores de transcrição vegetais compreendendo domínios AT-hook e o uso destes fatores de transcrição para produzir plantas tendo biomassa aumentada e tolerância a estresse aumentada. O pedido se relaciona a membros do clado G1073 de fatores de transcrição e states que, "Uso de promotores tecido específicos ou induzíveis mitiga efeitos morfológicos indesejáveis que podem estar associados com superexpressão constitutiva de membros de clado G1073 (e.g., quando tamanho aumentado é indesejável)." Os dados fornecidos neste pedido se referem a plantas dicotiledôneas.

Em contraste com estas instruções, foi verificado agora que expressão em uma planta monocotiledônea (monocotiledônea) de um ácido polinucleico codificando um fator de transcrição AT-hook compreendendo um domínio DUF296 (o qual inclui membros de clado G1073), gera plantas tendo pouco ou nenhum aumento em biomassa comparada com plantas de controle adequadas, independentemente de se tal expressão é dirigida por um promotor constitutivo ou de uma maneira tecido específica. Isto sugere que instruções referentes à expressão de tais fatores de transcrição em dicotiledôneas podem não ser prontamente aplicáveis para monocotiledôneas. Também foi encontrado agora que o grau ou natureza de qualquer aumento em produção de sementes obtido é dependente do promotor tecido específico

usado.

Fatores de transcrição DOF

Proteínas de domínio Dof são fatores de transcrição planta específicos com um domínio de ligação de DNA altamente conservado com um único dedo de zinco C_2-C_2. Durante a última década, numerosas proteínas de domínio Dof foram identificadas tanto em monocotiledôneas como dicotiledôneas incluindo milho, cevada, trigo, arroz, tabaco, *Arabidopsis*, abóbora, batata e ervilha. Proteínas de domínio Dof mostraram funcionar como ativadores ou repressores transcricionais em diversos processos biológicos planta específicos.

Inibidores de Quinase Dependentes de Ciclina (CKI)

A capacidade de aumentar produção de sementes de plantas, quer através de número de sementes, biomassa de semente, desenvolvimento de semente, enchimento de semente ou qualquer outra característica relacionada a sementes teria muitas aplicações em agricultura, e mesmo muitos usos não agrícolas tal como na produção biotecnológica de substâncias tal como produtos farmacêuticos, anticorpos ou vacinas. Uma abordagem para aumentar produção de sementes em plantas pode ser através de modificação dos mecanismos de crescimento inerentes de uma planta.

Os mecanismos de crescimento inerentes de uma planta residem em uma seqüência altamente ordenada de eventos coletivamente conhecida como o 'ciclo celular'. Progressão através do ciclo celular é fundamental para o crescimento e desenvolvimento de todos os organismos multicelulares e é crucial para proliferação celular. Os principais componentes do ciclo celular são altamente conservados em levedura, mamíferos, e plantas. O ciclo celular é tipicamente dividido nas seguintes fases seqüenciais: G0 – G1 – S – G2 – M. Replicação ou síntese de DNA geralmente ocorre durante a fase S ("S" é para síntese de DNA) e segregação mitótica dos cromossomos ocorre durante a fase M (o "M" é para mitose), com fases de intervalos

intervenientes, G1 (durante a qual células crescem antes de replicação de DNA) e G2 (um período depois de replicação de DNA durante o qual a célula se prepara para divisão). Divisão celular é completada depois de citocinese, a última etapa da fase M. Células que saíram do ciclo celular e que se tornaram quiescentes são ditas estarem na fase G0. Células nesta fase podem ser estimuladas para manter o ciclo celular na fase G1. O "G" em G1, G2 e G0 representa "lacuna". Conclusão do processo de ciclo celular permite que cada célula filha durante divisão celular receba uma cópia completa do genoma parental.

Divisão celular é controlada por dois eventos principais de ciclo celular, em outras palavras início de síntese de DNA e início de mitose. Cada transição para cada um destes eventos chave é controlada por um ponto de verificação representado por complexos protéicos específicos (envolvidos em replicação e divisão de DNA). A expressão de genes necessários para síntese de DNA no limite de G1/S é regulada pela família E2F de fatores de transcrição em mamíferos e células vegetais (La Thangue, 1994; Muller *et al.*, 2001; De Veylder *et al.*, 2002). Entrada no ciclo celular é regulada/desencadeada por um complexo E2F/Rb que integra sinais e permite ativação de transcrição de genes de ciclo celular. A transição entre as diferentes fases do ciclo celular, e portanto progressão através do ciclo celular, é dirigida pela formação e ativação de diferentes proteínas quinases heterodiméricas de serina/treonina, geralmente referidas como quinases dependentes de ciclina (CDKs). Um pré-requisito para atividade destas quinases é a associação física com uma ciclina específica, o momento de ativação sendo amplamente dependente de expressão de ciclina. Ligação de ciclina induz mudanças conformacionais no lobo N-terminal do CDK associante e contribui para a localização e especificidade de substrato do complexo. CDKs monoméricos são ativados quando eles estão associados com ciclinas e assim têm uma atividade de quinase. Níveis protéicos de

ciclina normalmente flutuam no ciclo celular e portanto representam um fator principal para determinar momento de ativação de CDK. A ativação periódica destes complexos contendo ciclinas e CDK durante ciclo celular media a regulação temporal de transições de ciclo celular (momentos de verificação). Outros fatores regulando atividade de CDK incluem Inibidores de Quinase Dependentes de Ciclina (CKIs ou ICKs, KIPs, CIPs, INKs), quinases que ativam CDK (CAKs), uma CDK fosfatase (Cdc25) e uma subunidade de CDK (CKS) (Mironov *et al.* 1999; Reed 1996).

A existência de um inibidor de CDKs mitóticos foi inferida a partir de experimentos com endosperma de semente de milho (Grafi and Larkins (1995) *Science* 269, 1262-1264). Desde então, diversos CKIs foram identificados em várias espécies vegetais, tal como *Arabidopsis* (Wang *et al.* (1997) Nature 386(6624): 451-2; De Veylder *et al.* (2001) Plant Cell 13: 1653–1668; Lui *et al.* (2000) Plant J 21: 379–385), tabaco (Jasinski et al. (2002) Plant physiol 2002 130(4): 871-82), *Chenopodium rubrum* (Fountain et al. (1999) Plant pH ys 120: 339) ou milho (Coelho *et al.* (2005) Plant physiol 138: 2323–2336). As proteínas codificadas são caracterizadas por um trecho de aproximadamente 45 aminoácidos carbóxi-terminais mostrando homologia com o domínio de ligação amino-terminal de ciclina/Cdk de CKIs animais dos tipos $p21^{Cip1}/p27^{Kip1}/p57^{Kip2}$. For a desta região carbóxi-terminal, CKIs vegetais mostram pouca homologia.

Pedido de patente Internacional publicado WO 2005/007829 no nome de Monsanto Technology LLC descreve várias moléculas de ácidos nucleicos isolados codificando polipeptídeos tendo atividade de inibidor de quinase dependente de ciclina.

Pedidos de patentes Internacionais publicados, WO 02/28893 e WO 99/14331, ambos no nome de CropDesign N.V., descrevem vários Inibidores de Quinase Dependentes de Ciclina vegetais. O uso destes inibidores para aumentar produção é mencionado nestes pedidos.

SUMÁRIO DA INVENÇÃO

Foi surpreendentemente verificado agora que aumentar atividade de uma proteína SYR e/ou expressão de um ácido nucleico codificando uma proteína SYR em plantas resulta em plantas tendo produção aumentada de sementes e ou taxa de crescimento aumentada, em relação a plantas tipo selvagem correspondentes. Também foi surpreendentemente verificado agora que superexpressão de *SYR* em camundongos aumenta primariamente produção de sementes, ao passo que biomassa de folhas e tempo de florescimento não são obviamente afetados (em contraste com os principais fenótipos de superexpressão de *ARGOS* em *Arabidopsis*, que mostraram ter biomassa de folhas aumentada e florescimento atrasado (Hu et al., Plant Cell 15, 1951-1961, 2003; US 2005/0108793)).

De acordo com uma forma de realização da presente invenção é fornecido um método para aumentar produção de semente e/ou taxa de crescimento de uma planta compreendendo aumentar atividade de um polipeptídeo SYR ou um homólogo deste em uma planta e/ou expressão de um ácido nucleico codificando uma tal proteína; e opcionalmente selecionar plantas tendo características de crescimento melhoradas.

Vantajosamente, desempenho dos métodos da invenção na medida em que eles se relacionam a SYR, resulta em plantas tendo uma variedade de características de crescimento melhoradas, tal como produção de sementes aumentada sem efeito na biomassa de partes vegetativas de planta, quando comparada com plantas de controle correspondentes, e um ciclo de vida comparável a plantas de controle correspondentes, sem atraso em época de florescimento. Ademais vantajosamente, desempenho dos métodos de acordo com a presente invenção resulta em plantas tendo tolerância melhorada para estresse abiótico em relação a plantas tipo selvagem correspondente (ou outro controle).

Foi surpreendentemente verificado agora que modular

atividade de uma proteína FG-GAP e/ou expressão de um ácido nucleico codificando uma proteína FG-GAP em plantas resulta em plantas tendo características de crescimento melhoradas, e em particular produção aumentada, em relação a plantas tipo selvagem correspondentes.

De acordo com outra forma de realização da presente invenção é fornecido um método para melhorar características de crescimento de uma planta compreendendo modular atividade de um polipeptídeo FG-GAP ou um homólogo deste e/ou modular expressão de um ácido nucleico codificando um polipeptídeo FG-GAP ou um homólogo deste em uma planta e opcionalmente selecionar plantas tendo características de crescimento melhoradas.

Vantajosamente, desempenho dos métodos de acordo com a presente invenção, na medida em que eles se relacionam a um polipeptídeo FG-GAP ou um homólogo deste, resulta em plantas tendo uma variedade de características de crescimento melhoradas, tal como crescimento melhorado, produção melhorada, biomassa melhorada, arquitetura melhorada ou divisão celular melhorada, cada em relação a plantas tipo selvagem correspondentes. Preferivelmente, as características de crescimento melhoradas compreendem pelo menos produção aumentada em relação a plantas tipo selvagem correspondentes.

Foi surpreendentemente verificado agora que aumentar expressão não constitutiva em uma planta de um ácido nucleico codificando um polipeptídeo CYP90B ou um homólogo deste gera plantas tendo produção aumentada em relação a plantas de controle adequadas.

De acordo com uma forma de realização adicional da presente invenção, é fornecido um método para aumentar produção vegetal compreendendo aumentar expressão não constitutiva em uma planta de um ácido nucleico codificando um polipeptídeo CYP90B ou um homólogo deste.

Foi verificado agora que aumentar preferencialmente expressão no tecido de meristema apical de broto de plantas de um ácido

nucleico codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo gera plantas tendo número de sementes aumentado em relação a plantas de controle adequadas.

A invenção portanto fornece um método para aumentar o número de sementes de plantas em relação àquele de plantas de controle adequadas, compreendendo aumentar preferencialmente expressão em tecido de meristema apical de broto vegetal de um ácido nucleico codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo.

Foi verificado agora que aumentar preferencialmente expressão de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 em tecido de endosperma de uma planta monocotiledônea gera plantas tendo produção aumentada de sementes em relação a plantas de controle adequadas.

Uma forma de realização adicional da presente invenção portanto fornece um método para aumentar produção de semente em plantas monocotiledôneas em relação a plantas de controle adequadas, compreendendo aumentar preferencialmente expressão em tecido de endosperma de uma planta monocotiledônea de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296.

Foi verificado agora que aumentar expressão em uma planta de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF gera plantas tendo produção aumentada em relação a plantas de controle adequadas.

De acordo com uma forma de realização adicional da presente invenção, é fornecido um método para aumentar produção vegetal compreendendo aumentar expressão em uma planta de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF.

Foi verificado agora que redução preferencial em expressão de um gene CKI endógeno em tecido de endosperma de uma planta gera plantas com melhor produção de sementes do que produção de sementes em plantas onde não existe redução preferencial em expressão de um gene CKI endógeno em tecido de endosperma de planta. A presente invenção portanto fornece um método para aumentar produção de semente em plantas em relação a plantas de controle adequadas, compreendendo reduzir preferencialmente expressão de um gene CKI endógeno em tecido de endosperma de uma planta.

DESCRIÇÃO DETALHADA DA INVENÇÃO

O termo "produção aumentada" como definido neste lugar é adotado para significar um aumento em biomassa (peso) de uma ou mais partes de uma planta (particularmente partes aptas à colheita) em relação a tipo selvagem correspondente ou outras plantas de controle, cujo aumento em biomassa pode ser acima da terra ou subterrâneo. Um aumento em biomassa subterrânea pode ser devido a um aumento na biomassa de partes vegetais, tal como tubérculos, rizomas, bulbos etc. Particularmente preferido é um aumento em qualquer um ou mais dos seguintes: biomassa de raiz aumentada, volume de raiz aumentado, número de raiz aumentado, diâmetro de raiz aumentado e comprimento de raiz aumentado. O termo produção aumentado também abrange um aumento em produção de sementes.

O termo "produção aumentada de sementes" como definido neste lugar é adotado para significar um aumento em qualquer um ou mais dos seguintes, cada em relação a plantas tipo selvagem correspondentes: (i) produção total de sementes aumentada, o que inclui um aumento em biomassa de semente (peso de semente) e que pode ser um aumento no peso de semente por planta ou em uma base de semente individual; (ii) número aumentado de flores ("floretes") por panícula (iii) número aumentado de sementes cheias; (iv) tamanho de semente aumentado; (v) volume de semente aumentado; (vi) área de semente individual aumentada; (vii) comprimento e/ou largura de

semente individual aumentada; (viii) índice de colheita aumentado, que é expresso como uma proporção da produção de partes aptas à colheita, tal como sementes, pela biomassa total; (ix) taxa de enchimento aumentada, (que é o número de grãos cheios dividido pelo número total de sementes e multiplicado por 100); e (x) peso de mil sementes (TKW) aumentado, que é extrapolado a partir do número de grãos cheios contado e seu peso total. Um TKW aumentado pode resultar de um tamanho de semente aumentado e/ou peso de semente. Um TKW aumentado pode resultar de um aumento em tamanho de embrião e/ou tamanho de endosperma.

Tomando milho como um exemplo, um aumento de produção pode ser manifestado como um ou mais dos seguintes: um aumento no número de orelhas por planta, um aumento no número de fileiras, número de sementes por fileira, peso de semente, TKW, comprimento/diâmetro de espiga, entre outros. Tomando arroz como um exemplo, um aumento de produção pode ser manifestado por um aumento em um ou mais dos seguintes: número de panículas por planta, número de espiguilhas por panícula, número de flores por panícula, aumento na taxa de enchimento de sementes, aumento em TKW, entre outros. Um aumento em produção também pode resultar em arquitetura modificada, ou pode ocorrer como um resultado de arquitetura modificada.

As características de crescimento melhoradas obtidas realizando os métodos da invenção, na medida em que eles se relacionam a uso de CDC27, resultam em plantas tendo número de sementes aumentado. Um número de sementes aumentado abrange um aumento no número total de sementes e/ou no número de grãos cheios e/ou um aumento na taxa de enchimento de sementes (que é o número de grãos cheios dividido pelo número total de sementes e multiplicado por 100), cada em relação a plantas de controle adequadas, cujo aumento pode ser por planta e/ou por hectare ou acre. Tomando milho como um exemplo, um aumento no número de

sementes é manifestado tipicamente por um aumento no número de orelhas por planta, um aumento no número de fileiras, número de sementes por fileira, aumento na taxa de enchimento de sementes, entre outros. Tomando arroz como um exemplo, um aumento no número de sementes é manifestado tipicamente por um aumento em número de panículas por planta, número de espiguilhas por panícula, número de flores (floretes) por panícula (que é expresso como uma proporção do número de grãos cheios pelo número de panículas primárias), aumento na taxa de enchimento de sementes.

A invenção portanto fornece um método para aumentar o número de sementes de plantas em relação àquele de plantas de controle adequadas, compreendendo aumentar preferencialmente expressão em tecido de meristema apical de broto vegetal de um ácido nucleico codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo.

Na medida em que os métodos da invenção se relacionam a SYR, preferivelmente desempenho dos métodos resulta em plantas tendo produção aumentada de sementes. Ainda preferivelmente, a produção aumentada de sementes compreende um aumento em um ou mais de número de grãos (cheios), peso total de semente, tamanho de semente, peso de mil sementes, taxa de enchimento e índice de colheita, cada em relação a plantas de controle. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar produção de sementes de planta, cujo método compreende aumentar atividade de um polipeptídeo SYR e/ou expressão em uma planta de um ácido nucleico codificando um polipeptídeo SYR ou um homólogo deste.

Na medida em que os métodos da invenção se relacionam a FG-GAP, preferivelmente desempenho dos métodos resulta em plantas tendo produção aumentada e, mais particularmente, biomassa aumentada e/ou produção aumentada de sementes. Preferivelmente, a produção aumentada de

sementes compreende um aumento em um ou mais de número de grãos (cheios), total peso de semente, tamanho de semente, peso de mil sementes e índice de colheita, cada em relação a plantas de controle. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar produção vegetal, particularmente, biomassa aumentada e/ou produção aumentada de sementes, cujo método compreende modular atividade de um polipeptídeo FG-GAP e/ou expressão em uma planta de um ácido nucleico codificando um polipeptídeo FG-GAP ou um homólogo deste.

Na medida em que os métodos da invenção se relacionam a CYP90B, preferivelmente a produção aumentada inclui um ou mais dos seguintes: HI aumentado, TKW aumentado, área de semente aumentada e comprimento de semente aumentado, cada em relação a plantas de controle adequadas. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar produção vegetal, particularmente produção de sementes, em relação a plantas de controle adequadas, cujo método compreende aumentar expressão não constitutiva em uma planta de um ácido nucleico codificando um polipeptídeo CYP90B ou um homólogo deste.

Na medida em que os métodos da invenção se relacionam a fatores de transcrição AT-hook, produção de sementes em plantas monocotiledôneas é aumentada. Portanto é fornecido um método para aumentar produção de semente em plantas monocotiledôneas em relação a plantas de controle adequadas, compreendendo aumentar preferencialmente expressão em tecido de endosperma de uma planta monocotiledônea de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296.

Na medida em que os métodos da invenção se relacionam a fatores de transcrição DOF, preferivelmente a produção aumentada é produção aumentada de sementes. De acordo com uma característica preferida da presente invenção, é fornecido um método para aumentar produção de

sementes de planta em relação à produção de sementes de plantas de controle adequadas, cujo método compreende aumentar expressão em uma planta de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF.

Na medida em que os métodos da invenção se relacionam a CKIs, a característica de crescimento melhorada é produção aumentada de sementes. A presente invenção portanto fornece um método para aumentar produção de semente em plantas em relação a plantas de controle adequadas, compreendendo reduzir preferencialmente expressão de um gene CKI endógeno em tecido de endosperma de uma planta.

Uma vez que as plantas melhoradas de acordo com a presente invenção têm produção aumentada (produção de sementes), é provável que estas plantas exibam uma taxa de crescimento aumentada (durante pelo menos parte de seu ciclo de vida), em relação à taxa de crescimento de plantas tipo selvagem correspondentes em um estágio correspondente em seu ciclo de vida. A taxa de crescimento aumentada pode ser específica para uma ou mais partes ou tipos de células de uma planta (incluindo sementes), ou pode ser ao longo de substancialmente toda a planta. Plantas tendo uma taxa de crescimento aumentada podem ter um ciclo de vida mais curto. O ciclo de vida de uma planta é adotado para significar o tempo necessário para crescer de uma semente madura seca até o estágio onde a planta produziu sementes maduras secas, similar ao material inicial. Este ciclo de vida pode ser influenciado por fatores tal como vigor precoce, taxa de crescimento, tempo de florescimento e velocidade de maturação de sementes. Um aumento em taxa de crescimento pode ocorrer em um ou mais estágios no ciclo de vida de uma planta ou durante substancialmente todo o ciclo de vida de planta. Taxa de crescimento aumentada durante os estágios precoces no ciclo de vida de uma planta pode refletir vigor estimulado. O aumento em taxa de crescimento pode alterar o ciclo de colheita de uma planta permitindo que plantas sejam semeadas mais tarde e/ou colhidas mais cedo do que seria de outra maneira

possível. Se a taxa de crescimento é suficientemente aumentada, ela pode permitir a semeadura de sementes adicionais da mesma espécie de planta (por exemplo semeadura e colheita de plantas de arroz seguido por 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, ela pode permitir a semeadura adicional de sementes de espécie de planta diferente (por exemplo a semeadura e colheita de plantas de arroz seguido por, por exemplo, a semeadura e colheita opcional de soja, batatas ou qualquer outra planta adequada). Momentos adicionais de colheita do mesmo porta-enxerto no caso de algumas plantas também podem ser possíveis. Alterar o ciclo de colheita de uma planta pode levar a um aumento em produção de biomassa anual por acre (devido a um aumento no número de vezes (diga em um ano) que qualquer planta particular pode ser crescida e colhida). Um aumento em taxa de crescimento também pode permitir o cultivo de plantas transgênicas em uma área geográfica maior do que suas contrapartes tipo selvagem, uma vez que as limitações territoriais para crescer uma safra são freqüentemente determinadas por condições ambientais adversas ou no momento de plantio (ciclo precoce) ou no momento de colheita (ciclo tardio). Tais condições adversas podem ser evitadas se o ciclo de colheita é encurtado. A taxa de crescimento pode ser determinada derivando vários parâmetros de curvas de crescimento plotando experimentos de crescimento, tais parâmetros podem ser: T-Mid (the tempo levado por plantas para atingir 50% de seu tamanho máximo) e T-90 (tempo levado por plantas para atingir 90% de seu tamanho máximo), entre outros. O termo "tempo de florescimento" como usado neste lugar deve significar o período de tempo entre o início de germinação de semente e o início de florescimento.

Desempenho dos métodos da invenção gera plantas tendo uma taxa de crescimento aumentada.

Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a taxa de crescimento de plantas, cujo método compreende aumentar atividade em uma planta de um polipeptídeo SYR ou um homólogo deste e/ou expressão de um ácido nucleico codificando uma tal proteína.

De acordo com a presente invenção, é fornecido um método para aumentar a taxa de crescimento de plantas, cujo método compreende modular (preferivelmente aumentar) atividade em uma planta de um polipeptídeo FG-GAP ou um homólogo deste e/ou modular (preferivelmente aumentar) expressão de um ácido nucleico codificando tal proteína.

De acordo com a presente invenção, é fornecido um método para aumentar a taxa de crescimento de plantas cujo método compreende aumentar expressão não constitutiva em uma planta de um ácido nucleico codificando um polipeptídeo CYP90B ou um homólogo deste.

De acordo com a presente invenção, é fornecido um método para aumentar a taxa de crescimento de plantas, cujo método compreende aumentar expressão em uma planta de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF.

De acordo com a presente invenção, é fornecido um método para aumentar a taxa de crescimento de plantas em relação a plantas de controle adequadas, cujo método compreende reduzir preferencialmente expressão de um gene endógeno de **I**nibidor de **Q**uinase Dependente de **C**iclina (CKI) em tecido de endosperma de uma planta.

Um aumento em produção e/ou produção de sementes e/ou taxa de crescimento ocorre se a planta está em condições sem estresse ou se a planta é exposta a vários estresses comparado com plantas de controle. Plantas tipicamente respondem à exposição a estresse crescendo mais lentamente. Em condições de estresse severo, a planta pode mesmo parar de crescer no geral. Estresse brando por outro lado é definido neste lugar como

sendo qualquer estresse ao qual uma planta é exposta que não resulta na planta parando de crescer no geral sem a capacidade de reiniciar crescimento. Estresse brando no sentido da invenção leva a uma redução no crescimento das plantas estressadas 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 com a planta de controle em condições sem estresse. Devido a avanços em práticas agrícolas (irrigação, fertilização, tratamentos com pesticida) estresses severos não são freqüentemente encontrados em cultivares cultivadas. Como uma conseqüência, o crescimento comprometido induzido por estresse brando freqüentemente é uma característica indesejável para agricultura. Estresses brandos são os estresses típicos aos quais uma planta pode ser exposta. Estes estresses podem ser os estresses bióticos e/ou abióticos (ambientais) cotidianos aos quais uma planta é exposta. Estresses abióticos ou ambientais típicos incluem estresses por temperatura causados por calor atípico ou temperaturas frias/congelantes; estresse salino; estresse hídrico (seca ou água em excesso), estresse anaeróbico, toxicidade química e estresse oxidativo. O estresse abiótico pode ser um estresse osmótico causado por um estresse hídrico (particularmente devido à seca), estresse salino, estresse oxidativo ou um estresse iônico. Produtos químicos também podem causar estresses abióticos (por exemplo concentrações muito altas ou muito baixas de minerais ou nutrientes). Estresses bióticos são tipicamente aqueles estresses causados por patógenos, tal como bactérias, vírus, fungos e insetos. O termo "condições sem estresse" como usado neste lugar são aquelas condições ambientais que não vão significantemente além das condições climáticas cotidianas e outras condições abióticas que plantas podem encontrar, e que permitem crescimento ótimo da planta. Pessoas versadas na técnica têm consciência de condições de solo e condições climáticas normais para uma dada localização geográfica.

Na medida em que os métodos da invenção se relacionam a

SYR, desempenho dos métodos resulta em plantas tendo tolerância aumentada para estresse abiótico. Como relatado em Wang *et al.* (Planta (2003) 218: 1-14), estresse abiótico leva a uma série de mudanças morfológicas, fisiológicas, bioquímicas e moleculares que afetam adversamente crescimento e produtividade vegetal. Seca, salinidade, temperaturas extremas e estresse oxidativo são conhecidos por serem interconectados e podem induzir crescimento e lesão celular através de mecanismos similares. Por exemplo, seca e/ou salinização são manifestadas primariamente como estresse osmótico, resultando no rompimento de homeostase e distribuição iônica na célula. Estresse oxidativo, que freqüentemente acompanha temperatura alta ou baixa, salinidade ou estresse por seca pode causar desnaturação de proteínas funcionais e estruturais. Como uma conseqüência, estes diversos estresses ambientais freqüentemente ativam vias de sinalização celular e respostas celulares similares, tal como a produção de proteínas de estresse, aumento de antioxidantes, acumulação de solutos compatíveis e interrupção de crescimento.

Uma vez que diversos estresses ambientais ativam vias similares, a exemplificação da presente invenção com estresse por seca (na medida em que a invenção se relaciona ao uso de polipeptídeos SYR e seus ácidos nucleicos de codificação) não deve ser vista como uma limitação a estresse por seca, mas mais como uma separação para indicar o envolvimento de polipeptídeos SYR ou homólogos destes em estresses abióticos em geral. Além disso, os métodos da presente invenção podem ser realizados em condições sem estresse ou em condições de seca branda para gerar plantas tendo características de crescimento melhoradas (particularmente produção aumentada) em relação a tipo selvagem correspondente ou outras plantas de controle.

Um grau particularmente alto de "interferência" é relatado entre estresse por seca e estresse por alta salinidade (Rabbani *et al.* (2003)

Plant physiol 133: 1755-1767). Portanto, seria perceptível que um polipeptídeo SYR ou um homólogo deste, junto com sua utilidade para conferir tolerância à seca em plantas, também iria encontrar uso para proteger a planta contra vários outros estresses abióticos. Similarmente, seria perceptível que uma proteína SYR (como definido neste lugar), junto com sua utilidade para conferir tolerância a sal em plantas, também iria encontrar uso para proteger a planta contra vários outros estresses abióticos. Além disso, Rabbani *et al.* (2003, Plant physiol 133: 1755-1767) relatam que mecanismos moleculares de tolerância a estresse e respostas similares existem entre dicotiledôneas e monocotiledôneas. Os métodos da invenção portanto são vantajosamente aplicáveis a qualquer planta.

O termo "estresse abiótico" como definido neste lugar é adotado para significar qualquer um ou mais de: estresse hídrico (devido à seca ou água em excesso), estresse anaeróbico, estresse salino, estresse por temperatura (devido a calor, frio ou temperaturas congelantes), estresse por toxicidade química e estresse oxidativo. De acordo com um aspecto da invenção, o estresse abiótico é um estresse osmótico, selecionado a partir de estresse hídrico, estresse salino, estresse oxidativo e estresse iônico. Preferivelmente, o estresse hídrico é estresse por seca. O termo estresse salino não é restrito a sal comum (NaCl), mas pode ser qualquer um ou mais de: NaCl, KCl, LiCl, $MgCl_2$, $CaCl_2$, entre outros.

Tolerância aumentada para estresse abiótico é manifestada por produção vegetal aumentada em condições de estresse abiótico. Na medida em que a invenção se relaciona ao uso de polipeptídeos SYR e seus ácidos nucleicos de codificação, tal produção aumentada pode incluir um ou mais dos seguintes: número aumentado de sementes cheias, produção total de sementes aumentada, número aumentado de flores por panícula, taxa aumentada de enchimento de sementes, índice de colheita aumentado, peso de mil sementes aumentado, comprimento de raiz aumentado ou diâmetro de raiz

aumentado, cada em relação a plantas tipo selvagem correspondentes.

Desempenho dos métodos da invenção gera plantas tendo tolerância aumentada para estresse abiótico. Desempenho dos métodos da invenção gera plantas crescidas em condições sem estresse ou em condições de seca branda com características de crescimento melhoradas (particularmente produção aumentada e/ou vigor de emergência aumentado (ou vigor precoce)) em relação a plantas tipo selvagem correspondentes ou outras plantas de controle crescidas em condições comparáveis.

De acordo com a presente invenção, é fornecido um método para aumentar tolerância a estresse abiótico em plantas cujo método compreende modular expressão em uma planta de um ácido nucleico codificando um polipeptídeo SYR ou um homólogo deste. De acordo com um aspecto da invenção, o estresse abiótico é estresse osmótico, selecionado a partir de um ou mais dos seguintes: estresse hídrico, estresse salino, estresse oxidativo e estresse iônico. Preferivelmente, o estresse hídrico é estresse por seca.

A presente invenção também fornece um método para melhorar tolerância a estresse abiótico em plantas, compreendendo aumentar atividade em uma planta de uma proteína SYR ou um homólogo desta.

Na medida em que os métodos da invenção se relacionam a fatores de transcrição DOF, os métodos podem ser realizados em condições de seca branda para gerar plantas tendo produção aumentada em relação a plantas de controle adequadas. Como relatado em Wang *et al.* (Planta (2003) 218: 1-14), estresse abiótico leva a uma série de mudanças morfológicas, fisiológicas, bioquímicas e moleculares que afetam adversamente crescimento e produtividade vegetal. Seca, salinidade, temperaturas extremas e estresse oxidativo são conhecidos por serem interconectados e podem induzir crescimento e lesão celular através de mecanismos similares. Rabbani *et al.* (Plant physiol (2003) 133: 1755-1767) descreve um grau particularmente alto

de "interferência" entre estresse por seca e estresse por alta salinidade. Por exemplo, seca e/ou salinização são manifestadas primariamente como estresse osmótico, resultando no rompimento de homeostase e distribuição iônica na célula. Estresse oxidativo, que freqüentemente acompanha temperatura alta ou baixa, salinidade ou estresse por seca, pode causar desnaturação de proteínas funcionais e estruturais. Como uma conseqüência, estes diversos estresses ambientais freqüentemente ativam vias de sinalização celular e respostas celulares similares, tal como a produção de proteínas de estresse, aumento de antioxidantes, acumulação de solutos compatíveis e interrupção de crescimento.

Desempenho dos métodos da invenção gera plantas crescidas em condições de seca branda com produção aumentada em relação a plantas de controle adequadas crescidas em condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar produção em plantas crescidas em condições de seca branda, cujo método compreende aumentar expressão em uma planta de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF.

As características de crescimento melhoradas mencionadas acima vantajosamente podem ser melhoradas em qualquer planta. Na medida em que os métodos da invenção se relacionam ao uso de fatores de transcrição AT-hook, os métodos são aplicáveis a plantas monocotiledôneas.

O termo "planta" como usado neste lugar abrange plantas inteiras, ancestrais e progênie das plantas e partes vegetais, incluindo sementes, brotos, caules, folhas, raízes (incluindo tubérculos), flores, e tecidos e órgãos, caracterizado pelo fato de que cada um dos mencionados acima compreende o gene/ácido nucleico de interesse ou a modificação genética no gene/ácido nucleico de interesse. O termo "planta" também abrange células vegetais, safras em suspensão, tecido de calo, embriões, regiões meristemáticas, gametófitos, esporófitos, pólen e micrósporos, de novo

caracterizado pelo fato de que cada um dos mencionados acima compreende o gene/ácido nucleico de interesse.

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 forragem ou legumes de forragem, plantas ornamentais, cultivos alimentares, árvores ou arbustos selecionados a partir da lista compreendendo *Acer* spp., *Actinidia* spp., *Abelmoschus* spp., *Agropyron* spp., *Allium* spp., *Amaranthus* spp., *Ananas comosus*, *Annona* spp., *Apium graveolens*, *Arachis* spp, *Artocarpus* spp., *Asparagus officinalis*, *Avena* spp. (e.g. *Avena sativa*, *Avena fatua*, *Avena byzantina*, *Avena fatua* var. sativa, *Avena hybrida*), *Averrhoa carambola*, *Benincasa hispida*, *Bertholletia excelsea*, *Beta vulgaris*, *Brassica* spp. (e.g. *Brassica napus*, *Brassica rapa* ssp. [canola, colza, nabo]), *Cadaba farinosa*, *Camellia sinensis*, *Canna indica*, *Capsicum* spp., *Carex elata*, *Carica papaya*, *Carissa macrocarpa*, *Carya* spp., *Carthamus tinctorius*, *Castanea* spp., *Cichorium endivia*, *Cinnamomum* spp., *Citrullus lanatus*, *Citrus* spp., *Cocos* spp., *Coffea* spp., *Colocasia esculenta*, *Cola* spp., *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 (e.g. *Elaeis guineensis*, *Elaeis oleifera*), *Eleusine coracana*, *Eriobotrya japonica*, *Eugenia uniflora*, *Fagopyrum* spp., *Fagus* spp., *Ficus carica*, *Fortunella* spp., *Fragaria* spp., *Ginkgo biloba*, *Glycine* spp. (e.g. *Glycine max*, *Soja hispida* ou *Soja max*), *Gossypium hirsutum*, *Helianthus* spp. (e.g. Helianthus annuus), *Hemerocallis fulva*, *Hibiscus* spp., *Hordeum* spp. (e.g. *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. (e.g. *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., *Momordica* spp., *Morus nigra*, *Musa* spp., *Nicotiana* spp., *Olea* spp., *Opuntia* spp., *Ornithopus* spp., *Oryza* spp. (e.g. *Oryza sativa*, *Oryza latifolia*), *Panicum miliaceum*, *Passiflora edulis*, *Pastinaca sativa*, *Persea* spp., *Petroselinum crispum*, pH *aseolus* spp., pH *oenix* spp., pH *ysalis* 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., *Sambucus* spp., *Secale cereale*, *Sesamum* spp., *Sinapis* sp., *Solanum* spp. (e.g. *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. (e.g. *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.

Preferivelmente, a planta é uma planta cultivável tal como soja, girassol, canola, alfafa, colza, algodão, tomate, batata ou tabaco. Ainda preferivelmente, a planta é uma planta monocotiledônea, tal como cana-de-açúcar. Mais preferivelmente a planta é um cereal, tal como arroz, milho, trigo, cevada, milheto, centeio, sorgo ou aveia.

Onde os métodos da invenção se relacionam a uso de um fator de transcrição AT-hook, a planta monocotiledônea é um cereal, tal como arroz, milho, cana-de-açúcar, trigo, cevada, milheto, centeio, sorgo, gramíneas ou aveia.

DEFINIÇÕES

Polipeptídeo

Os termos "polipeptídeo" e "proteína" são usados permutavelmente neste lugar e se referem a aminoácidos em uma forma polimérica de qualquer comprimento. Os termos "polinucleotídeo(s)", "seqüência(s) de ácidos nucleicos", "seqüência(s) de nucleotídeos" são usados permutavelmente neste lugar e se referem a nucleotídeos, ou ribonucleotídeos ou desoxirribonucleotídeos ou uma combinação de ambos, em uma forma polimérica de qualquer comprimento.

Planta de Controle

A escolha de plantas de controle adequadas é uma parte de rotina de uma estrutura experimental e pode incluir plantas tipo selvagem correspondentes ou plantas correspondentes sem o gene de interesse. A planta de controle é tipicamente da mesma espécie de planta ou mesmo a mesma variedade do que a planta a ser verificada. A planta de controle também pode ser um nulizigoto da planta a ser verificada. Uma "planta de controle" como usado neste lugar se refere não apenas a plantas inteiras, mas também a partes vegetais, incluindo sementes e partes de sementes.

Aumento, Melhora

Os termos "aumento", "melhorando" ou "melhora" são usados permutavelmente neste lugar e são adotados para significar pelo menos a 5%, 6%, 7%, 8%, 9% ou 10%, preferivelmente pelo menos 15% ou 20%, mais preferivelmente 25%, 30%, 35% ou 40% mais de produção e/ou crescimento em comparação com tipo selvagem correspondente ou outras plantas de controle como definido neste lugar.

Hibridização

O termo "hibridização" como definido neste lugar é um processo caracterizado pelo fato de que seqüências de nucleotídeos complementares substancialmente homólogas se anelam uma a outra. O processo de hibridização pode ocorrer inteiramente em solução, i.e. ambos ácidos nucleicos complementares estão em solução. O processo de

hibridização também pode ocorrer com um dos ácidos nucleicos complementares imobilizado a uma matriz tal como microesferas magnéticas, microesferas de Sepharose ou qualquer outra resina. O processo de hibridização além disso pode ocorrer com um dos ácidos nucleicos complementares imobilizado a um suporte sólido tal como a nitrocelulose ou membrana de náilon ou imobilizado por e.g. fotolitografia a, por exemplo, um suporte de vidro silicioso (o último conhecido como arranjos ou microarranjos de ácido nucleico ou como circuitos integrados de ácido nucleico). A fim de permitir que hibridização ocorra, as moléculas de ácidos nucleicos geralmente são termalmente ou quimicamente desnaturadas para derreter uma dupla fita em duas fitas simples e/ou para remover estruturas em grampo ou outras estruturas secundárias de ácidos nucleicos de fita simples. A estringência de hibridização é influenciada por condições tal como temperatura, concentração salina, força iônica e composição de tampão de hibridização.

"Condições de hibridização estringentes" e "condições de lavagem de hibridização estringentes" no contexto de experimentos de hibridização de ácido nucleico tal como hibridizações Southern e Northern são dependentes de seqüência e são diferentes em parâmetros ambientais diferentes. O artesão versado está ciente de vários parâmetros que podem ser alterados durante hibridização e lavagem e que irão ou manter ou mudar as condições de estringência.

A T_m é a temperatura em força iônica e pH definidos, ba qual 50% da seqüência alvo hibridiza com uma sonda perfeitamente compatível. A T_m é dependente das condições de solução e ad composição de base e comprimento da sonda. Por exemplo, seqüências maiores hibridizam especificamente em temperaturas superiores. A taxa máxima de hibridização é obtida a partir de cerca de 16°C até 32°C abaixo de T_m. A presença de cátions monovalentes na solução de hibridização reduz a repulsão eletrostática entre as duas fitas de ácidos nucleicos com isso promovendo formação de híbrido;

este efeito é visível para concentrações de sódio de até 0,4M. Formamida reduz a temperatura de fusão de duplexes DNA-DNA e DNA-RNA com 0,6 a 0,7°C para cada formamida percentual, e adição de 50% de formamida permite que hibridização seja realizada a 30 a 45°C, embora a taxa de hibridização seja diminuída. Pareamento incorreto de bases reduz a taxa de hibridização e a estabilidade térmica dos duplexes. Em média e para sondas grandes, a T_m diminui cerca de 1°C por% de pareamento incorreto de bases. A T_m pode ser calculada usando as seguintes equações, dependendo dos tipos de híbridos:

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

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

- Híbridos de DNA-RNA ou RNA-RNA:

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

- Híbridos de oligo-DNA ou oligo-RNA[d]:

  Para <20 nucleotídeos: $T_m = 2\ (l_n)$

  Para 20–35 nucleotídeos: $T_m = 22 + 1,46\ (l_n)$

  [a] ou para outro cátion monovalente, mas apenas acurado no intervalo 0,01–0,4 M.

  [b] apenas acurado para%GC no intervalo de 30% a 75%.

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

  [d] Oligo, oligonucleotídeo; l_n, comprimento efetivo de iniciador = (no. de G/C)+(no. de A/T).

Nota: para cada 1% de formamida, a T_m é reduzida em cerca de 0,6 a 0,7°C, enquanto que a presença de 6M de uréia reduz a T_m em cerca de 30°C

Especificidade de hibridização é tipicamente a função de lavagens pós-hibridização. Para remover antecedentes resultando de hibridização não específica, amostras são lavadas com soluções salinas

diluídas. Fatores críticos de tais lavagens incluem a força iônica e temperatura da solução de lavagem final: quanto menor a concentração salina e maior a temperatura de lavagem, maior é a estringência da lavagem. Condições de lavagem são tipicamente realizadas em ou abaixo de estringência de hibridização. Geralmente, condições estringentes adequadas para ensaios de hibridização de ácidos nucleicos ou procedimentos de detecção de amplificação de gene são como expostos acima. Condições mais ou menos estringentes também podem ser selecionadas. Geralmente, condições de baixa estringência são selecionadas para ser cerca de 50°C menores do que o ponto de fusão térmica (T_m) para a seqüência específica em uma força iônica e pH definidos. Condições de média estringência são quando a temperatura é 20°C abaixo de T_m, e condições de alta estringência são quando a temperatura é 10°C abaixo de T_m. Por exemplo, condições estringentes são aquelas que são pelo menos tão estringentes como, por exemplo, condições A-L; e condições de estringência reduzida são pelo menos tão estringentes como, por exemplo, condições M-R. Ligação não específica pode ser controlada usando qualquer uma das diversas técnicas conhecidas tal como, por exemplo, bloqueando a membrana com soluções contendo proteína, adições de RNA, DNA heterólogos, e SDS ao tampão de hibridização, e tratamento com Rnase.

Exemplos de condições de hibridização e lavagem são listados em Tabela 1:

Tabela 1:

Condição de Estringência	Híbrido de Poli-nucleotídeo ±	Comprimento de Híbrido (bp) ‡	Temperatura e Tampão de Hibridização †	Temperatura e Tampão de Lavagem †
A	DNA:DNA	> ou igual a 50	65°C 1×SSC; ou 42°C, 1×SSC e 50% de formamida	65°C; 0,3×SSC
B	DNA:DNA	<50	Tb*; 1×SSC	Tb*; 1×SSC
C	DNA:RNA	> ou igual a 50	67°C 1×SSC; ou 45°C, 1×SSC e 50% de formamida	67°C; 0,3×SSC
D	DNA:RNA	<50	Td*; 1×SSC	Td*; 1×SSC
E	RNA:RNA	> ou igual a 50	70°C 1×SSC; ou 50°C, 1×SSC e 50% de formamida	70°C; 0,3×SSC
F	RNA:RNA	<50	Tf*; 1×SSC	Tf*; 1×SSC
G	DNA:DNA	> ou igual a 50	65°C 4×SSC; ou 45°C, 4×SSC e 50% de formamida	65°C; 1×SSC

H	DNA:DNA	<50	Th*; 4 ×SSC	Th*; 4×SSC
I	DNA:RNA	> ou igual a 50	67°C 4×SSC; ou 45°C, 4×SSC e 50% de formamida	67°C; 1×SSC
J	DNA:RNA	<50	Tj*; 4 ×SSC	Tj*; 4 ×SSC
K	RNA:RNA	> ou igual a 50	70°C 4×SSC; ou 40°C, 6×SSC e 50% de formamida	67°C; 1×SSC
L	RNA:RNA	<50	Tl*; 2 ×SSC	Tl*; 2×SSC
M	DNA:DNA	> ou igual a 50	50°C 4×SSC; ou 40°C, 6×SSC e 50% de formamida	50°C; 2×SSC
N	DNA:DNA	<50	Tn*; 6 ×SSC	Tn*; 6×SSC
O	DNA:RNA	> ou igual a 50	55°C 4×SSC; ou 42°C, 6×SSC e 50% de formamida	55°C; 2×SSC
P	DNA:RNA	<50	Tp*; 6 ×SSC	Tp*; 6×SSC
Q	RNA:RNA	> ou igual a 50	60°C 4×SSC; ou 45°C, 6×SSC e 50% de formamida	60°C.; 2×SSC
R	RNA:RNA	<50	Tr*; 4 ×SSC	Tr*; 4×SSC

‡ O "Comprimento de Híbrido" é o comprimento antecipado para o ácido nucleico hibridizante. Quando ácidos nucleicos de seqüência conhecida são hibridizados, o comprimento de híbrido pode ser determinado alinhando as seqüências e identificando as regiões conservadas descritas neste lugar.

† SSPE (1×SSPE é 0,15M de NaCl, 10mM de NaH_2PO_4, e 1,25mM de EDTA, pH 7,4) pode ser substituído por SSC (1×SSC é 0,15M de NaCl e 15mM de citrato de sódio) nos tampões de hibridização e lavagem; lavagens são realizadas por 15 minutos depois de hibridização estar completa. As hibridizações e lavagens podem incluir adicionalmente 5 × de reagente de Denhardt, 0,5-1,0% de SDS, 100 μg/ml de DNA de esperma de salmão fragmentado desnaturado, 0,5% de pirofosfato de sódio, e até 50% de formamida.

* Tb-Tr: A temperatura de hibridização para híbridos antecipada para serem menores do que 50 pares de bases de comprimento deve ser 5-10°C menor do que a temperatura de fusão T_m dos híbridos; a T_m é determinada de acordo com as equações mencionadas acima.

± A presente invenção também abrange a substituição de qualquer um, ou mais parceiros de híbrido de DNA ou RNA ou com um PNA, ou um ácido nucleico modificado.

Para os propósitos de definir o nível de estringência, referência pode ser convenientemente feita a Sambrook *et al.* (2001) Molecular Cloning: a laboratory manual, 3ª Edição Cold Spring Harbor Laboratory Press, CSH, New York ou to Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989).

Identificação por Ativação de T-DNA

Identificação por ativação de T-DNA (Hayashi et al. Science (1992) 1350-1353) envolve inserção de T-DNA, normalmente contendo um promotor (também pode ser um estimulador de tradução ou um íntron), na região genômica do gene de interesse ou 10 kb antes ou depois da região de codificação de um gene em uma configuração de forma que o promotor dirige

expressão do gene dirigido. Tipicamente, regulação de expressão do gene dirigido por seu promotor natural é interrompida e o gene cai sob o controle do promotor recentemente introduzido. O promotor é tipicamente embebido em um T-DNA. Este T-DNA é aleatoriamente inserido no genoma vegetal, por exemplo, através de infecção por *Agrobacterium* e leva a superexpressão de genes perto do T-DNA inserido. As plantas transgênicas resultantes mostram fenótipos dominantes devido a superexpressão de genes perto do promotor introduzido. O promotor a ser introduzido pode ser qualquer promotor capaz de direcionar expressão de um gene no organismo desejado, neste caso uma planta. Por exemplo, promotores constitutivos, preferidos por tecido, preferidos por tipo celular e induzíveis são todos adequados para uso em ativação de T-DNA.

TILLING

TILLING (Lesões Locais Induzidas Dirigidas em Genomas) é uma tecnologia de mutagênese útil para gerar e/ou identificar e/ou para eventualmente isolar ácidos nucleicos de variante mutagenizada. TILLING também permite seleção de plantas carregando tais variantes mutantes. Estas variantes mutantes podem ainda exibir maior atividade do que aquela exibida pelo gene em sua forma natural. TILLING combina mutagênese de alta densidade com métodos de triagem rápida. As etapas tipicamente seguidas em TILLING são: (a) mutagênese 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 combinação de indivíduos; (c) amplificação por PCR de uma região de interesse; (d) desnaturação e anelamento para permitir

formação de heteroduplexes; (e) DHPLC, onde a presença de um heteroduplex em uma combinação é detectada como um pico extra no cromatograma; (f) identificação do indivíduo mutante; e (g) seqüenciamento do produto de PCR mutante. 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).

Mutagênese sítio dirigida

Mutagênese sítio dirigida pode ser usada para gerar variantes de ácidos nucleicos SYR. Estão disponíveis diversos métodos para atingir mutagênese sítio dirigida; o mais comum sendo métodos baseados em PCR (Current Protocols in Molecular Biology. Wiley Eds. http://www.4ulr.com/products/currentprotocols/index.html).

Mutagênese por transposon

Mutagênese por transposon é uma técnica de mutagênese baseada na inserção de transposons em genes, o que freqüentemente resulta em nocaute de gene. A técnica foi usada para diversas espécies de plantas, incluindo arroz (Greco et al., Plant physiol, 125, 1175-1177, 2001), milho (McCarty et al., Plant J. 44, 52-61, 2005) e *Arabidopsis* (Parinov and Sundaresan, Curr. Opin. Biotechnol. 11, 157-161, 2000).

Evolução direta

Evolução direta ou embaralhamento de gene consiste de repetições de embaralhamento de DNA seguido por triagem e/ou seleção apropriada para gerar ácidos nucleicos variantes ou porções destes, ou polipeptídeos ou homólogos destes tendo uma atividade biológica modificada (Castle et al., (2004) Science 304(5674): 1151-4; Patentes Norte-Americanas 5.811.238 e 6.395.547).

Recombinação homóloga

Recombinação homóloga permite introdução em um genoma de um ácido nucleico selecionado em uma posição selecionada definida.

Recombinação homóloga é uma tecnologia padrão usada rotineiramente em ciências biológicas para organismos inferiores tal como levedura ou o musgo pH *yscomitrella*. Métodos para realizar recombinação homóloga em plantas foram descritos não apenas para plantas modelo (Offringa et al. (1990) EMBO J 9(10): 3077-84) mas também para cultivares, por exemplo arroz (Terada et al. (2002) Nat Biotech 20(10): 1030-4; Iida and Terada (2004) Curr Opin Biotech 15(2):132-8). O ácido nucleico a ser dirigido (que pode ser qualquer dos ácidos nucleicos ou variantes definidos neste lugar) precisa ser dirigido para o locus gênico particular. O ácido nucleico a ser dirigido pode ser um alelo melhorado usado para substituir o gene endógeno ou pode ser introduzido em adição ao gene endógeno.

Homólogos

"Homólogos" de uma proteína abrangem peptídeos, oligopeptídeos, polipeptídeos, proteínas e enzimas tendo substituições, deleções e/ou inserções de aminoácidos em relação à proteína não modificada em questão e tendo atividade biológica e funcional similar à proteína não modificada da qual eles são derivados. Para produzir tais homólogos, aminoácidos da proteína podem ser substituídos por outros aminoácidos tendo propriedades similares (tal como similar hidrofobicidade, hidrofilicidade, antigenicidade, propensão para formar ou quebrar estruturas helicoidais α ou estruturas de folha β). Tabelas de Substituição Conservativa são bem conhecidas na técnica (veja por exemplo Creighton (1984) Proteins. W.H. Freeman and Company e Tabela 2 abaixo).

Ortólogos e Parálogos

Abrangidas pelo termo "homólogos" são seqüências ortólogas e seqüências parálogas, duas formas especiais de homologia que abrangem conceitos evolutivos usados para descrever relações ancestrais de genes.

O termo "parálogos" se refere a duplicações de gene dentro do genoma de uma espécie levando a genes parálogos. Parálogos podem ser

facilmente identificados realizando uma análise BLAST contra um conjunto de seqüências da mesma espécie que a seqüência de consulta.

O termo "ortólogos" se refere a genes homólogos em organismos diferentes devido à especiação. Ortólogos em, por exemplo, espécies de plantas dicotiledôneas podem ser facilmente encontrados realizando uma assim chamada busca blast recíproca. Isto pode ser feito por um primeiro blast envolvendo sobrecarregar uma seqüência de consulta (por exemplo, SEQ ID NO: 1 ou SEQ ID NO: 2) contra qualquer banco de dados de seqüências, tal como o banco de dados de NCBI publicamente disponível que pode ser encontrado em: http://www.ncbi.nlm.nih.gov. BLASTN ou TBLASTX (usando valores pré-determinados padrão) podem ser usados quando iniciando a partir de uma seqüência de nucleotídeos e BLASTP ou TBLASTN (usando valores pré-determinados padrão) podem ser usados quando iniciando a partir de uma seqüência de proteínas. Os resultados de BLAST opcionalmente podem ser filtrados. As seqüências de comprimento completo ou dos resultados filtrados ou resultados não filtrados são então BLASTed de volta (segundo BLAST) contra seqüências do organismo do qual a seqüência de consulta é derivada (onde a seqüência de consulta é SEQ ID NO: 1 ou SEQ ID NO: 2 o segundo blast iria portanto ser contra seqüências de *Oryza sativa)*. Os resultados do primeiro e segundo BLASTs são então comparados. Um parálogo é identificado se um acerto de alto ranqueamento do segundo blast é da mesma espécie de qual a seqüência de consulta é derivada; um ortólogo é identificado se um acerto de alto ranqueamento não é da mesma espécie de qual a seqüência de consulta é derivada. Acertos de alto ranqueamento são aqueles tendo um baixo valor E. Quanto menor o valor E, mais significante o escore (ou em outras palavras menor a probabilidade de que o acerto seja encontrado ao acaso). Computação do valor E é bem conhecida na técnica. No caso de famílias grandes, ClustalW pode ser usado, seguido por uma árvore de agrupamento de

vizinhos, para ajudar a visualizar agrupamento de genes relacionados e para identificar ortólogos e parálogos.

Um homólogo pode estar na forma de uma "variante substitucional" de uma proteína, i.e. onde pelo menos um resíduo em uma seqüência de aminoácidos foi removido e um resíduo diferente inserido em seu lugar. Substituições de aminoácidos tipicamente são de resíduos únicos, mas podem ser agrupadas dependendo de restrições funcionais colocadas no polipeptídeo; inserções normalmente serão da ordem de cerca de 1 a 10 resíduos de aminoácidos. Preferivelmente, substituições de aminoácidos compreendem substituições conservativas de aminoácidos. Substituições menos conservativas podem ser feitas no caso das propriedades de aminoácidos mencionadas acima não serem tão críticas. Tabelas de Substituição Conservativa estão prontamente disponíveis na técnica. A Tabela abaixo fornece exemplos de substituições conservadas de aminoácidos.

Tabela 2: Exemplos de substituições conservadas de aminoácidos:

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; pH e
His	Asn; Gln	Val	Ile; Leu
Ile	Leu, Val		

Um homólogo também pode estar na forma de uma "variante insercional" de uma proteína, i.e. onde um ou mais resíduos de aminoácidos são introduzidos em um sítio pré-determinado em uma proteína. Inserções podem compreender fusões N-terminais e/ou C-terminais assim como inserções intra-seqüência de aminoácidos únicos ou múltiplos. Geralmente, inserções dentro da seqüência de aminoácidos serão menores do que fusões N- ou C-terminais, da ordem de cerca de 1 a 10 resíduos. Exemplos de proteínas ou peptídeos de fusão N- ou C-terminais incluem o domínio de ligação ou domínio de ativação de um ativador transcricional como usado no

sistema de dois híbridos de levedura, proteínas de revestimento de fago, etiqueta de (histidina)-6, etiqueta de glutationa S-transferase, proteína A, proteína de ligação de maltose, diidrofolato redutase, epítopo Etiqueta·100, epítopo c-myc, epítopo FLAG®, lacZ, CMP (peptídeo de ligação calmodulina), epítopo HA, epítopo de proteína C e epítopo VSV.

Homólogos na forma de "variantes de deleção" de uma proteína são caracterizados pela remoção de um ou mais aminoácidos de uma proteína.

Variantes de aminoácidos de uma proteína podem ser prontamente feitos usando técnicas de peptídeo sintético bem conhecidas na técnica, tal como síntese de peptídeo em fase sólida e os semelhantes, ou por manipulações de DNA recombinante. Métodos para a manipulação de seqüências de DNA para produzir variantes de substituição, inserção ou deleção de uma proteína são bem conhecidos na técnica. Por exemplo, técnicas para fazer mutações por substituição em sítios pré-determinados em DNA são bem conhecidas por aqueles versados na técnica e incluem mutagênese M13, mutagênese in vitro T7-Gen (USB, Cleveland, OH), mutagênese sítio dirigida QuickChange (Stratagene, San Diego, CA), mutagênese sítio dirigida mediada por PCR ou outros protocolos de mutagênese sítio dirigida.

Derivados

"Derivados" são polipeptídeos ou proteínas que podem compreender resíduos de aminoácidos naturalmente modificados e/ou não naturalmente modificados comparados com a seqüência de aminoácidos de uma forma naturalmente ocorrente (isto é não tendo passado por modificações pós-traducionais) da proteína, por exemplo, como apresentada em SEQ ID NO: 2. "Derivados" de uma proteína abrangem polipeptídeos ou proteínas que podem compreender resíduos de aminoácidos naturalmente ocorrentes alterados, glicosilados, acilados, prenilados ou resíduos de aminoácidos não

naturalmente ocorrentes comparado com a seqüência de aminoácidos de uma forma naturalmente ocorrente do polipeptídeo. Um derivado também pode compreender um ou mais substituintes não aminoácidos comparado com a seqüência de aminoácidos da qual ele é derivado, por exemplo uma molécula repórter ou outro ligante, covalentemente ou não covalentemente ligado à seqüência de aminoácidos, tal como uma molécula repórter que é ligada para facilitar sua detecção, e resíduos de aminoácidos não naturalmente ocorrentes em relação à seqüência de aminoácidos de uma proteína não naturalmente ocorrente.

Variantes de união Alternativas

O termo "variante de união alternativa" como usado neste lugar abrange variantes de uma seqüência de ácidos nucleicos nas quais íntrons e/ou éxons selecionados foram cortados, substituídos ou adicionados, ou nas quais íntrons foram encurtados ou alongados. Tais variantes serão umas nas quais a atividade biológica da proteína é mantida, o que pode ser atingido mantendo seletivamente segmentos funcionais da proteína. Tais variantes de união podem ser encontradas na natureza ou podem ser feitas pelo homem. Métodos para fazer tais variantes de união são conhecidos na técnica.

Variante alélica

Variantes alélicas existem na natureza, e abrangido dentro dos métodos da presente invenção é o uso destes alelos naturais. Variantes alélicas abrangem Polimorfismos de Nucleotídeos Únicos (SNPs), assim como Polimorfismos de Pequena Inserção/Deleção (INDELs). O tamanho de INDELs normalmente é menor do que 100 bp. SNPs e INDELs formam o maior conjunto de variantes de seqüência em linhagens polimórficas naturalmente ocorrentes da maioria dos organismos.

Promotor

Os termos "elemento regulador", "seqüência de controle" e

"promotor" são todos usados permutavelmente neste lugar e devem ser adotados em um amplo contexto para se referir a seqüências reguladoras de ácidos nucleicos capazes de efetuar expressão das seqüências aos quais eles são ligados. Abrangidas pelos termos mencionados acima são seqüências reguladoras transcricionais derivadas de um gene genômico eucariótico clássico (incluindo a caixa TATA que é requerida para iniciação acurada de transcrição, com ou sem uma seqüência de caixa CCAAT) e elementos reguladores adicionais (i.e. seqüências situadas acima dos promotores gênicos, estimuladores e silenciadores) que alteram expressão gênica em resposta a desenvolvimento e/ou estímulos externos, ou de uma maneira tecido específica. Também incluída dentro do termo é uma seqüência reguladora transcricional de um gene procariótico clássico, em cujo caso ela pode incluir uma seqüência de caixa –35 e/ou seqüências reguladoras adicionais de caixa –10. O termo "elemento regulador" também abrange uma molécula de fusão sintética ou derivado que confere, ativa ou estimula expressão de uma molécula de ácido nucleico em uma célula, tecido ou órgão. O termo "operavelmente ligado" como usado neste lugar se refere a uma ligação funcional entre a seqüência de promotor e o gene de interesse, de forma que a seqüência de promotor é capaz de iniciar transcrição do gene de interesse.

O promotor pode ser um promotor induzível, i.e. tendo iniciação de transcrição induzida ou aumentada em resposta a um desenvolvimento, estímulo químico, ambiental ou físico.

Um promotor preferido por tecido ou tecido específico é um que é capaz de iniciar preferencialmente transcrição em certos tecidos, tal como as folhas, raízes, tecido de semente etc, ou mesmo em células específicas.

O termo "constitutivo" como definido neste lugar se refere a um promotor que é expresso predominantemente em pelo menos um tecido ou

órgão e predominantemente em qualquer estágio de vida da planta. Preferivelmente o promotor é expresso predominantemente ao longo de toda a planta.

Exemplos de outros promotores constitutivos são mostrados em Tabela 3 abaixo.

Tabela 3: Exemplos de promotores constitutivos

Fonte Gênica	**Referência**
Actina	McElroy et al, Plant Cell, 2: 163-171, 1990
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 de arroz	Buchholz et al, Plant Mol Biol. 25(5): 837-43, 1994
H3 histona de milho	Lepetit et al, Mol. Gen. Genet. 231:276-285, 1992
H3 histona de alfafa	Wu et al. Plant Mol. Biol. 11:641-649, 1988
Actina 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	Patente Norte-Americana 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

Tabela 4: Exemplos de promotores não constitutivos

Fonte gênica e nome	**Padrão de expressão**	**Referência**
RP6 de Arroz	Endosperma específica	Wen *et al.* (1993) Plant physiol 101(3): 1115-6
Cafirina de sorgo	Endosperma específica	DeRose *et al.* (1996) Plant Molec Biol 32: 1029-35
Zeína de milho	Endosperma específica	Matzke *et al.* (1990) Plant Mol Biol 14(3): 323-32
Oleosina 18 kDa de Arroz	Embrião (e aleurona) específica	Chuang *et al.* (1996) J Biochem 120(1): 74-81
Oleosina 16 kDa de Arroz	Embrião (e aleurona) específica	Chuang *et al.* (1996) J Biochem 120(1): 74-81

Beta-conglicinina de Soja	Embrião	Chiera *et al.* (2005) Plant Molec Biol 56(6): 895-904
Wsi18 de Arroz	Semente inteira	Joshee *et al.* (1998) Plant Cell physiol 39(1): 64-72.
Arroz	Semente inteira	Sasaki *et al.* (2002) NCBI número de acesso BAA85411
OSH1 de Arroz	Meristema apical de broto precoce	Sato et al. (1996) Proc Natl Acad Sci 93 (15): 8117-8122
Rcc2 de Arroz	Raiz específica	Xu *et al.* (1995) Plant Mol Biol 27(2): 237-48
Rcc3 de Arroz	Raiz específica	Xu *et al.* (1995) Plant Mol Biol 27(2): 237-48
Pyk10 de Arabidopsis	Raiz específica	Nitz *et al.* (2001) Plant Sci 161(2): 337-346

Tabela 5: Exemplos de promotores de meristema apical de broto precoce

Fonte gênica	Família de gene	Fonte vegetal	Referência
OSH1	homeobox de classe 1 de família KNOX	*Oryza sativa*	-Matsuoka *et al.*, (1993) Plant Cell 5: 1039-1048 -Sato *et al.*, (1996) PNAS 93: 8117-8122
Knotted1	homeobox de classe 1 de família KNOX	*Zea mays*	Hake *et al.*, (1989) EMBO Journal 8: 15-22
KNAT1	homeobox de classe 1 de família KNOX	*Arabidopsis thaliana*	Lincoln *et al.*, (1994) Plant Cell 6: 1859-1876
Oskn2	homeobox de classe 1 de família KNOX	*Oryza sativa*	Postma-Haarsma *et al.*, (1999) Plant Mol Biol 39(2): 257-71
Oskn3	homeobox de classe 1 de família KNOX	*Oryza sativa*	Postma-Haarsma *et al.*, (1999) Plant Mol Biol 39(2): 257-71

Tabela 6: Exemplos de promotores endosperma específicos para uso na presente invenção

Fonte gênica	Padrão de expressão	Referência
LMW e HMW glutenina-1 de trigo	endosperma	Mol Gen Genet 216:81-90, 1989; NAR 17: 461-2, 1989.
a, b, g-gliadinas de trigo	endosperma	EMBO 3:1409-15, 1984.
promotor *Itr1* de cevada	endosperma	
Hordeína B1, C, D, de cevada	endosperma	Theor Appl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750-60, 1996.
DOF de cevada	endosperma	Mena *et al.*, The Plant Journal, 116(1): 53-62, 1998.
blz2	endosperma	EP99106056.7
Promotor sintético	endosperma	Vicente-Carbajosa *et al.*, Plant J. *13:* 629-640, 1998.
prolamina NRP33 de	endosperma	Wu *et al.*, Plant Cell physiology 39(8)

arroz		885-889, 1998.
a-globulina Glb-1 de arroz	endosperma	Wu *et al.*, Plant Cell physiology 39(8) 885-889, 1998.
a-globulina REB/OHP-1 de arroz	endosperma	Nakase *et al.*, Plant Mol. Biol. 33: 513-522, 1997.
ADP-glucose PP e arroz	endosperma	Trans Res 6:157-68, 1997.
Família de gene ESR de milho	endosperma	Plant J 12:235-46, 1997.
g-cafirina de sorgo	endosperma	PMB 32:1029-35, 1996.

Tabela 7: Exemplos de promotores semente específicos para uso na presente invenção

Fonte gênica	**Padrão de expressão**	**Referência**
Genes semente específicos	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 castanha-do-Pará	semente	Pearson, *et al.*, Plant Mol. Biol. 18: 235-245, 1992.
legumina	semente	Ellis, *et al.*, Plant Mol. Biol. 10: 203-214, 1988.
glutelina (arroz)	semente	Takaiwa, *et al.*, Mol. Gen. Genet. 208: 15-22, 1986; Takaiwa, *et al.*, FEBS Letts. 221: 43-47, 1987.
zeína	semente	Matzke *et al.*, Plant Mol Biol, 14(3): 323-32, 1990.
napA	semente	Stalberg, *et al.*, Planta 199: 515-519, 1996.
glutenina-1LMW e HMW de trigo	endosperma	Mol Gen Genet 216:81-90, 1989; NAR 17: 461-2, 1989.
SPA de trigo	semente	Albani *et al.*, Plant Cell, 9: 171-184, 1997.
a, b, g-gliadinas de trigo	endosperma	EMBO 3:1409-15, 1984.
promotor *Itr1* de cevada	endosperma	
hordeína B1, C, D, de cevada	endosperma	Theor Appl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750-60, 1996.
DOF de cevada	endosperma	Mena *et al.*, The Plant Journal, 116(1): 53-62, 1998.
blz2	endosperma	EP99106056.7
promotor sintético	endosperma	Vicente-Carbajosa *et al.*, Plant J. *13:* 629-640, 1998.
prolamina NRP33 de arroz	endosperma	Wu *et al.*, Plant Cell physiology 39(8) 885-889, 1998.
a-globulina Glb-1 de arroz	endosperma	Wu *et al.*, Plant Cell physiology 39(8) 885-889, 1998.
OSH1 de arroz	embrião	Sato *et al.*, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996.
a-globulina REB/OHP-1 de arroz	endosperma	Nakase *et al.*, Plant Mol. Biol. 33: 513-522, 1997.

ADP-glucose PP de arroz	endosperma	Trans Res 6:157-68, 1997.
Família de gene ESR de milho	endosperma	Plant J 12:235-46, 1997.
g-cafirina de sorgo	endosperma	PMB 32:1029-35, 1996.
KNOX	embrião	Postma-Haarsma *et al.*, Plant Mol. Biol. 39:257-71, 1999.
Oleosina de arroz	embrião e aleurona	Wu *et al.*, J. Biochem., 123:386, 1998.
Oleosina de girassol	semente (embrião e semente seca)	Cummins *et al.*, Plant Mol. Biol. 19: 873-876, 1992.

Seqüência Terminadora

O termo "terminador" abrange a seqüência de controle que é uma seqüência de DNA na extremidade de uma unidade transcricional que sinaliza processamento 3' e poliadenilação de um transcrito primário e término de transcrição. Elementos reguladores adicionais podem incluir estimuladores transcricionais assim como traducionais. Aqueles versados na técnica estarão cientes de seqüências terminadoras e estimuladoras que podem ser adequadas para uso para realizar a invenção. Tais seqüências seriam conhecidas ou podem ser prontamente obtidas por uma pessoa versada na técnica.

Marcador Selecionável

O termo "gene marcador selecionável" como referido neste lugar inclui qualquer gene que confere um fenótipo em uma célula na qual ele é expresso para facilitar a identificação e/ou seleção de células que são transfectadas ou transformadas com uma construção de ácidos nucleicos da invenção. Marcadores adequados podem ser selecionados a partir de marcadores que conferem resistência a antibiótico ou herbicida, que introduzem uma nova característica metabólica ou que permitem seleção visual. Exemplos de genes marcadores selecionáveis incluem genes conferindo resistência a antibióticos (tal como *nptII* que fosforila neomicina e canamicina, ou *hpt*, fosforilando higromicina), a herbicidas (por exemplo *bar* que fornece resistência a Basta™; *aroA* ou *gox* fornecendo resistência contra glifosato), ou genes que fornecem uma característica metabólica (tal como

manA que permite que plantas usem manose como única fonte de carbono). Genes marcadores visuais resultam na formação de cor (por exemplo β-glucuronidase, GUS), luminescência (tal como luciferase) ou fluorescência (Proteína Fluorescente Verde, GFP, e derivados destas).

Transformação

O termo "transformação" como referido neste lugar abrange a transferência de um polinucleotídeo exógeno para uma célula hospedeira, independente do método usado para transferência. Tecido vegetal capaz de subseqüente propagação clonal, quer por organogênese ou embriogênese, pode ser transformado com uma construção genética da presente invenção e uma planta inteira regenerada a partir daí. O tecido particular escolhido irá variar dependendo dos sistemas de propagação clonal disponíveis para, e melhor adequados para, a espécie particular sendo transformada. Alvos de tecido exemplares incluem discos foliares, pólen, embriões, cotilédones, hipocótilos, megagametófitos, tecido de calo, tecido meristemático existente (e.g., meristema apical, gemas axilares, e meristemas de raiz), e tecido de meristema induzido (e.g., meristema de cotilédone e meristema de hipocótilo). O polinucleotídeo pode ser transientemente ou estavelmente introduzido em uma célula hospedeira e pode ser mantido não integrado, por exemplo, como um plasmídeo. Alternativamente, ele pode ser integrado no genoma. hospedeiro A célula vegetal transformada resultante pode então ser usada para regenerar uma planta transformada de uma maneira conhecida por pessoas versadas na técnica.

Transformação de espécies vegetais é agora uma técnica bastante rotineira. Vantajosamente, qualquer dos diversos métodos de transformação pode ser usado para introduzir o gene de interesse em uma célula ancestral adequada. Métodos de transformação incluem o uso de lipossomos, eletroporação, produtos químicos que aumentam absorção de DNA livre, injeção do DNA diretamente na planta, bombardeamento por

pistola de partícula, transformação usando vírus ou pólen e microprojeção. Métodos podem ser selecionados a partir do método de cálcio/polietilenoglicol para 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 revestida de DNA ou RNA (Klein TM et al., (1987) Nature 327: 70) infecção com (não integrativa) vírus e os semelhantes. Plantas transgênicas de arroz preferivelmente são produzidas através de transformação mediada por *Agrobacterium* usando qualquer dos métodos bem conhecidos para transformação de arroz, tal como descrito em qualquer dos seguintes: pedido de patente Européia publicado 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 descrições são incorporadas por referência neste lugar como se completamente expostas. No caso de transformação de milho, o método preferido é como descrito ou em Ishida et al. (Nat. Biotechnol 14(6): 745-50, 1996) ou Frame et al. (Plant physiol 129(1): 13-22, 2002), cujas descrições são incorporadas por referência neste lugar como se completamente expostas.

Geralmente depois de transformação, células vegetais ou agrupamentos celulares são selecionados para a presença de um ou mais marcadores que são codificados por genes exprimíveis em plantas co-transferidos com o gene de interesse, seguindo o que o material transformado é regenerado em uma planta inteira.

Seguindo transferência de DNA e regeneração, plantas putativamente transformadas podem ser avaliadas, por exemplo usando análise Southern, para a presença do gene de interesse, número de cópias e/ou organização genômica. Alternativamente ou adicionalmente, níveis de expressão do DNA recentemente introduzido podem ser monitorados usando

análise Northern e/ou Western, ambas técnicas sendo bem conhecidas por pessoas tendo habitual verso na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de maneiras, tal como por propagação clonal ou técnicas clássicas de melhoramento. Por exemplo, uma planta transformada de primeira geração (ou T1) pode ser reproduzida vegetativamente e transformantes homozigotos de segunda geração (ou T2) selecionados, e as plantas T2 podem então ser adicionalmente propagadas através de técnicas clássicas de criação.

Os organismos transformados gerados podem adotar uma variedade de formas. Por exemplo, eles podem ser quimeras de células transformadas e células não transformadas; transformantes clonais (e.g., todas as células transformadas para conter a gaveta de expressão); enxertos de tecidos transformados e não transformados (e.g., em plantas, um porta-enxerto transformado enxertado para um descendente não transformado).

Regulador de Produção de Sementes (SYR) de Descrição Detalhada

A atividade de uma proteína SYR pode ser aumentada por níveis crescentes do polipeptídeo SYR. Alternativamente, atividade também pode ser aumentada quando não existem mudanças em níveis de uma SYR, ou mesmo quando existe uma redução em níveis de uma proteína SYR. Isto pode ocorrer quando as propriedades intrínsecas do polipeptídeo são alteradas, por exemplo, produzindo um mutante ou selecionando uma variante que é mais ativa do que o tipo selvagem.

O termo "proteína SYR ou homólogo desta" como definido neste lugar se refere a um polipeptídeo de cerca de 65 a cerca de 200 aminoácidos, compreendendo (i) um domínio rico em leucina que parece um zíper de leucina na metade C-terminal da proteína, cujo domínio rico em leucina é (ii) precedido por um tripeptídeo com a seqüência YFS (motivo conservado 1a, SEQ ID NO: 6), ou YFT (motivo conservado 1b, SEQ ID NO:

7), ou YFG (motivo conservado 1c, SEQ ID NO: 8) ou YLG (motivo conservado 1d, SEQ ID NO: 9), e (iii) seguido por a motivo conservado 2 ((V/A/I)LAFMP(T/S), SEQ ID NO: 10). Preferivelmente, o motivo conservado 2 é (A/V)LAFMP(T/S), o mais preferivelmente, o motivo conservado é VLAFMPT. A "proteína SYR ou homólogo desta" preferivelmente também tem um peptídeo de término C conservado terminando com o motivo conservado 3 (SYL ou PYL, SEQ ID NO: 11). Domínio rico em leucina da proteína SYR ou seu homólogo é de cerca de 38 a 48 aminoácidos de comprimento, começando imediatamente atrás do motivo conservado 1 e parando imediatamente antes do motivo conservado 2, e compreende pelo menos 30% de leucina. O domínio rico em Leu preferivelmente tem um motivo que parece o motivo Zíper de Leucina (L-X_6-L-X_6-L-X_6-L, caracterizado pelo fato de que X_6 é uma seqüência de 6 aminoácidos consecutivos). Um exemplo preferido de uma proteína SYR é representado por SEQ ID NO: 2, uma visão geral de seus domínios é dada em Figura 1. Deve ser observado que o termo "proteína SYR ou homólogo desta" não abrange a proteína ARGOS de Arabidopsis thaliana (SEQ ID NO: 26).

Ainda preferivelmente, proteínas SYR têm dois domínios transmembrana, com a parte N-terminal e parte C-terminal da proteína localizadas dentro e a parte entre os domínios transmembrana localizada fora.

Alternativamente, o homólogo de uma proteína SYR tem em ordem crescente de 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% de identidade total de seqüência com o aminoácido representado por SEQ ID NO: 2, contanto que a proteína homóloga

compreenda os motivos conservados 1 (a, b, c ou d), 2 e 3, e domínio rico em leucina como descrito acima. A identidade total de seqüência é determinada usando um algoritmo de alinhamento global, tal como o algoritmo de Needleman Wunsch no programa GAP (GCG Wisconsin Package, Accelrys), preferivelmente com parâmetros padrão.

Os vários domínios estruturais em uma proteína SYR podem ser identificados usando bancos de dados especializados e.g. SMART (Schultz *et al.* (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic *et al.* (2002) Nucleic Acids Res 30, 242-244; http://smart.embl-heidelberg.de/), InterPro (Mulder *et al.*, (2003) Nucl. Acids. Res. 31, 315-318; http://www.ebi.ac.uk/interpro/), Prosite (Bucher and Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (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), http://www.expasy.org/prosite/) ou Pfam (Bateman *et al.*, Nucleic Acids Research 30(1):276-280 (2002), http://www.sanger.ac.uk/Software/Pfam/).

Métodos para a busca e identificação de homólogos de SYR estariam bem dentro do âmbito de pessoas versadas na técnica. Tais métodos compreendem comparação das seqüências representadas por SEQ ID NO: 1 ou 2, em um formato legível por computador, com seqüências que estão disponíveis em bancos de dados públicos tal como MIPS (http://mips.gsf.de/), GenBank (http://www.ncbi.nlm.nih.gov/Genbank/index.html) ou Banco de Dados de Seqüência de Nucleotídeos de EMBL (http://www.ebi.ac.uk/embl/index.html), usando algoritmos bem conhecidos na técnica para o alinhamento ou comparação de seqüências, tal como GAP (Needleman and Wunsch, J. Mol. Biol. 48; 443-453 (1970)), BESTFIT

(usando o algoritmo de homologia local de Smith and Waterman (Advances in Applied Mathematics 2; 482-489 (1981))), BLAST (Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J., J. Mol. Biol. 215:403-410 (1990)), FASTA e TFASTA (W. R. Pearson and D. J. Lipman Proc.Natl.Acad.Sci. USA 85:2444- 2448 (1988)). O aplicativo computacional para realizar análise BLAST está publicamente disponível através de National Centre for Biotechnology Information (NCBI).

Domínios transmembrana são de cerca de 15 a 30 aminoácidos de comprimento e normalmente são compostos de resíduos hidrofóbicos que formam uma alfa hélice. Eles normalmente são previstos 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, editores, Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology, páginas 175-182, Menlo Park, CA, 1998. AAAI Press.).

Exemplos de proteínas caindo sob a definição de "polipeptídeo SYR ou um homólogo deste" são listados em Tabela A de Exemplo 1 e incluem seqüências de várias plantas monocotiledôneas, tal como arroz (SEQ ID NO: 2, SEQ ID NO: 12 e SEQ ID NO: 13), milho (SEQ ID NO: 14 e SEQ ID NO: 44), trigo (SEQ ID NO: 15), cevada (SEQ ID NO: 16), cana-de-açúcar (SEQ ID NO: 17 e SEQ ID NO: 18), sorgo (SEQ ID NO: 19); e de plantas dicotiledôneas tal como *Arabidopsis* (SEQ ID NO: 20 e SEQ ID NO: 21), uva (SEQ ID NO: 22), citros (SEQ ID NO: 23) ou tomate (SEQ ID NO: 24 e SEQ ID NO: 25). É contemplado que o domínio rico em Leu é importante para a função da proteína, portanto proteínas com o domínio rico em Leu mas sem os motivos conservados 1 ou 2 podem ser úteis também nos métodos da presente invenção; exemplos de tais proteínas são dados em SEQ ID NO: 34 e 35.

Deve ser entendido que o termo "polipeptídeo SYR ou um

homólogo deste" não deve ser limitado à seqüência representada por SEQ ID NO: 2 ou aos homólogos listados como SEQ ID NO: 12 a SEQ ID NO: 25, mas que qualquer polipeptídeo de cerca de 65 a cerca de 200 aminoácidos reunindo os critérios de compreender um domínio rico em leucina como definido acima, precedido pelo motivo conservado de tripeptídeo 1 (a, b, c ou d) e seguido pelo motivo conservado 2 e preferivelmente também pelo motivo conservado 3; ou tendo pelo menos 38% de identidade de seqüência com a seqüência de SEQ ID NO: 2, pode ser adequado para uso nos métodos da invenção.

Em outra forma de realização, a presente invenção fornece uma proteína SYR selecionada a partir do grupo consistindo de:

(a) um polipeptídeo como dado em SEQ ID NO 44,

(b) um polipeptídeo com uma seqüência de aminoácidos que tem pelo menos, em ordem crescente de preferência, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ou 99% de identidade de seqüência com a seqüência de aminoácidos como dada em SEQ ID NO 44,

(c) um derivado de uma proteína como definido em (a) ou (b).

A seqüência representada por SEQ ID NO: 43 era até agora desconhecida como um gene codificando SYR. Portanto é fornecida uma seqüência isolada de ácidos nucleicos compreendendo:

(i) uma seqüência de ácidos nucleicos representada por SEQ ID NO: 43, ou a fita complementar desta;

(ii) uma seqüência de ácidos nucleicos codificando a seqüência de aminoácidos representada por SEQ ID NO: 44;

(iii) uma seqüência de ácidos nucleicos capaz de hibridizar (preferivelmente em condições estringentes) com uma seqüência de ácidos nucleicos de (i) ou (ii) acima, cuja seqüência hibridizante codifica preferivelmente uma proteína SYR;

(iv) um ácido nucleico que é uma variante alélica para as seqüências de ácidos nucleicos de acordo com (i) ou (ii);

(v) um ácido nucleico que é uma variante de união para as seqüências de ácidos nucleicos de acordo com (i) ou (ii);

(vi) uma seqüência de ácidos nucleicos que tem 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de seqüência com a seqüência definida em (i) ou (ii).

A atividade de uma proteína SYR ou homólogo desta pode ser ensaiada expressando a proteína SYR ou homólogo desta sob controle de um promotor GOS2 em *Oryza sativa*, o que resulta em plantas com produção aumentada de sementes sem um atraso em tempo de florescimento quando comparada com plantas tipo selvagem correspondentes. Este aumento em produção de sementes pode ser medido de diversas maneiras, por exemplo como um aumento de peso total de semente, número de grãos cheios ou índice de colheita.

Uma proteína SYR ou homólogo desta é codificada por um ácido nucleico/gene SYR. Portanto o termo "ácido nucleico/gene SYR" como definido neste lugar é qualquer ácido nucleico/gene codificando uma proteína SYR ou um homólogo deste como definido acima.

Exemplos de ácidos nucleicos *SYR* incluem mas não são limitados àqueles representados por qualquer um de SEQ ID NO: 1, SEQ ID NO: 27 a SEQ ID NO: 32, SEQ ID NO: 36 a 42 e SEQ ID NO: 44. Veja também a lista de ácidos nucleicos mencionados em Tabela A de Exemplo 1.

Ácidos nucleicos/genes *SYR* e variantes destes podem ser adequados para praticar os métodos da invenção. *Ácidos nucleicos*/genes *SYR* variantes incluem porções de um ácido nucleico/gene SYR e/ou ácidos nucleicos capazes de hibridizar com um ácido nucleico/gene SYR.

O termo porção como definido neste lugar se refere a um pedaço de DNA codificando um polipeptídeo de cerca de 65 a cerca de 200

aminoácidos, compreendendo um domínio rico em leucina como definido acima, precedido pelo motivo conservado de tripeptídeo 1 (a, b, c ou d) e seguido pelo motivo conservado 2 e preferivelmente também pelo motivo conservado 3. Preferivelmente, a porção compreende um ou mais dos motivos conservados definidos acima. Uma porção pode ser preparada, por exemplo, fazendo uma ou mais deleções para um ácido nucleico SYR. As porções podem ser usadas em forma isolada ou elas podem ser fundidas a outras seqüências de codificação (ou não codificantes) a fim de, por exemplo, produzir uma proteína que combina diversas atividades. Quando fundido a outras seqüências de codificação, o polipeptídeo resultante produzido em tradução pode ser maior do que aquele previsto para o fragmento SYR. Preferivelmente, a porção é uma porção de um ácido nucleico como representado por qualquer um de SEQ ID NO: 1, SEQ ID NO: 27 a SEQ ID NO: 32, SEQ ID NO: 36 a SEQ ID NO: 42 e SEQ ID NO: 44. Mais preferivelmente a porção de um ácido nucleico é como representada por SEQ ID NO: 1.

Outra variante de um ácido nucleico/gene SYR é um ácido nucleico capaz de hibridizar em condições de estringência reduzida, preferivelmente em condições estringentes, com um ácido nucleico/gene SYR como definido acima, cuja seqüência hibridizante codifica um polipeptídeo de cerca de 65 a cerca de 200 aminoácidos, compreendendo um domínio rico em leucina como definido acima, precedido pelo motivo conservado de tripeptídeo 1 (a, b, c ou d) e seguido pelo motivo conservado 2 e preferivelmente também pelo motivo conservado 3; ou tendo pelo menos 38% de identidade de seqüência com a seqüência de SEQ ID NO: 2.

Preferivelmente, a seqüência hibridizante é um que é capaz de hibridizar com um ácido nucleico como representado por SEQ ID NO: 1, SEQ ID NO: 27 a SEQ ID NO: 32, SEQ ID NO: 36 a SEQ ID NO: 42 e SEQ ID NO: 44, ou com uma porção de qualquer das seqüências mencionadas

acima. Mais preferivelmente a seqüência hibridizante é capaz de hibridizar com SEQ ID NO: 1. O termo "hibridização" é como definido neste lugar.

O ácido nucleico SYR ou variante deste pode ser derivado de qualquer fonte natural ou artificial. O ácido nucleico/gene ou variante deste pode ser isolado de uma fonte microbiana, tal como levedura ou fungos, ou de uma fonte vegetal, algal ou animal (incluindo humano). Este ácido nucleico pode ser modificado a partir de sua forma nativa em composição e/ou ambiente genômico através de manipulação humana deliberada. O ácido nucleico é preferivelmente de origem vegetal, quer da mesma espécie de planta (por exemplo que aquela na qual ele é introduzido) ou quer de uma espécie diferente de planta. O ácido nucleico pode ser isolado de uma espécie monocotiledônea, preferivelmente da família Poaceae, ainda preferivelmente de *Oryza sativa*. Mais preferivelmente, o ácido nucleico SYR é isolado de *Oryza sativa* e é representado por SEQ ID NO: 1, e a seqüência de aminoácidos SYR é como representada por SEQ ID NO: 2.

A expressão de um ácido nucleico codificando um polipeptídeo SYR ou um homólogo deste pode ser modulada introduzindo uma modificação genética (preferivelmente no locus de um gene SYR). O locus de um gene como definido neste lugar é adotado para significar uma região genômica, a qual inclui o gene de interesse e 10 kb antes ou depois da região de codificação.

A modificação genética pode ser introduzida, por exemplo, por qualquer um (ou mais) dos seguintes métodos: ativação de T-DNA, TILLING, mutagênese sítio dirigida, mutagênese por transposon, evolução direta e recombinação homóloga ou introduzindo e expressando em uma planta um ácido nucleico codificando um polipeptídeo SYR ou um homólogo deste. Os métodos mencionados acima são definidos neste lugar na seção intitulada "Definições". Seguindo introdução da modificação genética, segue uma etapa de selecionar para expressão modificada de um ácido nucleico

codificando um polipeptídeo SYR ou um homólogo deste, cuja modificação em expressão gera plantas tendo produção aumentada de sementes.

Ativação de T-DNA, TILLING, mutagênese sítio dirigida, mutagênese por transposon e evolução direta são exemplos de tecnologias que possibilitam a geração de novos alelos e variantes SYR.

Um método preferido para introduzir uma modificação genética (que neste caso não precisa ser no locus de um gene SYR) é introduzir e expressar em uma planta um ácido nucleico codificando um polipeptídeo SYR ou um homólogo deste, como definido neste lugar. O ácido nucleico a ser introduzido em uma planta pode ser um ácido nucleico de comprimento completo ou pode ser uma porção ou uma seqüência hibridizante como definido acima.

"Homólogos" de uma proteína são definidos neste lugar na seção intitulada "Definições". O polipeptídeo SYR ou homólogo deste pode ser um derivado. Para uma definição do termo "derivado" veja a seção neste lugar intitulada "Definições".

O polipeptídeo SYR ou homólogo deste pode ser codificado por uma variante de união alternativa de um ácido nucleico/gene SYR. O termo "variante de união alternativa" é definido na seção "Definições". Variantes de união preferidas são variantes de união do ácido nucleico codificando um polipeptídeo de cerca de 65 a cerca de 200 aminoácidos, compreendendo um domínio rico em leucina como definido acima, precedido pelo motivo conservado de tripeptídeo 1 (a, b, c ou d) e seguido pelo motivo conservado 2 e preferivelmente também pelo motivo conservado 3; ou tendo pelo menos 38% de identidade de seqüência com a seqüência de SEQ ID NO: 2. Ainda preferidas são variantes de união representadas por SEQ ID NO: 1, SEQ ID NO: 27 a SEQ ID NO: 32, SEQ ID NO: 36 a SEQ ID NO: 42 e SEQ ID NO: 44. A mais preferida é a variante de união representada por SEQ ID NO: 1.

O homólogo também pode ser codificado por uma variante alélica de um ácido nucleico codificando um polipeptídeo SYR ou um homólogo deste, preferivelmente uma variante alélica de um ácido nucleico codificando um polipeptídeo de cerca de 65 a cerca de 200 aminoácidos, compreendendo um domínio rico em leucina como definido acima, precedido pelo motivo conservado de tripeptídeo 1 (a, b, c ou d) e seguido pelo motivo conservado 2 e preferivelmente também pelo motivo conservado 3; ou tendo pelo menos 38% de identidade de seqüência com a seqüência de SEQ ID NO: 2. Ainda preferivelmente, a variante alélica codificando o polipeptídeo SYR é representada por qualquer uma de SEQ ID NO: 1, ou SEQ ID NO: 12 a SEQ ID NO: 25. Mais preferivelmente, a variante alélica codificando o polipeptídeo SYR é como representada por SEQ ID NO: 1. O termo "variante alélica" é definido na seção "Definições".

De acordo com um aspecto preferido da presente invenção, expressão aumentada do ácido nucleico SYR ou variante deste é contemplada. Métodos para aumentar expressão de genes ou produtos de gene são bem documentados na técnica e incluem, por exemplo, superexpressão dirigida por promotores apropriados, o uso de estimuladores de transcrição ou estimuladores de tradução. Ácidos nucleicos isolados que servem como elementos promotores ou estimuladores podem ser introduzido em uma posição apropriada (tipicamente antes) de uma forma não heteróloga de um polinucleotídeo de forma a aumentar expressão de um ácido nucleico SYR ou variante deste. Por exemplo, promotores endógenos podem ser alterados in vivo por mutação, deleção, e/ou substituição (veja, Kmiec, Patente Norte-Americana No. 5.565.350; Zarling et al., PCT/US93/03868), ou promotores isolados podem ser introduzidos em uma célula vegetal na orientação e distância apropriada de um gene da presente invenção de forma a controlar a expressão do gene. Métodos para reduzir a expressão de genes ou produtos de gene são bem documentados na técnica.

Se expressão de polipeptídeo é desejada, geralmente é desejável incluir uma região de poliadenilação na extremidade 3' de uma região de codificação de polinucleotídeo. A região de poliadenilação pode ser derivada do gene natural, de uma variedade de outros 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 outro gene vegetal, ou menos preferivelmente de qualquer outro gene eucariótico.

Uma seqüência intrônica também pode ser adicionada à região 5' não traduzida ou à seqüência de codificação da seqüência de codificação parcial para aumentar a quantidade da mensagem madura que se acumula no citossol. Inclusão de um íntron processável na unidade de transcrição tanto em construções vegetais como animais mostrou aumentar expressão gênica tanto no nível de mRNA como protéico em até 1000 vezes, Buchman and Berg, Mol. Cell biol. 8:4395-4405 (1988); Callis et al., Genes Dev. 1:1183-1200 (1987). Tal estimulação intrônica de expressão gênica tipicamente é maior quando colocada perto da extremidade 5' da unidade de transcrição. Uso dos íntrons de milho íntron Adh1-S 1, 2, e 6, o íntron Bronze-1 são conhecidos na técnica. Veja geralmente, The Maize Handbook, capítulo 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

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

Portanto, é fornecida uma construção gênica compreendendo:

(i) um ácido nucleico SYR ou variante deste, como definido acima;

(ii) uma ou mais seqüências de controle capazes de dirigir expressão da seqüência de ácidos nucleicos de (i); e opcionalmente

(iii) uma seqüência de término de transcrição;

com a condição de que a construção gênica não compreenda uma seqüência de ácidos nucleicos codificando a proteína de SEQ ID NO: 26.

Construções úteis nos métodos de acordo com a presente invenção podem ser construídas usando tecnologia de DNA recombinante bem conhecida por pessoas versadas na técnica. As construções gênicas podem ser inseridas em vetores, que podem estar comercialmente disponíveis, adequados para transformar em plantas e adequados para expressão do gene de interesse nas células transformadas.

Plantas são transformadas com um vetor compreendendo a seqüência de interesse (i.e., um ácido nucleico codificando um polipeptídeo SYR ou homólogo deste). A seqüência de interesse é operavelmente ligada a uma ou mais seqüências de controle (pelo menos a um promotor). Os termos "elemento regulador", "seqüência de controle" e "promotor" são todos usados permutavelmente neste lugar e são definidos neste lugar na seção intitulada "Definições".

Vantajosamente, qualquer tipo de promotor pode ser usado para dirigir expressão da seqüência de ácidos nucleicos. Preferivelmente, o ácido nucleico SYR ou variante funcional deste é operavelmente ligado a um promotor constitutivo. Preferivelmente, o promotor constitutivo capaz de expressar preferencialmente o ácido nucleico ao longo de toda a planta tem um perfil de expressão comparável a um promotor GOS2. Mais preferivelmente, o promotor constitutivo tem o mesmo perfil de expressão do que o promotor GOS2 de arroz, o mais preferivelmente, o promotor capaz de expressar preferencialmente o ácido nucleico ao longo de toda a planta é o promotor GOS2 de arroz (SEQ ID NO: 5).

Deve ser claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico SYR representado por SEQ ID NO: 1, nem é a aplicabilidade da invenção restrita a expressão de um ácido nucleico SYR quando dirigida por um promotor GOS2. Um promotor constitutivo

alternativo que é útil nos métodos da presente invenção é o promotor de Proteína de Grupo de Alta Mobilidade (HMGP) (SEQ ID NO: 33). Exemplos de outros promotores constitutivos que também podem ser usados para dirigir expressão de um ácido nucleico SYR são mostrados em Tabela 3 na seção intitulada "Definições".

Opcionalmente, uma ou mais seqüências terminadoras também podem ser usadas na construção introduzida em uma planta. O termo "terminador" é definido na seção "Definições".

As construções genéticas da invenção podem incluir adicionalmente uma seqüência de origem de replicação que é requerida para manutenção e/ou replicação em um tipo celular específico. Um exemplo é quando uma construção genética é requerida ser mantida em uma célula bacteriana como um elemento genético episssomal (e.g. molécula de plasmídeo ou cosmídeo). Origens de replicação preferidas incluem, mas não são limitadas a, a f1-ori e colE1.

A construção genética pode compreender opcionalmente um gene marcador selecionável, como definido na seção "Definições".

A presente invenção também abrange plantas obteníveis pelos métodos de acordo com a presente invenção. A presente invenção portanto fornece plantas obteníveis pelo método de acordo com a presente invenção, cujas plantas têm introduzido nelas um ácido nucleico SYR ou variante deste, como definido acima.

A invenção também fornece um método para a produção de plantas transgênicas tendo produção aumentada de sementes, compreendendo introdução e expressão em uma planta de um ácido nucleico SYR ou uma variante deste como definido acima.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo produção aumentada de sementes, cujo método compreende:

(i) introduzir e expressar em uma planta ou célula vegetal um ácido nucleico SYR ou variante deste, e

(ii) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal;

com a condição de que o ácido nucleico SYR ou variante deste não seja uma seqüência de ácidos nucleicos codificando a proteína de SEQ ID NO: 26.

O ácido nucleico pode ser introduzido diretamente em uma célula vegetal ou na própria planta (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" é definido na seção "Definições".

A presente invenção claramente se estende a qualquer célula vegetal ou planta produzida por qualquer dos métodos descritos neste lugar, e a todas as partes vegetais e propágulos destas. A presente invenção adicionalmente de estende para abranger a progênie de uma célula primária transformada ou transfectada, tecido, órgão ou planta inteira que foi produzida por qualquer dos métodos mencionados acima, o único requerimento sendo que a progênie exiba a(s) mesma(s) característica(s) genotípica(s) e/ou fenotípica(s) que aquelas produzidas pelo ancestral nos métodos de acordo com a invenção. A invenção também inclui células hospedeiras contendo um ácido nucleico SYR isolado ou variante deste. Células hospedeiras preferidas de acordo com a invenção são células vegetais. A invenção também se estende a partes aptas à colheita de uma planta tal como, mas não limitado a sementes, folhas, frutos, flores, safras de caule, rizomas, tubérculos e bulbos. A invenção além disso se refere a produtos diretamente derivados de uma parte apta à colheita de uma tal planta, tal como pelotas secas ou pós, óleo, gordura e ácidos graxos, amido ou proteínas.

A presente invenção também abrange uso de ácidos nucleicos

SYR ou variantes destes e uso de polipeptídeos SYR ou homólogos destes.

Um tal uso se refere a melhorar as características de crescimento de plantas, em particular para melhorar produção de sementes. A produção de sementes pode incluir um ou mais dos seguintes: peso total aumentado de sementes, número aumentado de sementes cheias, taxa de enchimento e índice de colheita aumentado.

Ácidos nucleicos *SYR* ou variantes destes, ou polipeptídeos SYR ou homólogos destes podem encontrar uso em programas de melhoramento nos quais é identificado um marcador de DNA que pode ser geneticamente ligado a um gene SYR ou variante deste. Os ácidos nucleicos/genes *SYR* ou variantes destes, ou polipeptídeos SYR ou homólogos destes podem ser usados para definir um marcador molecular. Este marcador de DNA ou proteína pode então ser usado em programas de melhoramento para selecionar plantas tendo produção aumentada de sementes. O gene *SYR* ou variante deste, por exemplo, pode ser um ácido nucleico como representado por qualquer um de SEQ ID NO: 1, SEQ ID NO: 27 a SEQ ID NO: 32, SEQ ID NO: 36 a SEQ ID NO: 42 e SEQ ID NO: 44.

Variantes alélicas de um ácido nucleico/gene SYR também podem encontrar uso em programas de melhoramento auxiliados por marcador. Tais programas de melhoramento algumas vezes requerem introdução de variação alélica por tratamento mutagênico das plantas, usando por exemplo mutagênese EMS; alternativamente, o programa pode começar com uma coleção de variantes alélicas de assim chamada origem "natural" causada não intencionalmente. Identificação de variantes alélicas então ocorre, por exemplo, por PCR. Isto é seguido por uma etapa para seleção de variantes alélicas superiores da seqüência em questão e que gera produção aumentada de sementes. Seleção tipicamente é realizada monitorando desempenho de crescimento de plantas contendo diferentes variantes alélicas da seqüência em questão, por exemplo, variantes alélicas diferentes de

qualquer uma de SEQ ID NO: 1, SEQ ID NO: 27 a SEQ ID NO: 32, SEQ ID NO: 36 a SEQ ID NO: 42 e SEQ ID NO: 44. Desempenho de crescimento pode ser monitorado em uma casa de vegetação ou no campo. Etapas opcionais adicionais incluem cruzar plantas, nas quais a variante alélica superior foi identificada, com outra planta. Isto pode ser usado, por exemplo, para fazer uma combinação de características fenotípicas interessantes.

Um ácido nucleico SYR ou variante deste também pode ser usado como sondas para mapear geneticamente e fisicamente os genes dos quais eles são uma parte de, e como marcadores para características ligadas a estes genes. Tal informação pode ser útil em melhoramento vegetal a fim de desenvolver linhagens com fenótipos desejados. Tal uso de ácidos nucleicos *SYR* ou variantes destes requer apenas uma seqüência de ácidos nucleicos de pelo menos 15 nucleotídeos de comprimento. Os ácidos nucleicos *SYR* ou variantes destes 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 os ácidos nucleicos *SYR* ou variantes destes. Os padrões de bandeamento resultantes então podem ser sujeitos a análises genéticas usando programas de computador tal como MapMaker (Lander et al. (1987) Genomics 1: 174-181) a fim de construir um mapa genético. Em adição, os ácidos nucleicos podem ser usados para sondar Southern blots contendo DNAs genômicos tratados com endonuclease de restrição de um conjunto de indivíduos representando ancestral e progênie de um cruzamento genético definido. Segregação dos polimorfismos de DNA é observada e usada para calcular a posição do ácido nucleico SYR ou variante deste no mapa genético obtido previamente usando esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32: 314-331).

A produção e uso de sondas derivadas de gene vegetal para uso em mapeamento genético são descritos em Bernatzky and Tanksley

(GENETICS 112 (4): 887-898, 1986). Numerosas publicações descrevem mapeamento genético de clones de cDNA específicos usando a metodologia descrita acima ou variações desta. Por exemplo, populações de intercruzamento de F2, populações endocruzadas, populações cruzadas aleatoriamente, linhagens isogênicas próximas, e outros conjuntos de indivíduos podem ser usados para mapeamento. Tais metodologias são bem conhecidas por aqueles versados na técnica.

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

Em outra forma de realização, as sondas de ácido nucleico podem ser usadas em mapeamento direto por hibridização in situ por fluorescência (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora métodos atuais de mapeamento por FISH favoreçam uso de clones grandes (diversos kb a diversas centenas de kb; veja Laan et al. (1995) Genome Res. 5:13-20), melhoramentos em sensibilidade podem permitir desempenho de mapeamento por FISH usando sondas menores.

Uma variedade de métodos baseados em amplificação de ácido nucleico para mapeamento genético e físico pode ser realizada usando os ácidos nucleicos. Exemplos incluem amplificação alelo específica (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332), ligação alelo específica (Landegren *et al.* (1988) Science 241:1077-1080), reações de extensão de nucleotídeo (Sokolov (1990) Nucleic Acid Res. 18:3671), Mapeamento de Híbrido de Radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e Mapeamento Apropriado (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 iniciadores para uso na

reação de amplificação ou em reações de extensão de iniciador. O projeto de tais iniciadores é bem conhecido por aqueles versados na técnica. Em métodos empregando mapeamento genético baseado em PCR, pode ser necessário identificar diferenças de seqüência de DNA entre os ancestrais do cruzamento de mapeamento na região correspondente à seqüência imediata de ácidos nucleicos. Isto, entretanto, geralmente não é necessário para métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo produção aumentada de sementes, como descrito acima. Estas características de crescimento vantajosas também podem ser combinadas com outras características economicamente vantajosas, tal como adicionalmente características que estimulam produção, tolerância a vários estresses em adição à resistência a estresse abiótico, características modificando vários atributos arquitetônicos e/ou atributos bioquímicos e/ou fisiológicos.

FG-GAP de Descrição Detalhada

A atividade de uma proteína FG-GAP pode ser modulada modulando níveis do polipeptídeo FG-GAP. Alternativamente, atividade também pode ser modulada quando não existem mudanças em níveis de uma FG-GAP. Isto pode ocorrer quando as propriedades intrínsecas do polipeptídeo são alteradas, por exemplo, produzindo um mutante ou selecionando uma variante que é mais ativa ou menos ativa do que o tipo selvagem.

O termo "proteína FG-GAP ou homólogo desta" como definido neste lugar se refere a um polipeptídeo compreendendo (i) um peptídeo sinal de secreção N-terminal, (ii) um ou mais domínios FG-GAP seguido por (iii) um domínio transmembrana na metade C-terminal da proteína. Um exemplo é dado em Figura 6.

Peptídeos sinal são típicos para proteínas que são dirigidas para a via secretora. A presença de um sinal de secreção pode ser facilmente

prevista usando algoritmos computacionais (por exemplo SignalP 3.0, Bendtsen et al., J. Mol. Biol., 340:783-795, 2004). Um sinal de secreção típico consiste de uma n-região carregada positivamente, seguido por uma n-região hidrofóbica e uma polar c-região polar neutra. Além disso, os resíduos de aminoácidos em posição –3 e –1 em relação ao sítio de clivagem normalmente são pequenos e neutros.

Domínios transmembrana são de cerca de 15 a 30 aminoácidos de comprimento e normalmente são compostos de resíduos hidrofóbicos que formam uma alfa hélice. Eles normalmente são previstos 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, editores, Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology, pages 175-182, Menlo Park, CA, 1998. AAAI Press.).

O domínio FG-GAP (número de acesso de Pfam PF01839, INTERPRO entrada IPR000413) tipicamente é encontrado em integrinas onde ele está presente como uma repetição (até 7 cópias) na parte extracelular da proteína. Até agora, apenas integrinas de origem animal foram bem caracterizadas. A seqüência consenso para o domínio FG-GAP é dada em SEQ ID NO: 53:

fgssvaagDlnGDGrpDlvvgaPgadggtdgsvyll,

caracterizada pelo fato de que as letras maiúsculas representam o código de aminoácido de única letra para aminoácidos altamente conservados e as outras letras representam o código de aminoácido de única letra para aminoácidos menos conservados. O domínio freqüentemente compreende um motivo Phe-Gly-X_n-Gly-Ala-Pro caracterizado pelo fato de que X_n representa um número variável de aminoácidos. Devido ao fato desta seqüência consenso ser derivada de proteínas animais, ela não é inteiramente compatível com as seqüências de domínio FG-GAP vegetal. Por exemplo, o hexapeptídeo

"Pgadgg" pode não estar presente em domínios FG-GAP vegetais. Portanto, o termo "domínio FG-GAP" como usado neste lugar abrange SEQ ID NO: 53 e seqüências que têm pelo menos 40% de similaridade de seqüência com SEQ ID NO: 53, sob alinhamento de SEQ ID NO: 53 e a seqüência compatível correspondente, usando o algoritmo de Needleman & Wunsch com uma penalidade de abertura de lacuna de 10 com uma penalidade de extensão de lacuna de 0,5.

O domínio FG-GAP também pode compreender um sítio de ligação de Ca^{2+}.

Preferivelmente, a proteína FG-GAP também compreende um motivo 1 FDGYLYLI(D/E)G (SEQ ID NO: 50). Mais preferivelmente, o motivo conservado 1 é FDGYLYLIDG.

Adicionalmente e/ou alternativamente, a proteína FG-GAP pode compreender um ou mais motivos DGXX(D/E) (motivo conservado 2, SEQ ID NO: 51), caracterizado pelo fato de que X pode ser qualquer aminoácido. Este motivo conservado pode ser parte de um motivo maior DXDXDGXX(D/E) (motivo conservado 3, SEQ ID NO: 52), caracterizado pelo fato de que X pode ser qualquer aminoácido. Assim, a proteína FG-GAP preferivelmente compreende uma ou mais cópias do motivo conservado 3.

Alternativamente, o homólogo de uma proteína FG-GAP tem em ordem crescente de preferência 50%, 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% de identidade total de seqüência com o aminoácido representado por SEQ ID NO: 46, contanto que a proteína homóloga compreenda uma seqüência de peptídeo sinal, um ou mais domínios FG-GAP, e um domínio transmembrana na metade C-terminal da proteína, e preferivelmente também um ou mais dos motivos conservados 1, 2 ou 3. A identidade total de seqüência é determinada

usando um algoritmo de alinhamento global, tal como o algoritmo de Needleman Wunsch no programa GAP (GCG Wisconsin Package, Accelrys), preferivelmente com parâmetros padrão e seqüências protéicas de comprimento completo.

Os vários domínios estruturais em uma proteína FG-GAP podem ser identificados usando bancos de dados especializados e.g. 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), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp. 53-61, 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),).

Métodos para a busca e identificação de homólogos de FG-GAP estariam bem dentro do âmbito de pessoas versadas na técnica. Tais métodos compreendem comparação das seqüências representadas por SEQ ID NO: 45 ou 46, em um formato legível por computador, com seqüências que estão disponíveis em bancos de dados públicos tal como MIPS, GenBank ou Banco de Dados de Seqüências de Nucleotídeos de EMBL, usando algoritmos bem conhecidos na técnica para o alinhamento ou comparação de seqüências, tal como GAP (Needleman and Wunsch, J. Mol. Biol. 48; 443-453 (1970)), BESTFIT (usando o algoritmo de homologia local de Smith and Waterman (Advances in Applied Mathematics 2; 482-489 (1981))), BLAST (Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J., J. Mol. Biol. 215:403-410 (1990)), FASTA e TFASTA (W. R. Pearson and D. J. Lipman Proc.Natl.Acad.Sci. USA 85:2444- 2448 (1988)). O aplicativo computacional

para realizar análise BLAST está publicamente disponível através de National Centre for Biotechnology Information (NCBI).

Exemplos de proteínas caindo sob a definição de "polipeptídeo FG-GAP ou um homólogo deste" incluem uma proteína de *Arabidopsis* (SEQ ID NO: 55) e duas proteínas de arroz (SEQ ID NO: 57 e 59). A presença de proteínas FG-GAP também foi demonstrada em outras espécies de planta das Magnoliophyta, incluindo *Triticum aestivum*, *Zea mays*, *Solanum tuberosum*, *Aquilegia* sp., *Brassica napus*, *Citrus sinensis*, *Asparagus officinalis*, *Populus* sp., *Euphorbia esula* e também em outros taxa vegetais tal como samambaias (*Ceratopteris richardii*) ou em *Welwitschia mirabilis*. Uma lista não limitante de exemplos de EST's codificando proteínas FG-GAP é dada em Tabela 8:

Tabela 8:

Espécie	Acesso de GenBank	SEQ ID NO:
Triticum aestivum	CK207217	16
Zea mays	AY111316	17
Solanum tuberosum	BG598275	18
Aquilegia sp.	DT735817	19
Brassica napus	CX192752	20
Citrus sinensis	CX674859	21
Asparagus officinalis	CV288972	22
Populus sp.	CN520999	23
Populus sp.	CX176799	24
Euphorbia esula	DV130386	25
Ceratopteris richardii	CV736049	26
Welwitschia mirabilis	DT601669	27

As proteínas codificadas pelos genes dos quais estas EST's são derivadas também são úteis para praticar os métodos da presente invenção e caem dentro do escopo desta invenção. Uma pessoa versada na técnica pode ser capaz de isolar a seqüência de codificação de comprimento completo destes genes usando métodos padrão.

A invenção além disso fornece uma proteína FG-GAP isolada selecionada a partir do grupo consistindo de:

(a) uma proteína codificada pelo ácido nucleico de SEQ ID NO: 72;

(b) uma proteína compreendendo uma seqüência sinal, um ou

mais domínios FG-GAP e um domínio transmembrana localizado na metade C-terminal da proteína, caracterizado pelo fato de que dita proteína compreende pelo menos uma de SEQ ID NO: 73 a SEQ ID NO: 72;

(c) um fragmento ativo de uma seqüência de aminoácidos como definido em (a) ou (b), cujo fragmento ativo compreende uma seqüência sinal, um ou mais domínios FG-GAP e um domínio transmembrana localizado na metade C-terminal da proteína.

Deve ser entendido que o termo "polipeptídeo FG-GAP ou um homólogo deste" não deve ser limitado à seqüência representada por SEQ ID NO: 46 ou aos homólogos listados como SEQ ID NO: 55, 57 e 59, mas que qualquer polipeptídeo reunindo os critérios de compreendendo um peptídeo sinal, um ou mais domínios FG-GAP e um domínio transmembrana localizado na metade C-terminal da proteína, e preferivelmente também um ou mais dos motivos conservados de SEQ ID NO: 50 a 52; ou tendo pelo menos 50% de identidade de seqüência com a seqüência de SEQ ID NO: 46, pode ser adequado para uso nos métodos da invenção.

Proteínas FG-GAP vegetais desempenham um papel durante desenvolvimento de pólen (Paxson-Sowders et al. 2001). Em plantas mutantes dex1, deposição de primexina é atrasada e significantemente reduzida. A ondulação normal da membrana plasmática e produção de espaçadores observados em plantas tipo selvagem também é ausente no mutante. Proteínas FG-GAP são capazes de complementar esta mutação e de restaurar o fenótipo normal.

Alternativamente, a atividade de uma proteína FG-GAP ou homólogo desta pode ser ensaiada expressando a proteína FG-GAP ou homólogo desta sob controle de um promotor constitutivo em *Oryza sativa*, o que resulta em plantas com biomassa acima da terra aumentada e/ou produção aumentada de sementes comparado com plantas tipo selvagem correspondentes. Este aumento em produção de sementes pode ser medido de

diversas maneiras, por exemplo como um aumento de total peso de semente, número de grãos cheios ou número total de sementes.

Uma proteína FG-GAP ou homólogo desta é codificada por um ácido nucleico/gene FG-GAP. Portanto o termo "ácido nucleico/gene *FG-GAP*" como definido neste lugar é qualquer ácido nucleico/gene codificando uma proteína FG-GAP ou um homólogo deste como definido acima.

Exemplos de ácidos nucleicos *FG-GAP* incluem mas não são limitados àqueles representados por qualquer um de SEQ ID NO: 45, SEQ ID NO: 54, SEQ ID NO: 56 ou SEQ ID NO: 58. Exemplos de ácidos nucleicos *FG-GAP* parciais são listados em Tabela 8.

A invenção também fornece um ácido nucleico isolado codificando uma proteína FG-GAP, selecionado a partir do grupo consistindo de:

(i) o ácido nucleico como representado em SEQ ID NO: 72;

(ii) um ácido nucleico codificando uma proteína como definido em (a) a (c) acima;

(iii) uma seqüência de ácidos nucleicos capaz de hibridizar (preferivelmente em condições estringentes) com uma seqüência de ácidos nucleicos de (i) ou (ii) acima, cuja seqüência hibridizante preferivelmente codifica uma proteína compreendendo um peptídeo sinal, um ou mais domínios FG-GAP e um domínio transmembrana localizado na metade C-terminal da proteína;

(iv) um ácido nucleico que é uma variante alélica para as seqüências de ácidos nucleicos de acordo com (i) a (iii);

(v) um ácido nucleico que é uma variante alternativa de união para as seqüências de ácidos nucleicos de acordo com (i) a (iii);

(vi) uma porção de uma seqüência de ácidos nucleicos de acordo com qualquer de (i) a (v) acima, cuja porção preferivelmente codifica uma proteína compreendendo um peptídeo sinal, um ou mais domínios FG-

GAP e um domínio transmembrana localizado na metade C-terminal da proteína.

Ácidos nucleicos/genes *FG-GAP* e variantes destes podem ser adequados para praticar os métodos da invenção. Ácidos nucleicos/genes *FG-GAP* variantes incluem porções de um ácido nucleico/gene FG-GAP, variantes alélicas, variantes de união e/ou ácidos nucleicos capazes de hibridizar com um ácido nucleico/gene FG-GAP.

O termo porção como definido neste lugar se refere a um pedaço de DNA codificando um polipeptídeo compreendendo um peptídeo sinal, um ou mais domínios FG-GAP e um domínio transmembrana localizado na metade C-terminal da proteína, e preferivelmente também um ou mais dos motivos conservados de SEQ ID NO: 50 a 52. Preferivelmente, a porção compreende um ou mais dos motivos conservados definidos acima. Uma porção pode ser preparada, por exemplo, fazendo uma ou mais deleções para um ácido nucleico FG-GAP. As porções podem ser usadas em forma isolada ou elas podem ser fundidas a outras seqüências de codificação (ou não codificantes) a fim de, por exemplo, produzir uma proteína que combina diversas atividades. Quando fundido a outras seqüências de codificação, o polipeptídeo resultante produzido em tradução pode ser maior do que aquele previsto para o fragmento FG-GAP. Preferivelmente, a porção é uma porção de um ácido nucleico como representado por qualquer um de SEQ ID NO: 45, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58 ou SEQ ID NO: 72. A porção também pode ser uma porção das seqüências de codificação das quais as seqüências de Tabela 8 são derivadas. Mais preferivelmente a porção de um ácido nucleico é como representada por SEQ ID NO: 45.

Outra variante de um ácido nucleico/gene FG-GAP é um ácido nucleico capaz de hibridizar em condições de estringência reduzida, preferivelmente em condições estringentes, com um ácido nucleico/gene FG-GAP como definido acima, cuja seqüência hibridizante codifica um

polipeptídeo compreendendo um peptídeo sinal, um ou mais domínios FG-GAP e um domínio transmembrana localizado na metade C-terminal da proteína, e preferivelmente também um ou mais dos motivos conservados de SEQ ID NO: 50 a 52.

Preferivelmente, a seqüência hibridizante é uma que é capaz de hibridizar com um ácido nucleico como representado por SEQ ID NO: 45, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58 ou SEQ ID NO: 72, ou com uma porção de qualquer das seqüências mencionadas acima, incluindo as EST's listadas em Tabela 8. Mais preferivelmente a seqüência hibridizante é capaz de hibridizar com SEQ ID NO: 45. O termo "hibridização" é como definido na seção intitulada "Definições".

O ácido nucleico *FG-GAP* ou variante deste pode ser derivado de qualquer fonte natural ou artificial. O ácido nucleico/gene ou variante deste pode ser isolado de uma fonte microbiana, tal como levedura ou fungos, ou de uma fonte vegetal, algal ou animal (incluindo humano). Este ácido nucleico pode ser modificado a partir de sua forma nativa em composição e/ou ambiente genômico através de manipulação humana deliberada. O ácido nucleico é preferivelmente de origem vegetal, quer da mesma espécie de planta (por exemplo que aquela na qual ele é introduzido) ou quer de uma espécie diferente de planta. O ácido nucleico pode ser isolado de uma espécie dicotiledônea, preferivelmente da família Brassicaceae, ainda preferivelmente de *Arabidopsis thaliana*. Mais preferivelmente, o ácido nucleico *FG-GAP* é isolado de *Arabidopsis thaliana* e é representado por SEQ ID NO: 45, e a seqüência de aminoácidos FG-GAP é como representada por SEQ ID NO: 46.

A expressão de um ácido nucleico codificando um polipeptídeo FG-GAP ou um homólogo deste pode ser modulada introduzindo uma modificação genética (preferivelmente no locus de um gene FG-GAP). O locus de um gene como definido neste lugar é adotado para significar uma região genômica, a qual inclui o gene de interesse e 10 kb

antes ou depois da região de codificação.

A modificação genética pode ser introduzida, por exemplo, por qualquer um (ou mais) dos seguintes métodos: ativação de T-DNA, TILLING, mutagênese sítio dirigida, mutagênese por transposon, evolução direta e recombinação homóloga ou introduzindo e expressando em uma planta um ácido nucleico codificando um polipeptídeo FG-GAP ou um homólogo deste. Estes métodos são definidos na seção intitulada "Definições". Seguindo introdução da modificação genética, segue uma etapa de selecionar para expressão modificada de um ácido nucleico codificando um polipeptídeo FG-GAP ou um homólogo deste, cuja modificação em expressão gera plantas tendo produção aumentada.

Ativação de T-DNA, TILLING, mutagênese sítio dirigida, mutagênese por transposon e evolução direta são exemplos de tecnologias que possibilitam a geração de novos alelos e variantes de FG-GAP.

Um método preferido para introduzir uma modificação genética (que neste caso não precisa ser no locus de um gene FG-GAP) é introduzir e expressar em uma planta um ácido nucleico codificando um polipeptídeo FG-GAP ou um homólogo deste, como definido acima. O ácido nucleico a ser introduzido em uma planta pode ser um ácido nucleico de comprimento completo ou pode ser uma porção ou uma seqüência hibridizante como definido acima. Preferivelmente, a planta na qual a modificação genética é introduzida não é uma planta mutante *dex1*, na qual o gene *DEX1* não é funcional (Paxson-Sowders et al. 2001).

"Homólogos" de uma proteína são definidos na seção intitulada "Definições". O polipeptídeo FG-GAP ou homólogo deste pode ser um derivado, como definido na seção "Definições".

O polipeptídeo FG-GAP ou homólogo desta pode ser codificado por uma variante de união alternativa de um ácido nucleico/gene FG-GAP. O termo "variante de união alternativa" é como definido neste

lugar. São preferidas variantes de união do ácido nucleico codificando um polipeptídeo compreendendo um peptídeo sinal, um ou mais domínios FG-GAP e um domínio transmembrana localizado na metade C-terminal da proteína, e preferivelmente também um ou mais dos motivos conservados de SEQ ID NO: 50 a 52. Ainda preferidas são variantes de união representadas por SEQ ID NO: 45, SEQ ID NO: 54, SEQ ID NO: 56 ou SEQ ID NO: 58, ou uma variante de união do ácido nucleico representado por SEQ ID NO: 72, ou uma variante de união de um dos genes dos quais as seqüências em Tabela 8 são derivadas. A mais preferida é a variante de união representada por SEQ ID NO: 45.

O homólogo também pode ser codificado por uma variante alélica de um ácido nucleico codificando um polipeptídeo FG-GAP ou um homólogo deste, preferivelmente uma variante alélica de um ácido nucleico codificando um polipeptídeo compreendendo um peptídeo sinal, um ou mais domínios FG-GAP e um domínio transmembrana localizado na metade C-terminal da proteína, e preferivelmente também um ou mais dos motivos conservados de SEQ ID NO: 50 a 52. Ainda preferivelmente, a variante alélica codificando o polipeptídeo FG-GAP é representada por qualquer uma de SEQ ID NO: 45, SEQ ID NO: 54, SEQ ID NO: 56 ou SEQ ID NO: 58. Mais preferivelmente, a variante alélica codificando o polipeptídeo FG-GAP é como representado por SEQ ID NO: 45. Variantes alélicas são definidas na seção "Definições".

De acordo com um aspecto preferido da presente invenção, expressão modulada do ácido nucleico FG-GAP ou variante deste é contemplada. Preferivelmente, a expressão modulada é superexpressão. Métodos para superexpressão de genes ou produtos de gene são bem documentados na técnica e incluem, por exemplo, superexpressão dirigida por promotores apropriados, o uso de estimuladores de transcrição ou estimuladores de tradução. Ácidos nucleicos isolados que servem como

elementos promotores ou estimuladores podem ser introduzidos em uma posição apropriada (tipicamente antes) de uma forma não heteróloga de um polinucleotídeo de forma a aumentar expressão de um ácido nucleico FG-GAP ou variante deste. Por exemplo, promotores endógenos podem ser alterados in vivo por mutação, deleção, e/ou substituição (veja, Kmiec, Patente Norte-Americana No. 5.565.350; Zarling et al., PCT/US93/03868), ou promotores isolados podem ser introduzidos em uma célula vegetal na orientação e distância apropriada de um gene da presente invenção de forma a controlar a expressão do gene. Métodos para reduzir a expressão de genes ou produtos de gene também são bem documentados na técnica.

Se expressão de polipeptídeo é desejada, geralmente é desejável incluir uma região de poliadenilação na extremidade 3' de uma região de codificação de polinucleotídeo. A região de poliadenilação pode ser derivada do gene natural, de uma variedade de outros 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 outro gene vegetal, ou menos preferivelmente de qualquer outro gene eucariótico.

Uma seqüência intrônica também pode ser adicionada à região 5' não traduzida ou à seqüência de codificação da seqüência de codificação parcial para aumentar a quantidade da mensagem madura que se acumula no citossol. Inclusão de um íntron processável na unidade de transcrição tanto em construções vegetais como animais mostrou aumentar expressão gênica tanto no nível de mRNA como protéico em até 1000 vezes, Buchman and Berg, Mol. Cell biol. 8:4395-4405 (1988); Callis et al., Genes Dev. 1:1183-1200 (1987). Tal estimulação intrônica de expressão gênica tipicamente é maior quando localizada perto da extremidade 5' da unidade de transcrição. Uso dos íntrons de milho íntron Adh1-S 1, 2, e 6, o íntron Bronze-1 são conhecidos na técnica. Veja geralmente, The Maize Handbook, capítulo 116, Freeling and

Walbot, Eds., Springer, N.Y. (1994).

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

Portanto, é fornecida uma construção gênica compreendendo:

(i) um ácido nucleico FG-GAP ou variante deste, como definido acima;

(ii) uma ou mais seqüências de controle capazes de dirigir expressão da seqüência de ácidos nucleicos de (i); e opcionalmente

(iii) uma seqüência de término de transcrição;

com a condição de que a construção gênica não é uma construção gênica tipo pPZP como descrito por Hajdukiewicz et al. (Plant Mol. Biol. 25, 989-994) e Paxson-Sowders (2001).

Construções úteis nos métodos de acordo com a presente invenção podem ser construídas usando tecnologia de DNA recombinante bem conhecida por pessoas versadas na técnica. As construções gênicas podem ser inseridas em vetores, que podem estar comercialmente disponíveis, adequados para transformar em plantas e adequados para expressão do gene de interesse nas células transformadas.

Plantas são transformadas com um vetor compreendendo a seqüência de interesse (i.e., um ácido nucleico codificando um polipeptídeo FG-GAP ou homólogo desta). A seqüência de interesse é operavelmente ligada a uma ou mais seqüências de controle (pelo menos a um promotor). Os termos "elemento regulador", "seqüência de controle" e "promotor" são todos usados permutavelmente neste lugar e são definidos na seção intitulada "Definições".

Vantajosamente, qualquer tipo de promotor pode ser usado para dirigir expressão da seqüência de ácidos nucleicos. Preferivelmente, o ácido nucleico FG-GAP ou variante funcional deste é operavelmente ligado a

um promotor constitutivo. O termo "constitutivo" é como definido neste lugar. Preferivelmente, o promotor constitutivo capaz de expressar preferencialmente o ácido nucleico ao longo de toda a planta tem um perfil de expressão comparável a um promotor GOS2. Mais preferivelmente, o promotor constitutivo tem o mesmo perfil de expressão do que o promotor GOS2 de arroz, o mais preferivelmente, o promotor capaz de expressar preferencialmente o ácido nucleico ao longo de toda a planta é o promotor GOS2 de arroz (nucleotídeos 1 a 2193 da seqüência representada em SEQ ID NO: 49). Deve ser claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico FG-GAP representado por SEQ ID NO: 45, nem é a aplicabilidade da invenção restrita a expressão de um ácido nucleico FG-GAP quando dirigida por um promotor GOS2. Exemplos de outros promotores constitutivos que também podem ser usados para dirigir expressão de um ácido nucleico FG-GAP são mostrados em Tabela 3 na seção "Definições".

Opcionalmente, uma ou mais seqüências terminadoras também podem ser usadas na construção introduzida em uma planta. O termo "terminador" sendo definido na seção "Definições".

As construções genéticas da invenção podem incluir adicionalmente uma seqüência de origem de replicação que é requerida para manutenção e/ou replicação em um tipo celular específico. Um exemplo é quando uma construção genética é requerida ser mantida em uma célula bacteriana como um elemento genético episssomal (e.g. molécula de plasmídeo ou cosmídeo). Origens de replicação preferidas incluem, mas não são limitadas a, a f1-ori e colE1.

A construção genética pode compreender opcionalmente um gene marcador selecionável como definido na seção "Definições" neste lugar.

A presente invenção também abrange plantas obteníveis pelos métodos de acordo com a presente invenção. A presente invenção portanto fornece plantas obteníveis pelo método de acordo com a presente invenção,

cujas plantas têm introduzido nelas um ácido nucleico FG-GAP ou variante deste, como definido acima.

A invenção também fornece um método para a produção de plantas transgênicas tendo produção aumentada, compreendendo introdução e expressão em uma planta de um ácido nucleico FG-GAP ou uma variante deste como definido acima.

Mais especificamente, a presente invenção fornece a método para a produção de plantas transgênicas tendo produção aumentada, cujo método compreende:

(i) introduzir e expressar em a planta ou célula vegetal um ácido nucleico FG-GAP ou variante deste; e

(ii) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

O ácido nucleico pode ser introduzido diretamente em uma célula vegetal ou na própria planta (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" é como definido na seção "Definições".

A presente invenção claramente se estende a qualquer célula vegetal ou planta produzida por qualquer dos métodos descritos neste lugar, e a todas as partes vegetais e propágulos destas. A presente invenção adicionalmente de estende para abranger a progênie de uma célula primária transformada ou transfectada, tecido, órgão ou planta inteira que foi produzida por qualquer dos métodos mencionados acima, o único requerimento sendo que a progênie exiba a(s) mesma(s) característica(s) genotípica(s) e/ou fenotípica(s) que aquelas produzidas pelo ancestral nos métodos de acordo com a invenção. A invenção também inclui células hospedeiras contendo um

ácido nucleico FG-GAP isolado ou variante deste. Células hospedeiras preferidas de acordo com a invenção são células vegetais. A invenção também se estende a partes aptas à colheita de uma planta tal como, mas não limitada a sementes, folhas, frutos, flores, safras de caule, rizomas, tubérculos e bulbos. A invenção além disso se refere a produtos derivados, preferivelmente diretamente derivados, de uma parte apta à colheita de uma tal planta, tal como pelotas secas ou pós, óleo, gordura e ácidos graxos, amido e proteínas.

A presente invenção também abrange uso de ácidos nucleicos *FG-GAP* ou variantes destes e uso de polipeptídeos FG-GAP ou homólogos destes.

Um tal uso se refere a melhorar as características de crescimento de plantas, em particular para melhorar produção, especialmente produção de sementes. A produção de sementes pode incluir um ou mais dos seguintes: peso total aumentado de sementes, número aumentado de sementes cheias e número total aumentado de sementes.

Ácidos nucleicos *FG-GAP* ou variantes destes, ou polipeptídeos FG-GAP ou homólogos destes podem encontrar uso em programas de melhoramento nos quais é identificado um marcador de DNA que pode ser geneticamente ligado a um gene *FG-GAP* ou variante deste. Os ácidos nucleicos/genes FG-GAP ou variantes destes, ou polipeptídeos FG-GAP ou homólogos destes pode ser usado para definir um marcador molecular. Este marcador de DNA ou proteína pode então ser usado em programas de melhoramento para selecionar plantas tendo produção aumentada. O gene *FG-GAP* ou variante deste, por exemplo, pode ser um ácido nucleico como representado por qualquer um de SEQ ID NO: 45, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, e SEQ ID NO: 72, ou genes dos quais as seqüências listadas em Tabela 8 foram derivadas.

Variantes alélicas de um ácido nucleico/gene FG-GAP também podem encontrar uso em programas de melhoramento auxiliados por

marcador. Tais programas de melhoramento algumas vezes requerem introdução de variação alélica por tratamento mutagênico das plantas, usando por exemplo mutagênese EMS; alternativamente, o programa pode começar com uma coleção de variantes alélicas de assim chamada origem "natural" causada não intencionalmente. Identificação de variantes alélicas então ocorre, por exemplo, por PCR. Isto é seguido por uma etapa para seleção de variantes alélicas superiores da seqüência em questão e que gera produção aumentada. Seleção tipicamente é realizada monitorando desempenho de crescimento de plantas contendo diferentes variantes alélicas da seqüência em questão, por exemplo, variantes alélicas diferentes de qualquer um de SEQ ID NO: 45, SEQ ID NO: 54, SEQ ID NO: 56, SEQ ID NO: 58, e SEQ ID NO: 72, ou de uma das seqüências de codificação de quais as seqüências listadas em Tabela 8 foram derivadas. Desempenho de crescimento pode ser monitorado em uma casa de vegetação ou no campo. Etapas opcionais adicionais incluem cruzar plantas, nas quais a variante alélica superior foi identificada, com outra planta. Isto pode ser usado, por exemplo, para fazer uma combinação de características fenotípicas interessantes.

Um ácido nucleico FG-GAP ou variante deste também pode ser usado como sondas para mapear geneticamente e fisicamente os genes dos quais eles são uma parte de, e como marcadores para características ligadas a estes genes. Tal informação pode ser útil em melhoramento vegetal a fim de desenvolver linhagens com fenótipos desejados. Tal uso de ácidos nucleicos *FG-GAP* ou variantes destes requer apenas uma seqüência de ácidos nucleicos de pelo menos 15 nucleotídeos de comprimento. Os ácidos nucleicos FG-GAP ou variantes destes 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 os ácidos nucleicos *FG-GAP* ou variantes destes. Os padrões de

bandeamento resultantes então podem ser sujeitos a análises genéticas usando programas de computador tal como MapMaker (Lander et al. (1987) Genomics 1: 174-181) a fim de construir um mapa genético. Em adição, os ácidos nucleicos podem ser usados para sondar Southern blots contendo DNAs genômicos tratados com endonuclease de restrição de um conjunto de indivíduos representando ancestral e progênie de um cruzamento genético definido. Segregação dos polimorfismos de DNA é observada e usada para calcular a posição do ácido nucleico FG-GAP ou variante deste no mapa genético obtido previamente usando esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32: 314-331).

A produção e uso de sondas derivadas de gene vegetal para uso em mapeamento genético são descritos em Bernatzky and Tanksley (Plant Mol. Biol. Reporter 4: 37-41, 1986). Numerosas publicações descrevem mapeamento genético de clones de cDNA específicos usando a metodologia descrita acima ou variações desta. Por exemplo, populações de intercruzamento de F2, populações endocruzadas, populações cruzadas aleatoriamente, linhagens isogênicas próximas, e outros conjuntos de indivíduos podem ser usados para mapeamento. Tais metodologias são bem conhecidas por aqueles versados na técnica.

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

Em outra forma de realização, as sondas de ácido nucleico podem ser usadas em mapeamento direto por hibridização in situ por fluorescência (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora métodos atuais de mapeamento por FISH favoreçam uso de clones grandes (diversos kb a diversas centenas kb; veja Laan et al. (1995) Genome Res. 5:13-20), melhoramentos em sensibilidade podem permitir desempenho de

mapeamento por FISH usando sondas menores.

Uma variedade de métodos baseados em amplificação de ácido nucleico para mapeamento genético e físico pode ser realizada usando os ácidos nucleicos. Exemplos incluem amplificação alelo específica (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332), ligação alelo específica (Landegren *et al.* (1988) Science 241:1077-1080), reações de extensão de nucleotídeo (Sokolov (1990) Nucleic Acid Res. 18:3671), Mapeamento de Híbrido de Radiação (Walter et al. (1997) Nat. Genet. 7:22-28) e Mapeamento Apropriado (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 iniciadores para uso na reação de amplificação ou em reações de extensão de iniciador. O projeto de tais iniciadores é bem conhecido por aqueles versados na técnica. Em métodos empregando mapeamento genético baseado em PCR, pode ser necessário identificar diferenças de seqüência de DNA entre os ancestrais do cruzamento de mapeamento na região correspondente à seqüência imediata de ácidos nucleicos. Isto, entretanto, geralmente não é necessário para métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo produção aumentada, como descrito acima. Estas características de crescimento vantajosas também podem ser combinadas com outras características economicamente vantajosas, tal como características adicionais que estimulam produção, tolerância a vários estresses, características modificando vários atributos arquitetônicos e/ou atributos bioquímicos e/ou fisiológicos.

CYP90B de Descrição Detalhada

O termo "polipeptídeo CYP90B ou homólogo desta" como definido neste lugar se refere a um polipeptídeo compreendendo os seguintes:

(a) domínios CYP A a D; (b) um domínio âncora hidrofóbico N-terminal; (c) um domínio de transição; e (d) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição.

Além disso, o polipeptídeo CYP90B ou homólogo desta pode compreender adicionalmente (i) uma seqüência com mais do que 50% de identidade com SEQ ID NO: 78 e (ii) atividade enzimática de esteróide 22 alfa hidroxilase.

Exemplos de um polipeptídeo CYP90B como definido acima são dados em Tabela 9a neste lugar.

Um polipeptídeo CYP90B ou homólogo deste é codificado por um ácido nucleico/gene CYP90B. Portanto o termo "ácido nucleico/gene CYP90B" como definido neste lugar é qualquer ácido nucleico/gene codificando um polipeptídeo CYP90B ou um homólogo deste como definido acima.

Os vários domínios estruturais encontrados na superfamília CYP de proteínas, incluindo em polipeptídeos CYP90B da presente invenção, são bem conhecidos na técnica e podem ser identificados usando bancos de dados gerais e.g. SMART (Schultz *et al.* (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic *et al.* (2002) Nucleic Acids Res 30, 242-244; http://smart.embl-heidelberg.de/), InterPro (Mulder *et al.*, (2003) Nucl. Acids. Res. 31, 315-318; http://www.ebi.ac.uk/interpro/), Prosite (Bucher and Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation, *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), http://www.expasy.org/prosite/) ou Pfam (Bateman *et al.*, Nucleic Acids Research 30(1):276-280 (2002),

http://www.sanger.ac.uk/Software/Pfam/).

Bancos de dados especializados também podem ser procurados em http://arabidopsis-P450.biotec.uiuc.edu/cgi-bin/p450.pl para Arabidopsis, ou mais geralmente na Página da Internet de CYP http://drnelson.utmem.edu/CytochromeP450.html. Domínios estruturais típicos encontrados em CYP podem ser os quatro domínios A a D como descrito originalmente por Kalb & Loper ((1988) Proc Natl Acad Sci 85: 7221-7225). O domínio A (também chamado hélice I) compreende a seqüência consenso Ala/Gly-Gly-X-Asp/Glu-Thr-Thr/Ser, e é proposto se ligar a dioxigênio. O domínio B é o domínio de ligação de esteróide. O domínio D corresponde ao domínio de ligação heme e compreende a mais característica seqüência consenso de aminoácidos CYP (Phe-X-X-Gly-X-Arg-X-Cys-X-Gly) (Figuras 10 e 13).

A presença de seqüências consenso pode ser identificada usando métodos para o alinhamento de seqüências para comparação como descrito acima. Em algumas instâncias, os parâmetros padrão podem ser ajustados para modificar a estringência da busca. Por exemplo usando BLAST, o limiar de significância estatística (chamado valor "esperado") para relatar identidades contra seqüências de banco de dados pode ser aumentado para mostrar identidades menos estringentes. Desta maneira, identidades curtas praticamente exatas podem ser identificadas. A seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser dentro do domínio A do polipeptídeo CYP90B (compreendendo a seqüência consenso Ala/Gly-Gly-X-Asp/Glu-Thr-Thr/Ser como definida acima) como definido neste lugar pode ser identificada desta maneira, como uma pessoa versada na técnica deve estar bem ciente.

Outro domínio identificado em proteínas CYP P450, e em particular no polipeptídeo CYP90B da invenção, pode ser o domínio âncora no término N da proteína para direcionamento de membrana, rico em resíduos

hidrofóbicos tal como Leu, Ile, Val, pH e Ala. O domínio âncora N-terminal é tipicamente entre 20 a 40 aminoácidos de comprimento, mas pode ser menor (até 10 aminoácidos) ou maior (até 100 aminoácidos). O domínio âncora N-terminal é separado do resto da proteína (domínio globular) por um domínio de transição compreendendo um agrupamento de resíduos básicos (pelo menos dois, ou Lys ou Arg, chamado o sinal de interrupção de transferência) precedendo um agrupamento de prolina que forma uma dobradiça entre o domínio âncora mencionado acima e o domínio globular da proteína. Uma seqüência consenso típica para o domínio de transição é Lys/Arg-Lys/Arg-$(X)_{3-9}$-Pro-Pro-Gly (Figuras 10 e 13). Uma tal seqüência consenso pode ser identificada como mencionado acima.

A presença de um domínio âncora hidrofóbico N-terminal pode ser prontamente identificada. Composição primária de aminoácidos (em%) para determinar se um domínio de polipeptídeo é rico em aminoácidos específicos pode ser calculada usando programas de aplicativos computacionais do servidor ExPASy, em particular a ferramenta ProtParam (Gasteiger E *et al.* (2003) ExPASy: o servidor de proteômicos para conhecimento e análise detalhada de proteína. Nucleic Acids Res 31:3784-3788). A composição da proteína de interesse pode ser então comparada com a composição média de aminoácidos (em%) no banco de dados Swiss-Prot Protein Sequence. Dentro deste banco de dados, a adição das médias de Leu (L), Ile (I), Val (V), pH e (F) e Ala (A) é de 34,04%. Como um exemplo, o domínio âncora hidrofóbico N-terminal de SEQ ID NO: 78 contém 62,5% dos mesmos resíduos hidrofóbicos. Como definido neste lugar, um domínio âncora hidrofóbico N-terminal tem um conteúdo de aminoácidos hidrofóbicos (em termos de%) acima do que aquele na composição média de aminoácidos (em termos de%) das proteínas no banco de dados Swiss-Prot Protein Sequence.

Aplicativos computacionais especiais tal como ProtScale

(Gasteiger et al. (2005) Protein Identification and Analysis Tools on the ExPASy Server. *In* John M. Walker, ed: The Proteomics Protocols Handbook, Humana Press pp. 571-607) computam e representam o perfil produzido por qualquer escala de aminoácidos em uma proteína selecionada. Uma escala de aminoácidos é definida por um valor numérico assinalado a cada tipo de aminoácido. As escalas mais freqüentemente usadas são as escalas de hidrofobicidade ou hidrofilicidade e as escalas de parâmetros conformacionais de estrutura secundária. Uma das mais freqüentemente usadas escalas de hidrofobicidade de aminoácidos foi produzida por Kyte & Doolittle ((1982) J. Mol. Biol. 157:105-132), na qual aminoácidos hidrofóbicos tiveram num número positivo atribuído, e aminoácidos hidrofílicos um número negativo. Por exemplo, o perfil de resultado de ProtScale para hidrofobicidade do polipeptídeo CYP90B da invenção mostra claramente que aproximadamente os 34 primeiros aminoácidos N-terminais representam um domínio hidrofóbico, pois estes estão localizados acima da linha delimitando zero (Figura 12). Esta região corresponde ao domínio âncora N-terminal. Uma pessoa versada na técnica deve estar bem ciente de tais análises.

Polipeptídeos CYP90B ou homólogos destes podem ser prontamente identificados usando técnicas de rotina bem conhecidas na técnica, tal como por alinhamento de seqüências. 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 de duas seqüências completas que maximiza o número de identidades e minimiza o número de lacunas. O algoritmo BLAST (Altschul *et al.* (1990) J Mol Biol 215: 403-10) calcula identidade de seqüência percentual e realiza uma análise estatística da similaridade entre as duas seqüências. O aplicativo computacional para realizar análise BLAST está publicamente disponível através de National Centre for Biotechnology

Information. Homólogos de CYP90B compreendendo uma seqüência com mais do que 50% de identidade com SEQ ID NO: 78 podem ser prontamente identificados usando, por exemplo, o algoritmo de alinhamento de múltiplas seqüências ClustalW (versão 1.83) disponível em http://clustalw.genome.jp/sit-bin/nph-ClustalW, com os parâmetros padrão de alinhamento pareado, e um método de pontuação em porcentagem. Edição manual secundária pode ser realizada para otimizar alinhamento entre motivos conservados, como deve ser perceptível para uma pessoa versada na técnica.

Exemplos de polipeptídeos CYP90B ou homólogos destes (codificados por seqüência de polinucleotídeos número de acesso em parênteses) são dados em Tabela 9a. Tabela 9b sustenta seqüências CYP90B parciais codificando quadros de leitura aberta (ORF) CYP90B parciais.

Tabela 9: a) Exemplos de homólogos de CYP90B

Nome	Número de acesso de nucleotídeo NCBI ou TIGR	Nucleotídeo SEQ ID NO	Polipeptídeo traduzido SEQ ID NO	Estado de inserto	Fonte
Orysa_CYP90B	AB206579.1	77	78	ORF de comprimento completo	*Oryza sativa*
Arath_CYP90B1	NM_114926.2	79	80	ORF de comprimento completo	*Arabidopsis thaliana*
Sacof_CYP90B**	CA092707.1 CF574030,1 CA217329.1	81	82	ORF de comprimento completo	*Saccharum officinarum*
Allce_CYP90B	TC2113	83	84	ORF de comprimento completo	*Allium cepa*
Zinel_CYP90B	AB231155	85	86	ORF de comprimento completo	*Zinnia elegans*
Medtr_CYP90B*	AC147964.10	87	88	ORF de comprimento completo	*Medicago trunculata*
Poptr_CYP90B****	CK090847.1 CV280598.1 DT503533.1	89	90	ORF de comprimento completo	*Populus trichocarpa*

Tabela 9: b) Exemplos de CYP90B com um quadro aberto de leitura (ORF) parcial

Nome	Número de acesso de nucleotídeo NCBI ou TIGR	Nucleotídeo SEQ ID NO	Polipeptídeo traduzido SEQ ID NO	Estado do Inserto	Fonte
Aqufo_CYP90B**	DR940523.1 DR940522.1	91	92	ORF parcial	*Aquilegia formosa x Aquilegia pubescens*
Triae_CYP90B extremidade 5'	BQ620306.1	93	94	ORF parcial	*Triticum aestivum*
Triae_CYP90B extremidade 3' **	BQ619714.1 CA715360,1	95	96	ORF parcial	*Triticum aestivum*
Eupes_CYP90B	DV141872.1	97	98	ORF parcial	*Euphorbia esula*
Goshi_CYP90B extremidade 5' **	CO125422 DT568185.1	99	100	ORF parcial	*Gossypium hirsutum*
Lyces_CYP90B extremidade 5' **	BF050501 AW221826.1 BM409833	101	102	ORF parcial	*Lycopersicon esculentum*
Soltu_CYP90B extremidade 5' **	BQ045917 BQ114367	103	104	ORF parcial	*Solanum tuberosum*
Soltu_CYP90B extremidade 3' **	BQ114368	105	106	ORF parcial	*Solanum tuberosum*

* Processamento manual a partir de clone genômico

**Agrupamento de seqüências compilado a partir de diversos acessos EST (os principais mostrados); qualidade de seqüenciamento de EST sendo normalmente inferior, umas poucas substituições de aminoácidos podem ser esperadas.

Deve ser entendido que seqüências caindo sob a definição de "polipeptídeo CYP90B ou homólogo desta" não devem ser limitadas às seqüências representadas por SEQ ID NO: 78, SEQ ID NO: 80, SEQ ID NO: 82, SEQ ID NO: 84, SEQ ID NO: 86, SEQ ID NO: 88 ou SEQ ID NO: 90, mas que qualquer polipeptídeo compreendendo os seguintes: (a) domínios CYP A a D; (b) um domínio âncora hidrofóbico N-terminal; (c) um domínio de transição; e (d) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer

posição pode ser adequado para uso em desempenho da invenção.

As seqüências caindo sob a definição de "polipeptídeo CYP90B ou homólogo desta" podem compreender adicionalmente (i) uma seqüência com mais do que 50% de identidade com SEQ ID NO: 78 e (ii) atividade enzimática de esteróide 22 alfa hidroxilase.

Polipeptídeos CYP90B ou homólogos destes têm atividade enzimática de 22-alfa hidroxilase, a qual pode ser determinado por teste de complementação usando plantas tendo uma mutação em DWF4. Tais plantas mutantes são descritas em *Arabidopsis* (*dwf4* mutante) por Choe *et al.* ((1998) Plant Cell 10:231-243) e em arroz (Tos2091 mutante) por Tanaka et al (US2004/0060079). O tamanho destas plantas mutantes é diversas vezes menor do que aquele de seus tipo selvagens correspondentes, i.e., as plantas mutantes são super definhadas. O polipeptídeo isolado é colocado sob o controle de um promotor capaz de expressar este polipeptídeo em plantas, em um vetor de DNA recombinante adequado para transformação vegetal. As plantas mutantes são então transformada com este vetor, usando técnicas que são bem conhecidas na técnica. Se as plantas transformadas não apresentam mais o fenótipo super definhado que é indicativo de que o polipeptídeo isolado é capaz de apresentar atividade enzimática de 22-alfa hidroxilase. Um tal polipeptídeo pode ser adequado para uso em desempenho dos métodos da invenção.

Exemplos de ácidos nucleicos CYP90B incluem mas não são limitados àqueles representado por qualquer um de SEQ ID NO: 77, SEQ ID NO: 79, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87 ou SEQ ID NO: 89. Ácidos nucleicos/genes CYP90B e variantes destes podem ser adequados para praticar os métodos da invenção. Variantes de ácidos nucleicos/genes CYP90B incluem porções de um ácido nucleico/gene CYP90B e/ou ácidos nucleicos capazes de hibridizar com um ácido nucleico/gene CYP90B.

O termo porção como definido neste lugar se refere a um pedaço de DNA codificando um polipeptídeo compreendendo os seguintes: (a) domínios CYP P450 A a D; (b) um domínio âncora hidrofóbico N-terminal; (c) um domínio de transição; e (d) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição. Uma porção pode ser preparada, por exemplo, fazendo uma ou mais deleções para um ácido nucleico CYP90B. As porções podem ser usadas em forma isolada ou elas podem ser fundidas a outras seqüências de codificação (ou não codificantes) a fim de, por exemplo, produzir uma proteína que combina diversas atividades. Quando fundido a outras seqüências de codificação, o polipeptídeo resultante produzido em tradução pode ser maior do que aquele previsto para a porção CYP90B. Preferivelmente, a porção é uma porção de um ácido nucleico como representado por qualquer um de SEQ ID NO: 77, SEQ ID NO: 79, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87 e SEQ ID NO: 89. Mais preferivelmente a porção é uma porção de um ácido nucleico como representada por SEQ ID NO: 77.

Outra variante de um ácido nucleico/gene CYP90B é um ácido nucleico capaz de hibridizar em condições de estringência reduzida, preferivelmente em condições estringentes, com um ácido nucleico/gene CYP90B como definido acima, cuja seqüência hibridizante codifica um polipeptídeo compreendendo os seguintes: (a) domínios CYP A a D; (b) um domínio âncora hidrofóbico N-terminal; (c) um domínio de transição; e (d) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição. Preferivelmente, a seqüência hibridizante é um que é capaz de hibridizar com um ácido nucleico como representado por qualquer um de SEQ ID NO: 77, SEQ ID NO: 79, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87 e SEQ ID NO: 89, ou com uma porção de qualquer das seqüências

mencionadas acima como definido acima. Mais preferivelmente a seqüência hibridizante é uma que é capaz de hibridizar com um ácido nucleico como representado por SEQ ID NO: 77. O termo "hibridização" é como definido neste lugar na seção "Definições".

O ácido nucleico CYP90B ou variante deste pode ser derivado de qualquer fonte natural ou artificial. O ácido nucleico/gene ou variante deste pode ser isolado de uma fonte microbiana, tal como levedura ou fungos, ou de uma fonte vegetal, algal ou animal (incluindo humano). Este ácido nucleico pode ser modificado a partir de sua forma nativa em composição e/ou ambiente genômico através de manipulação humana deliberada. O ácido nucleico é preferivelmente de origem vegetal, quer da mesma espécie de planta (por exemplo que aquela na qual ele é introduzido) ou quer de uma espécie diferente de planta. O ácido nucleico pode ser isolado de uma espécie monocotiledônea, preferivelmente da família Poaceae, ainda preferivelmente de gênero *Oryza*, o mais preferivelmente de *Oryza sativa*. Mais preferivelmente, o ácido nucleico CYP90B isolado de *Oryza sativa* é representado por SEQ ID NO: 77 e a seqüência de aminoácidos CYP90B é como representado por SEQ ID NO: 78.

A invenção além disso fornece uma proteína CYP90B isolada selecionada a partir do grupo consistindo de:

(a) uma proteína codificada pelo ácido nucleico de SEQ ID NO: 117;

(b) uma proteína compreendendo os seguintes: (i) domínios CYP A a D; (ii) um domínio âncora hidrofóbico N-terminal; (iii) um domínio de transição; e (iv) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição, e tendo em ordem crescente de preferência pelo menos 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% de identidade com a seqüência de aminoácidos de SEQ ID NO: 118.

A invenção também fornece um ácido nucleico isolado codificando uma proteína CYP90B, selecionado a partir do grupo consistindo de:

(i) um ácido nucleico como representado por SEQ ID NO: 117;

(ii) um ácido nucleico codificando uma proteína como definido em (a) e (b) acima;

(iii) um ácido nucleico tendo em ordem crescente de preferência pelo menos 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais de identidade com o ácido nucleico representado por SEQ ID NO: 117;

(iv) uma seqüência de ácidos nucleicos capaz de hibridizar em condições estringentes com uma seqüência de ácidos nucleicos de (i) a (iii) acima, cuja seqüência hibridizante codifica uma proteína compreendendo (a) domínios CYP A a D; (b) um domínio âncora hidrofóbico N-terminal; (c) um domínio de transição; e (d) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição, e tendo em ordem crescente de preferência pelo menos 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais de identidade com a seqüência de aminoácidos de SEQ ID NO: 118;

(v) um ácido nucleico que é uma variante alélica ou uma variante de união das seqüências de ácidos nucleicos de acordo com (i) a (iv);

(vi) uma porção de uma seqüência de ácidos nucleicos de acordo com qualquer de (i) a (v) acima, cuja porção codifica uma proteína compreendendo: (i) domínios CYP A a D; (ii) um domínio âncora hidrofóbico N-terminal; (iii) um domínio de transição; e (iv) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição, e tendo em ordem crescente de

preferência pelo menos 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% com a seqüência de aminoácidos de SEQ ID NO: 118.

Além disso, o polipeptídeo CYP90B ou homólogo deste pode compreender adicionalmente (i) uma seqüência com mais do que 50% de identidade com SEQ ID NO: 78 e (ii) atividade enzimática de esteróide 22 alfa hidroxilase.

A expressão de um ácido nucleico codificando um polipeptídeo CYP90B ou um homólogo deste pode ser aumentada não constitutiva introduzindo uma modificação genética (preferivelmente no locus de um gene CYP90B). O locus de um gene como definido neste lugar é adotado para significar uma região genômica, a qual inclui o gene de interesse e 10 kb antes ou depois da região de codificação.

A modificação genética pode ser introduzida, por exemplo, por qualquer um (ou mais) dos seguintes métodos: ativação de T-DNA, TILLING, mutagênese sítio dirigida, evolução direta e recombinação homóloga ou introduzindo e expressando em uma planta um ácido nucleico codificando um polipeptídeo CYP90B ou um homólogo deste. Os métodos mencionados acima são definidos na seção "Definições". Seguindo introdução da modificação genética, segue uma etapa de selecionar para expressão não constitutiva aumentada de um ácido nucleico codificando um polipeptídeo CYP90B ou um homólogo deste, cujo aumento em expressão não constitutiva gera plantas tendo produção aumentada.

Ativação de T-DNA, TILLING, mutagênese sítio dirigida e evolução direta são exemplos de tecnologias que possibilitam a geração de novos alelos e variantes de CYP90B.

Um método preferido para introduzir uma modificação genética (que neste caso não precisa ser no locus de um gene CYP90B) é introduzir e expressar em uma planta um ácido nucleico codificando um

polipeptídeo CYP90B ou um homólogo deste. Um polipeptídeo CYP90B ou um homólogo deste é definido como polipeptídeo compreendendo os seguintes: (a) domínios CYP A a D; (b) um domínio âncora hidrofóbico N-terminal; (c) um domínio de transição; e (d) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição. O ácido nucleico a ser introduzido em uma planta pode ser um ácido nucleico de comprimento completo ou pode ser uma porção ou uma seqüência hibridizante como definido acima. Além disso, o ácido nucleico codificando um polipeptídeo CYP90B ou um homólogo deste pode compreender adicionalmente (i) uma seqüência com mais do que 50% de identidade com SEQ ID NO: 78 e (ii) atividade enzimática de esteróide 22 alfa hidroxilase.

"Homólogos" de uma proteína são definidos neste lugar na seção "Definições". O polipeptídeo CYP90B ou homólogo deste pode ser um derivado, como definido na seção "Definições".

O polipeptídeo CYP90B ou homólogo deste pode ser codificado por uma variante de união alternativa de um ácido nucleico/gene CYP90B. O termo "variante de união alternativa" é definido na seção "Definições". Variantes de união preferidas são variantes de união do ácido nucleico codificando um polipeptídeo compreendendo os seguintes: (a) domínios CYP A a D; (b) um domínio âncora hidrofóbico N-terminal; (c) um domínio de transição; e (d) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição. Adicionalmente, o polipeptídeo CYP90B ou um homólogo deste pode compreender adicionalmente (i) uma seqüência com mais do que 50% de identidade com SEQ ID NO: 78 e (ii) atividade enzimática de esteróide 22 alfa hidroxilase. Ainda preferidas são variantes de união de seqüências de ácidos nucleicos representadas por SEQ ID NO: 77, SEQ ID NO: 79, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87 e

SEQ ID NO: 89. A mais preferida é uma variante de união de uma seqüência de ácidos nucleicos como representada por SEQ ID NO: 77.

O homólogo também pode ser codificado por uma variante alélica de um ácido nucleico codificando um polipeptídeo CYP90B ou um homólogo deste, preferivelmente uma variante alélica do ácido nucleico codificando um polipeptídeo compreendendo os seguintes: (a) domínios CYP A a D; (b) um domínio âncora hidrofóbico N-terminal; (c) um domínio de transição; e (d) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição. Adicionalmente, o polipeptídeo CYP90B ou um homólogo deste pode compreender adicionalmente (i) uma seqüência com mais do que 50% de identidade com SEQ ID NO: 78 e (ii) atividade enzimática de esteróide 22 alfa hidroxilase. Ainda preferidas são variantes alélicas de seqüências de ácidos nucleicos representadas por SEQ ID NO: 77, SEQ ID NO: 79, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87 e SEQ ID NO: 89. A mais preferida é uma variante alélica de uma seqüência de ácidos nucleicos como representada por SEQ ID NO: 77. Variantes alélicas também são definidas na seção "Definições".

De acordo com um aspecto preferido da presente invenção, expressão não constitutiva aumentada do ácido nucleico CYP90B ou variante deste é contemplada. Métodos para aumentar expressão de genes ou produtos de gene são bem documentados na técnica e incluem, por exemplo, superexpressão dirigida por promotores apropriados, o uso de estimuladores de transcrição ou estimuladores de tradução. Ácidos nucleicos isolados que servem como elementos promotores ou estimuladores podem ser introduzidos em uma posição apropriada (tipicamente antes) de uma forma não heteróloga de um polinucleotídeo de forma a aumentar expressão de um ácido nucleico CYP90B ou variante deste. Por exemplo, promotores endógenos podem ser alterados *in vivo* por mutação, deleção, e/ou substituição (veja, Kmiec,

Patente Norte-Americana No. 5.565.350; Zarling *et al.*, PCT/US93/03868), ou promotores isolados podem ser introduzidos em uma célula vegetal na orientação e distância apropriada de um gene da presente invenção de forma a controlar a expressão do gene. Métodos para reduzir a expressão de genes ou produtos de gene são bem documentados na técnica.

Se expressão de polipeptídeo é desejada, geralmente é desejável incluir uma região de poliadenilação na extremidade 3' de uma região de codificação de polinucleotídeo. A região de poliadenilação pode ser derivada do gene natural, de uma variedade de outros 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 outro gene vegetal, ou menos preferivelmente de qualquer outro gene eucariótico.

Uma seqüência intrônica também pode ser adicionada à região 5' não traduzida ou à seqüência de codificação da seqüência de codificação parcial para aumentar a quantidade da mensagem madura que se acumula no citossol. Inclusão de um íntron processável na unidade de transcrição tanto em construções vegetais como animais mostrou aumentar expressão gênica tanto no nível de mRNA como protéico em até 1000 vezes (Buchman and Berg (1988) Mol. Cell biol. 8: 4395-4405; Callis *et al.* (1987) Genes Dev 1:1183-1200). Tal estimulação intrônica de expressão gênica tipicamente é maior quando localizada perto da extremidade 5' da unidade de transcrição. Uso dos íntrons de milho íntron Adh1-S 1, 2, e 6, o íntron Bronze-1 são conhecidos na técnica. Veja geralmente, The Maize Handbook, capítulo 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

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

Portanto, é fornecida uma construção gênica compreendendo:

(i) Um ácido nucleico CYP90B ou variante deste, como definido acima;

(ii) Uma ou mais seqüências de controle capazes de dirigir expressão não constitutiva da seqüência de ácidos nucleicos de (i); e opcionalmente

(iii) Uma seqüência de término de transcrição.

Construções úteis nos métodos de acordo com a presente invenção podem ser construídas usando tecnologia de DNA recombinante bem conhecida por pessoas versadas na técnica. As construções gênicas podem ser inseridas em vetores, que podem estar comercialmente disponíveis, adequados para transformar em plantas e adequados para expressão do gene de interesse nas células transformadas. A invenção portanto fornece uso de uma construção gênica como definido acima nos métodos da invenção.

Plantas são transformadas com um vetor compreendendo a seqüência de interesse (i.e., um ácido nucleico codificando um polipeptídeo CYP90B ou homólogo deste). A seqüência de interesse é operavelmente ligada a uma ou mais seqüências de controle (pelo menos a um promotor). Os termos "elemento regulador", "seqüência de controle" e "promotor" são todos usados permutavelmente neste lugar e são definidos na seção "Definições".

Vantajosamente, qualquer tipo não constitutivo de promotor pode ser usado para dirigir expressão da seqüência de ácidos nucleicos. O promotor não constitutivo pode ser um promotor induzível, i.e. tendo iniciação de transcrição induzida ou aumentada em resposta a um desenvolvimento, estímulo químico, ambiental ou físico. Um exemplo de um promotor induzível sendo um promotor induzível por estresse, i.e. um promotor ativado quando uma planta é exposta a várias condições de estresse. O promotor não constitutivo pode ser um promotor preferido por tecido, i.e. um que é capaz de iniciar preferencialmente transcrição em certos tecidos, tal como as folhas, raízes, tecido de semente etc. Promotores capazes de iniciar

transcrição em certos tecidos são referidos neste lugar apenas como "tecido específicos".

De acordo com os métodos da invenção, o ácido nucleico CYP90B ou variante deste é operavelmente ligado a um promotor não constitutivo. Um promotor não constitutivo é transcricionalmente ativo apenas durante algumas fases de crescimento e desenvolvimento vegetal e não é ubiqüitariamente expresso. O promotor não constitutivo pode ser por exemplo um promotor semente específico, ou um promotor raiz específico. O promotor semente específico pode ser um promotor endosperma específico e/ou embrião/aleurona específico, i.e., transcricionalmente ativo no endosperma de semente e/ou embrião e aleurona de semente, respectivamente. O promotor endosperma específico é preferivelmente um promotor de proteína de armazenamento de semente, ainda preferivelmente o promotor endosperma específico é um promotor de prolamina, mais preferivelmente o promotor endosperma específico é um promotor de prolamina RP6 de arroz, ainda mais preferivelmente o promotor endosperma específico é representado por uma seqüência de ácidos nucleicos substancialmente similar a SEQ ID NO: 109, o mais preferivelmente o promotor endosperma específico é como representado por SEQ ID NO: 109. O promotor embrião/aleurona específico é preferivelmente um promotor de proteína de armazenamento de semente, ainda preferivelmente o promotor embrião/aleurona específico é um promotor de oleosina, mais preferivelmente o promotor embrião/aleurona específico é um promotor de 18 kDa de oleosina de arroz, ainda mais preferivelmente o promotor embrião/aleurona específico é representado por uma seqüência de ácidos nucleicos substancialmente similar a SEQ ID NO: 110, o mais preferivelmente o promotor embrião/aleurona específico é como representado por SEQ ID NO: 110. O promotor raiz específico é preferivelmente um promotor Rcc3, o promotor raiz específico é preferivelmente um promotor Rcc3 de arroz (Xu et al. (1995)

Plant Mol Biol 27(2):237-48).

Deve ser claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico CYP90B representado por SEQ ID NO: 77, nem é a aplicabilidade da invenção restrita a expressão de um ácido nucleico CYP90B quando dirigida por um promotor de prolamina RP6 ou oleosina de 18 kDa. Exemplos de outros promotores não constitutivos que também podem ser usados para realizar os métodos da invenção são mostrados em Tabela 4 na seção "Definições".

Em contraste com os promotores descritos acima, um promotor constitutivo é transcricionalmente ativo durante a maioria das fases de crescimento e desenvolvimento vegetal e é substancialmente ubiqüitariamente expresso na planta. Tais promotores constitutivos devem ser excluídos para desempenho dos métodos da invenção. Exemplos de tais promotores também podem ser encontrados na seção "Definições" (veja Tabela 3).

Opcionalmente, uma ou mais seqüências terminadoras também podem ser usadas na construção introduzida em uma planta. O termo "terminador" é definido na seção "Definições".

As construções genéticas da invenção podem incluir adicionalmente uma seqüência de origem de replicação que é requerida para manutenção e/ou replicação em um tipo celular específico. Um exemplo é quando uma construção genética é requerida ser mantida em uma célula bacteriana como um elemento genético episssomal (e.g. molécula de plasmídeo ou cosmídeo). Origens de replicação preferidas incluem, mas não são limitadas a, a f1-ori e colE1.

A construção genética pode compreender opcionalmente um gene marcador selecionável como definido na seção "Definições".

Em uma forma de realização preferida, é fornecida uma construção gênica compreendendo:

(i) Um ácido nucleico CYP90B ou variante deste, como definido acima;

(ii) Um promotor capaz de dirigir expressão não constitutiva da seqüência de ácidos nucleicos de (i); e opcionalmente

(iii) Uma seqüência de término de transcrição.

O promotor não constitutivo é preferivelmente um promotor semente específico. O promotor semente específico pode ser um promotor endosperma específico e/ou embrião/aleurona específico, i.e., transcricionalmente ativo no endosperma de semente e/ou embrião e aleurona de semente, respectivamente. O promotor endosperma específico é preferivelmente um promotor de proteína de armazenamento de semente, ainda preferivelmente o promotor endosperma específico é um promotor de prolamina, mais preferivelmente o promotor endosperma específico é um promotor de prolamina RP6 de arroz, mais preferivelmente o promotor endosperma específico é representado por uma seqüência de ácidos nucleicos substancialmente similar a SEQ ID NO: 109, o mais preferivelmente o promotor endosperma específico é como representado por SEQ ID NO: 109. O promotor embrião/aleurona específico é preferivelmente um promotor de proteína de armazenamento de semente, ainda preferivelmente o promotor embrião/aleurona específico é um promotor de oleosina, mais preferivelmente o promotor embrião/aleurona específico é um promotor de 18 kDa de oleosina de arroz, mais preferivelmente o promotor embrião/aleurona específico é representado por uma seqüência de ácidos nucleicos substancialmente similar a SEQ ID NO: 110, o mais preferivelmente o promotor embrião/aleurona específico é como representado por SEQ ID NO: 110. A invenção adicionalmente fornece uso de uma construção como definido acima nos métodos da invenção.

A presente invenção também abrange plantas obteníveis pelos métodos de acordo com a presente invenção. A presente invenção portanto

fornece plantas, partes vegetais ou células vegetais destas obteníveis pelo método de acordo com a presente invenção, cujas plantas ou partes ou células destas compreendem um ácido nucleico CYP90B transgênico ou variante deste.

A invenção também fornece um método para a produção de plantas transgênicas tendo produção aumentada em relação a plantas de controle adequadas compreendendo introdução e expressão não constitutiva em uma planta de um ácido nucleico CYP90B ou uma variante deste.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo produção aumentada cujo método compreende:

(i) introduzir e expressar não constitutivamente em uma planta, parte vegetal ou célula vegetal um ácido nucleico CYP90B ou variante deste; e

(ii) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

O ácido nucleico pode ser introduzido diretamente em uma célula vegetal ou na própria planta (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" é como definido na seção "Definições".

A presente invenção claramente se estende a qualquer célula vegetal ou planta produzida por qualquer dos métodos descritos neste lugar, e a todas as partes vegetais e propágulos destas. A presente invenção adicionalmente de estende para abranger a progênie de uma célula primária transformada ou transfectada, tecido, órgão ou planta inteira que foi produzida por qualquer dos métodos mencionados acima, o único requerimento sendo

que a progênie exiba a(s) mesma(s) característica(s) genotípica(s) e/ou fenotípica(s) que aquelas produzidas pelo ancestral nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucleico CYP90B isolado ou variante deste, expresso não constitutivamente. Células hospedeiras preferidas de acordo com a invenção são células vegetais.

A invenção também se estende a partes aptas à colheita de uma planta tal como, mas não limitada a sementes, folhas, frutos, flores, caules, rizomas, tubérculos e bulbos. A invenção além disso se refere a produtos derivados de uma parte apta à colheita de uma tal planta, tal como pelotas secas ou pós, óleo, gordura e ácidos graxos, amido ou proteínas.

A presente invenção também abrange uso de ácidos nucleicos CYP90B ou variantes destes e uso de polipeptídeos CYP90B ou homólogos destes. Tais usos se referem a aumentar produção vegetal como definido acima nos métodos da invenção.

Ácidos nucleicos CYP90B ou variantes destes, ou polipeptídeos CYP90B ou homólogos destes podem encontrar uso em programas de melhoramento nos quais é identificado um marcador de DNA que pode ser geneticamente ligado a um gene CYP90B ou variante deste. Os ácidos nucleicos/genes CYP90B ou variantes destes, ou polipeptídeos CYP90B ou homólogos destes podem ser usados para definir um marcador molecular. Este marcador de DNA ou proteína pode então ser usado em programas de melhoramento para selecionar plantas tendo produção aumentada como definido acima nos métodos da invenção. O gene CYP90B ou variante deste, por exemplo, pode ser um ácido nucleico como representado por qualquer um de SEQ ID NO: 77, SEQ ID NO: 79, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87 e SEQ ID NO: 89.

Variantes alélicas de um ácido nucleico/gene CYP90B

também podem encontrar uso em programas de melhoramento auxiliados por marcador. Tais programas de melhoramento algumas vezes requerem introdução de variação alélica por tratamento mutagênico das plantas, usando por exemplo mutagênese EMS; alternativamente, o programa pode começar com uma coleção de variantes alélicas de assim chamada origem "natural" causada não intencionalmente. Identificação de variantes alélicas então ocorre, por exemplo, por PCR. Isto é seguido por uma etapa para seleção de variantes alélicas superiores da seqüência em questão e que gera produção aumentada. Seleção tipicamente é realizada monitorando desempenho de crescimento de plantas contendo diferentes variantes alélicas da seqüência em questão, por exemplo, variantes alélicas diferentes de qualquer uma de SEQ ID NO: 77, SEQ ID NO: 79, SEQ ID NO: 81, SEQ ID NO: 83, SEQ ID NO: 85, SEQ ID NO: 87 e SEQ ID NO: 89. Desempenho de crescimento pode ser monitorado em uma casa de vegetação ou no campo. Etapas opcionais adicionais incluem cruzar plantas nas quais a variante alélica superior foi identificada com outra planta. Isto pode ser usado, por exemplo, para fazer uma combinação de características fenotípicas interessantes.

Um ácido nucleico CYP90B ou variante deste também pode ser usado como sondas para mapear geneticamente e fisicamente os genes dos quais eles são uma parte de, e como marcadores para características ligadas a estes genes. Tal informação pode ser útil em melhoramento vegetal a fim de desenvolver linhagens com fenótipos desejados. Tal uso de ácidos nucleicos CYP90B ou variantes destes requer apenas uma seqüência de ácidos nucleicos de pelo menos 15 nucleotídeos de comprimento. Os ácidos nucleicos CYP90B ou variantes destes 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 os ácidos nucleicos CYP90B ou variantes destes. Os padrões

de bandeamento resultantes então podem ser sujeitos a análises genéticas usando programas de computador tal como MapMaker (Lander *et al.* (1987) Genomics 1: 174-181) a fim de construir um mapa genético. Em adição, os ácidos nucleicos podem ser usados para sondar Southern blots contendo DNAs genômicos tratados com endonuclease de restrição de um conjunto de indivíduos representando ancestral e progênie de um cruzamento genético definido. Segregação dos polimorfismos de DNA é observada e usada para calcular a posição do ácido nucleico CYP90B ou variante deste no mapa genético obtido previamente usando esta população (Botstein *et al.* (1980) Am. J. Hum. Genet. 32:314-331).

A produção e uso de sondas derivadas de gene vegetal para uso em mapeamento genético são descritos em Bematzky and Tanksley (1986) (GENETICS 112 (4): 887-898). Numerosas publicações descrevem mapeamento genético de clones de cDNA específicos usando a metodologia descrita acima ou variações desta. Por exemplo, populações de intercruzamento de F2, populações endocruzadas, populações cruzadas aleatoriamente, linhagens isogênicas próximas, e outros conjuntos de indivíduos podem ser usados para mapeamento. Tais metodologias são bem conhecidas por aqueles versados na técnica.

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

Em outra forma de realização, as sondas de ácido nucleico podem ser usadas em mapeamento direto por hibridização *in situ* por fluorescência (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora métodos atuais de mapeamento por FISH favoreçam uso de clones maiores (diversos kb a diversas centenas kb; veja Laan *et al.* (1995) Genome Res. 5:13-20), melhoramentos em sensibilidade podem permitir desempenho de

mapeamento por FISH usando sondas menores.

Uma variedade de métodos baseados em amplificação de ácido nucleico para mapeamento genético e físico pode ser realizada usando os ácidos nucleicos. Exemplos incluem amplificação alelo específica (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield *et al.* (1993) Genomics 16:325-332), ligação alelo específica (Landegren *et al.* (1988) Science 241:1077-1080), reações de extensão de nucleotídeo (Sokolov (1990) Nucleic Acid Res. 18:3671), Mapeamento de Híbrido de Radiação (Walter *et al.* (1997) Nat. Genet. 7:22-28) e Mapeamento Apropriado (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 iniciadores para uso na reação de amplificação ou em reações de extensão de iniciador. O projeto de tais iniciadores é bem conhecido por aqueles versados na técnica. Em métodos empregando mapeamento genético baseado em PCR, pode ser necessário identificar diferenças de seqüência de DNA entre os ancestrais do cruzamento de mapeamento na região correspondente à seqüência imediata de ácidos nucleicos. Isto, entretanto, geralmente não é necessário para métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo produção aumentada, como descrito acima. Esta produção aumentada também pode ser combinada com outras características economicamente vantajosas, tal como características adicionais que estimulam produção, tolerância a outros estresses abióticos e bióticos, características modificando vários atributos arquitetônicos e/ou atributos bioquímicos e/ou fisiológicos.

CDC27 de Descrição Detalhada

Polipeptídeos CDC27 são bem conhecidos na técnica e são facilmente identificáveis pela presença de uma região NH_2 terminal

conservada (veja Figura 16) e de pelo menos 5 domínios TPR com pelo menos um domínio TPR na região NH_2 terminal. Além disso, o polipeptídeo CDC27 pode compreender adicionalmente uma seqüência com mais do que 30% de identidade com SEQ ID NO: 130.

Motivos TPR estão presentes em uma ampla variedade de proteínas funcionais em levedura e eucariotos superiores em mitose (incluindo os componentes de proteínas APC CDC16, CDC23 e CDC27), transcrição, processamento, importação de proteínas e neurogênese (Goebl and Yanagida 1991, Trends Biochem Sci 16, 173-177). Uma seqüência consenso mínima sugerida do motivo TPR é: X_3-W-X_2-L-G-X_2-Y-X_8-A-X_3-F-X_2-A-X_4-P-X_2, onde X = qualquer aminoácido (Lamb et al. 1994, EMBO J 13, 4321-4328). Os resíduos consenso podem exibir degenerescência significante e os resíduos não consenso exibem pouca ou nenhuma homologia. É a hidrofobicidade e o tamanho dos resíduos consenso, ao invés de sua identidade, que parece ser importante. Em uma proteína CDC27 nativa, o TPR forma uma estrutura α–helicoidal, repetições em tandem se organizam em uma estrutura super-helicoidal idealmente adequada como interfaces para reconhecimento de proteínas (Groves and Barford 1999, Curr Opin Struct Biol 9, 383-389). Dentro da α–hélice, dois domínios anfipáticos estão normalmente presentes, um na região NH_2 terminal e o outro perto da região COOH-terminal (Sikorski et al. 1990, Cell 60, 307-317). Também motivos TPR individuais podem estar dispersos ao longo da seqüência protéica.

Um CDC27 nativo de comprimento completo tipicamente compreende pelo menos 5 TPRs, preferivelmente 6 TPRs, mais preferivelmente 7 TPRs, a maioria destes TPRs sendo localizada na região COOH terminal. Como mostrado em Figura 16, existe tipicamente um domínio TPR na região NH_2 terminal de um polipeptídeo CDC27 nativo, embora seqüências variantes de CDC27 possam existir ou possam ser criadas para compreender mais do que um TPR na região NH_2 terminal.

Qualquer polipeptídeo CDC27 pode ser tornado útil nos métodos da invenção por inativação de pelo menos um domínio TPR na região NH_2 terminal do polipeptídeo. Métodos para inativação são bem conhecidos na técnica e incluem: remoção ou substituição de aminoácidos, neste caso, remoção ou substituição de aminoácidos de pelo menos um domínio TPR na região NH_2 terminal; ou técnicas de mutação, tal como substituir aminoácidos conservados por alanina ou substituir aminoácidos fosforiláveis (tal como serina, treonina ou tirosina) por aminoácidos não fosforiláveis ou vice versa (dependendo se a proteína fosforilada está ativa ou inativa); ou qualquer outro método para inativação.

Para os propósitos deste pedido, a região NH_2 terminal de uma proteína CDC27 é considerada como sendo a primeira metade de uma seqüência CDC27 de comprimento completo (de NH_2 terminal para COOH terminal) (veja Figura 16); preferivelmente a região NH_2 terminal de uma proteína CDC27 é considerada como sendo o primeiro terço de uma seqüência CDC27 de comprimento completo (de NH_2 terminal para COOH terminal); e de acordo com outra característica preferida da presente invenção, a região N-terminal de uma proteína CDC27 é considerada como sendo os primeiros 166 aminoácidos (de NH_2 terminal para COOH terminal) de uma seqüência CDC27 de comprimento completo.

Um exemplo de um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal é o polipeptídeo representado por SEQ ID NO: 130, com seqüência de ácidos nucleicos de codificação representada por SEQ ID NO: 129.

Tabela 10 abaixo fornece alguns exemplos de seqüências CDC27; estas seqüências podem ser tornadas úteis nos métodos da invenção por inativação de pelo menos um domínio TPR na região NH_2 terminal do polipeptídeo, por exemplo usando qualquer dos métodos de inativação discutidos acima.

Tabela 10: Exemplos de polipeptídeos CDC27

Nome	Número de acesso de nucleotídeo NCBI	Nucleotídeo SEQ ID NO	Polipeptídeo traduzido SEQ ID NO	Fonte
CDC27B	AC006081	129	130	*Arabidopsis thaliana*
CDC27B/ Hobbit	AJ487669	131	132	*Arabidopsis thaliana*
CDC27a	NM_112503.2	133	134	*Arabidopsis thaliana*
CDC27	AP003539.3	135	136	*Oryza sativa*
CDC27	BG887406.1* BG590616.1 DN939130,1 CV470643.1	137	139	*Solanum tuberosum*
CDC27/nuc2+	NM_001020032 .1	139	140	*Schizosaccharomyces pombe*
CDC27/BimA	X59269.1	141	142	*Aspergillus niger*
CDC27	NM_001256.2	143	144	*Homo sapiens*
CDC27 5'	CA102186.1* CA279358.1	145	146	*Saccharum officinarum*
CDC27 3'	CA197669.1* CA197670,1 CA203636.1 CA232307.1	147	148	*Saccharum officinarum*

*Agrupamento de seqüências compilado a partir de diversos acessos de EST (os principais mostrados); qualidade de seqüenciamento de EST sendo normalmente inferior, umas poucas substituições de ácidos nucleicos podem ser esperadas.

As seqüências descritas em Tabela 10 são dadas como forma de Exemplo apenas. Exemplos adicionais são dados em Figura 19, codificando polipeptídeos ou de comprimento completo ou parciais (que podem ser usados para obter a seqüência de comprimento completo usando métodos de rotina). Deve ser entendido que qualquer seqüência de

polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo, ou um ácido nucleico/gene codificando um tal polipeptídeo, pode ser adequado para uso para realizar os métodos da invenção.

Outros polipeptídeos CDC27 podem ser prontamente identificados usando técnicas de rotina bem conhecidas na técnica, tal como por alinhamento de seqüências. Seqüências assim identificadas podem ser subseqüentemente tornadas úteis nos métodos da invenção por inativação de pelo menos um domínio TPR na região NH_2 terminal do polipeptídeo, por exemplo usando qualquer dos métodos de inativação discutidos acima. 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 de duas seqüências completas que maximiza o número de identidades e minimiza o número de lacunas. O algoritmo BLAST (Altschul *et al.* (1990) J Mol Biol 215: 403-10) calcula identidade de seqüência percentual e realiza uma análise estatística da similaridade entre as duas seqüências. O aplicativo computacional para realizar análise BLAST está publicamente disponível através de National Centre for Biotechnology Information. Homólogos de um CDC27 podem ser prontamente identificados usando, por exemplo, o algoritmo de alinhamento de múltiplas seqüências ClustalW (versão 1.83) disponível em http://clustalw.genome.jp/sit-bin/nph-ClustalW, com os parâmetros padrão de alinhamento pareado, e um método de pontuação em porcentagem. Edição manual secundária pode ser realizada para otimizar alinhamento entre motivos conservados, como deve ser perceptível para uma pessoa versada na técnica.

Vários domínios estruturais em uma proteína CDC27, tal como domínios TPR, podem ser identificados usando bancos de dados

especializados e.g. SMART (Schultz *et al.* (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic *et al.* (2002) Nucleic Acids Res 30, 242-244; http://smart.embl-heidelberg.de/), InterPro (Mulder *et al.*, (2003) Nucl. Acids. Res. 31, 315-318; http://www.ebi.ac.uk/interpro/), Prosite (Bucher and Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (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), http://www.expasy.org/prosite/), Pfam (Bateman *et al.*, Nucleic Acids Research 30(1): 276-280 (2002), http://www.sanger.ac.uk/Software/Pfam/) ou ProDom (Servant F, Bru C, Carrère S, Courcelle E, Gouzy J, Peyruc D, Kahn D (2002) ProDom: Automated clustering of homologous domains. Briefings in Bioinformatics. vol 3, no 3:246-251).

As seqüências mencionadas em Tabela 10 e Figura 19 podem ser consideradas homólogos de um polipeptídeo CDC27. "Homólogos" de uma proteína são definidos na seção "Definições" neste lugar. Homólogos preferidos são seqüências de aminoácidos tendo em ordem crescente de preferência pelo menos 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% ou mais de identidade de seqüência com a proteína CDC27 de comprimento completo representada por SEQ ID NO: 132.

Homólogos, ortólogos e parálogos podem ser tornados úteis nos métodos da invenção por inativação de pelo menos um domínio TPR na região NH_2 terminal do polipeptídeo, por exemplo usando qualquer dos métodos de inativação discutidos acima.

Polipeptídeos CDC27 humanos e de levedura mostraram interagir com duas outras proteínas do complexo APC, CDC16 e CDC23, *in vivo* através de análise de dois híbridos de levedura, e *in vitro* através de co-

imunoprecipitação (Lam et al. (1994) EMBO J 13(18): 4321-4328; Ollendorf & Donoghue (1997) J Biol Chem 272(51): 32011-32018). Uma tal interação pode ser útil para identificar polipeptídeos CDC27 a serem tornados úteis nos métodos da invenção por inativação de pelo menos um domínio TPR na região NH_2 terminal do polipeptídeo, por exemplo usando qualquer dos métodos de inativação discutidos acima

Um polipeptídeo CDC27 tendo pelo menos um domínio TRP inativo na região NH_2 terminal do polipeptídeo é codificado por um assim chamado ácido nucleico/gene CDC27 modificado. Portanto, o termo "ácido nucleico/gene CDC27 modificado" como definido neste lugar é qualquer ácido nucleico/gene codificando um polipeptídeo CDC27 tendo pelo menos um domínio TRP inativo na região NH_2 terminal do polipeptídeo.

O ácido nucleico CDC27 ou ácido nucleico/gene CDC27 modificado pode ser derivado de qualquer fonte natural ou artificial. O ácido nucleico/gene pode ser isolado de uma fonte microbiana, tal como levedura ou fungos, ou de uma fonte vegetal, algal ou animal. Este ácido nucleico pode ser modificado a partir de sua forma nativa em composição e/ou ambiente genômico através de manipulação humana deliberada. O ácido nucleico é preferivelmente de origem vegetal, quer da mesma espécie de planta (por exemplo que aquela na qual ele é introduzido) ou quer de uma espécie diferente de planta. O ácido nucleico pode ser isolado de uma espécie dicotiledônea, preferivelmente da família Brassicaceae, ainda preferivelmente de *Arabidopsis thaliana*. Mais preferivelmente, o ácido nucleico CDC27 modificado isolado de *Arabidopsis thaliana* é representado por SEQ ID NO: 129 e o CDC27 tendo pelo menos um TPR inativo na região NH_2 terminal do aminoácido é como representado por SEQ ID NO: 130.

Um ácido nucleico/gene CDC27 é um ácido nucleico capaz de hibridizar em condições de estringência reduzida, preferivelmente em condições estringentes, com um ácido nucleico/gene CDC27 como

representado por qualquer um de SEQ ID NO: 129, SEQ ID NO: 131, SEQ ID NO: 133, SEQ ID NO: 135, SEQ ID NO: 137 ou SEQ ID NO: 141. Mais preferivelmente a seqüência hibridizante é uma que é capaz de hibridizar com um ácido nucleico como representado por SEQ ID NO: 129 ou SEQ ID NO: 131. Tais seqüências hibridizantes podem ser tornadas úteis nos métodos da invenção por inativação de pelo menos um domínio TPR na região NH_2 terminal do polipeptídeo codificado, por exemplo usando qualquer dos métodos de inativação discutidos acima.

O termo "hibridização" é como definido neste lugar na seção "Definições".

O ácido nucleico CDC27 ou ácido nucleico/gene CDC27 modificado pode estar na forma de uma variante de união alternativa. Uma variante de união alternativa é definida na seção "Definições". São preferidas variantes de união de qualquer das seqüências de ácidos nucleicos CDC27 mencionadas acima, em outras palavras SEQ ID NO: 129, SEQ ID NO: 131, SEQ ID NO: 133, SEQ ID NO: 135, SEQ ID NO: 137 ou SEQ ID NO: 141. A mais preferida é uma variante de união de uma seqüência de ácidos nucleicos como representada por SEQ ID NO: 129 ou SEQ ID NO: 131. Tais variantes de união podem ser tornadas úteis nos métodos de inativação da invenção de pelo menos um domínio TPR na região NH_2 terminal do polipeptídeo CDC27 codificado, por exemplo usando qualquer dos métodos de inativação discutidos acima.

O ácido nucleico CDC27 ou ácido nucleico/gene CDC27 modificado pode estar na forma de uma variante alélica de um ácido nucleico codificando um polipeptídeo CDC27 truncado compreendendo pelo menos um domínio TPR inativado na região NH2 terminal. São preferidas variantes alélicas de seqüências de ácidos nucleicos representadas por SEQ ID NO: 129, SEQ ID NO: 131, SEQ ID NO: 133, SEQ ID NO: 135, SEQ ID NO: 137 ou SEQ ID NO: 141. A mais preferida é uma variante alélica de uma

seqüência de ácidos nucleicos como representada por SEQ ID NO: 129 ou SEQ ID NO: 131. Variantes alélicas existem na natureza, e abrangido dentro dos métodos da presente invenção é o uso destes alelos naturais. Variantes alélicas abrangem Polimorfismos de Nucleotídeo Único (SNPs), assim como Polimorfismos de Pequena Inserção/Deleção (INDELs). O tamanho de INDELs normalmente é menor do que 100 bp. SNPs e INDELs formam o maior conjunto de variantes de seqüência em linhagens polimórficas naturalmente ocorrentes da maioria dos organismos. Tais variantes alélicas podem ser tornadas úteis nos métodos de inativação da invenção de pelo menos um domínio TPR na região NH_2 terminal do polipeptídeo CDC27 codificado, por exemplo usando qualquer dos métodos de inativação discutidos acima.

O ácido nucleico CDC27 ou ácido nucleico/gene CDC27 modificado pode ser gerado por mutagênese sítio dirigida. Estão disponíveis diversos métodos para atingir mutagênese sítio dirigida, o mais comum sendo métodos baseados em PCR (Current Protocols in Molecular Biology, Wiley Eds http://www.4ulr.com/products/currentprotocols/index.html).

O ácido nucleico CDC27 ou ácido nucleico/gene CDC27 modificado também pode ser gerado por evolução direta (veja seção "Definições" para detalhes adicionais).

Tais variantes produzidas por mutagênese sítio dirigida ou por evolução direta podem ser tornadas úteis nos métodos de inativação da invenção de pelo menos um domínio TPR na região NH_2 terminal do polipeptídeo CDC27 codificado, por exemplo usando qualquer dos métodos de inativação discutidos acima.

A expressão de um ácido nucleico/gene CDC27 modificado codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo pode ser aumentada introduzindo uma modificação genética (preferivelmente no locus de um gene

CDC27). O locus de um gene como definido neste lugar é adotado para significar uma região genômica, a qual inclui o gene de interesse e 10 KB antes ou depois da região de codificação.

A modificação genética é preferivelmente introduzida introduzindo e expressando em uma planta um ácido nucleico codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo. Seguindo introdução da modificação genética, segue uma etapa adicional de selecionar para expressão aumentada (em tecido de meristema apical de broto) de um ácido nucleico modificado codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo, cujo aumento em expressão gera plantas tendo produção aumentada.

De acordo com um aspecto preferido da presente invenção, expressão aumentada do ácido nucleico CDC27 é contemplada. Métodos para aumentar expressão de genes ou produtos de gene são bem documentados na técnica e incluem, superexpressão dirigida por promotores apropriados, o uso de estimuladores de transcrição ou estimuladores de tradução. Ácidos nucleicos isolados que servem como elementos promotores ou estimuladores podem ser introduzidos em uma posição apropriada (tipicamente antes) de uma forma não heteróloga de um polinucleotídeo de forma a aumentar expressão de um ácido nucleico CDC27. Por exemplo, promotores endógenos podem ser alterados *in vivo* por mutação, deleção, e/ou substituição (veja, Kmiec, Patente Norte-Americana No. 5.565.350; Zarling *et al.*, PCT/US93/03868), ou promotores isolados podem ser introduzidos em uma célula vegetal na orientação e distância apropriada de um gene da presente invenção para controlar a expressão do gene.

Se expressão de polipeptídeo é desejada, geralmente é desejável incluir uma região de poliadenilação na extremidade 3' de uma região de codificação de polinucleotídeo. A região de poliadenilação pode ser

derivada do gene natural, de uma variedade de outros 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 outro gene vegetal, ou menos preferivelmente de qualquer outro gene eucariótico.

Uma seqüência intrônica também pode ser adicionada à região 5' não traduzida ou à seqüência de codificação da seqüência de codificação parcial para aumentar a quantidade da mensagem madura que se acumula no citossol. Inclusão de um íntron processável na unidade de transcrição tanto em construções vegetais como animais mostrou aumentar expressão gênica tanto no nível de mRNA como protéico em até 1000 vezes (Buchman and Berg (1988) Mol. Cell biol. 8: 4395-4405; Callis *et al.* (1987) Genes Dev 1:1183-1200). Tal estimulação intrônica de expressão gênica tipicamente é maior quando localizada perto da extremidade 5' da unidade de transcrição. Uso dos íntrons de milho íntron Adh1-S 1, 2, e 6, o íntron Bronze-1 são conhecidos na técnica. Veja geralmente, The Maize Handbook, capítulo 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

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

Portanto, é fornecida uma construção gênica compreendendo:

(i) Um ácido nucleico CDC27 codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativado na região NH_2 terminal do polipeptídeo;

(ii) Uma ou mais seqüências de controle capazes de dirigir preferencialmente expressão da seqüência de ácidos nucleicos de (i) em tecido de meristema apical de broto; e opcionalmente

(iii) Uma seqüência de término de transcrição.

Tais construções genéticas podem ser construídas usando

tecnologia de DNA recombinante bem conhecida por pessoas versadas na técnica. As construções gênicas podem ser inseridas em vetores, que podem estar comercialmente disponíveis, adequados para transformar em plantas e adequados para expressão do gene de interesse nas células transformadas. A invenção portanto fornece uso de uma construção gênica como definido acima nos métodos da invenção.

Plantas são transformadas com um vetor compreendendo a seqüência de interesse (i.e., um ácido nucleico codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo). A seqüência de interesse é operavelmente ligada a uma ou mais seqüências de controle (pelo menos a um promotor) capaz de dirigir preferencialmente expressão em tecido de meristema apical de broto de uma planta. Os termos "elemento regulador", "seqüência de controle" e "promotor" são todos usados permutavelmente neste lugar e são definidos na seção "Definições".

O ácido nucleico CDC27 codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo ou variante é operavelmente ligado a um promotor de meristema apical de broto, preferivelmente a um promotor de meristema apical de broto precoce. Um "promotor de meristema apical de broto precoce" como definido neste lugar é um promotor que é transcricionalmente ativo no meristema apical de broto do estágio de embrião globular até o estágio de plântula jovem, estes estágios sendo bem conhecidos por pessoas versadas na técnica. Referência neste lugar a aumentar preferencialmente expressão em tecido de meristema apical de broto é adotada para significar aumentar expressão em tecido de meristema apical de broto substancialmente para a exclusão de expressão em qualquer outro lugar na planta, sem considerar qualquer expressão residual devido a promotores permeáveis. Preferivelmente, o promotor de meristema apical de broto precoce é um promotor OSH1 (de

arroz; SEQ ID NO: 151 (Matsuoka *et al.*, (1993) Plant Cell 5: 1039-1048; Sato *et al.*, (1996) Proc Natl Acad Sci U S A 93(15): 8117-22)). Deve ser claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico CDC27 modificado representado por SEQ ID NO: 129, nem é a aplicabilidade da invenção restrita a expressão de um ácido nucleico CDC27 modificado quando dirigida por um promotor OSH1. Exemplos de outros promotores de meristema apical de broto precoce são mostrados em Tabela 5 na seção "Definições". Estes são membros de homeobox de classe 1 da família KNOX, de genes parálogos ou ortólogos. Deve ser entendido que a lista abaixo não é exaustiva.

Opcionalmente, uma ou mais seqüências terminadoras também podem ser usadas na construção introduzida em uma planta. O termo "terminador" é definido neste lugar na seção "Definições".

As construções genéticas da invenção podem incluir adicionalmente uma seqüência de origem de replicação que é requerida para manutenção e/ou replicação em um tipo celular específico. Um exemplo é quando uma construção genética é requerida ser mantida em uma célula bacteriana como um elemento genético episssomal (e.g. molécula de plasmídeo ou cosmídeo). Origens de replicação preferidas incluem, mas não são limitadas a, a f1-ori e colE1.

A construção genética pode compreender opcionalmente um gene marcador selecionável como definido na seção "Definições".

A presente invenção também abrange plantas obteníveis pelos métodos de acordo com a presente invenção. A presente invenção portanto fornece plantas ou partes destas, incluindo células vegetais, obteníveis pelo método de acordo com a presente invenção, cujas plantas ou partes vegetais compreendem um ácido nucleico CDC27 codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo e cujo ácido nucleico é operavelmente ligado a um promotor de

meristema apical de broto.

A invenção também fornece um método para a produção de plantas transgênicas tendo número de sementes aumentado em relação a plantas de controle adequadas, compreendendo introdução e expressão em uma planta de um ácido nucleico CDC27 codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo, cujo ácido nucleico CDC27 está sob o controle de um promotor de meristema apical de broto.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo número de sementes aumentado em relação a plantas de controle adequadas, cujo método compreende:

(i) introduzir e expressar em a planta, parte vegetal ou célula vegetal um ácido nucleico CDC27 codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo, cujo ácido nucleico está sob o controle de um promotor de meristema apical de broto; e

(ii) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

O ácido nucleico pode ser introduzido diretamente em uma célula vegetal ou na própria planta (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" é definido na seção "Definições".

A presente invenção claramente se estende a qualquer célula vegetal ou planta produzida por qualquer dos métodos descritos neste lugar, e a todas as partes vegetais e propágulos destas. A presente invenção adicionalmente de estende para abranger a progênie de uma célula primária transformada ou transfectada, tecido, órgão ou planta inteira que foi produzida

por qualquer dos métodos mencionados acima, o único requerimento sendo que a progênie exiba a(s) mesma(s) característica(s) genotípica(s) e/ou fenotípica(s) que aquelas produzidas pelo ancestral nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucleico CDC27 isolado codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo e cujo ácido nucleico está sob o controle de um promotor de meristema apical de broto. Células hospedeiras preferidas de acordo com a invenção são células vegetais.

A invenção também se estende a partes aptas à colheita de uma planta tal como, mas não limitada a sementes, folhas, frutos, flores, caules, rizomas, tubérculos e bulbos. A invenção além disso se refere a produtos derivados, preferivelmente diretamente derivados, de uma parte apta à colheita de uma tal planta, tal como pelotas secas ou pós, óleo, gordura e ácidos graxos, amido ou proteínas.

A presente invenção também abrange uso de ácidos nucleicos CDC27 codificando polipeptídeos CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo, cujos ácidos nucleicos estão sob o controle de um promotor de meristema apical de broto. Tais usos se referem a aumentar produção vegetal como definida acima nos métodos da invenção.

Desempenho dos métodos de acordo com a presente invenção resulta em plantas tendo número de sementes aumentado em relação a plantas de controle adequadas. Este aumento em número de sementes também pode ser combinado com outras características economicamente vantajosas, tal como características adicionais que estimulam produção, tolerância a outros estresses abióticos e bióticos, características modificando vários atributos arquitetônicos e/ou atributos bioquímicos e/ou fisiológicos.

AT-hook de Descrição Detalhada

Domínios AT-hook são bem conhecidos na técnica e são tipicamente encontrados em polipeptídeos pertencendo a uma família de fatores de transcrição associados com remodelagem de Cromatina. O motivo AT-hook é constituído de 13 ou mais (algumas vezes cerca de 9) aminoácidos que participam em ligação de DNA e que têm uma preferência por regiões ricas em A/T. Em *Arabidopsis* existem pelo menos 34 proteínas contendo domínios AT-hook. Estas proteínas compartilham homologia ao longo da maioria da seqüência, com o domínio AT-hook sendo uma região particularmente altamente conservada. O domínio AT-hook é ilustrado em Figura 23 e Tabela 11 a seguir; veja também a anotação apropriada de SEQ ID NO: 153, SEQ ID NO: 155, SEQ ID NO: 157, SEQ ID NO: 159, SEQ ID NO: 161, SEQ ID NO: 163, SEQ ID NO: 165, SEQ ID NO: 167, SEQ ID NO: 169 e SEQ ID NO: 171 onde a posição do domínio AT-hook é especificada. Como mostrado no alinhamento de Figura 23, alguma variação dentro do domínio AT-hook é permitida. Tipicamente, um ou dois domínios AT-hook precedem o domínio DUF296. Referência neste lugar a um domínio AT-hook é adotada para significar uma seqüência de polipeptídeos tendo em ordem crescente de preferência pelo menos 70%, 75%, 80%, 85%, 90% ou 95% de identidade com o domínio AT-hook de SEQ ID NO: 153, que é repetido aqui para conveniência: RRPRGRPAGSKNK (domínio AT-hook de SEQ ID NO: 153).

Domínios DUF296 (referidos em Interpro como IPR005175) também são bem conhecidos na técnica. O domínio DUF296 é ilustrado em Figura 23 e Tabela 11 a seguir; veja também a anotação apropriada de SEQ ID NO: 153, SEQ ID NO: 155, SEQ ID NO: 157, SEQ ID NO: 159, SEQ ID NO: 161, SEQ ID NO: 163, SEQ ID NO: 165, SEQ ID NO: 167, SEQ ID NO: 169 and SEQ ID NO: 171, onde a posição do domínio DUF296 é especificada. Como mostrado no alinhamento de Figura 23, variação dentro

do domínio DUF296 é permitida enquanto que ainda sendo facilmente identificado como um domínio DUF296 devido à presença de alguns resíduos de aminoácidos altamente conservados. Tipicamente, o domínio DUF296 é precedido por um ou dois domínios AT-hook.

De acordo com uma característica preferida da presente invenção, polipeptídeos compreendendo um domínio AT-hook e um domínio DUF296 adicionalmente compreendem um dos seguintes motivos:

Motivo 1 (SEQ ID NO: 190): QGQ V/I GG; ou

Motivo 2 (SEQ ID NO: 191): ILSLSGSFLPPPAPP; ou

Motivo 3 (SEQ ID NO: 192): NATYERLP; ou

Motivo 4 (SEQ ID NO: 193): SFTNVAYERLPL com nenhuma ou uma mudança de aminoácido em qualquer posição; ou

Motivo 5 (SEQ ID NO: 194): GRFEILSLTGSFLPGPAPPGSTGLTIYLAGGQGQVVGGSVVG com nenhuma, uma ou duas mudança de aminoácido em qualquer posição.

De acordo com uma característica preferida da presente invenção, seqüências adequadas para uso nos métodos da invenção são polipeptídeos compreendendo um domínio AT-hook (como definido acima) e um domínio DUF296 (como definido acima) e Motivo 2 (como definido acima), ou ácidos nucleicos codificando tais polipeptídeos.

Deve ser entendido que as seqüências detalhadas em Tabela 1 e aquelas mostradas no alinhamento de Figura 23 são apenas exemplos de seqüências úteis nos métodos da invenção e que qualquer polipeptídeo tendo um domínio AT-hook e um domínio DUF296, ou qualquer ácido nucleico codificando o mesmo, pode ser adequado para uso para realizar os métodos da invenção.

Tabela 11: Exemplos de seqüências de aminoácidos compreendendo um domínio AT-hook e um domínio DUF296 com detalhes das seqüências destes domínios e suas respectivas posições

	SEQ ID NO	Espécie	Posição de	Seqüência de	Posição de	Seqüência de domínio

			domínio AT Hook	domínio AT Hook	domínio Duf296	DUF296
1	SEQ ID NO: 153 CDS3129	ORYSA	97-109	rrprgrpagsknk	124-241	lrthvmevaggcdisesittfarrr qrgvcvlsgagtvtnvtlrqpasq gavvalhgrfeilslsgsflpppap peatgltvylaggqgqvvggsvv galtaagpvvimaasfanavy
2	SEQ ID NO: 155 CDS3128	ORYSA	97-109	RRPRGRPPG SKNK	109-227	lrahilevgsgcdvfecvstyarrr qrgvcvlsgsgvvtnvtlrqpsap agavvslhgrfeilslsgsflpppa ppgatsltiflaggqgqvvggnv vgalyaagpviviaasfanvay
3	SEQ ID NO: 157 LOTCO-AP006863.1 (gi68264919)	LOTUS	81-93	rrprgrpagsknk	108-225	lkthvmevadgcdivdsvsnfar rrqrgvcimsgtgvtnvtlrqpa ssgavvtlhgrfeilslagsflppp appaasgltiylaggqgqvvggs vvgaliasgpvvimaasfsnaay
4	SEQ ID NO: 159 NP_192942	ARATH	119-131	RRPRGRPA GSKNK	145-263	lrthvmeigdgcdivdcmatfar rrqrgvcvmsgtgsvtnvtirqp gsppgsvvslhgrfeilslsgsflp ppappaatglsvylaggqgqvv ggsvvgpllcsgpvvvmaasfs naay
5	SEQ ID NO: 161 NP_194012	ARATH	105-117	rrprgrpagsknk	132-252	farrrqrgvcvmsgtgnvtnvtir qpgshpspgsvvslhgrfeilsls gsflpppapptatglsvylaggqg qvvggsvvgpllcagpvvvma asfsna
6	SEQ ID NO: 163 NP_182067	ARATH	89-101	rrprgrpagsknk	116-237	lkshvmevangcdvmesvtvf arrrqrgicvlsgngavtnvtirqp asvpgggssvvnlhgrfeilslsg sflpppappaasgltiylaggqgq vvggsvvgplmasgpvvimaa sfgnaay
7	SEQ ID NO: 165 At3g60870/At_NP _191646	ARATH	59-71	rrprgrpagsknk	86-201	frchvmeitnacdvmeslavfar rrqrgvcvltgngavtnvtvrqpg ggvvslhgrfeilslsgsflpppap paasglkvylaggqgqviggsv vgpltasspvvvmaasfgnasy
8	SEQ ID NO: 167 CDS0185	ARATH	88-100	rrprgrppgsknk	115-233	lqshvleiatgadvaeslnafarrr grgvsvlsgsglvtnvtlrqpaas ggvvslrgqfeilsmcgaflptsg spaaaagltiylagaqgqvvggg vagpliasgpviviaatfcnaty
9	SEQ ID NO: 169 PROT codificado por AK107405	ORYSA	111-123	rrprgrpagsknk	138-256	lrahvlevasgcdlvdsvatfarrr qvgvcvlsatgavtnvsvrqpga gpgavvnltgrfdilslsgsflppp appsatgltvyvsggqgqvvggt vagpliavgpvvimaasfgnaa y
10	SEQ ID NO: 171 NP_912386.1	ORYSA	45-57	rrprgrppgsknk	72-190	mrshvleiasgadiveaiagfsrrr qrgvsvlsgsgavtnvtlrqpagt gaaavalrgrfeilsmsgaflpap appgatglavylaggqgqvvgg svmgeliasgpvmviaatfgnat y
11	SEQ ID NO: 173 Le_BT013387	LYCES	54-66	rrprgrpagsknk	81-198	lrahilevssghdvfesvatyarkr qrgicilsgsgtvnnvtirqpqaa gsvvtlhgrfeilslsgsflpppap pgatsltiylaggqgqvvggnvv galiasgpviviassftnvay
12	SEQ ID NO: 175 CDS3125	ARATH	34-46	rrprgrpagsknk	61-180	lrshvlevtsgsdiseavstyatrrg cgvciisgtgavtnvtirqpaapa gggvitlhgrfdilsltgtalpppa ppgaggltvylaggqgqvvggn vagsliasgpvvlmaasfanavy
13	SEQ ID NO: 177 CDS3399	ARATH	80-92	rrprgrpagsknk	107-232	lkshvmeiasgtdvietlatfarrr qrgicilsgngtvanvtlrqpstaa

						vaaapggaavlalqgrfeilsltgs flpgpappgstgltiylaggqgqv vggsvvgplmaagpvmliaatf snaty
14	SEQ ID NO: 179 PRO AK110263	ORYSA	35-47	rrprgrppgsknk	62-179	lrshvmevaggadvaesiahfar rrqrgvcvlsgagtvtdvalrqpa apsavvalrgrfeilsltgtflpgpa ppgstgltvylaggqgqvvggsv vgtltaagpvmv
15	SEQ ID NO: 181 At4g14465/NP_56 7432	ARATH	67-79	rrprgrppgsknk	94-211	lrshvleisdgsdvadtiahfsrrr qrgvcvlsgtgsvanvtlrqaaap ggvvslqgrfeilsltgaflpgpsp pgstgltvylagvqgqvvggsvv gpllaigsvmviaatfsnaty
16	SEQ ID NO: 183 CDS4145	ARATH	82-94	rrprgrppgsknk	109-226	lrahilevtngcdvfdcvatyarrr qrgicvlsgsgtvtnvsirqpsaa gavvtlqgtfeilslsgsflpppap pgatsltiflaggqgqvvggsvvg eltaagpviviaasftnvay
17	SEQ ID NO: 185 XP_473716	ORYSA	82-94	rrprgrppgsknk	109-227	lrahilevgsgcdvfecvstyarrr qrgvcvlsgsgvvtnvtlrqpsap agavvslhgrfeilslsgsflpppa ppgatsltiflaggqgqvvggnv vgalyaagpviviaasfanvay
18	SEQ ID NO: 187 NP_181070	ARATH	78-90	rrprgrpagsknk	105-222	lrahilevgsgcdvfecistyarrr qrgicvlsgtgtvtnvsirqptaag avvtlrgtfeilslsgsflpppapp gatsltiflagaqgqvvggnvvge lmaagpvmvmaasftnvay
19	SEQ ID NO: 189 TC102931	TC102931	86-98	rrprgrpagsknk	113-230	lrshvmevangcdimesvtvfar rrqrgvcilsgsgtvtnvtlrqpas pgavvtlhgrfeilslsgsflpppa ppaasglaiylaggqgqvvggsv vgpllasgpvvimaasfgnaay

Uma pessoa versada na técnica será prontamente capaz de identificar polipeptídeos compreendendo um domínio AT-hook e um domínio DUF296 usando técnicas e ferramentas bem conhecidas na técnica. Tal identificação pode ser por alinhamento de seqüências para comparação de seqüências usando 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 de duas seqüências completas que maximiza o número de identidades e minimiza o número de lacunas. O algoritmo BLAST (Altschul *et al.* (1990) J Mol Biol 215: 403-10) calcula identidade de seqüência percentual e realiza uma análise estatística da similaridade entre as duas seqüências. O aplicativo computacional para realizar análise BLAST está publicamente disponível através de National Centre for Biotechnology Information. Polipeptídeos compreendendo um domínio AT-hook e um domínio DUF296 podem ser prontamente identificados usando, por exemplo,

o algoritmo de alinhamento de múltiplas seqüências ClustalW (versão 1.83) disponível em http://clustalw.genome.jp/sit-bin/nph-ClustalW, com os parâmetros padrão de alinhamento pareado, e um método de pontuação em porcentagem. Edição manual secundária pode ser realizada para otimizar alinhamento entre motivos conservados, como deve ser perceptível para uma pessoa versada na técnica.

O domínio AT-hook e o domínio DUF296 podem ser identificados usando bancos de dados especializados e.g. SMART (Schultz *et al.* (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic *et al.* (2002) Nucleic Acids Res 30, 242-244; http://smart.embl-heidelberg.de/), InterPro (Mulder *et al.*, (2003) Nucl. Acids. Res. 31, 315-318; http://www.ebi.ac.uk/interpro/), Prosite (Bucher and Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (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), http://www.expasy.org/prosite/) ou Pfam (Bateman *et al.*, Nucleic Acids Research 30(1):276-280 (2002), http://www.sanger.ac.uk/Software/Pfam/).

As seqüências mencionadas em Tabela 11, ou como identificadas usando as técnicas mencionadas acima (tal como alinhamento de seqüências), podem ser consideradas homólogos de um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296, cujos homólogos também compreendem um domínio AT-hook e um domínio DUF296 mas que podem variar em qualquer outro lugar na seqüência. "Homólogos" de uma proteína são definidos na seção "Definições" neste lugar. Homólogos preferidos são seqüências de aminoácidos tendo em ordem crescente de preferência pelo menos 30%, 40%, 50%, 60%, 70%, 80%, 85%,

90%, 95%, 96%, 97%, 98% ou 99% ou mais de identidade de seqüência com uma seqüência de aminoácidos representada por SEQ ID NO: 153, cujos homólogos compreendem um domínio AT-hook e um domínio DUF296 e ainda preferivelmente compreendem Motivo 2.

O polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296, ou um homólogo de tal polipeptídeo, pode ser um derivado, como definido na seção "Definições" neste lugar.

Qualquer ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 pode ser adequado para uso nos métodos da invenção. Exemplos de tais seqüências incluem aquelas seqüências de nucleotídeos representadas por SEQ ID NO: 152, SEQ ID NO: 154, SEQ ID NO: 156, SEQ ID NO: 158, SEQ ID NO: 160, SEQ ID NO: 162, SEQ ID NO: 164, SEQ ID NO: 166, SEQ ID NO: 168 e SEQ ID NO: 170.

Variantes de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 também podem ser adequadas para uso para praticar os métodos da invenção contanto que as variantes codifiquem polipeptídeos compreendendo um domínio AT-hook e um domínio DUF296. Tais variantes de ácidos nucleicos podem ser porções de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 e/ou ácidos nucleicos capazes de hibridizar com um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296.

Uma porção pode ser preparada, por exemplo, fazendo uma ou mais deleções para um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296. As porções podem ser usadas em forma isolada ou elas podem ser fundidas a outras seqüências de codificação (ou não codificantes) a fim de, por exemplo, produzir uma proteína que combina diversas atividades. Quando fundido a

outras seqüências de codificação, o polipeptídeo resultante produzido em tradução pode ser maior do que aquele previsto para a porção. Preferivelmente, a porção é uma porção de um ácido nucleico como representada por qualquer um de SEQ ID NO: 152, SEQ ID NO: 154, SEQ ID NO: 156, SEQ ID NO: 158, SEQ ID NO: 160, SEQ ID NO: 162, SEQ ID NO: 164, SEQ ID NO: 166, SEQ ID NO: 168 e SEQ ID NO: 170. Mais preferivelmente a porção é uma porção de um ácido nucleico como representada por SEQ ID NO: 152, cuja porção codifica um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 e ainda preferivelmente compreende Motivo 2.

Outra variante de ácido nucleico é um ácido nucleico capaz de hibridizar em condições de estringência reduzida, preferivelmente em condições estringentes, com um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296. Preferivelmente, a seqüência hibridizante é uma que é capaz de hibridizar com um ácido nucleico como representado por qualquer um de SEQ ID NO: 152, SEQ ID NO: 154, SEQ ID NO: 156, SEQ ID NO: 158, SEQ ID NO: 160, SEQ ID NO: 162, SEQ ID NO: 164, SEQ ID NO: 166, SEQ ID NO: 168 e SEQ ID NO: 170, ou com uma porção de qualquer das seqüências mencionadas acima como definido acima. Mais preferivelmente, a seqüência hibridizante é uma que é capaz de hibridizar com um ácido nucleico como representado por SEQ ID NO: 152, cuja seqüência hibridizante codifica um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 e ainda preferivelmente compreende Motivo 2.

O termo "hibridização" é como definido neste lugar na seção “Definições”.

Outra variante de ácido nucleico é uma variante de união alternativa, como definido na seção “Definições”. São preferidas variantes de união de seqüências de ácidos nucleicos representadas por SEQ ID NO: 152,

SEQ ID NO: 154, SEQ ID NO: 156, SEQ ID NO: 158, SEQ ID NO: 160, SEQ ID NO: 162, SEQ ID NO: 164, SEQ ID NO: 166, SEQ ID NO: 168 e SEQ ID NO: 170. A mais preferida é uma variante de união de uma seqüência de ácidos nucleicos como representada por SEQ ID NO: 152, cuja variante de união codifica um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 e ainda preferivelmente compreendendo Motivo 2.

Outra variante de ácido nucleico é uma variante alélica como definido na seção "Definições". São preferidas variantes alélicas de seqüências de ácidos nucleicos representadas por SEQ ID NO: 152, SEQ ID NO: 154, SEQ ID NO: 156, SEQ ID NO: 158, SEQ ID NO: 160, SEQ ID NO: 162, SEQ ID NO: 164, SEQ ID NO: 166, SEQ ID NO: 168 e SEQ ID NO: 170. A mais preferida é uma variante alélica de uma seqüência de ácidos nucleicos como representada por SEQ ID NO: 152, cuja variante alélica codifica um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 e ainda preferivelmente compreende Motivo 2.

Variantes de ácidos nucleicos também podem ser obtidas através de evolução direta (veja seção "Definições").

Mutagênese sítio dirigida também pode ser usada para gerar variantes de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296. Veja seção "Definições".

O ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 pode ser derivado de qualquer fonte natural ou artificial. O ácido nucleico/gene ou variante deste pode ser isolado de uma fonte microbiana, tal como levedura ou fungos, ou de uma fonte vegetal, algal ou animal. Este ácido nucleico pode ser modificado a partir de sua forma nativa em composição e/ou ambiente genômico através de manipulação humana deliberada. O ácido nucleico é preferivelmente de origem vegetal, quer da mesma espécie de planta (por exemplo que aquela na qual ele é introduzido) ou quer de uma espécie diferente de planta. O ácido

nucleico pode ser isolado de uma espécie dicotiledônea, preferivelmente de uma espécie monocotiledônea tal como arroz. Mais preferivelmente, o ácido nucleico de arroz codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 é representado por SEQ ID NO: 152 e o polipeptídeo codificado é como representado por SEQ ID NO: 153.

A expressão de um ácido nucleico codificando AT-hook pode ser modulada introduzindo uma modificação genética (preferivelmente no locus de um gene codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296). O locus de um gene como definido neste lugar é adotado para significar uma região genômica, a qual inclui o gene de interesse and 10 kb antes ou depois da região de codificação.

A modificação genética pode ser introduzida, por exemplo, por qualquer um (ou mais) dos seguintes métodos: ativação de T-DNA, TILLING, recombinação homóloga e introduzindo e expressando em uma planta monocotiledônea um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296. Veja a seção "Definições" para detalhes de ativação de T-DNA, TILLING e recombinação homóloga. Seguindo introdução da modificação genética, pode seguir uma etapa de selecionar para expressão aumentada em tecido de endosperma de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296, cuja expressão direcionada gera plantas tendo produção aumentada de sementes.

A escolha de promotor para identificação por ativação de T-DNA no caso da presente invenção pode ser qualquer promotor capaz de dirigir preferencialmente expressão em tecido de endosperma de uma planta monocotiledônea.

Ativação de T-DNA e TILLING são exemplos de tecnologias que possibilitam a geração de novos alelos e variantes de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um

domínio DUF296.

Um método preferido para introduzir uma modificação genética (que neste caso não precisa ser no locus de um ácido nucleico/gene codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296) é introduzir e expressar em uma planta um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296. O ácido nucleico a ser introduzido em uma planta pode ser um ácido nucleico de comprimento completo ou pode ser uma porção ou qualquer outro ácido nucleico variante contanto que o ácido nucleico variante codifique um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296.

Os métodos da presente invenção recorrem a aumentar preferencialmente expressão em tecido de endosperma de uma planta monocotiledônea de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296. Isto pode ser atingido por superexpressão dirigida por promotores apropriados, o uso de estimuladores de transcrição ou estimuladores de tradução. Ácidos nucleicos isolados que servem como elementos promotores ou estimuladores podem ser introduzidos em uma posição apropriada (tipicamente antes) de uma forma não heteróloga de um polinucleotídeo de forma a aumentar expressão de um gene/ácido nucleico ou variante deste codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296. Por exemplo, promotores endógenos podem ser alterados *in vivo* por mutação, deleção, e/ou substituição (veja, Kmiec, Patente Norte-Americana No. 5.565.350; Zarling *et al.*, PCT/US93/03868), ou promotores isolados podem ser introduzidos em uma célula vegetal na orientação e distância apropriada de um gene da presente invenção para controlar expressão do gene.

Se expressão de polipeptídeo é desejada, geralmente é desejável incluir uma região de poliadenilação na extremidade 3' de uma

região de codificação de polinucleotídeo. A região de poliadenilação pode ser derivada do gene natural, de uma variedade de outros 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 outro gene vegetal, ou menos preferivelmente de qualquer outro gene eucariótico.

Uma seqüência intrônica também pode ser adicionada à região 5' não traduzida ou à seqüência de codificação da seqüência de codificação parcial para aumentar a quantidade da mensagem madura que se acumula no citossol. Inclusão de um íntron processável na unidade de transcrição tanto em construções vegetais como animais mostrou aumentar expressão gênica tanto no nível de mRNA como protéico em até 1000 vezes (Buchman and Berg (1988) Mol. Cell biol. 8: 4395-4405; Callis *et al.* (1987) Genes Dev 1:1183-1200). Tal estimulação intrônica de expressão gênica tipicamente é maior quando localizada perto da extremidade 5' da unidade de transcrição. Uso dos íntrons de milho íntron Adh1-S 1, 2, e 6, o íntron Bronze-1 são conhecidos na técnica. Veja geralmente, The Maize Handbook, capítulo 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

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

Portanto, é fornecida uma construção gênica compreendendo:

(i) Um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296;

(ii) Uma ou mais seqüências de controle capazes de dirigir expressão da seqüência de ácidos nucleicos de (i) em tecido de endosperma de uma planta monocotiledônea; e opcionalmente

(iii) Uma seqüência de término de transcrição.

A invenção também fornece uso de uma construção como

definido acima em métodos para aumentar produção de sementes de uma planta monocotiledônea.

Construções úteis nos métodos de acordo com a presente invenção podem ser construídas usando tecnologia de DNA recombinante bem conhecida por pessoas versadas na técnica. As construções gênicas podem ser inseridas em vetores, que podem estar comercialmente disponíveis, adequados para transformar em plantas e adequados para expressão do gene de interesse nas células transformadas. A invenção também fornece uso de uma construção como definido acima em métodos para aumentar produção de sementes em uma planta monocotiledônea.

Plantas monocotiledôneas são transformadas com um vetor compreendendo a seqüência de interesse (i.e., um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296). A seqüência de interesse é operavelmente ligada a uma ou mais seqüências de controle (pelo menos a um promotor) capaz de aumentar preferencialmente expressão em tecido de endosperma de uma planta monocotiledônea. Os termos "elemento regulador", "seqüência de controle" e "promotor" são todos usados permutavelmente neste lugar e são definidos na seção "Definições".

Um promotor endosperma específico se refere a qualquer promotor capaz de dirigir preferencialmente expressão do gene de interesse em tecido de endosperma. Referência neste lugar para aumentar preferencialmente expressão em tecido de endosperma é adotada para significar aumentar expressão em tecido de endosperma substancialmente para a exclusão de expressão em qualquer outro lugar na planta, sem considerar qualquer expressão residual devido a promotores permeáveis. Por exemplo, o promotor de prolamina mostra forte expressão no endosperma, com permeabilidade em meristema, mais especificamente o meristema de broto e/ou centro de discriminação no meristema.

Preferivelmente, o promotor endosperma específico é um promotor isolado de um gene de prolamina, tal como um promotor de prolamina RP6 de arroz (Wen *et al.*, (1993) Plant physiol 101(3):1115-6) como representado por SEQ ID NO: 195 ou um promotor de força similar e/ou um promotor com um padrão de expressão similar que o promotor de prolamina de arroz. Força similar e/ou padrão de expressão similar podem ser analisados, por exemplo, acoplando os promotores a um gene repórter e verificando a função do gene repórter em tecidos da planta. Um gene repórter bem conhecido é beta-glucuronidase e o corante colorimétrico GUS usado para visualizar atividade de beta-glucuronidase em tecido vegetal. Deve ser claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico representado por SEQ ID NO: 152, nem é a aplicabilidade da invenção restrita a expressão de um ácido nucleico codificando um domínio AT-hook e um domínio DUF296 quando dirigida por um promotor de prolamina. Exemplos de outros promotores endosperma específicos que também podem ser usados para realizar os métodos da invenção são mostrados em Tabela 6 na seção "Definições".

Opcionalmente, uma ou mais seqüências terminadoras também podem ser usadas na construção introduzida em uma planta. O termo "terminador" é definido na seção "Definições".

As construções genéticas da invenção podem incluir adicionalmente uma seqüência de origem de replicação que é requerida para manutenção e/ou replicação em um tipo celular específico. Um exemplo é quando uma construção genética é requerida ser mantida em uma célula bacteriana como um elemento genético episssomal (e.g. molécula de plasmídeo ou cosmídeo). Origens de replicação preferidas incluem, mas não são limitadas a, a f1-ori e colE1.

A construção genética pode compreender opcionalmente um gene marcador selecionável como definido neste lugar.

Em uma forma de realização preferida, é fornecida uma construção gênica compreendendo:

(i) Um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296;

(ii) Um promotor de prolamina capaz de dirigir preferencialmente expressão da seqüência de ácidos nucleicos de (i) em tecido de endosperma de uma planta monocotiledônea; e opcionalmente

(iii) Uma seqüência de término de transcrição.

A presente invenção também abrange plantas monocotiledôneas obteníveis pelos métodos de acordo com a presente invenção. A presente invenção portanto fornece plantas monocotiledôneas, partes destas (incluindo células vegetais) obteníveis pelos métodos de acordo com a presente invenção, cujas plantas ou partes destas compreendem um transgene codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 operavelmente ligado a um promotor endosperma específico, preferivelmente a um promotor de prolamina.

A invenção também fornece um método para a produção de plantas monocotiledôneas transgênicas tendo produção aumentada de sementes em relação a plantas de controle adequadas, compreendendo introdução e expressão em uma planta monocotiledônea de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296, caracterizado pelo fato de que dita expressão é preferencialmente aumentada em tecido de endosperma de uma planta monocotiledônea.

Mais especificamente, a presente invenção fornece um método para a produção de plantas monocotiledôneas transgênicas tendo produção aumentada de sementes cujo método compreende:

(i) introduzir e aumentar preferencialmente expressão em tecido de endosperma de uma planta monocotiledônea de um ácido nucleico

codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296; e

(ii) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

O ácido nucleico pode ser introduzido diretamente em uma célula vegetal de uma planta monocotiledônea ou na própria planta (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" é definido na seção "Definições" neste lugar.

A presente invenção claramente se estende a qualquer célula vegetal ou planta produzida por qualquer dos métodos descritos neste lugar, e a todas as partes vegetais e propágulos destas. A presente invenção adicionalmente de estende para abranger a progênie de uma célula primária transformada ou transfectada, tecido, órgão ou planta inteira que foi produzida por qualquer dos métodos mencionados acima, o único requerimento sendo que a progênie exiba a(s) mesma(s) característica(s) genotípica(s) e/ou fenotípica(s) que aquelas produzidas pelo ancestral nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 operavelmente ligado a um promotor endosperma específico. Células hospedeiras preferidas de acordo com a invenção são células de planta monocotiledônea.

A invenção também se estende a partes aptas à colheita de uma planta monocotiledônea tal como, mas não limitada a sementes, folhas, frutos, flores, caules, rizomas, tubérculos e bulbos. A invenção além disso se refere a produtos derivados de, preferivelmente diretamente derivados de, uma parte

apta à colheita de uma tal planta, tal como pelotas secas ou pós, óleo, gordura e ácidos graxos, amido ou proteínas.

A presente invenção também abrange uso de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 para aumentar produção de sementes de uma planta monocotiledônea usando os métodos da invenção.

Fatores de transcrição DOF de Descrição Detalhada

O termo "polipeptídeo de fator de transcrição DOF" como definido neste lugar se refere a qualquer polipeptídeo compreendendo característica (i) como segue, e adicionalmente ou característica (ii) ou (iii) como seguem:

(i) em ordem crescente de preferência pelo menos 60%, 65%, 70%, 75%, 80%, 85%, 90% ou 95% de identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e

(ii) em ordem crescente de preferência pelo menos 70%, 75%, 80%, 85%, 90% ou 95% de identidade de seqüência com o domínio DOF representado por SEQ ID NO: 200; ou

(iii) Motivo I: KALKKPDKILP (SEQ ID NO: 229) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

Motivo II: DDPGIKLFGKTIPF (SEQ ID NO: 230) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição.

Adicionalmente, polipeptídeos compreendendo característica (i) e característica (iii) acima podem compreender qualquer um, quaisquer dois ou todos os três dos seguintes motivos:

- Motivo III: SPTLGKHSRDE (SEQ ID NO: 231) sem

nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

- Motivo IV: LQANPAALSRSQNFQE (SEQ ID NO: 232) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

- Motivo V: KGEGCLWVPKTLRID DPDEAAKSSIWT TLGIK (SEQ ID NO: 233) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas, três, quatro ou cinco mudança(s) não-conservativa(s) em qualquer posição.

Um polipeptídeo preferido compreendendo característica (i) e característica (iii) acima compreende tanto Motivo I como II.

Além disso, polipeptídeos de fator de transcrição DOF (pelo menos na sua forma nativa) tipicamente têm atividade de ligação de DNA e têm um domínio de ativação. A presença de um domínio de ativação e atividade de ligação de DNA pode ser facilmente determinada por uma pessoa versada na técnica usando técnicas e procedimentos de rotina.

SEQ ID NO: 199 (codificada por SEQ ID NO: 198) é um exemplo de um polipeptídeo de fator de transcrição DOF compreendendo características (i) e (ii) como definido acima, i.e. pelo menos 60% de identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e pelo menos 70% de identidade de seqüência com o domínio DOF representado por SEQ ID NO: 200. Exemplos adicionais de polipeptídeos de fator de transcrição DOF compreendendo características (i) e (ii) como definido acima são dados em SEQ ID NO: 202 (codificada por SEQ ID NO: 201), SEQ ID NO: 204 (codificada por SEQ ID NO: 203), SEQ ID NO: 206 (codificada por SEQ ID NO: 205), SEQ ID NO: 208 (codificada por SEQ ID NO: 207), SEQ ID NO: 210 (codificada por SEQ ID NO: 209),

SEQ ID NO: 212 (codificada por SEQ ID NO: 211), SEQ ID NO: 214 (codificada por SEQ ID NO: 213), SEQ ID NO: 216 (codificada por SEQ ID NO: 215), SEQ ID NO: 218 (codificada por SEQ ID NO: 217), SEQ ID NO: 220 (codificada por SEQ ID NO: 219), SEQ ID NO: 222 (codificada por SEQ ID NO: 221).

SEQ ID NO: 227 (codificada por SEQ ID NO: 226) é um exemplo de um polipeptídeo de fator de transcrição DOF compreendendo características (i) e (iii) como definido acima, i.e. pelo menos 60% de identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e Motivo I e/ou Motivo II como definido acima. Exemplos adicionais de polipeptídeos de fator de transcrição DOF compreendendo características (i) e (iii) como definido acima são dados em SEQ ID NO: 235 (codificada por SEQ ID NO: 234), SEQ ID NO: 237 (codificada por SEQ ID NO: 236), SEQ ID NO: 239 (codificada por SEQ ID NO: 238), SEQ ID NO: 241 (codificada por SEQ ID NO: 240), SEQ ID NO: 243 (codificada por SEQ ID NO: 242), SEQ ID NO: 245 (codificada por SEQ ID NO: 244), SEQ ID NO: 247 (codificada por SEQ ID NO: 246), SEQ ID NO: 249 (codificada por SEQ ID NO: 248), SEQ ID NO: 251 (codificada por SEQ ID NO: 250), SEQ ID NO: 253 (codificada por SEQ ID NO: 252), SEQ ID NO: 255 (codificada por SEQ ID NO: 254).

Os exemplos adicionais representados por SEQ ID NO: 202, SEQ ID NO: 204, SEQ ID NO: 206, SEQ ID NO: 208, SEQ ID NO: 210, SEQ ID NO: 212, SEQ ID NO: 214, SEQ ID NO: 216, SEQ ID NO: 218, SEQ ID NO: 220, SEQ ID NO: 222 são exemplos de "homólogos" de um polipeptídeo de fator de transcrição DOF representado por SEQ ID NO: 199.

Os exemplos adicionais representados por SEQ ID NO: 235, SEQ ID NO: 237, SEQ ID NO: 239, SEQ ID NO: 241, SEQ ID NO: 243, SEQ ID NO: 245, SEQ ID NO: 247, SEQ ID NO: 249, SEQ ID NO: 251, SEQ ID NO: 253, SEQ ID NO: 255 são exemplos de "homólogos" de um

polipeptídeo de fator de transcrição DOF representado por SEQ ID NO: 227.

"Homólogos" de uma proteína são como definido neste lugar na seção "Definições".

O polipeptídeo de fator de transcrição DOF ou homólogo deste pode ser um derivado. "Derivados" são definidos na seção "Definições" neste lugar.

Os vários domínios estruturais em uma proteína de fator de transcrição DOF, tal como o domínio DOF, podem ser identificados usando bancos de dados especializados e.g. SMART (Schultz *et al.* (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic *et al.* (2002) Nucleic Acids Res 30, 242-244; http://smart.embl-heidelberg.de/), InterPro (Mulder *et al.*, (2003) Nucl. Acids. Res. 31, 315-318; http://www.ebi.ac.uk/interpro/), Prosite (Bucher and Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (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), http://www.expasy.org/prosite/) ou Pfam (Bateman *et al.*, Nucleic Acids Research 30(1): 276-280 (2002), http://www.sanger.ac.uk/Software/Pfam/).

Exemplos de ácidos nucleicos codificando polipeptídeos de fator de transcrição DOF (e homólogos destes) incluem aqueles representados por qualquer um de: SEQ ID NO: 198, SEQ ID NO: 201, SEQ ID NO: 203, SEQ ID NO: 205, SEQ ID NO: 207, SEQ ID NO: 209, SEQ ID NO: 211, SEQ ID NO: 213, SEQ ID NO: 215, SEQ ID NO: 217, SEQ ID NO: 219, SEQ ID NO: 221, SEQ ID NO: 226, SEQ ID NO: 234, SEQ ID NO: 236, SEQ ID NO: 238, SEQ ID NO: 240, SEQ ID NO: 242, SEQ ID NO: 244, SEQ ID NO: 246, SEQ ID NO: 248, SEQ ID NO: 250, SEQ ID NO: 252 e SEQ ID NO: 254. Variantes de ácidos nucleicos codificando polipeptídeos de

fator de transcrição DOF podem ser adequadas para uso nos métodos da invenção. Variantes adequadas incluem porções de ácidos nucleicos codificando polipeptídeos de fator de transcrição DOF e/ou ácidos nucleicos capazes de hibridizar com ácidos nucleicos/genes codificando polipeptídeos de fator de transcrição DOF. Variantes adicionais incluem variantes de união e variantes alélicas de ácidos nucleicos codificando polipeptídeos de fator de transcrição DOF (e homólogos destes).

O termo "porção" como definido neste lugar se refere a um pedaço de DNA codificando um polipeptídeo compreendendo característica (i) como segue, e adicionalmente ou característica (ii) ou (iii) como seguem:

(i) em ordem crescente de preferência pelo menos 60%, 65%, 70%, 75%, 80%, 85%, 90% ou 95% de identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e

(ii) em ordem crescente de preferência pelo menos 70%, 75%, 80%, 85%, 90% ou 95% de identidade de seqüência com o domínio DOF representado por SEQ ID NO: 200; ou

(iii) Motivo I: KALKKPDKILP (SEQ ID NO: 229) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

Motivo II: DDPGIKLFGKTIPF (SEQ ID NO: 230) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição.

Adicionalmente característica (iii) acima pode compreender qualquer um, quaisquer dois ou todos os três dos seguintes motivos:

- Motivo III: SPTLGKHSRDE (SEQ ID NO: 231) sem nenhuma mudança; ou uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em

qualquer posição; e/ou

- Motivo IV: LQANPAALSRSQNFQE (SEQ ID NO: 232) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

- Motivo V: KGEGCLWVPKTLRIDDPDE AAKSSIWTTL GIK (SEQ ID NO: 233) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas, três, quatro ou cinco mudança(s) não-conservativa(s) em qualquer posição.

Uma porção pode ser preparada, por exemplo, fazendo uma ou mais deleções para um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF. As porções podem ser usadas em forma isolada ou elas podem ser fundidas a outras seqüências de codificação (ou não codificantes) a fim de, por exemplo, produzir uma proteína que combina diversas atividades. Quando fundido a outras seqüências de codificação, o polipeptídeo resultante produzido em tradução pode ser maior do que aquele previsto para a porção de fator de transcrição DOF.

Porções de ácidos nucleicos codificando polipeptídeos de fator de transcrição DOF compreendendo características (i) e (ii) como definido acima são preferivelmente porções de um ácido nucleico como representado por qualquer um de: SEQ ID NO: 198, SEQ ID NO: 201, SEQ ID NO: 203, SEQ ID NO: 205, SEQ ID NO: 207, SEQ ID NO: 209, SEQ ID NO: 211, SEQ ID NO: 213, SEQ ID NO: 215, SEQ ID NO: 217, SEQ ID NO: 219 e SEQ ID NO: 221.

Porções de ácidos nucleicos codificando polipeptídeos de fator de transcrição DOF compreendendo características (i) e (iii) como definido acima são preferivelmente porções de um ácido nucleico como representado por qualquer um de: SEQ ID NO: 226, SEQ ID NO: 234, SEQ ID NO: 236, SEQ ID NO: 238, SEQ ID NO: 240, SEQ ID NO: 242, SEQ ID NO: 244,

SEQ ID NO: 246, SEQ ID NO: 248, SEQ ID NO: 250, SEQ ID NO: 252 e SEQ ID NO: 254.

Outra variante de um ácido nucleico/gene de fator de transcrição DOF é um ácido nucleico capaz de hibridizar em condições de estringência reduzida, preferivelmente em condições estringentes, com um ácido nucleico/gene de fator de transcrição DOF como definido acima, cuja seqüência hibridizante codifica um polipeptídeo compreendendo característica (i) como segue e adicionalmente ou característica (ii) ou (iii) como seguem:

(i) em ordem crescente de preferência pelo menos 60%, 65%, 70%, 75%, 80%, 85%, 90% ou 95% de identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e

(ii) em ordem crescente de preferência pelo menos 70%, 75%, 80%, 85%, 90% ou 95% de identidade de seqüência com o domínio DOF representado por SEQ ID NO: 200; ou

(iii) Motivo I: KALKKPDKILP (SEQ ID NO: 229) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

Motivo II: DDPGIKLFGKTIPF (SEQ ID NO: 230) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição.

Adicionalmente característica (iii) acima pode compreender qualquer um, quaisquer dois ou todos os três dos seguintes motivos:

- Motivo III: SPTLGKHSRDE (SEQ ID NO: 231) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

- Motivo IV: LQANPAALSRSQNFQE (SEQ ID NO: 232)

sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

- Motivo V: KGEGCLWVPKTLRIDDPDEAAK SSIWTTL GIK (SEQ ID NO: 233) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas, três, quatro ou cinco mudança(s) não-conservativa(s) em qualquer posição.

Preferivelmente, a seqüência hibridizante codificando polipeptídeos de fator de transcrição DOF compreendendo características (i) e (ii) como definido acima é uma seqüência capaz de hibridizar com um ácido nucleico como representado por qualquer um de: SEQ ID NO: 201, SEQ ID NO: 203, SEQ ID NO: 205, SEQ ID NO: 207, SEQ ID NO: 209, SEQ ID NO: 211, SEQ ID NO: 213, SEQ ID NO: 215, SEQ ID NO: 217, SEQ ID NO: 219 e SEQ ID NO: 221.

Preferivelmente, a seqüência hibridizante codificando polipeptídeos de fator de transcrição DOF compreendendo características (i) e (iii) como definido acima é uma seqüência capaz de hibridizar com um ácido nucleico como representado por qualquer um de: SEQ ID NO: 234, SEQ ID NO: 236, SEQ ID NO: 238, SEQ ID NO: 240, SEQ ID NO: 242, SEQ ID NO: 244, SEQ ID NO: 246, SEQ ID NO: 248, SEQ ID NO: 250, SEQ ID NO: 252 e SEQ ID NO: 254.

O termo "hibridização" é como definido neste lugar na seção "Definições".

O polipeptídeo de fator de transcrição DOF pode ser codificado por uma variante de união alternativa. O termo "variante de união alternativa" é como definido na seção "Definições" neste lugar.

Variantes de união preferidas são variantes de união do ácido nucleico codificando um polipeptídeo compreendendo característica (i) como segue e adicionalmente ou característica (ii) ou (iii) como seguem:

(i) em ordem crescente de preferência pelo menos 60%, 65%, 70%, 75%, 80%, 85%, 90% ou 95% de identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e

(ii) em ordem crescente de preferência pelo menos 70%, 75%, 80%, 85%, 90% ou 95% de identidade de seqüência com o domínio DOF representado por SEQ ID NO: 200; ou

(iii) Motivo I: KALKKPDKILP (SEQ ID NO: 229) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

Motivo II: DDPGIKLFGKTIPF (SEQ ID NO: 230) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição.

Variantes de união preferidas de ácidos nucleicos codificando polipeptídeos de fator de transcrição DOF compreendendo características (i) e (ii) como definido acima são variantes de união de um ácido nucleico como representado por qualquer um de: SEQ ID NO: 201, SEQ ID NO: 203, SEQ ID NO: 205, SEQ ID NO: 207, SEQ ID NO: 209, SEQ ID NO: 211, SEQ ID NO: 213, SEQ ID NO: 215, SEQ ID NO: 217, SEQ ID NO: 219 e SEQ ID NO: 221.

Variantes de união preferidas de ácidos nucleicos codificando polipeptídeos de fator de transcrição DOF compreendendo características (i) e (iii) como definido acima são preferivelmente variantes de união de um ácido nucleico como representado por qualquer um de: SEQ ID NO: 234, SEQ ID NO: 236, SEQ ID NO: 238, SEQ ID NO: 240, SEQ ID NO: 242, SEQ ID NO: 244, SEQ ID NO: 246, SEQ ID NO: 248, SEQ ID NO: 250, SEQ ID NO: 252 e SEQ ID NO: 254.

O polipeptídeo de fator de transcrição DOF também pode ser

codificado por uma variante alélica, que também são definidas na seção "Definições" neste lugar.

Variantes alélicas preferidas são variantes alélicas do ácido nucleico codificando um polipeptídeo compreendendo característica (i) como segue e adicionalmente ou característica (ii) ou (iii) como seguem:

(i) em ordem crescente de preferência pelo menos 60%, 65%, 70%, 75%, 80%, 85%, 90% ou 95% de identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e

(ii) em ordem crescente de preferência pelo menos 70%, 75%, 80%, 85%, 90% ou 95% de identidade de seqüência com o domínio DOF representado por SEQ ID NO: 200; ou

(iii) Motivo I: KALKKPDKILP (SEQ ID NO: 229) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

Motivo II: DDPGIKLFGKTIPF (SEQ ID NO: 230) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição.

Variantes alélicas preferidas de ácidos nucleicos codificando polipeptídeos de fator de transcrição DOF compreendendo características (i) e (ii) como definido acima são variantes de união de um ácido nucleico como representado por qualquer um de: SEQ ID NO: 201, SEQ ID NO: 203, SEQ ID NO: 205, SEQ ID NO: 207, SEQ ID NO: 209, SEQ ID NO: 211, SEQ ID NO: 213, SEQ ID NO: 215, SEQ ID NO: 217, SEQ ID NO: 219 e SEQ ID NO: 221.

Variantes alélicas preferidas de ácidos nucleicos codificando polipeptídeos de fator de transcrição DOF compreendendo características (i) e (iii) como definido acima são preferivelmente porções de um ácido nucleico

como representado por qualquer um de: SEQ ID NO: 234, SEQ ID NO: 236, SEQ ID NO: 238, SEQ ID NO: 240, SEQ ID NO: 242, SEQ ID NO: 244, SEQ ID NO: 246, SEQ ID NO: 248, SEQ ID NO: 250, SEQ ID NO: 252 e SEQ ID NO: 254.

Variantes adicionais de ácidos nucleicos codificando polipeptídeos de fator de transcrição DOF como definido acima podem ser geradas usando, por exemplo, mutagênese sítio dirigida como definido na seção "Definições" neste lugar.

Evolução direta (ou embaralhamento de gene) também pode ser usada para gerar variantes de ácidos nucleicos codificando polipeptídeos de fator de transcrição DOF. Veja seção "definições".

Polipeptídeos de fator de transcrição DOF são planta específicos. Ácidos nucleicos codificando os mesmos podem ser derivados de qualquer fonte natural ou artificial. O ácido nucleico ou variante deste pode ser modificado a partir de sua forma nativa em composição e/ou ambiente genômico através de manipulação humana deliberada. Preferivelmente o ácido nucleico de fator de transcrição DOF ou variante deste é de uma planta dicotiledônea, ainda preferivelmente da família *Brassicaceae*, mais preferivelmente o ácido nucleico é de *Arabidopsis thaliana*.

A expressão de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF pode ser aumentada introduzindo uma modificação genética (preferivelmente no locus de um gene de fator de transcrição DOF). O locus de um gene como definido neste lugar é adotado para significar uma região genômica, a qual inclui o gene de interesse e 10 KB antes ou depois da região de codificação.

A modificação genética pode ser introduzida, por exemplo, por qualquer um (ou mais) dos seguintes métodos: ativação de T-DNA, TILLING e recombinação homóloga ou introduzindo e expressando em uma planta um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF. Os

métodos de ativação de T-DNA, TILLING e recombinação homóloga são como definidos na seção "Definições" neste lugar. Seguindo introdução da modificação genética, segue uma etapa adicional de selecionar para expressão aumentada de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF, cuja expressão aumentada gera plantas tendo produção aumentada.

Ativação de T-DNA e TILLING são exemplos de tecnologias que possibilitam a geração de novos alelos e variantes de fator de transcrição DOF.

Um método preferido para introduzir uma modificação genética (que neste caso não precisa ser no locus de um gene de fator de transcrição DOF) é introduzir e expressar em uma planta um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF como definido acima. O ácido nucleico a ser introduzido em uma planta pode ser um ácido nucleico de comprimento completo ou pode ser uma porção ou uma seqüência hibridizante ou outra variante de ácido nucleico como definido acima.

Os métodos da invenção recorrem à expressão aumentada de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF. Métodos para aumentar expressão de genes ou produtos de gene são bem documentados na técnica e incluem, por exemplo, superexpressão dirigida por promotores apropriados, o uso de estimuladores de transcrição ou estimuladores de tradução. Ácidos nucleicos isolados que servem como elementos promotores ou estimuladores podem ser introduzidos em uma posição apropriada (tipicamente antes) de uma forma não heteróloga de um polinucleotídeo de forma a aumentar expressão de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF. Por exemplo, promotores endógenos podem ser alterados *in vivo* por mutação, deleção, e/ou substituição (veja, Kmiec, Patente Norte-Americana No. 5.565.350; Zarling *et al.*, PCT/US93/03868), ou promotores isolados podem ser introduzidos em

uma célula vegetal na orientação e distância apropriada de um gene da presente invenção de forma a controlar a expressão do gene.

Se expressão de polipeptídeo é desejada, geralmente é desejável incluir uma região de poliadenilação na extremidade 3' de uma região de codificação de polinucleotídeo. A região de poliadenilação pode ser derivada do gene natural, de uma variedade de outros 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 outro gene vegetal, ou menos preferivelmente de qualquer outro gene eucariótico.

Uma seqüência intrônica também pode ser adicionada à região 5' não traduzida ou à seqüência de codificação da seqüência de codificação parcial para aumentar a quantidade da mensagem madura que se acumula no citossol. Inclusão de um íntron processável na unidade de transcrição tanto em construções vegetais como animais mostrou aumentar expressão gênica tanto no nível de mRNA como protéico em até 1000 vezes (Buchman and Berg (1988) Mol. Cell biol. 8: 4395-4405; Callis *et al.* (1987) Genes Dev 1:1183-1200). Tal estimulação intrônica de expressão gênica tipicamente é maior quando localizada perto da extremidade 5' da unidade de transcrição. Uso dos íntrons de milho íntron Adh1-S 1, 2, e 6, o íntron Bronze-1 são conhecidos na técnica. Para informação geral veja: The Maize Handbook, Capítulo 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

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

Portanto, é fornecida uma construção gênica compreendendo:

(i) Um ácido nucleico ou variante deste codificando um polipeptídeo de fator de transcrição DOF como definido acima;

(ii) Uma ou mais seqüências de controle capazes de dirigir

expressão da seqüência de ácidos nucleicos de (i); e opcionalmente

(iii) Uma seqüência de término de transcrição.

Construções úteis nos métodos de acordo com a presente invenção podem ser construídas usando tecnologia de DNA recombinante bem conhecida por pessoas versadas na técnica. As construções gênicas podem ser inseridas em vetores, que podem estar comercialmente disponíveis, adequados para transformar em plantas e adequados para expressão do gene de interesse nas células transformadas. A invenção portanto fornece uso de uma construção gênica como definido acima nos métodos da invenção.

Plantas são transformadas com um vetor compreendendo a seqüência de interesse (i.e., um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF). A seqüência de interesse é operavelmente ligada a uma ou mais seqüências de controle (pelo menos a um promotor). Os termos "elemento regulador", "seqüência de controle" e "promotor" são todos usados permutavelmente neste lugar e são definidos na seção "Definições" neste lugar.

Vantajosamente, qualquer tipo de promotor, quer natural ou sintético, pode ser usado para dirigir expressão da seqüência de ácidos nucleicos.

De acordo com uma característica preferida da invenção, o ácido nucleico de fator de transcrição DOF ou variante deste é operavelmente ligado a um promotor constitutivo como definido na seção "Definições" neste lugar. O promotor constitutivo é preferivelmente um promotor GOS2, mais preferivelmente o promotor constitutivo é um promotor GOS2 de arroz, ainda preferivelmente o promotor constitutivo é representado por uma seqüência de ácidos nucleicos substancialmente similar a SEQ ID NO: 225, o mais preferivelmente o promotor constitutivo é como representado por SEQ ID NO: 225. É preferido o uso de um promotor constitutivo para dirigir expressão de um ácido nucleico codificando um polipeptídeo de fator de

transcrição DOF compreendendo características (i) e (ii) como definido acima, i.e. pelo menos 60% de identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e pelo menos 70% de identidade de seqüência com o domínio DOF representado por SEQ ID NO: 200.

Deve ser claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico de fator de transcrição DOF representado por SEQ ID NO: 198, nem é a aplicabilidade da invenção restrita a expressão de um ácido nucleico de fator de transcrição DOF quando dirigida por um promotor GOS2. Exemplos de outros promotores constitutivos que também podem ser usados para realizar os métodos da invenção são mostrados em Tabela 3 na seção "Definições" neste lugar.

De acordo com outra característica preferida da invenção, o ácido nucleico codificando um polipeptídeo de fator de transcrição DOF é operavelmente ligado a um promotor semente específico, i.e. um promotor que é expresso predominantemente em tecido de semente, mas que pode ter expressão residual em qualquer outro lugar na planta devido à expressão de promotor permeável. Ainda preferivelmente, o promotor semente específico é isolado de um gene codificando uma proteína de armazenamento de semente, especialmente um promotor endosperma específico. Mais preferivelmente o promotor endosperma específico é isolado de um gene de prolamina, tal como um promotor de prolamina RP6 de arroz (Wen *et al.*, (1993) Plant physiol 101(3): 1115-6) como representado por SEQ ID NO: 258, ou um promotor de força similar e/ou um promotor com um padrão de expressão similar que o promotor de prolamina de arroz. Força similar e/ou padrão de expressão similar podem ser analisados, por exemplo, acoplando os promotores a um gene repórter e verificando a função do gene repórter em tecidos da planta. Um gene repórter bem conhecido é beta-glucuronidase e o corante colorimétrico GUS usado para visualizar atividade de beta-glucuronidase em

tecido vegetal. O promotor de prolamina mostra forte expressão no endosperma, com permeabilidade em meristema, mais especificamente o meristema de broto e/ou centro de discriminação no meristema.

Preferido de acordo com a invenção é o uso de um promotor semente específico, especialmente um promotor endosperma específico, para dirigir expressão de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF compreendendo características (i) e (iii) como definido acima, i.e. pelo menos 60% de identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e Motivo I e/ou Motivo II.

Deve ser claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico de fator de transcrição DOF representado por SEQ ID NO: 226, nem é a aplicabilidade da invenção restrita a expressão de um ácido nucleico de fator de transcrição DOF quando dirigida por um promotor de prolamina.

Exemplos de promotores semente específicos são apresentados em Tabela 7 na seção "Definições" neste lugar, cujos promotores ou derivados destes são úteis para realizar os métodos da presente invenção.

Opcionalmente, uma ou mais seqüências terminadoras também podem ser usadas na construção introduzida em uma planta. O termo "terminador" é como definido na seção "Definições" neste lugar.

As construções genéticas da invenção podem incluir adicionalmente uma seqüência de origem de replicação que é requerida para manutenção e/ou replicação em um tipo celular específico. Um exemplo é quando uma construção genética é requerida ser mantida em uma célula bacteriana como um elemento genético episssomal (e.g. molécula de plasmídeo ou cosmídeo). Origens de replicação preferidas incluem, mas não são limitadas a, a f1-ori e colE1.

A construção genética pode compreender opcionalmente um

gene marcador selecionável como definido neste lugar na seção "Definições".

A presente invenção também abrange plantas obteníveis pelos métodos de acordo com a presente invenção. A presente invenção portanto fornece plantas, partes vegetais ou células vegetais destas obteníveis pelo método de acordo com a presente invenção, cujas plantas ou partes ou células destas compreendem um ácido nucleico transgene (ou variante deste como definido acima) codificando um polipeptídeo de fator de transcrição DOF.

A invenção também fornece um método para a produção de plantas transgênicas tendo produção aumentada em relação a plantas de controle adequadas, compreendendo introdução e expressão em uma planta de um ácido nucleico ou uma variante deste codificando um polipeptídeo de fator de transcrição DOF.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo produção aumentada cujo método compreende:

(i) introduzir e expressar em uma planta, parte vegetal ou célula vegetal um ácido nucleico ou variante deste codificando um polipeptídeo de fator de transcrição DOF; e

(ii) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

O ácido nucleico pode ser introduzido diretamente em uma célula vegetal ou na própria planta (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" é como definido neste lugar na seção "Definições".

A presente invenção claramente se estende a qualquer célula vegetal ou planta produzida por qualquer dos métodos descritos neste lugar, e

a todas as partes vegetais e propágulos destas. A presente invenção adicionalmente de estende para abranger a progênie de uma célula primária transformada ou transfectada, tecido, órgão ou planta inteira que foi produzida por qualquer dos métodos mencionados acima, o único requerimento sendo que a progênie exiba a(s) mesma(s) característica(s) genotípica(s) e/ou fenotípica(s) que aquelas produzidas pelo ancestral nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucleico isolado ou variante deste codificando um polipeptídeo de fator de transcrição DOF. Células hospedeiras preferidas de acordo com a invenção são células vegetais.

A invenção também se estende a partes aptas à colheita de uma planta tal como, mas não limitada a sementes, folhas, frutos, flores, caules, rizomas, tubérculos e bulbos. A invenção além disso se refere a produtos derivados, preferivelmente diretamente derivados, de uma parte apta à colheita de uma tal planta, tal como pelotas secas ou pós, óleo, gordura e ácidos graxos, amido ou proteínas.

A presente invenção também abrange uso de ácidos nucleicos ou variantes destes codificando polipeptídeos de fator de transcrição DOF e uso de polipeptídeos de fator de transcrição DOF para aumentar produção vegetal como definido acima nos métodos da invenção.

Ácidos nucleicos ou variantes destes codificando polipeptídeos de fator de transcrição DOF, ou polipeptídeos de fator de transcrição DOF, podem encontrar uso em programas de melhoramento nos quais é identificado um marcador de DNA que pode ser geneticamente ligado a um gene de fator de transcrição DOF ou variante deste. Os ácidos nucleicos/genes ou variantes destes, ou os polipeptídeos de fator de transcrição DOF podem ser usados para definir um marcador molecular. Este marcador de DNA ou proteína pode então ser usado em programas de melhoramento para selecionar plantas tendo

produção aumentada como definido acima nos métodos da invenção.

Variantes alélicas de um ácido nucleico/gene de fator de transcrição DOF também podem encontrar uso em programas de melhoramento auxiliados por marcador. Tais programas de melhoramento algumas vezes requerem introdução de variação alélica por tratamento mutagênico das plantas, usando por exemplo mutagênese EMS; alternativamente, o programa pode começar com uma coleção de variantes alélicas de assim chamada origem "natural" causada não intencionalmente. Identificação de variantes alélicas então ocorre, por exemplo, por PCR. Isto é seguido por uma etapa para seleção de variantes alélicas superiores da seqüência em questão e que gera produção aumentada. Seleção tipicamente é realizada monitorando desempenho de crescimento de plantas contendo diferentes variantes alélicas da seqüência em questão. Desempenho de crescimento pode ser monitorado em uma casa de vegetação ou no campo. Etapas opcionais adicionais incluem cruzar plantas nas quais a variante alélica superior foi identificada com outra planta. Isto pode ser usado, por exemplo, para fazer uma combinação de características fenotípicas interessantes.

Um ácido nucleico ou variante deste codificando um polipeptídeo de fator de transcrição DOF também pode ser usado como sondas para mapear geneticamente e fisicamente os genes dos quais eles são uma parte de, e como marcadores para características ligadas a estes genes. Tal informação pode ser útil em melhoramento vegetal a fim de desenvolver linhagens com fenótipos desejados. Tal uso de ácidos nucleicos de fator de transcrição DOF ou variantes destes requer apenas uma seqüência de ácidos nucleicos de pelo menos 15 nucleotídeos de comprimento. Os ácidos nucleicos de fator de transcrição DOF ou variantes destes 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 de fator de transcrição DOF ou variantes destes. Os padrões de bandeamento resultantes então podem ser sujeitos a análises genéticas usando programas de computador tal como MapMaker (Lander *et al.* (1987) Genomics 1: 174-181) a fim de construir um mapa genético. Em adição, os ácidos nucleicos podem ser usados para sondar Southern blots contendo DNAs genômicos tratados com endonuclease de restrição de um conjunto de indivíduos representando ancestral e progênie de um cruzamento genético definido. Segregação dos polimorfismos de DNA é observada e usada para calcular a posição do ácido nucleico de fator de transcrição DOF ou variante deste no mapa genético obtido previamente usando esta população (Botstein *et al.* (1980) Am. J. Hum. Genet. 32:314-331).

A produção e uso de sondas derivadas de gene vegetal para uso em mapeamento genético são descritos em Bematzky and Tanksley (1986) Plant Mol. Biol. Reporter 4: 37-41. Numerosas publicações descrevem mapeamento genético de clones de cDNA específicos usando a metodologia descrita acima ou variações desta. Por exemplo, populações de intercruzamento de F2, populações endocruzadas, populações cruzadas aleatoriamente, linhagens isogênicas próximas, e outros conjuntos de indivíduos podem ser usados para mapeamento. Tais metodologias são bem conhecidas por aqueles versados na técnica.

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

Em outra forma de realização, as sondas de ácido nucleico podem ser usadas em mapeamento direto por hibridização *in situ* por fluorescência (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora métodos atuais de mapeamento por FISH favoreçam uso de clones grandes

(diversos kb a diversas centenas kb; veja Laan *et al.* (1995) Genome Res. 5:13-20), melhoramentos em sensibilidade podem permitir desempenho de mapeamento por FISH usando sondas menores.

Uma variedade de métodos baseados em amplificação de ácido nucleico para mapeamento genético e físico pode ser realizada usando os ácidos nucleicos. Exemplos incluem amplificação alelo específica (Kazazian (1989) J. Lab. Clin. Med 11:95-96), polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield *et al.* (1993) Genomics 16:325-332), ligação alelo específica (Landegren *et al.* (1988) Science 241:1077-1080), reações de extensão de nucleotídeo (Sokolov (1990) Nucleic Acid Res. 18:3671), Mapeamento de Híbrido de Radiação (Walter *et al.* (1997) Nat. Genet. 7:22-28) e Mapeamento Apropriado (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 iniciadores para uso na reação de amplificação ou em reações de extensão de iniciador. O projeto de tais iniciadores é bem conhecido por aqueles versados na técnica. Em métodos empregando mapeamento genético baseado em PCR, pode ser necessário identificar diferenças de seqüência de DNA entre os ancestrais do cruzamento de mapeamento na região correspondente à seqüência imediata de ácidos nucleicos. Isto, entretanto, geralmente não é necessário para métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo produção aumentada, como descrito acima. Esta produção aumentada também pode ser combinada com outras características economicamente vantajosas, tal como características adicionais que estimulam produção, tolerância a outros estresses abióticos e bióticos, características modificando vários atributos arquitetônicos e/ou atributos bioquímicos e/ou fisiológicos.

CKI de Descrição Detalhada

Referência neste lugar a uma "redução" preferencial em expressão de um gene CKI endógeno em tecido de endosperma de uma planta é adotada para significar uma redução ou eliminação substancial de expressão de um gene CKI endógeno (em tecido de endosperma) em relação a níveis de expressão de gene CKI endógeno encontrados em tecido de endosperma de plantas tipo selvagem. Esta redução ou eliminação substancial de expressão de gene CKI endógeno pode resultar em níveis protéicos de CKI e/ou atividade em tecido de endosperma de uma planta reduzidos ou substancialmente eliminados.

Referência neste lugar a um gene CKI "endógeno" não se refere apenas a genes CKI como encontrado em uma planta em sua forma natural (i.e., sem haver qualquer intervenção humana), mas também se refere a genes CKI isolados subseqüentemente introduzidos em uma planta. Por exemplo, uma planta transgênica contendo um transgene CKI pode encontrar uma redução ou eliminação substancial do transgene CKI e/ou uma redução ou eliminação substancial de um gene CKI endógeno (em tecido de endosperma).

Esta redução (ou eliminação substancial) de expressão de gene CKI endógeno pode ser atingida usando qualquer um ou mais de diversos métodos de silenciamento de genes bem conhecidos. "Silenciamento de genes" ou "diminuição" de expressão, como usado neste lugar, se refere a uma redução ou à eliminação substancial de expressão gênica de CKI e/ou níveis de polipeptídeo CKI e/ou atividade de polipeptídeo CKI.

Um tal método para redução ou eliminação substancial de expressão de gene CKI endógeno é regulação negativa mediada por RNA de expressão gênica (silenciamento de RNA). Silenciamento neste caso é desencadeado em uma planta por uma molécula de RNA de dupla fita (dsRNA) que é substancialmente homóloga a um gene CKI alvo. Este dsRNA é adicionalmente processado pela planta em cerca de 21 a cerca de 26

nucleotídeos chamado RNAs de baixa interferência (siRNAs). Estes siRNAs são incorporados em um complexo de silenciamento induzido por (RISC) que cliva o mRNA de um gene CKI alvo, com isso reduzindo ou substancialmente eliminando o número de mRNAs de CKI a serem traduzidos em uma proteína CKI.

Um exemplo de um método de silenciamento de RNA envolve a introdução de seqüências de codificação ou partes destas em uma orientação sentido em uma planta. "Orientação sentido" se refere a DNA que é homólogo a um transcrito de mRNA deste. Deve ser introduzido em uma planta portanto pelo menos uma cópia adicional (completa ou em parte) de um gene CKI já presente na planta hospedeira. O gene adicional, ou parte deste, irá silenciar um gene CKI endógeno, gerando um fenômeno conhecido como co-supressão. A redução de expressão gênica de CKI será mais pronunciada se diversas cópias adicionais são introduzidas na planta, pois existe uma correlação positiva entre altos níveis de transcritos e o desencadeamento de co-supressão.

Outro exemplo de um método de silenciamento de RNA envolve o uso de seqüências de ácidos nucleicos CKI anti-sentido. Um ácido nucleico "anti-sentido" compreende uma seqüência de nucleotídeos que é complementar a um ácido nucleico "sentido" codificando uma proteína, *e.g.*, complementar à fita de codificação de uma molécula de cDNA dupla fita ou complementar a uma seqüência de mRNA. Por conseguinte, um ácido nucleico anti-sentido pode ser ligado por ponte de hidrogênio a um ácido nucleico sentido. O ácido nucleico anti-sentido pode ser complementar a uma fita de codificação inteira de CKI ou apenas a uma porção desta. A molécula de ácido nucleico anti-sentido pode ser anti-sentido para uma "região de codificação" ou anti-sentido para uma "região não codificante" da fita de codificação de uma seqüência de nucleotídeos codificando CKI. O termo "região de codificação" se refere à região da seqüência de nucleotídeos

compreendendo códons que são traduzidos em resíduos de aminoácidos. O termo "região não codificante" se refere a seqüências 5' e 3' que flanqueiam a região de codificação que não são traduzidas em aminoácidos (*i.e.*, também referidas como regiões 5' e 3' não traduzidas).

Ácidos nucleicos anti-sentido podem ser projetados de acordo com as regras de pareamento de bases de Watson e Crick. A molécula de ácido nucleico anti-sentido pode ser complementar à região de codificação inteira de mRNA de CKI, mas é preferivelmente um oligonucleotídeo que é anti-sentido para apenas uma porção da região de codificação ou não codificante de mRNA de CKI. Por exemplo, o oligonucleotídeo anti-sentido pode ser complementar à região ao redor do sítio de início de tradução de mRNA de CKI. O comprimento de um oligonucleotídeo anti-sentido adequado deve ser conhecido na técnica e pode começar de cerca de 20 nucleotídeos de comprimento ou menos. Um ácido nucleico anti-sentido da invenção pode ser construído usando síntese química e reações de ligação enzimática usando procedimentos conhecidos na técnica. Por exemplo, um ácido nucleico anti-sentido (*e.g.*, um oligonucleotídeo anti-sentido) pode ser quimicamente sintetizado usando nucleotídeos naturalmente ocorrentes ou nucleotídeos variadamente modificados projetados para aumentar a estabilidade biológica das moléculas ou para aumentar a estabilidade física do duplex formado entre os ácidos nucleicos anti-sentido e sentido, *e.g.*, derivados de fosforotioato e nucleotídeos de acridina substituídos podem ser usados. Exemplos de nucleotídeos modificados que podem ser usados para gerar o ácido nucleico anti-sentido são bem conhecidos na técnica.

Outras modificações de nucleotídeos conhecidas incluem metilação, ciclização e 'caps' e substituição de um ou mais dos nucleotídeos naturalmente ocorrentes por um análogo tal como inosina. Outras modificações de nucleotídeos são bem conhecidas por uma pessoa versada na técnica.

Alternativamente, o ácido nucleico anti-sentido pode ser produzido biologicamente usando um vetor de expressão no qual um ácido nucleico foi subclonado em uma orientação anti-sentido (*i.e.*, RNA transcrito do ácido nucleico inserido será de uma orientação anti-sentido para um ácido nucleico alvo de interesse, descrito posteriormente na subseção seguinte). Preferivelmente, produção de ácidos nucleicos anti-sentido em plantas ocorre por meio de um transgene estavelmente integrado compreendendo um promotor operativo para expressão preferencial em plantas com tecido de endosperma, um oligonucleotídeo anti-sentido, e um terminador.

Um método preferido para redução ou eliminação substancial de expressão de gene CKI endógeno através de silenciamento de RNA é usando um vetor de expressão no qual um gene CKI ou fragmento deste foi clonado como uma repetição invertida (em parte ou completamente), separado por um espaçador (DNA não codificante). Depois de transcrição da repetição invertida, um RNA quimérico de CKI com uma estrutura autocomplementar é formado (parcial ou completo). Esta estrutura de RNA dupla fita é referida como o RNA de estrutura em grampo (hpRNA). O hpRNA é processado pela planta em siRNAs que são incorporados em um RISC. O RISC posteriormente cliva o mRNA de um gene CKI alvo, com isso reduzindo ou substancialmente eliminando o número de mRNAs de CKI a serem traduzidos em uma proteína CKI. Veja por exemplo, Grierson *et al.* (1998) WO 98/53083; Waterhouse *et al.* (1999) WO 99/53050).

As moléculas de ácidos nucleicos usadas para silenciamento nos métodos da invenção (quer introduzidas em uma planta ou geradas *in situ*) hibridizam com ou se ligam a mRNA celular e/ou DNA genômico codificando uma proteína CKI para com isso inibir expressão da proteína, *e.g.*, inibindo transcrição e/ou tradução. A hibridização pode ser por complementaridade convencional de nucleotídeos para formar um duplex estável, ou, por exemplo, no caso de uma molécula de ácido nucleico anti-

sentido que se liga a duplexes de DNA, através de interações específicas na principal fenda da dupla hélice. Moléculas de ácido nucleico anti-sentido podem ser introduzidas em uma planta por transformação ou injeção direta em um sítio de tecido específico. Alternativamente, moléculas de ácido nucleico anti-sentido podem ser modificadas para direcionar células selecionadas e então sistematicamente administradas. Por exemplo, para administração sistêmica, moléculas anti-sentido podem ser modificadas de forma que elas se liguem especificamente a receptores ou antígenos expressos em uma superfície celular selecionada, *e.g.*, ligando as moléculas de ácido nucleico anti-sentido a peptídeos ou anticorpos que se ligam a receptores ou antígenos de superfície celular. As moléculas de ácido nucleico anti-sentido também podem ser liberadas para células usando os vetores descritos neste lugar.

De acordo com um aspecto adicional, o ácido nucleico anti-sentido é uma molécula de ácido nucleico a-anomérica. Uma molécula de ácido nucleico a-anomérica forma híbridos dupla fita específicos com RNA complementar nos quais, ao contrário das unidades b usuais, as fitas correm paralelas uma a outra (Gaultier *et al.* (1987) *Nucleic Acids. Res.* 15:6625-6641). A molécula de ácido nucleico anti-sentido também pode compreender um 2'-o-metilribonucleotídeo (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) ou um análogo quimérico de RNA-DNA (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

Em ainda outra forma de realização, um ácido nucleico anti-sentido da invenção é uma ribozima. Ribozimas são moléculas catalíticas de RNA com atividade de ribonuclease que são capazes de clivar um ácido nucleico de fita simples, tal como um mRNA, ao qual eles têm uma região complementar. Assim, ribozimas (*e.g.*, ribozimas cabeça-de-martelo (descritas em Haselhoff and Gerlach (1988) *Nature* 334:585-591)) podem ser usadas para clivar cataliticamente transcritos de mRNA de CKI para com isso

inibir tradução de mRNA de CKI. Uma ribozima tendo especificidade para um ácido nucleico codificando CKI pode ser projetada baseado na seqüência de nucleotídeos de um cDNA de CKI. Por exemplo, pode ser construído um derivado de um RNA de *Tetrahymena* L-19 IVS no qual a seqüência de nucleotídeos do sítio ativo é complementar à seqüência de nucleotídeos a ser clivada em um mRNA codificando CKI. Veja, *e.g.*, Cech *et al.* Patente Norte-Americana No. 4.987.071; e Cech *et al.* Patente Norte-Americana No. 5.116.742. Alternativamente, mRNA de CKI pode ser usado para selecionar um RNA catalítico tendo uma atividade específica de ribonuclease a partir de um conjunto de moléculas de RNA. Veja, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418. O uso de ribozimas para silenciamento de genes em plantas é conhecido na técnica (*e.g.*, 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).

Silenciamento de genes também pode ser atingido por mutagênese por inserção (por exemplo, inserção de T-DNA ou inserção de transposon) ou por estratégias de silenciamento de genes como descrito por, entre outros, Angell and Baulcombe 1998 (Amplicon VIGS WO 98/36083); Baulcombe (WO 99/15682).

Silenciamento de genes também pode ocorrer se existe uma mutação no gene CKI endógeno e/ou uma mutação em um gene CKI isolado subseqüentemente introduzido em uma planta. A redução ou eliminação substancial de expressão de CKI pode ser causada por um CKI não funcional. CKI se liga tanto a CDK como ciclinas (Verkest *et al.*, (2005) Plant Cell 17: 1723–1736). Por exemplo, mutação do sítio de ligação de ciclina dentro de um CKI, sustenta um CKI que ainda pode se ligar a um CDK mas que não pode inibir o complexo ativo de CDK-ciclina.

Uma abordagem adicional para silenciamento de genes é

direcionando seqüências de nucleotídeos complementares à região reguladora do CKI (*e.g.*, o promotor e/ou estimuladores de CKI) para formar estruturas de tripla hélice que previnem transcrição do gene CKI em células alvo. Veja Helene, C. (1991) *Anticancer Drug Des.* 6(6): 569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12): 807-15.

São descritos acima exemplos de vários métodos para silenciamento de genes (para a redução ou eliminação substancial de expressão de gene CKI endógeno). Os métodos da invenção se baseiam na redução preferencial de expressão de um gene CKI endógeno em tecido de endosperma de uma planta. Uma pessoa versada na técnica deve ser prontamente capaz de adaptar os métodos mencionados acima para silenciamento de forma a atingir silenciamento preferencial de genes em tecido de endosperma, através do uso de um promotor apropriado, por exemplo.

Deve ser observado que a essência da presente invenção reside nos resultados vantajosos e surpreendentes encontrados mediante redução ou eliminação substancial de expressão de gene CKI endógeno em tecido de endosperma de uma planta, e não é limitada a qualquer método particular para tal redução ou eliminação substancial de expressão de gene CKI endógeno. Outros tais métodos serão conhecidos pelo homem versado.

Para desempenho ótimo, as técnicas de silenciamento de genes usadas para a redução ou eliminação substancial de expressão de gene CKI endógeno requerem o uso de seqüências de ácidos nucleicos CKI de plantas monocotiledôneas para transformação em plantas monocotiledôneas. Preferivelmente, um ácido nucleico de CKI de qualquer espécie vegetal dada é introduzido naquela mesma espécie. Por exemplo, um ácido nucleico de CKI de arroz (seja ele uma seqüência CKI de comprimento completo ou um fragmento) é transformado em uma planta de arroz. O ácido nucleico CKI não

precisa ser introduzido na mesma variedade de planta.

Referência neste lugar a um "gene CKI" ou um ácido nucleico de CKI" é adotada para significar uma forma polimérica de um desoxiribonucleotídeo ou um polímero de ribonucleotídeo de qualquer comprimento, ou de dupla fita ou fita simples, ou análogos destes, que têm as características essenciais de um ribonucleotídeo natural com as quais eles podem hibridizar com ácidos nucleicos de uma maneira similar a polinucleotídeos naturalmente ocorrentes. Um "gene CKI" ou um ácido nucleico de CKI" se refere a um comprimento suficiente de nucleotídeos substancialmente contíguos de um gene codificando CKI para realizar silenciamento de genes; isto pode ser tão pouco quanto 20 ou menos nucleotídeos. Um gene codificando uma proteína (funcional) não é um requerimento para os vários métodos discutidos acima para a redução ou eliminação substancial de expressão de um gene CKI endógeno.

Os métodos da invenção podem ser realizados usando um comprimento suficiente de nucleotídeos substancialmente contíguos de um gene/ácido nucleico CKI, o que pode consistir de 20 ou menos nucleotídeos, o que pode ser de qualquer parte do gene/ácido nucleico CKI, tal como a extremidade 3' da região de codificação que é bem conservada dentro da família de genes CKI.

Genes CKI são bem conhecidos na técnica e úteis nos métodos da invenção são nucleotídeos substancialmente contíguos de qualquer dos genes/ácidos nucleicos CKI vegetais descritos em pedido de patente Internacional publicado WO 2005/007829 no nome de Monsanto Technology LLC e Pedidos de patente Internacional publicados, WO 02/28893 e WO 99/14331 no nome de CropDesign N.V, cujos genes/seqüências de nucleotídeos CKI são incorporados neste lugar como se completamente expostos.

Outros genes/seqüências de ácidos nucleicos CKI também

podem ser usados nos métodos da invenção, e podem ser prontamente identificados por uma pessoa versada na técnica. Polipeptídeos CKI podem ser identificados pela presença de um ou mais de diversas características bem conhecidas (veja abaixo). Mediante identificação de um polipeptídeo CKI, uma pessoa versada na técnica pode facilmente derivar, usando técnicas de rotina, a seqüência de ácidos nucleicos de codificação correspondente e usar um comprimento suficiente de nucleotídeos contíguos da mesma para realizar qualquer um ou mais dos métodos de silenciamento de genes descritos acima (para a redução ou eliminação substancial de expressão de um gene CKI endógeno, no endosperma).

Uma característica diferenciada de um polipeptídeo CKI é uma região C-terminal compreendendo entre cerca de 40 e cerca de 55 aminoácidos altamente conservados. Como uma orientação, polipeptídeos compreendendo em ordem crescente de preferência pelo menos 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%, 99% de identidade com a região C-terminal de um CKI como representado por SEQ ID NO: 262 podem ser considerados como sendo homólogos de CKI. Uma pessoa versada na técnica pode facilmente derivar o ácido nucleico correspondente codificando tais homólogos, e usar um comprimento suficiente de nucleotídeos contíguos do mesmo para realizar qualquer um ou mais dos métodos de silenciamento de genes descritos acima (para a redução ou eliminação substancial de expressão de um gene CKI endógeno).

Uma pessoa versada na técnica estará ciente do que é pretendido por um "C-terminal" de uma proteína; para os propósitos deste pedido, a região C-terminal de um CKI pode ser considerada como sendo a segunda metade (de N-terminal para C-terminal) de um polipeptídeo CKI de

comprimento completo.

Homólogos, como definido acima, i.e. polipeptídeos compreendendo pelo menos 50% de identidade com a região C-terminal de um CKI como representado por SEQ ID NO: 262, podem ser prontamente identificados usando técnicas de rotina bem conhecidas na técnica, tal como por alinhamento de seqüências. 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 de duas seqüências completas que maximiza o número de identidades e minimiza o número de lacunas. O algoritmo BLAST (Altschul *et al.* (1990) J Mol Biol 215: 403-10) calcula identidade de seqüência percentual e realiza uma análise estatística da similaridade entre as duas seqüências. O aplicativo computacional para realizar análise BLAST está publicamente disponível através de National Centre for Biotechnology Information. Seqüências homólogas podem ser prontamente identificadas usando, por exemplo, o algoritmo de alinhamento de múltiplas seqüências ClustalW (versão 1.83) disponível em http://clustalw.genome.jp/sit-bin/nph-ClustalW, com os parâmetros padrão de alinhamento pareado, e um método de pontuação em porcentagem. Edição manual secundária pode ser realizada para otimizar alinhamento entre motivos conservados (veja abaixo), como deve ser perceptível para uma pessoa versada na técnica.

Polipeptídeos CKI vegetais também podem ser identificados pela presença de certos motivos conservados (veja Tabela 12 abaixo). A presença destes motivos conservados pode ser identificada usando métodos para o alinhamento de seqüências para comparação como descrito acima. Em algumas instâncias, os parâmetros padrão podem ser ajustados para modificar a estringência da busca. Por exemplo usando BLAST, o limiar de significância estatística (chamado valor "esperado") para relatar identidades

contra seqüências de banco de dados pode ser aumentado para mostrar identidades menos estringentes. Desta maneira, identidades curtas praticamente exatas podem ser identificadas. Mediante identificação de um polipeptídeo CKI pela presença destes motivos, uma pessoa versada na técnica pode facilmente derivar o ácido nucleico correspondente codificando o polipeptídeo compreendendo os motivos relevantes, e usar um comprimento suficiente de nucleotídeos contíguos dos mesmos para realizar qualquer um ou mais dos métodos de silenciamento de genes descritos acima (para a redução ou eliminação substancial de expressão de um gene CKI endógeno).

Tipicamente, a presença de pelo menos um dos motivos 1 a 5 (por exemplo motivo 2 é particularmente bem conservado) deve ser suficiente para identificar qualquer seqüência de consulta como um CKI, entretanto para certeza aumentada, a presença de pelo menos Motivos 1, 2 e 3 é preferida. A seqüência consenso fornecida é baseada nas seqüências apresentadas em Tabela 12 abaixo. Uma pessoa versada na técnica deve estar bem ciente de que a seqüência consenso pode variar de alguma maneira se seqüências adicionais ou diferentes foram usadas para comparação.

Motivo 1: FXXKYNFD (SEQ ID NO: 261), caracterizado pelo fato de que X é qualquer aminoácido

Motivo 2: [P/L]LXGRYEW (SEQ ID NO: 262), caracterizado pelo fato de que X é qualquer aminoácido e [P/L] significa que ou uma prolina ou uma leucina aparecem na posição indicada

Motivo 3: EXE[D/E]FFXXXE (SEQ ID NO: 263), caracterizado pelo fato de que X é qualquer aminoácido e [D/E] significa que ou um aspartato ou um glutamato aparecem na posição indicada

Motivo 4: YXQLRSRR (SEQ ID NO: 264), caracterizado pelo fato de que X é qualquer aminoácido

Motivo 5: MGKY[M/I][K/R]KX[K/R] (SEQ ID NO: 265), caracterizado pelo fato de que X é qualquer aminoácido, [M/I] significa que

ou uma metionina ou uma isoleucina aparecem na posição indicada, e [K/R] significa que ou uma lisina ou uma arginina aparecem na posição indicada

Motivo 6: SXGVRTRA (SEQ ID NO: 266), caracterizado pelo fato de que X é qualquer aminoácido

Motivos 1, 2, e 3 são tipicamente encontrados na região carboxil-terminal de proteínas CKI vegetais. Acredita-se que esta região esteja envolvida na interação de CKIs tanto com CDKs como ciclinas (Chen *et al.* (1996) *Mol.Cell Biol* 16, 4673-4682, Matsuoka *et al.* (1995) *Genes Dev.* 9, 650-662, e Nakayama and Nakayama (1998) *Bioessays* 20, 1020-1029). Motivos 4, 5, e 6 são tipicamente encontrados na região amino-terminal de proteínas CKI vegetais.

Proteínas CKI de plantas monocotiledôneas, particularmente arroz, são caracterizadas por extensivos trechos α-helicoidais especialmente entre motivos 5 e 6 e entre motivos 6 e 4.

Tabela 12. Motivos conservados em proteínas CKI vegetais. CKI1 a CKI7 denotam CKIs de *Arabidopsis thaliana*. Os: *Oryza sativa*, Zm: *Zea mays*, Sb: *Sorghum bicolor*

	Motivo 1	**Motivo 2**	**Motivo 3**	**Motivo 4**	**Motivo 5**	**Motivo 6**
Aalfalfa CKI	198-FMEKYNFD	211-PLPGRYET	182-EFEEFCAKHE	74 - YLQLRNRR	1-MGKYMKKLK	45-SDGVRTRA
CKI1 AC003040	167-FKKKYNFD	180-PLEGRYEW	151-EIEDFFVEAE	20 - YMQLRSRR		
CKI2 AL132979	183-CSMKYNFD	197-LGGGRYEW	164-ELEDFFQVAE			
CKI3 AB012242	197-FMEKYNFD	210-PLSGRYEW	181-EMEEFFAYAE	58 - YLQLRSRR	1-MGKYMKKSK	26-SPGVRTRA
CKI4 AC003974	264-FIEKYNFD	277-PLPGRFEW	248-EMDEFFSGAE	102-YLQLRSRR	1-MGKYIRKSK	44-SLGVLTRA
CKI5 AB028609	164-FIQKYNFD	177-PLPGRYEW	148-EIEDFFASAE	54-YLQLRSRR	1-MGKYIKKSK	24-ALGFRTRA
CKI6 AP000419	173-FIEKYNFD	186-PLEGRYKW	155-EIEDLFSELE			
CKI7 AC011807	170-FTEKYNYD	183-PLEGRYQW	154-ELDDFFSAAE			
Chenopodium CKI AJ002173	171-FSEKYNFD	184-PLKGRYDW	155-EIEEFFAVAE	25-IPQLRSRR		
OsCKI2	233-FAAKYNFD	247-LDAGRFEW	217-EIEAFFAAAE	75-YLQLRSRM	1-MGKYMRKFR	24-VVGVRTRS
OsCKI1	----YNYD	PLQGRYEW				
OsCKI3	FAEKY---		EIEAFFAAAE			
OsCKI4	170-FIDKYNFD	183-PLPGRFEW	154-ELEAFFAAEE	48-YLELRSRR	1-MGKYMRKAK	28-PLGVRTRA
OsCKI5	196-FAAKYNFD	209-PLDAGGAGRFEW	180-EIEEFLAAAE	63-YLRLRSRR	1-MGKKKKRDG	20-VGGVRTRA
ZmCKI1	FASKYNFD	LDAGRFEW	EIQEFFAAAE			
ZmCKI2	FIDKYNFD	PLPGRFEW	EMNEYFAAEQ			

SbCKI	FAEAYNYD	PLEGRFEW	EIEAFFAAAE			
CONSENSUS	FX_2KYNFD	[P/L]LXG R[Y/F]EW	EXE[D/E] FFX_3E	YXQLRSRR	MGKY[M/I][K/R]KX [K/R]	SXGVRTRA

Em adição às características mencionadas acima, uma proteína CKI também pode compreender qualquer um ou mais dos seguintes: uma caixa Cy, uma seqüência de localização nuclear e uma seqüência PEST.

O termo "Caixa Cy" se refere a uma seqüência de aminoácidos de cerca de 5 resíduos de aminoácidos de comprimento tendo a seqüência consenso RXHuF, caracterizada pelo fato de que X é qualquer aminoácido e Hu é um aminoácido não carregado hidrofóbico, tal como M, I, L ou V. Caixas Cy tipicamente estão envolvidas na interação de CKIs com ciclinas.

A "seqüência de localização nuclear" se refere a uma seqüência de aminoácidos de cerca de 4-20 resíduos de aminoácidos de comprimento, que serve para dirigir uma proteína para o núcleo. Tipicamente, a seqüência de localização nuclear é rica em aminoácidos básicos, tal como arginina (R) e lisina (K). Sinais de localização nuclear são descritos em, por exemplo, Gorlich D. (1998) EMBO 5.17:2721-7. A proteína Os CKI4 compreende múltiplas seqüências de localização nuclear.

Uma "seqüência PEST" se refere a uma seqüência de aminoácidos que é enriquecida nos resíduos de aminoácidos prolina (P), glutamato (E), serina (S) e treonina (T) e que está presente em proteínas com uma alta taxa de renovação proteolítica. Seqüências PEST são descritas em, por exemplo, Rogers *et al.* (1986) *Science* 234, 364-368.

Os vários domínios estruturais em uma proteína CKI podem ser identificados usando bancos de dados especializados e.g. SMART (Schultz *et al.* (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic *et al.* (2002) Nucleic Acids Res 30, 242-244; http://smart.embl-heidelberg.de/), InterPro (Mulder *et al.*, (2003) Nucl. Acids. Res. 31, 315-318; http://www.ebi.ac.uk/interpro/), Prosite (Bucher and Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (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), http://www.expasy.org/prosite/) ou Pfam (Bateman *et al.*, Nucleic Acids Research 30(1): 276-280 (2002), http://www.sanger.ac.uk/Software/Pfam/).

Além disso, uma proteína CKI também pode ser identificável por sua capacidade de inibir a atividade de uma Quinase Dependente de Ciclina (CDK), *e.g.*, uma CDK vegetal. CDKs são um grupo de quinases de serina/treonina que regulam a progressão do ciclo celular em eucariotos, *e.g.*, plantas. CDKs tipicamente são complexadas com ciclinas formando um complexo enzimático, CDK sendo a subunidade catalítica e ciclina sendo a subunidade reguladora do complexo enzimático (Wang, H. (1997) *The Plant Journal* 15(4): 501-510).

Portanto mediante identificação de um polipeptídeo CKI usando uma ou diversas das características descritas acima, uma pessoa versada na técnica pode facilmente derivar o ácido nucleico correspondente codificando o polipeptídeo, e usar um comprimento suficiente de nucleotídeos substancialmente contíguos do mesmo para realizar qualquer um ou mais dos métodos de silenciamento de genes descritos acima (para a redução ou eliminação substancial de expressão de um gene CKI endógeno).

Preferido para uso nos métodos da invenção é um comprimento suficiente de nucleotídeos substancialmente contíguos de SEQ ID NO: 267 (OsCKI4), ou o uso de um comprimento suficiente de nucleotídeos substancialmente contíguos de uma seqüência de ácidos nucleicos codificando um ortólogo ou parálogo de OsCKI4 (SEQ ID NO: 267). Exemplos de tais ortólogos e parálogos de OsCKI4 são fornecidos em Tabela 13 abaixo.

Ortólogos e parálogos são homólogos que abrangem conceitos

evolutivos usados para descrever relações ancestrais de genes. Parálogos são genes dentro da mesma espécie que se originaram através de duplicação de um gene ancestral e ortólogos são genes de organismos diferentes que se originaram através de especiação.

Ortólogos em, por exemplo, espécies de plantas monocotiledôneas podem ser facilmente encontrados realizando uma assim chamada busca blast recíproca. Isto pode ser feito por um primeiro blast envolvendo blasting uma seqüência de consulta (por exemplo, SEQ ID NO: 267 ou SEQ ID NO: 268) contra qualquer banco de dados de seqüências, tal como o banco de dados NCBI publicamente disponível que pode ser encontrado em: http://www.ncbi.nlm.nih.gov. BLASTN ou TBLASTX (usando valores pré-determinados padrão) podem ser usados quando iniciando a partir de uma seqüência de nucleotídeos e BLASTP ou TBLASTN (usando valores pré-determinados padrão) podem ser usados quando iniciando a partir de uma seqüência de proteínas. Os resultados de BLAST opcionalmente podem ser filtrados. As seqüências de comprimento completo ou dos resultados filtrados ou resultados não filtrados são então BLASTed de volta (segundo BLAST) contra seqüências do organismo do qual a seqüência de consulta é derivada (onde a seqüência de consulta é SEQ ID NO: 267 ou SEQ ID NO: 268 o segundo blast portanto deve ser contra seqüências de arroz). Os resultados do primeiro e segundo BLASTs são então comparados. Um parálogo é identificado se um acerto de alto ranqueamento do segundo blast é da mesma espécie de qual a seqüência de consulta é derivada; um ortólogo é identificado se um acerto de alto ranqueamento não é da mesma espécie de qual a seqüência de consulta é derivada. Acertos de alto ranqueamento são aqueles tendo um baixo valor E. Quanto menor o valor E, mais significante é o escore (ou em outras palavras menor a chance de que o acerto tenha sido encontrado ao acaso). Computação do valor E é bem conhecida na técnica. No caso de famílias grandes, ClustalW pode ser usado, seguido por uma árvore

de agrupamento de vizinhos, para ajudar a visualizar agrupamento de genes relacionados e para identificar ortólogos e parálogos.

Tabela 13 Ortólogos e parálogos de OsCKI4 (SEQ ID NO: 267 e 268)

Nome	NCBI	Seqüência de ácidos nucleicos SEQ ID	Seqüência de polipeptídeos SEQ ID	Fonte
Zeama_tipo CKI4	AY986792	269	270	*Zea mays*
Triae_tipo CKI4	Agrupamento de seqüências de BG908519.1 e CA640135.1	271	272	*Triticum aestivum*
Orysa_CKI3	AK064723.1	273	274	*Oryza sativa*
Zeama_CKI3 like	DV174570,1	275	276	*Zea mays*
Sorbi_tipo CKI3	contig de CN152732.1 e CD224882.1	277	278	*Sorghum bicolor*
Sacof_tipo CKI4	CO373621.1	279	280	*Saccharum officinarum*

A fonte dos nucleotídeos substancialmente contíguos de um gene/ácido nucleico CKI pode ser qualquer fonte vegetal ou fonte artificial. Para desempenho ótimo, as técnicas de silenciamento de genes usadas para a redução ou eliminação substancial de expressão de gene CKI endógeno requerem o uso de seqüências CKI de plantas monocotiledôneas para transformação em plantas monocotiledôneas. Preferivelmente, seqüências CKI da família Poaceae são transformadas em plantas da família Poaceae. Ainda preferivelmente, um ácido nucleico de CKI de arroz (seja ele uma seqüência CKI de comprimento completo ou um fragmento) é transformado em uma planta de arroz. O ácido nucleico CKI não precisa ser introduzido na mesma variedade de planta. Mais preferivelmente, o ácido nucleico CKI de arroz é um comprimento suficiente de nucleotídeos substancialmente contíguos de SEQ ID NO: 267 (OsCKI4) ou um comprimento suficiente de nucleotídeos substancialmente contíguos de uma seqüência de ácidos nucleicos codificando um ortólogo ou parálogo de OsCKI4 (SEQ ID NO:

267). Como mencionado acima, uma pessoa versada na técnica deve estar bem ciente do que deve constituir um comprimento suficiente de nucleotídeos substancialmente contíguos para realizar qualquer dos métodos de silenciamento de genes definidos acima, isto pode ser tão pouco quanto 20 ou menos nucleotídeos substancialmente contíguos em alguns casos.

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

Portanto, é fornecida uma construção gênica compreendendo uma ou mais seqüências de controle capaz de dirigir preferencialmente expressão de uma seqüência de ácidos nucleicos CKI sentido e/ou anti-sentido em tecido de endosperma de planta de forma a silenciar um gene CKI endógeno em tecido de endosperma de uma planta; e opcionalmente uma seqüência de término de transcrição.

Uma construção preferida para silenciamento de genes é uma compreendendo uma repetição invertida de um gene CKI ou fragmento deste, preferivelmente capaz de formar uma estrutura em grampo, cuja repetição invertida está sob o controle de um promotor endosperma específico.

Construções úteis nos métodos de acordo com a presente invenção podem ser construídas usando tecnologia de DNA recombinante bem conhecida por pessoas versadas na técnica. As construções gênicas podem ser inseridas em vetores, que podem estar comercialmente disponíveis, adequados para transformar em plantas e adequados para expressão do gene de interesse nas células transformadas. A invenção portanto fornece uso de uma construção gênica como definido acima nos métodos da invenção.

A seqüência de interesse é operavelmente ligada a uma ou mais seqüências de controle (pelo menos a um promotor) capaz de aumentar preferencialmente expressão em tecido de endosperma de uma planta. Os termos "elemento regulador", "seqüência de controle" e "promotor" são todos

usados permutavelmente neste lugar e são definidos na seção "Definições" neste lugar.

Um promotor endosperma específico se refere a qualquer promotor capaz de dirigir preferencialmente expressão do gene de interesse em tecido de endosperma. Referência neste lugar a dirigir "preferencialmente" expressão em tecido de endosperma é adotada para significar dirigir expressão de qualquer seqüência operavelmente ligada a este em tecido de endosperma substancialmente para a exclusão de dirigir expressão em qualquer outro lugar na planta, sem considerar qualquer expressão residual devido à expressão de promotor permeável. Por exemplo, o promotor de prolamina mostra forte expressão no endosperma, com permeabilidade em meristema, mais especificamente o meristema de broto e/ou centro de discriminação no meristema.

Preferivelmente, o promotor endosperma específico é um promotor isolado de um gene de prolamina, tal como um promotor de prolamina RP6 de arroz (Wen *et al.*, (1993) Plant physiol 101(3): 1115-6) como representado por SEQ ID NO: 281 ou um promotor de força similar e/ou um promotor com um padrão de expressão similar que o promotor de prolamina de arroz. Força similar e/ou padrão de expressão similar podem ser analisados, por exemplo, acoplando os promotores a um gene repórter e verificando a função do gene repórter em tecidos da planta. Um gene repórter bem conhecido é beta-glucuronidase e o corante colorimétrico GUS usado para visualizar atividade de beta-glucuronidase em tecido vegetal. Exemplos de outros promotores endosperma específicos que também podem ser usados para realizar os métodos da invenção são mostrados em Tabela 6 na seção "Definições" neste lugar.

Opcionalmente, uma ou mais seqüências terminadoras também podem ser usadas na construção introduzida em uma planta. O termo "terminador" é como definido neste lugar na seção "Definições".

As construções genéticas da invenção podem incluir adicionalmente uma seqüência de origem de replicação que é requerida para manutenção e/ou replicação em um tipo celular específico. Um exemplo é quando uma construção genética é requerida ser mantida em uma célula bacteriana como um elemento genético episssomal (e.g. molécula de plasmídeo ou cosmídeo). Origens de replicação preferidas incluem, mas não são limitadas a, a f1-ori e colE1.

A construção genética pode compreender opcionalmente um gene marcador selecionável como definido neste lugar na seção "Definições".

A presente invenção também abrange plantas incluindo partes vegetais obteníveis pelos métodos de acordo com a presente invenção tendo produção aumentada de sementes em relação a plantas de controle adequadas e que têm expressão reduzida ou substancialmente eliminada de um gene CKI endógeno em tecido de endosperma de planta.

A invenção também fornece um método para a produção de plantas transgênicas tendo produção aumentada de sementes em relação a plantas de controle adequadas, cujas plantas transgênicas têm expressão reduzida ou substancialmente eliminada de um gene CKI endógeno em tecido de endosperma de planta.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo produção aumentada de sementes cujo método compreende:

(i) introduzir e expressar em uma planta, parte vegetal ou célula vegetal a construção gênica compreendendo uma ou mais seqüências de controle capaz de dirigir preferencialmente expressão de uma seqüência de ácidos nucleicos CKI sentido e/ou anti-sentido em tecido de endosperma de planta de forma a silenciar um gene CKI endógeno em tecido de endosperma de uma planta; e

(ii) cultivar a planta, parte vegetal ou célula vegetal em

condições promovendo crescimento e desenvolvimento vegetal.

Preferivelmente, a construção introduzida em uma planta é uma compreendendo uma repetição invertida (em parte ou completa) de um gene CKI ou fragmento deste, preferivelmente capaz de formar uma estrutura em grampo.

De acordo com uma característica preferida da presente invenção, a construção é introduzida em uma planta por transformação.

O termo "transformação" é como definido na seção "Definições" neste lugar.

A presente invenção claramente se estende a qualquer célula vegetal ou planta produzida por qualquer dos métodos descritos neste lugar, e a todas partes vegetais e propágulos destas. A presente invenção adicionalmente de estende para abranger a progênie de uma célula primária transformada ou transfectada, tecido, órgão ou planta inteira que foi produzida por qualquer of os métodos mencionados acima, o único requerimento sendo que a progênie exiba a(s) mesma(s) característica(s) genotípica(s) e/ou fenotípica(s) que aquelas produzidas pelo ancestral nos métodos de acordo com a invenção.

A invenção também se estende a partes aptas à colheita de uma planta tal como sementes e produtos derivados, preferivelmente diretamente derivados, de uma parte apta à colheita de uma tal planta, tal como pelotas secas ou pós, óleo, gordura e ácidos graxos, amido ou proteínas.

A presente invenção também abrange uso de ácidos nucleicos CKI para a redução ou eliminação substancial de expressão de gene CKI endógeno em tecido de endosperma de planta para aumentar produção de sementes de planta como definido acima.

Descrição de figuras

A presente invenção será agora descrita co referência às seguintes figuras nas quais:

Fig. 1 dá uma visão geral dos motivos conservados presentes em SEQ ID NO: 2. Domínio rico em leucina é sublinhado, os motivos conservados 1, 2 e 3 são indicados em negrito e a seqüência em itálico representa o provável sítio de N-glicosilação com o provável sítio de fosforilação de proteína quinase C.

Fig. 2 mostra um alinhamento múltiplo de várias proteínas SYR. Os asteriscos indicam resíduos de aminoácidos idênticos, os dois pontos representam substituições altamente conservadas e os pontos representam substituições menos conservativas. Com a informação de Figura 1, os vários domínios e motivos conservados em SEQ ID NO: 2 podem ser facilmente identificados nas outras proteínas SYR.

Fig. 3 mostra vetores binários para transformação e expressão em *Oryza sativa* de um ácido nucleico SYR de *Oryza sativa*. Em pGOS2::SYR, a seqüência de codificação de *SYR* está sob o controle de um promotor GOS2 de arroz.

Fig. 4 mostra vetores binários para transformação e expressão em *Oryza sativa* de um ácido nucleico SYR de *Oryza sativa*. Em pH MGP::SYR, a seqüência de codificação de *SYR* está sob o controle de um promotor HMGP de arroz (SEQ ID NO: 18 em WO 2004/070039, cuja SEQ ID NO: 18 de WO 2004/070039 é incorporada neste lugar como se completamente exposta),.

Fig. 5 detalha exemplos de seqüências úteis para realizar os métodos de acordo com a presente invenção. SEQ ID NO: 1 e SEQ ID NO: 2 representam a seqüência de nucleotídeos e proteínas de SYR usado nos exemplos. Os códons de início e término em SEQ ID NO: 1 são dados em negrito. SEQ ID NO: 3 e SEQ ID NO: 4 são seQüências de iniciadores usados para isolar o ácido nucleico SYR. SEQ ID NO: 5 é a seqüência do promotor GOS2 e SEQ ID NO: 33 do promotor PRO0170 como usado nos exemplos, SEQ ID NO: 6 a SEQ ID NO: 11 representam seqüências consenso de partes

conservadas nas proteínas SYR. SEQ ID NO: 12 a 25, 27 a 32 e 36 a 42 são seqüências de nucleotídeos (comprimento completo ou parcial) e proteínas de homólogos do gene SYR e proteína como dado em SEQ ID NO: 1 e SEQ ID NO: 2. SEQ ID NO: 26 representa a seqüência de proteína ARGOS (acesso de GenBank AY305869).

Fig. 6 dá uma visão geral de domínios protéicos FG-GAP. A proteína de SEQ ID NO: 46 compreende sinal de secreção (parte N-terminal encaixotada), um domínio FG-GAP começando em P73 e terminando com L98, indicado em negrito e sublinhado, e um domínio transmembrana (negrito e encaixotado). O motivo conservado DXDXDGXX(D/E) é encaixotado e sublinhado, caracterizado pelo fato de que o motivo DGXX(D/E) está em itálico. O domínio conservado FDGYLYLID está sublinhado.

Fig. 7 mostra um alinhamento múltiplo de proteínas FG-GAP de comprimento completo (SEQ ID NO: 46, SEQ ID NO: 55, SEQ ID NO: 57 e SEQ ID NO: 59), os asteriscos indicam aminoácidos idênticos, os dois pontos indicam substituições altamente conservadas e os pontos indicam substituições menos conservativas. As seqüências parciais listadas em Tabela G de Exemplo 12 podem ser úteis em um tal alinhamento múltiplo para a identificação de motivos adicionais.

Fig. 8 mostra um vetor binário para transformação e expressão em *Oryza sativa* de um ácido nucleico codificando *FG-GAP* de *Arabidopsis thaliana* sob o controle de um promotor GOS2 de arroz.

Fig. 9 detalha exemplos de seqüências úteis para realizar os métodos de acordo com a presente invenção. SEQ ID NO: 45 e SEQ ID NO: 46 representam a seqüência de nucleotídeos e proteínas de FG-GAP usados nos exemplos; os códons de início e término em SEQ ID NO: 45 são dados em negrito. SEQ ID NO: 47 e SEQ ID NO: 48 são seqüências de iniciadores usados para isolar o ácido nucleico FG-GAP. SEQ ID NO: 49 é a seqüência da combinação promotor-gene como usada nos exemplos, SEQ ID NO: 50 a

SEQ ID NO: 53 representam seqüências consenso de partes conservadas nas proteínas FG-GAP. SEQ ID NO: 54 a 71 são seqüências de nucleotídeos (comprimento completo ou parcial) e proteínas de homólogos do gene FG-GAP e proteína como dada em SEQ ID NO: 45 e SEQ ID NO: 46. SEQ ID NO: 72 é a seqüência genômica codificando uma proteína FG-GAP de *Medicago sativa* cuja proteína compreende as seqüências de peptídeos representadas por SEQ ID NO: 72 a 76.

Fig. 10 mostra as características importantes encontradas em polipeptídeos CYP90B ou homólogos destes: o domínio hidrofóbico N-terminal, o domínio de transição (com o K/R-K/R-X3-9-P-G-G, os domínios A a D. Dentro do domínio A a seqüência consenso Ala/Gly-Gly-X-Asp/Glu-Thr-Thr/Ser é identificada. A seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser dos polipeptídeos CYP90B compreende esta seqüência consenso Ala/Gly-Gly-X-Asp/Glu-Thr-Thr/Ser.

Fig. 11 mostra a via biossintética ramificada de brassinoesteróide. Em *Arabidopsis*, o polipeptídeo CYP90B1/DWF4 compreende a atividade enzimática de esteróide 22 alfa hidroxilase.

Fig. 12 mostra o perfil de produção de ProtScale para hidrofobicidade do polipeptídeo CYP90B da invenção. Os primeiros 34 aminoácidos N-terminais (encaixotados) representam um domínio hidrofóbico, pois estes estão localizados acima da linha delimitando zero. Esta região corresponde ao domínio âncora N-terminal.

Fig. 13 mostra um alinhamento múltiplo de diversos polipeptídeos CYP90B vegetais, usando o programa de alinhamento múltiplo VNTI AlignX, baseado em um algoritmo de ClustalW modificado (InforMax, Bethesda, MD, http://www.informaxinc.com), com configurações padrão para penalidade de abertura de lacuna de 10 e uma extensão de lacuna de 0,05. O domínio hidrofóbico N-terminal, o domínio de transição (com o K/R-K/R-X_{3-9}-P-G-G) e os domínios A a D estão indicados. A seqüência consenso Phe-

Ala-Gly-His-Glu-Thr-Ser-Ser é encaixotada dentro do domínio A. Os números de acesso dos polipeptídeos CYP90B podem ser encontrados em Tabela 9a e 9b. Os Arath_CYP90A1_CPD (At5g05690), Arath_CYP90C1_ROT3 (At4g36380) e Arath_CYP90D1 (At3g13730) de *Arabidopsis* são mostrados como não polipeptídeos CYP90B.

Fig. 14 mostra um vetor de transformação vegetal para expressão em *Oryza sativa* de um ácido nucleico *CYP90B* de *Oryza sativa* sob o controle de um promotor vegetal, o qual pode ser um promotor não constitutivo (tal como endosperma ou embrião/aleurona específico) ou um promotor constitutivo (tal como GOS2 e HMGB1).

Fig. 15 detalha exemplos de seqüências úteis para realizar os métodos de acordo com a presente invenção. Diversas seqüências resultam de agrupamentos públicos de EST (veja Tabela 9a), com seqüenciamento de menos qualidade. Como uma conseqüência, umas poucas substituições de ácidos nucleicos podem ser esperadas. Os códons de início (ATG) e término delimitam as seqüências de ácidos nucleicos quando estas são de comprimento completo.

Fig. 16 representa uma figura esquemática de um polipeptídeo CDC27 de comprimento completo (mais especificamente o polipeptídeo hobbit CDC27B de *Arabidopsis thaliana*). As repetições de tetratricopeptídeo (TPR) são representadas como caixas pretas. A região NH_2 terminal do polipeptídeo é representada como uma barra preta.

Fig. 17 mostra o alinhamento múltiplo de polipeptídeos CDC27 de diferentes fontes, usando o programa de alinhamento múltiplo VNTI AlignX, baseado em um algoritmo de ClustalW modificado (InforMax, Bethesda, MD, http://www.informaxinc.com), com configurações padrão para penalidade de abertura de lacuna de 10 e uma extensão de lacuna de 0,05. As repetições de tetratricopeptídeo (TPR) são encaixotadas através do alinhamento. O domínio NH_2 conservado PD011373 (como definido em

ProDom, http://ribosome.toulouse.inra.fr/prodom/current/cgi-bin/ProDomBlast3.pl) é duplamente sublinhado.

Fig. 18 mostra um vetor binário pOSH1::CDC27 para expressão em *Oryza sativa* de um ácido nucleico *CDC27* modificado de *Arabidopsis thaliana* sob o controle de um promotor vegetal que é um promotor de meristema apical de broto.

Fig. 19 mostra uma Tabela listando ortólogos e parálogos de CDC27 parciais e de comprimento completo de diferentes fontes, produzidos por TIGR (Institute for Genomic Research at http://www.tigr.org). TC895803 pode ser encontrado em http://www.tigr.org/tigr-scripts/tgi/ego/ego_report.pl?ego=895803.

Fig. 20 detalha exemplos de seqüências úteis para realizar os métodos de acordo com a presente invenção, ou úteis para isolar tais seqüências. Diversas seqüências resultam de assembléias públicas de EST (veja Tabela 10), com seqüenciamento de menos qualidade. Como uma conseqüência, umas poucas substituições de ácidos nucleicos podem ser esperadas. Os códons de início (ATG) e término delimitam as seqüências de ácidos nucleicos quando estes codificam polipeptídeos CDC27 de comprimento completo.

Fig. 21 mostra uma árvore filogenética de várias seqüências de polipeptídeos compreendendo um domínio AT-hook e um domínio DUF296. A árvore filogenética foi feita usando programa de alinhamento múltiplo VNTI AlignX, baseado em um algoritmo de ClustalW modificado (InforMax, Bethesda, MD, http://www.informaxinc.com), com configurações padrão para penalidade de abertura de lacuna de 10 e uma extensão de lacuna de 0,05.

Fig. 22 mostra um vetor binário pPROLAMINA::AT-hook, para expressão em *Oryza sativa* de um ácido nucleico de *Oryza sativa* codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 e Motivo 2 sob o controle de um promotor de prolamina.

Fig. 23 mostra um alinhamento múltiplo de um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296, preparado usando programa de alinhamento múltiplo VNTI AlignX, baseado em um algoritmo de ClustalW modificado (InforMax, Bethesda, MD, http://www.informaxinc.com), com configurações padrão para penalidade de abertura de lacuna de 10 e uma extensão de lacuna de 0,05. É mostrado no alinhamento o domínio AT-hook e o domínio DUF296 e Motivo 2 em ***negrito, itálico e sublinhado***..

Fig. 24 detalha exemplos de seqüências úteis para realizar os métodos de acordo com a presente invenção.

Fig. 25 mostra uma árvore filogenética de fatores de transcrição DOF. A caixa mais perto do topo mostra o principal agrupamento de seqüências compartilhando homologia com SEQ ID NO: 227 (e compreendendo características (i) e (iii) como definido acima, i.e. pelo menos 60% de identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e Motivo I e/ou Motivo II como definido acima). A caixa mais perto do fundo mostra o principal agrupamento de seqüências compartilhando homologia com SEQ ID NO: 199 (e compreendendo características (i) e (ii) como definido acima, i.e. pelo menos 60% de identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e pelo menos 70% de identidade de seqüência com o domínio DOF representado por SEQ ID NO: 200).

Fig. 26 mostra um vetor binário pGOS2::DOF, para expressão em *Oryza sativa* de um fator de transcrição *DOF* de *Arabidopsis thaliana* sob o controle de um promotor GOS2.

Fig. 27 mostra um vetor binário pPROLAMINA::DOF, para expressão em *Oryza sativa* de um fator de transcrição *DOF* de *Arabidopsis thaliana* sob o controle de um promotor de prolamina.

Fig. 28 detalha exemplos de seqüências úteis para realizar os

métodos de acordo com a presente invenção.

Fig. 29 é uma representação esquemática de um polipeptídeo CKI vegetal de comprimento completo. Os motivos típicos 1 a 5 (SEQ ID NO: 261 a SEQ ID NO: 265) úteis para identificar CKIs são encaixotados e numerados por conseguinte (motivo 6 não mostrado).

Fig. 30 mostra uma árvore de agrupamento de vizinhos de um alinhamento múltiplo de polipeptídeos CKI de diferentes fontes, e feito usando o aplicativo computacional público ClustalW disponível em http://clustalw.genome.jp, com as configurações padrão. Um subgrupo de CKI4s de monocotiledôneas e dicotiledôneas é indicado pelo colchete grande. Dentro deste subgrupo, CKIs de monocotiledôneas se agrupam juntos, como indicado pelo colchete médio. O galho de CKI4 de monocotiledôneas é indicado pelo colchete pequeno.

Fig. 31 é um alinhamento múltiplo de polipeptídeos CKI de diferentes fontes vegetais, feito usando programa de alinhamento múltiplo VNTI AlignX, baseado em um algoritmo de ClustalW modificado (InforMax, Bethesda, MD, http://www.informaxinc.com), com configurações padrão para penalidade de abertura de lacuna de 10 e uma extensão de lacuna de 0,05. A extremidade C-terminal conservada de CKIs é encaixotada, assim como motivos 1 a 5 (SEQ ID NO: 261 a SEQ ID NO: 265) úteis para identificar CKIs vegetais (motivo 6 não mostrado).

Fig. 32 mostra um vetor binário para silenciamento de RNA de CKI em *Oryza sativa*, usando uma construção com estrutura em grampo, sob o controle de um promotor endosperma específico e sob o controle de um promotor broto específico.

Fig. 33 detalha exemplos de seqüências úteis para realizar os métodos de acordo com a presente invenção, ou úteis para isolar tais seqüências. Diversas seqüências resultam de assembléias públicas de EST, com seqüenciamento de menos qualidade. Como uma conseqüência, umas

poucas substituições de ácidos nucleicos podem ser esperadas. Os códons de início (ATG) e término delimitam as seqüências de ácidos nucleicos quando estas codificam polipeptídeos CKI de comprimento completo. Entretanto tanto UTR 5' como 3' também podem ser usadas para realizar os métodos da invenção.

Exemplos

A presente invenção será agora descrita com referência aos seguintes exemplos, que estão como forma de ilustração apenas. Os seguintes exemplos não são pretendidos para definir completamente ou de outra maneira limitar o escopo da invenção.

Manipulação de DNA

A menos que de outra maneira determinado, técnicas de DNA recombinante são realizadas de acordo com protocolos padrão descritos em (Sambrook (2001) Molecular Cloning: a laboratory manual, 3ª Edição, 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 (http://www.4ulr.com/products/currentprotocols/index.html). Materiais e métodos padrão para trabalho molecular com plantas 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).

Análise estatística

Uma ANOVA de dois fatores (análises de variância) corrigida para o modelo não equilibrado foi usada como modelo estatístico para a avaliação geral de características fenotípicas vegetais. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformadas com aquele gene. O teste F foi realizado para verificar por um efeito do gene sobre todos os eventos de transformação e para verificar por um efeito geral do gene, também chamado neste lugar "efeito

gênico global". Se o valor do teste F mostra que os dados são significantes, então é concluído que existe um efeito "gênico", significando que não apenas presença ou a posição do gene está causando o efeito. O limiar para significância para um verdadeiro efeito gênico global é ajustado em 5% de nível de probabilidade para o teste F.

Para verificar por um efeito dos genes dentro de um evento, i.e., por um efeito linhagem específico, um teste t foi realizado dentro de cada evento usando conjuntos de dados das plantas transgênicas e das plantas nulas correspondentes. "Plantas nulas" ou "segregantes nulos" ou "nulizigotos" são as plantas tratadas da mesma maneira que a planta transgênica, mas das quais o transgene foi segregado. Plantas nulas também podem ser descritas como as plantas transformadas negativas homozigotas. O limiar para significância para o teste t é ajustado em 10% de nível de probabilidade. Os resultados para alguns eventos podem estar acima ou abaixo deste limiar. Isto é baseado na hipótese de que um gene pode apenas ter um efeito em certas posições no genoma, e que a ocorrência deste efeito dependente de posição não é comum. Este tipo de efeito gênico é também chamado neste lugar um "efeito de linhagem do gene". O valor de p é obtido comparando o valor de t com a distribuição de t ou alternativamente, comparando o valor de F com a distribuição de F. O valor de p então gera a probabilidade de que a hipótese nula (i.e., que não existe nenhum efeito do transgene) esteja correta.

EXEMPLO A: SYR

Exemplo 1: Identificação de seqüências relacionadas a SEQ ID NO: 1 e SEQ ID NO: 2

Seqüências (cDNA de comprimento completo, ESTs ou genômicas) relacionadas a SEQ ID NO: 1 e/ou seqüências de proteínas relacionadas a SEQ ID NO: 2 foram identificadas entre aquelas mantidas no banco de dados Entrez Nucleotides no National Center for Biotechnology Information (NCBI) usando ferramentas de busca de seqüência de banco de

dados, tal como a Ferramenta Básica de Alinhamento Local (BLAST) (Altschul *et al.* (1990) J. Mol. Biol. 215:403-410; e Altschul *et al.* (1997) Nucleic Acids Res. 25:3389-3402). O programa foi usado para encontrar regiões de similaridade local entre seqüências comparando seqüências de ácidos nucleicos ou polipeptídeos com banco de dados de seqüências e calculando a significância estatística de identidades. O polipeptídeo codificado por SEQ ID NO: 1 foi usado para o algoritmo TBLASTN, com configurações padrão e o filtro para ignorar contrapeso de seqüências de baixa complexidade. O resultado da análise foi visto por comparação pareada, e ranqueada de acordo com o escore de probabilidade (valor E), onde o escore reflete a probabilidade de que um alinhamento particular ocorra ao acaso (quanto menor o valor E, mais significante o acerto). Em adição a valores E, comparações também foram pontuadas por identidade de porcentagem. Identidade de porcentagem se refere ao número de nucleotídeos (ou aminoácidos) idênticos entre as duas seqüências de ácidos nucleicos (ou polipeptídeos) comparadas em um comprimento particular. Em algumas instâncias, os parâmetros padrão foram ajustados para modificar a estringência da busca.

Em adição às seqüências de ácidos nucleicos disponíveis publicamente disponíveis em NCBI, outros bancos de dados de seqüências também podem ser buscados seguindo o mesmo procedimento que descrito acima.

Tabela A fornece uma lista de seqüências de ácidos nucleicos e proteínas relacionadas à seqüência de ácidos nucleicos como representada por SEQ ID NO: 1 e à seqüência de proteínas representada por SEQ ID NO: 2.

Tabela A: Seqüências de ácidos nucleicos relacionadas à seqüência de ácidos nucleicos (SEQ ID NO: 1) úteis nos métodos da presente invenção, e os polipeptídeos deduzidos correspondentes.

Nome	Organismo fonte	Ácido nucleico SEQ ID NO:	Polipeptídeo SEQ ID NO:	Número de acesso de banco de dados	Estado
OsSYR	*Oryza sativa*	1	2	/	Comprimento completo ou parcial
homólogo 1 de SYR de arroz	*Oryza sativa*	12	27	XP_472637	Comprimento completo
homólogo 2 de SYR de arroz	*Oryza sativa*	13		AP008218	Comprimento completo
homólogo de SYR de milho	*Zea mays*	14	28	AY110705	parcial
homólogo de SYR de trigo	*Triticum aestivum*	15		/	Comprimento completo
homólogo de SYR de cevada	*Hordeum vulgare*	16	36	CB871444	Comprimento completo
homólogo 1 de SYR de cana-de-açúcar	*Saccharum officinarum*	17	37	CA165713	parcial
homólogo 2 de SYR de cana-de-açúcar	*Saccharum officinarum*	18	38	CA242805	Comprimento completo
homólogo de SYR de sorgo	*Sorghum bicolor*	19	39	CX611532	Comprimento completo
Homólogo 1 de AtSYR	*Arabidopsis thaliana*	20	40	NM_115853	Comprimento completo
Homólogo 2 de AtSYR	*Arabidopsis thaliana*	21	41	NM_180078	Comprimento completo
homólogo de SYR de uva	*Vitis vinifera*	22	29	CF404276	Comprimento completo
Homólogo de SYR de citrus	*Citrus reticulata*	23	30	CF830612	parcial
homólogo 1 de SYR de tomate	*Lycopersicon esculentum*	24	32	AI774560	Comprimento completo
homólogo 2 de SYR de tomate	*Lycopersicon esculentum*	25	31	BG125370	Comprimento completo

Exemplo 2: Alinhamento de seqüências de polipeptídeos relevantes

AlignX do Vetor NTI (Invitrogen) é baseado no popular algoritmo de Clustal de alinhamento progressivo (Thompson *et al.* (1997) Nucleic Acids Res 25:4876-4882; Chenna *et al.* (2003). Nucleic Acids Res 31:3497-3500). Uma árvore filogenética pode ser construída usando um algoritmo de agrupamento de vizinhos. Valores padrão são para a penalidade de abertura de lacuna de 10, para a penalidade de extensão de lacuna de 0,1 e a matriz de peso selecionada é Blosum 62 (se polipeptídeos estão alinhados).

O resultado do alinhamento múltiplo de seqüências usando polipeptídeos relevantes para identificar aqueles úteis para realizar os métodos da invenção é mostrado em Figura 2. A repetição rica em leucina e

os motivos conservados podem ser facilmente discriminados nas várias seqüências.

Exemplo 3: Cálculo de identidade de porcentagem global entre seqüências de polipeptídeos úteis para realizar os métodos da invenção

Porcentagens globais de similaridade e identidade entre seqüências de polipeptídeos de comprimento completo úteis para realizar os métodos da invenção foram determinadas usando um dos métodos disponíveis na técnica, o aplicativo computacional MatGAT (Ferramenta de Alinhamento Global de Matriz) (BMC Bioinformatics. 2003 4:29. MatGAT: an application that generates similarity/identity matrices using protein ou DNA sequences. Campanella JJ, Bitincka L, Smalley J; aplicativo computacional hospedado por Ledion Bitincka). Aplicativo computacional MatGAT gera matrizes de similaridade/identidade para seqüências de DNA ou proteína sem necessitar de pré-alinhamento dos dados. O programa realiza uma série de alinhamentos pareados usando o algoritmo de alinhamento global de Myers e Miller (com uma penalidade de abertura de lacuna de 12, e uma penalidade de extensão de lacuna de 2), calcula similaridade e identidade usando por exemplo Blosum 62 (para polipeptídeos), e então coloca os resultados em uma matriz de distância. Similaridade de seqüências é mostrada na metade de baixo da linha divisora e identidade de seqüências é mostrada na metade de cima da linha divisora diagonal.

Parâmetros usados na comparação foram:

Matriz de pontuação: Blosum62

Primeira Lacuna: 12

Lacuna de extensão: 2

Resultados da análise de aplicativo computacional são mostrados em Tabela B para a similaridade e identidade global no comprimento completo das seqüências de polipeptídeos (excluindo as seqüências de polipeptídeos parciais). Identidade de porcentagem é dada

acima da diagonal e similaridade de porcentagem é dada abaixo da diagonal.

A identidade de porcentagem entre as seqüências de polipeptídeos úteis para realizar os métodos da invenção pode ser tão baixa quanto 27% de identidade de aminoácidos comparada com SEQ ID NO: 2.

Tabela B: Resultados de MatGAT para similaridade e identidade global no comprimento completo das seqüências de polipeptídeos.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. SEQID2		**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. SEQID12	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. SEQID13	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. SEQID14	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. SEQID15	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. SEQID16	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. SEQID17	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. SEQID18	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. SEQID19	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. SEQID20	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. SEQID21	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. SEQID22	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. SEQID23	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. SEQID24	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. SEQID25	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. SEQID34	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. SEQID35	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 4: Previsão de topologia do polipeptídeo seqüências úteis para realizar os métodos da invenção

TargetP 1.1 foi usado para prever a localização subcelular de proteínas eucariotas. De acordo com o programa, a designação de localização é baseada na presença prevista de qualquer das pré-seqüências N-terminais: peptídeo de trânsito de cloroplasto (cTP), peptídeo de direcionamento mitocondrial (mTP) ou peptídeo sinal de via secretora (SP). Escores nos quais a previsão final está baseada não são probabilidades reais, e eles não necessariamente adicionam a um. Entretanto, a localização com o maior escore é a mais provável de acordo com TargetP, e a relação entre os escores (a classe de confiança) pode ser uma indicação de quão certa é a previsão. A classe de confiança (RC) varia de 1 a 5, onde 1 indica a previsão mais forte. TargetP é mantido no servidor da Technical University of Denmark.

Para as seqüências previstas conter uma pré-seqüência N-terminal um sítio de clivagem potencial também pode estar presente.

Diversos parâmetros foram selecionados, tal como grupo de organismo (não vegetal ou vegetal), conjuntos de atalhos (nenhum, conjunto pré-definido de atalhos, ou conjunto de atalhos especificado por usuário), e o cálculo de previsão de sítios de clivagem (sim ou não).

Os resultados de análise TargetP 1.1 da seqüência de polipeptídeos como representada por SEQ ID NO: 2 são apresentados na Tabela C abaixo. O grupo de organismo "vegetal" foi selecionado, nenhum atalho definido, e o comprimento previsto do peptídeo de trânsito requisitado. De acordo com os resultados, a localização subcelular da seqüência de polipeptídeos como representada por SEQ ID NO: 2 pode ser a mitocôndria; entretanto a classe de confiança de 5 (i.e. a classe de confiança mais baixa) deve ser considerada.

Tabela C: Análise TargetP 1.1 da seqüência de polipeptídeos como representada por SEQ ID NO: 2

Comprimento (AA)	105
Peptídeo de trânsito cloroplástico	0,025
Peptídeo de trânsito mitocondriano	0,552
Peptídeo sinal de via secretora	0,009
Outro direcionamento subcelular	0,416
Localização Prevista	mitocôndria
Classe de confiança	5

Dois domínios transmembrana foram identificados pelo programa TMHMM, hospedado no servidor do Center for Biological Sequence Analysis, Technical University of Denmark. Os resultados abaixo mostram que a probabilidade de que o término N esteja localizado dentro é 0,997. Detalhes adicionais na orientação são dados em Tabela D abaixo.

Tabela D: resultados de TMHMM 2.0

Orientação	**Resíduo de começo – fim**
dentro	1 42
Hélice de TM	43 65
fora	66 74
Hélice de TM	75 92
dentro	93 105

Muitos outros algoritmos podem ser usados para realizar tais análises, incluindo:

- ChloroP 1.1 hospedado no servidor da Technical University of Denmark;
- Protein Prowler Subcellular Localisation Predictor versão 1.2 hospedado no servidor do Institute for Molecular Bioscience, University of Queensland, Brisbane, Austrália;
- PENCE Proteome Analyst PA-GOSUB 2.5 hospedado no servidor da University of Alberta, Edmonton, Alberta, Canadá;

Exemplo 5: Clonagem de gene

O gene *SYR* de *Oryza sativa* foi amplificado por PCR usando como modelo uma biblioteca de cDNA de plântula de *Oryza sativa* (Invitrogen, Paisley, UK). Depois de transcrição reversa de RNA extraído de plântulas, os cDNAs foram clonados em pCMV Sport 6.0. Tamanho médio de

inserto do banco foi 1,5 kb e o número original de clones foi da ordem de 1,59 x 10^7 cfu. Título original foi determinado como sendo 9,6 x 10^5 cfu/ml depois de primeira amplificação de 6 x 10^{11} cfu/ml. Depois de extração de plasmídeo, 200 ng de modelo foram usados em uma mistura de PCR de 50 µl. Iniciadores prm08170 (SEQ ID NO: 3; sentido, códon de início em negrito, sítio AttB1 em itálico: 5'-*ggggacaagtttgtacaaaaaagcag gcttaaaca***atg**gaaggtgtaggtgctagg-3') e prm08171 (SEQ ID NO: 4; reverso, complementar, sítio AttB2 em itálico: 5'-*ggggaccactttgtacaagaaagctgggt*caaaaacaaaaataaattcccc-3'), os quais incluem os sítios AttB para recombinação Gateway, foram usados para amplificação por PCR. PCR foi realizado usando DNA polimerase Hifi Taq em condições padrão. Um fragmento de PCR do tamanho correto foi amplificado e purificado também usando métodos padrão. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante a qual o fragmento de PCR se recombina in vivo com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada", pSYR. Plasmídeo pDONR201 foi adquirido a partir de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 6: Construção de Vetor

O clone de entrada pSYR foi subseqüentemente usado em uma reação LR com um vetor de destinação usado para transformação de *Oryza sativa*. Este vetor contém como elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; uma gaveta de expressão de marcador triável; e uma gaveta Gateway pretendida para recombinação in vivo LR com a seqüência de interesse já clonada no clone de entrada. Um promotor GOS2 de arroz (SEQ ID NO: 5) para expressão constitutiva foi localizado antes desta gaveta Gateway. Uma construção de vetor similar foi preparada, mas com o promotor de Proteína de Grupo de Alta Mobilidade (HMGP, SEQ ID NO: 33) ao invés do promotor GOS.

Depois da etapa de recombinação LR, os vetores de expressão resultantes, pGOS2::SYR (com o promotor GOS2) e pH MGP::SYR (com o promotor HMGP), ambos para expressão constitutiva de *SYR* (Figura 2) foram transformados em linhagem LBA4044 de *Agrobacterium* e subseqüentemente para plantas de *Oryza sativa*.

Exemplo 7: Transformação de arroz

A *Agrobacterium* contendo o vetor de expressão foi usada para transformar plantas de *Oryza sativa*. Sementes maduras secas da cultivar japonesa de arroz Nipponbare foram descascadas. Esterilização foi realizada incubando por um minuto em etanol a 70%, seguido por 30 minutos em 0,2% de $HgCl_2$, seguido por uma lavagem de 15 minutos por 6 vezes com água destilada estéril. As sementes estéreis foram então germinadas em um meio contendo 2,4-D (meio de indução de calo). Depois de incubação no escuro por quatro semanas, calos embriogênicos derivados de escutelo foram excisados e propagados no mesmo meio. Depois de duas semanas, os calos foram multiplicados ou propagados por subsafra no mesmo meio por outras 2 semanas. Pedaços de calo embriogênico foram sub-cultivados em meio fresco 3 dias antes de co-cultivo (para estimular atividade de divisão celular).

Linhagem LBA4404 de *Agrobacterium* contendo o vetor de expressão foi usada para co-cultivo. *Agrobacterium* foi inoculada em meio AB com os antibióticos apropriados e cultivada por 3 dias a 28°C. As bactérias foram então coletadas e suspensas em meio de co-cultivo líquido para uma densidade (OD_{600}) de cerca de 1. A suspensão foi então transferida para uma placa de Petri e os calos imersos na suspensão por 15 minutos. Os tecidos de calos foram então secados em um papel filtro e transferidos para um meio de co-cultivo solidificado e incubados por 3 dias no escuro a 25°C. Calos co-cultivados foram crescidos em meio contendo 2,4-D- por 4 semanas no escuro a 28°C na presença de um agente de seleção. Durante este período, ilhas de calos resistentes que crescem rapidamente se desenvolveram. Depois

de transferência deste material para um meio de regeneração e incubação na luz, o potencial embriogênico foi liberado e brotos se desenvolveram nas próximas quatro a cinco semanas. Brotos foram excisados dos calos e incubados por 2 a 3 semanas em um meio contendo auxina a partir do qual eles foram transferidos para o solo. Brotos endurecidos foram crescidos em alta umidade e dias curtos em uma casa de vegetação.

Aproximadamente 35 transformantes T0 de arroz independentes foram gerados para uma construção. Os transformantes primários foram transferidos de uma câmera de safra de tecido para uma casa de vegetação. Depois de uma análise de PCR quantitativo para verificar número de cópias do inserto de T-DNA, apenas plantas transgênicas de cópia única que exibem tolerância ao agente de seleção foram mantidas para colheita de semente T1. Sementes foram então coletadas três a cinco meses depois de transplante. O método produziu transformantes de locus único em uma taxa de até 50% (Aldemita and Hodges1996, Chan *et al.* 1993, Hiei *et al.* 1994).

Para transformação de outras safras veja Exemplo 40.

Exemplo 8: Métodos de avaliação de plantas transformada com SYR sob o controle do promotor GOS2 de arroz ou do promotor HMGP

Configuração de avaliação

Aproximadamente 15 a 20 transformantes T0 de arroz independentes foram gerados. Os transformantes primários foram transferidos de uma câmera de safra de tecido para uma casa de vegetação para crescimento e colheita de semente T1. Oito eventos, dos quais a progênie T1 segregou 3:1 para presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 plântulas T1 contendo o transgene (hetero e homo-zigotos) e aproximadamente 10 plântulas T1 carecendo do transgene (nulizigotos) foram selecionadas monitorando expressão de marcador visual. As plantas T1 selecionadas foram transferidas para uma casa

de vegetação. Cada planta recebeu uma etiqueta de código de barras exclusiva para ligar inequivocamente os dados de fenotipagem à planta correspondente. As plantas T1 selecionadas foram crescidas no solo em potes de 10 cm de diâmetro nas seguintes configurações ambientais: fotoperíodo= 11,5 h, intensidade de luz do dia= 30.000 lux ou mais, temperatura de dia= 28°C ou superior, temperatura de noite= 22°C, umidade relativa= 60-70%. Plantas transgênicas e os nulizigotos correspondentes foram crescidos lado a lado em posições aleatórias. Do estágio de semeadura até o estágio de maturidade as plantas foram passadas diversas vezes através de um armário para formação de imagem digital. Em cada momento imagens digitais (2048x1536 pixéis, 16 milhões de cores) foram tiradas de cada planta de pelo menos 6 ângulos diferentes.

Triagem de estresse salino

Plantas de 4 eventos (sementes T2) foram crescidas em um substrato feito de fibras de coco e argex (proporção de 3 para 1). Uma solução nutritiva normal foi usada durante as primeiras duas semanas depois de transplantar as plântulas na casa de vegetação. Depois das primeiras duas semanas, 25 mM de sal (NaCl) foi adicionado à solução nutritiva, até as plantas terem sido coletadas.

Triagem de seca

Plantas de cinco eventos (sementes T2) foram crescidas em solo de vaso em condições normais até elas atingirem o estágio de espigamento. Elas foram então transferidas para uma seção "seca" onde irrigação foi retida. Sondas de umidade foram inseridas em potes aleatoriamente escolhidos para monitorar o teor de água no solo (SWC). Quando SWC passou abaixo de certos limites, as plantas foram automaticamente re-irrigadas continuamente até que um nível normal fosse atingido novamente. As plantas foram então re-transferidas novamente para condições normais. O resto do cultivo (maturação de planta, coleta de

semente) foi o mesmo que para plantas não crescidas em condições de estresse abiótico. Uma sessão de confirmação foi realizada consistindo de repetir a triagem com sementes T2 não coletadas de plantas da primeira triagem de seca, mas de plantas crescidas em condições normais.

Parâmetros medidos

A parte aérea de planta (ou biomassa de folhas) foi determinada contando o número total de pixéis nas imagens digitais de partes aéreas de planta discriminado do fundo. Foi tirada a média deste valor para as fotos tiradas no mesmo momento dos ângulos diferentes e foi convertido para um valor de superfície física expresso em mm quadrados por calibração. Experimentos mostram que a parte aérea de planta medida desta maneira se correlaciona com a biomassa de partes aéreas de planta. A Areamax é a parte aérea no momento no qual a planta atingiu sua biomassa de folhas máxima.

As panículas primárias maduras foram coletadas, ensacadas, etiquetadas com código de barras e então secas por três dias no forno a 37°C. As panículas foram então trilhadas e todas as sementes coletadas. As cascas cheias foram separadas daquelas vazias usando um dispositivo que sopra ar. Depois de separação, ambos os lotes de sementes foram então contados usando uma máquina de contagem disponível comercialmente. As cascas vazias foram descartadas. As cascas cheias foram pesadas em uma balança analítica e a área seccional cruzada das sementes foi medida usando formação de imagem digital. Este procedimento resultou no conjunto dos seguintes parâmetros relacionados a sementes:

As flores por panícula estimam o número médio de floretes por panícula em uma planta, derivado do número de sementes total dividido pelo número de primeiras panículas. A panícula mais alta e todas as panículas que se sobrepõe com a panícula mais alta quando alinhadas verticalmente, foram consideradas as primeiras panículas e foram contadas manualmente. O número de grãos cheios foi determinado contando the número de cascas

cheias que permaneceram depois da etapa de separação. A produção total de sementes (peso total de semente) foi medida pesando todas as cascas cheias coletadas de uma planta. Número total de sementes por planta foi medido contando o número de cascas coletadas de uma planta e corresponde ao número de floretes por planta. Peso de Mil Sementes (TKW) é extrapolado a partir do número de grãos cheios contadas e seu peso total. Índice de colheita é definido como a proporção entre o peso total de semente e a área superficial (mm^2), multiplicada por um fator 10^6. O parâmetro EmerVigor é uma indicação do vigor de plântula. Ele é calculado a partir da área (em mm^2) coberta por biomassa de folhas na primeira formação de imagem. A taxa de enchimento de sementes (taxa de enchimento) é uma indicação do preenchimento das sementes. Ela é expressa como uma proporção (em%) do número de grãos cheios pelo número de floretes (no. total de sementes).

Estes parâmetros foram derivados de uma maneira automatizada a partir das imagens digitais usando aplicativo computacional de análise de imagem e foram analisados estatisticamente. Parâmetros individuais de sementes (incluindo largura, comprimento, área, peso) foram medidos usando um dispositivo feito sob encomenda consistindo de dois componentes principais, um dispositivo de pesagem e formação de imagem, acoplado a aplicativo computacional para análise de imagem.

Exemplo 9: mensuração de parâmetros relacionados a produção para transformantes pGOS2::SYR crescidos em condições normais de crescimento:

Mediante análise das sementes como descrito acima, os requerentes encontraram que plantas transformada com a construção gênica pGOS2::*SYR* tiveram uma maior produção de sementes, expressa como número de grãos cheios, peso total de sementes e índice de colheita, comparado com plantas carecendo do transgene *SYR*. Os valores de p mostram que os aumentos foram significantes. Métodos para análise

estatística são como dados na seção introdutória para os Exemplos.

Os resultados obtidos para plantas na geração T1 estão resumidos Tabela E, a qual representa os valores médios para todas as linhagens testadas:

Tabela E:

	% de diferença	valor de p
Número de sementes cheias	+47	0,0000
Peso total de sementes	+52	0,0000
Índice de colheita	+54	0,0000

Os dados obtidos para *SYR* no primeiro experimento foram confirmados em um segundo experimento com plantas T2. Quatro linhagens que tiveram o padrão de expressão correto foram selecionadas para análise adicional. Grupos de sementes das plantas positivas (tanto hetero como homozigotas) em T1 foram triados monitorando expressão de marcador. Para cada evento escolhido, os grupos de sementes heterozigotas foram então retidos para avaliação de T2. Dentro de cada grupo de sementes um número igual de plantas positivas e negativas foi crescido na casa de vegetação para avaliação. Mensuração dos parâmetros de produção de sementes novamente mostrou aumento em número de grãos cheios, peso total de sementes e índice de colheita, comparado com plantas carecendo do transgene *SYR*.

Exemplo 10: mensuração de parâmetros relacionados a produção para transformantes pGOS2::SYR crescidas em condições de estresse:

Mediante análise das sementes como descrito acima, os requerentes encontraram que plantas transformadas com a construção gênica pGOS2::*SYR* e crescidas sob estresse salino, tiveram uma maior produção de sementes, expressa como número de grãos cheios, peso total de sementes, taxa de enchimento e índice de colheita, comparada com plantas carecendo do transgene *SYR*. Além disso, estas plantas sal-estressadas tiveram um maior vigor de plântula comparado com as plantas de controle. Quando as plantas foram crescidas sob estresse por seca, as plantas transgênicas tiveram um maior peso total de sementes e um índice de colheita aumentado comparado

com plantas carecendo do transgene *SYR*. Estas diferenças foram significantes, com um valor de P do teste F abaixo de 0,05.

Exemplo 11: mensuração de parâmetros relacionados a produção para transformantes pH MGP::SYR:

Similarmente como para as plantas transformadas com a construção gênica pGOS2::*SYR*, os requerentes encontraram que plantas transformadas com a construção gênica pH MGP::*SYR* tiveram uma maior produção de sementes, expressa como número de grãos cheios, peso total de sementes e índice de colheita, comparada com plantas carecendo do transgene *SYR*. Os valores de p mostram que os aumentos foram significantes.

Os resultados obtidos para plantas na geração T1 estão resumidos Tabela F, a qual representa os valores médios para todas as linhagens testadas:

Tabela F:

	% de diferença	valor de p
Número de sementes cheias	+34	0,0000
Peso total de sementes	+33	0,0000
Índice de colheita	+37	0,0000

EXEMPLO B: FG-GAP

Exemplo 12: Identificação de seqüências relacionadas a SEQ ID NO: 45 e SEQ ID NO: 46

Seqüências (cDNA de comprimento completo, ESTs ou genômicas) relacionadas a SEQ ID NO: 45 e/ou seqüências de proteínas relacionadas a SEQ ID NO: 46 foram identificadas entre aquelas mantidas no banco de dados Entrez Nucleotides no National Center for Biotechnology Information (NCBI) usando ferramentas de busca de seqüência de banco de dados, tal como a Ferramenta Básica de Alinhamento Local (BLAST) (Altschul *et al.* (1990) J. Mol. Biol. 215:403-410; e Altschul *et al.* (1997) Nucleic Acids Res. 25:3389-3402). O programa foi usado para encontrar regiões de similaridade local entre seqüências comparando seqüências de ácidos nucleicos ou polipeptídeos com banco de dados de seqüências e

calculando a significância estatística de identidades. O polipeptídeo codificado por SEQ ID NO: 45 foi usado para o algoritmo TBLASTN, com configurações padrão e o filtro para ignorar contrapeso de seqüências de baixa complexidade. O resultado da análise foi visto por comparação pareada, e ranqueado de acordo com o escore de probabilidade (valor E), onde o escore reflete a probabilidade de que um alinhamento particular ocorra ao acaso (quanto menor o valor E, mais significante o acerto). Em adição a valores E, comparações também foram pontuadas por identidade de porcentagem. Identidade de porcentagem se refere ao número de nucleotídeos (ou aminoácidos) idênticos entre as duas seqüências de ácidos nucleicos (ou polipeptídeos) comparadas em um comprimento particular. Em algumas instâncias, os parâmetros padrão podem ser ajustados para modificar a estringência da busca.

Em adição às seqüências de ácidos nucleicos disponíveis publicamente disponíveis em NCBI, outros bancos de dados de seqüências também podem ser buscados seguindo o mesmo procedimento que descrito acima.

Tabela G fornece uma lista de seqüências de ácidos nucleicos e proteínas relacionadas à seqüência de ácidos nucleicos como representada por SEQ ID NO: 45 e à seqüência de proteínas representada por SEQ ID NO: 46.

Tabela G: Seqüências de ácidos nucleicos relacionadas à seqüência de ácidos nucleicos (SEQ ID NO: 45) úteis nos métodos da presente invenção, e os polipeptídeos deduzidos correspondentes.

Nome	Organismo fonte	Ácido nucleico SEQ ID NO:	Polipeptídeo SEQ ID NO:	Número de acesso de banco de dados	Estado
AtFG-GAP	*Arabidopsis thaliana*	45	46		Comprimento completo
Homólogo de AtFG-GAP	*Arabidopsis thaliana*	54	55	NM_114965	Comprimento completo
Homólogo 1 de OsFG-GAP	*Oryza sativa*	56	57	NM_185137	Comprimento completo
Homólogo 2 de OsFG-GAP	*Oryza sativa*	58	59	AK068943	Comprimento completo
Homólogo de TaFG-GAP	*Triticum aestivum*	60	/	CK207217	parcial
Homólogo de ZmFG-GAP	*Zea mays*	61	/	AY111316	Parcial
Homólogo de StFG-GAP	*Solanum tuberosum*	62	/	BG598275	Parcial
Homólogo de AFG-GAP	*Aquilegia sp.*	63	/	DT735817	Parcial
Homólogo de BnFG-GAP	*Brassica napus*	64	/	CX192752	Parcial
Homólogo de CsFG-GAP	*Citrus sinensis*	65	/	CX674859	Parcial
Homólogo de AoFG-GAP	*Asparagus officinalis*	66	/	CV288972	Parcial
Homólogo 1 de PFG-GAP	*Populus sp.*	67	/	CN520999	Parcial
Homólogo 2 de PFG-GAP	*Populus sp.*	68	/	CX176799	Parcial
Homólogo de EeFG-GAP	*Euphorbia esula*	69	/	DV130386	Parcial
Homólogo de CrFG-GAP	*Ceratopteris richardii*	70	/	CV736049	Parcial
Homólogo de WmFG-GAP	*Welwitschia mirabilis*	71	/	DT601669	Parcial
Homólogo de Ms FG-GAP	*Medicago sativa*	72	SEQ ID NO: 73 to SEQ ID NO: 76		parcial

Exemplo 13: Alinhamento de seqüências de polipeptídeos relevantes

AlignX do Vetor NTI (Invitrogen) é baseado no popular algoritmo de Clustal de alinhamento progressivo (Thompson *et al.* (1997) Nucleic Acids Res 25:4876-4882; Chenna *et al.* (2003). Nucleic Acids Res 31:3497-3500). Uma árvore filogenética pode ser construída usando um

algoritmo de agrupamento de vizinhos. Valores padrão são para a penalidade de abertura de lacuna de 10, para a penalidade de extensão de lacuna de 0,1 e a matriz de peso selecionada é Blosum 62 (se polipeptídeos estão alinhados).

O resultado do alinhamento múltiplo de seqüências usando polipeptídeos relevantes para identificar aqueles úteis para realizar os métodos da invenção é mostrado em Figura 7. Qualquer um pode ver claramente que apesar de algumas lacunas no alinhamento, conservação de seqüência é encontrada ao longo da maioria da seqüência de proteínas.

Exemplo 14: Cálculo de identidade de porcentagem global entre seqüências de polipeptídeos úteis para realizar os métodos da invenção

Porcentagens globais de similaridade e identidade entre seqüências de polipeptídeos de comprimento completo úteis para realizar os métodos da invenção foram determinadas usando um dos métodos disponíveis na técnica, o aplicativo computacional MatGAT (Ferramenta de Alinhamento Global de Matriz) (BMC Bioinformatics. 2003 4:29. MatGAT: an application that generates similarity/identity matrices using protein ou DNA sequences. Campanella JJ, Bitincka L, Smalley J; aplicativo computacional hospedado por Ledion Bitincka). Aplicativo computacional MatGAT gera matrizes de similaridade/identidade para seqüências de DNA ou proteína sem precisar de pré-alinhamento dos dados. O programa realiza uma série de alinhamentos pareados usando o algoritmo de alinhamento global de Myers e Miller (com uma penalidade de abertura de lacuna de 12, e uma penalidade de extensão de lacuna de 2), calcula similaridade e identidade usando por exemplo Blosum 62 (para polipeptídeos), e então coloca os resultados em uma matriz de distância. Similaridade de seqüências é mostrada na metade de baixo da linha divisora e identidade de seqüências é mostrada na metade de cima da linha divisora diagonal.

Parâmetros usados na comparação foram:

Matriz de pontuação: Blosum62

Primeira Lacuna: 12

Lacuna de extensão: 2

Resultados da análise de aplicativo computacional são mostrados em Tabela H para a similaridade e identidade global no comprimento completo das seqüências de polipeptídeos (excluindo as seqüências de polipeptídeos parciais). Identidade de porcentagem é dada acima da diagonal e similaridade de porcentagem é dada abaixo da diagonal.

A identidade de porcentagem entre as seqüências de polipeptídeos úteis para realizar os métodos da invenção pode ser tão baixa quanto 17% de identidade de aminoácidos comparada com SEQ ID NO: 46.

Tabela H: Resultados de MatGAT para similaridade e identidade global no comprimento completo das seqüências de polipeptídeos.

	1	2	3	4
1. AtFGAP1		**18,1**	**65,5**	**17,4**
2. AtFGGAP2	31,4		**17,9**	**67,7**
3. OsFGGAP1	76,7	33,5		**16,9**
4. OsFGGAP2	32,8	83,6	33	

Exemplo 15: Identificação de domínios compreendidos em seqüências de polipeptídeos úteis para realizar os métodos da invenção

O banco de dados Integrated Resource of Protein Families, Domains and Sites (InterPro) é uma interface integrada para os bancos de dados de assinatura comumente usados para buscas baseadas em texto e seqüência. O banco de dados InterPro combina estes bancos de dados, que usam diferentes metodologias e graus variados de informação biológica sobre proteínas bem caracterizadas para derivar assinaturas de proteínas. Bancos de dados colaboradores incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. Interpro é hospedado no European Bioinformatics Institute no Reino Unido.

Os resultados da varredura de InterPro da seqüência de polipeptídeos como representada por SEQ ID NO: 46 são apresentados em

Tabela I.

Tabela I: Resultados de varredura de InterPro da seqüência de polipeptídeos como representada por SEQ ID NO: 46

Banco de dados	Número de acesso	Nome de acesso
Pfam	PF01839	FG-GAP
INTERPRO	IPR013517	FG-GAP
INTERPRO	IPR000413	Cadeia alfa de integrinas

Exemplo 16: Previsão de topologia das seqüências de polipeptídeos úteis para realizar os métodos da invenção

TargetP 1.1 prevê a localização subcelular de proteínas eucariotas. A designação de localização é baseada na presença prevista de qualquer das pré-seqüências N-terminais: peptídeo de trânsito de cloroplasto (cTP), peptídeo de direcionamento mitocondrial (mTP) ou peptídeo sinal de via secretora (SP). Escores nos quais a previsão final está baseada não são probabilidades reais, e eles não necessariamente adicionam a um. Entretanto, a localização com o maior escore é a mais provável de acordo com TargetP, e a relação entre os escores (a classe de confiança) pode ser uma indicação de quão certa é a previsão. A classe de confiança (RC) varia de 1 a 5, onde 1 indica a previsão mais forte. TargetP é mantido no servidor da Technical University of Denmark.

Para as seqüências previstas conterem uma pré-seqüência N-terminal um sítio de clivagem potencial também pode ser previsto.

Diversos parâmetros foram selecionados, tal como grupo de organismo (não vegetal ou vegetal), conjuntos de atalhos (nenhum, conjunto pré-definido de atalhos, ou conjunto de atalhos especificado por usuário), e o cálculo de previsão de sítios de clivagem (sim ou não).

Os resultados de análise TargetP 1.1 da seqüência de polipeptídeos como representada por SEQ ID NO: 46 são apresentados em Tabela J. O grupo de organismo "vegetal" foi selecionado, nenhum atalho definido, e o comprimento previsto do peptídeo de trânsito requisitado. A localização subcelular da seqüência de polipeptídeos como representada por

SEQ ID NO: 46 provavelmente não é intracelular, existe uma leve preferência pela via secretora (embora com um escore de confiança de 5) e o comprimento previsto do peptídeo de trânsito putativo é de 24 aminoácidos começando do término N (não tão confiável quanto à previsão da própria localização subcelular, pode variar em comprimento de uns poucos aminoácidos).

Tabela J: Análise TargetP 1.1 da seqüência de polipeptídeos como representada por SEQ ID NO: 46

Comprimento (AA)	896
Peptídeo de trânsito cloroplástico	0,010
Peptídeo de trânsito mitocondriano	0,546
Peptídeo sinal de via secretora	0,643
Outro direcionamento subcelular	0,038
Localização Prevista	secretada
Classe de confiança	5
Comprimento previsto de peptídeo de trânsito	24

Quando analisada com SignalP (Bendtsen et al., J. Mol. Biol., 340:783-795, 2004), existe uma identificação positiva confiável (probabilidade de 0.998) para a presença de um peptídeo sinal de secreção N-terminal com um comprimento de 24 aminoácidos. Além disso, quando usando o algoritmo THMM (Center for Biological Sequence Analysis, Technical University of Denmark), a proteína é prevista ser localizada no lado exterior da célula com apenas uma cauda C-terminal no citoplasma: resíduos 1-859: fora; resíduos 860-879: domínio transmembrana, resíduos 880-896: dentro.

Muitos outros algoritmos podem ser usados para realizar tais análises, incluindo:

- ChloroP 1.1 hospedado no servidor do Technical University of Denmark;
- Protein Prowler Subcellular Localisation Predictor versão 1.2 hospedado no servidor do Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia;

- PENCE Proteome Analyst PA-GOSUB 2.5 hospedado no servidor da University of Alberta, Edmonton, Alberta, Canada;

Exemplo 17: Clonagem de gene

O gene *FG-GAP* de *Arabidopsis thaliana* foi amplificado por PCR usando como modelo uma biblioteca de cDNA de plântula de *Arabidopsis thaliana* (Invitrogen, Paisley, UK). Depois de transcrição reversa de RNA extraído de plântulas, os cDNAs foram clonados em pCMV Sport 6.0. Tamanho médio de inserto do banco foi de 1,5 kb e o número original de clones foi da ordem de 1,59 x 10^7 cfu. Título original foi determinado como sendo 9,6 x 10^5 cfu/ml depois de primeira amplificação de 6 x 10^{11} cfu/ml. Depois de extração de plasmídeo, 200 ng de modelo foram usados em uma mistura de PCR de 50 µl. Iniciadores prm06643 (SEQ ID NO: 47; sentido, códon de início em negrito, sítio AttB1 em itálico: 5'-*ggggacaagtttgtacaaaaaagcaggcttaaaca***atg**aaatctcgagcgagg-3') e prm06644 (SEQ ID NO: 48; reverso, complementar, sítio AttB2 em itálico: 5'-*ggggaccactttgtacaagaaagctgggt*cctg tttacagatggtacctagt-3'), os quais incluem os sítios AttB para recombinação Gateway, foram usados para amplificação por PCR. PCR foi realizado usando DNA polimerase Hifi Taq em condições padrão. Um fragmento de PCR de 3,2 kb (incluindo sítios attB) foi amplificado e purificado também usando métodos padrão. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante a qual o fragmento de PCR se recombina in vivo com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada", pFG-GAP. Plasmídeo pDONR201 foi adquirido a partir de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 18: Construção de Vetor

O clone de entrada pFG-GAP foi subseqüentemente usado em uma reação LR com pGOS2, um vetor de destinação usado para transformação de Oryza sativa. Este vetor contém como elementos funcionais

dentro dos limites de T-DNA: um marcador selecionável vegetal; uma gaveta de expressão de marcador triável; e uma gaveta Gateway pretendida para recombinação in vivo LR com a seqüência de interesse já clonada no clone de entrada. Um promotor GOS2 de arroz (nucleotídeos 1 a 2193 de SEQ ID NO: 49, a combinação promotor-gene) para expressão constitutiva foi localizado antes desta gaveta Gateway.

Depois da etapa de recombinação LR, o vetor de expressão resultante, pGOS2::FG-GAP para *FG-GAP* (Figura 7) foi transformado em linhagem LBA4044 de *Agrobacterium* e subseqüentemente para plantas de *Oryza sativa*. Plantas transformadas de arroz foram permitidas crescer e foram então avaliadas para os parâmetros descritos em Exemplo 19.

Para transformação de outras safras veja Exemplo 40.

Exemplo 19: Métodos de avaliação para plantas transformadas com FG-GAP sob o controle do promotor GOS2 de arroz

Aproximadamente 15 a 20 transformantes T0 de arroz independentes foram gerados. Os transformantes primários foram transferidos de uma câmera de safra de tecido para uma casa de vegetação para crescimento e colheita de semente T1. Sete eventos, dos quais a progênie T1 segregou 3:1 para presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 plântulas T1 contendo o transgene (hetero e homo-zigotas) e aproximadamente 10 plântulas T1 carecendo do transgene (nulizigotos) foram selecionadas monitorando expressão de marcador visual. As plantas T1 selecionadas foram transferidas para uma casa de vegetação. Cada planta recebeu uma etiqueta com código de barras exclusivo para ligar inequivocamente os dados de fenotipagem à planta correspondente. As plantas T1 selecionadas foram crescidas no solo em potes de 10 cm de diâmetro nas seguintes configurações ambientais: fotoperíodo= 11,5 h, intensidade de luz do dia= 30.000 lux ou mais, temperatura de dia= 28°C ou superior, temperatura de noite= 22°C, umidade relativa= 60-70%.

Plantas transgênicas e os nulizigotos correspondentes foram crescidos lado a lado em posições aleatórias. Do estágio de semeadura até o estágio de maturidade as plantas foram passadas diversas vezes através de uma cabine de formação de imagem digital. Em cada momento imagens digitais (2048x1536 pixéis, 16 milhões de cores) foram tiradas de cada planta de pelo menos 6 ângulos diferentes.

A parte aérea de planta (ou biomassa de folhas) foi determinada contando o número total de pixéis nas imagens digitais de partes vegetais superficiais discriminado a partir do fundo. Foi tirada a média deste valor para as fotos tiradas no mesmo momento dos ângulos diferentes e foi convertido para um valor de superfície física expresso em mm quadrados por calibração. Experimentos mostram que a parte aérea de planta medida desta maneira se correlaciona com a biomassa de partes aéreas de planta. A Areamax é a parte aérea no momento no qual a planta atingiu sua biomassa de folhas máxima.

As panículas primárias maduras foram coletadas, ensacadas, etiquetadas com código de barras e então secas por três dias no forno a 37°C. As panículas foram então trilhadas e todas as sementes coletadas. As cascas cheias foram separadas daquelas vazias usando um dispositivo que sopra ar. Depois de separação, ambos os lotes de sementes foram então contados usando uma máquina de contagem disponível comercialmente. As cascas vazias foram descartadas. As cascas cheias foram pesadas em uma balança analítica e a área seccional cruzada das sementes foi medida usando formação de imagem digital. Este procedimento resultou no conjunto dos seguintes parâmetros relacionados a sementes:

As flores por panícula é um parâmetro estimando o número médio de floretes por panícula em uma planta, derivado do número de sementes total dividido pelo número de primeiras panículas. A panícula mais alta e todas as panículas que se sobrepõe com a panícula mais alta quando

alinhadas verticalmente, foram consideradas as primeiras panículas e foram contadas manualmente. O número de grãos cheios foi determinado contando o número de cascas cheias que permaneceram depois da etapa de separação. A produção total de sementes (peso total de semente) foi medida pesando todas as cascas cheias coletadas de uma planta. Número total de sementes por planta foi medido contando o número de cascas coletadas de uma planta e corresponde ao número de floretes por planta. Peso de Mil Sementes (TKW) é extrapolado a partir do número de grãos cheios contadas e seu peso total. Índice de colheita é definido como a proporção entre o peso total de semente e a área superficial (mm^2), multiplicada por um fator 10^6. Estes parâmetros foram derivados de uma maneira automatizada a partir das imagens digitais usando aplicativo computacional de análise de imagem e foram analisados estatisticamente. Parâmetros individuais de sementes (incluindo largura, comprimento, área, peso) foram medidos usando um dispositivo feito sob encomenda consistindo de dois componentes principais, um dispositivo de pesagem e formação de imagem, acoplado a aplicativo computacional para análise de imagem.

Uma ANOVA de dois fatores (análises de variância) corrigida para o modelo não equilibrado foi usada como modelo estatístico para a avaliação geral de características fenotípicas vegetais. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformadas com aquele gene. O teste F foi realizado para verificar por um efeito do gene sobre todos os eventos de transformação e para verificar por um efeito geral do gene, também referido neste lugar como um "efeito gênico global". Se o valor do teste F mostrou que os dados foram significantes, então foi concluído que houve um efeito "gênico", significando que não foi apenas presença ou a posição do gene que estava causando o efeito. O limiar para significância para um verdadeiro efeito gênico global foi ajustado em um nível de probabilidade de 5% para o teste F.

Para verificar por um efeito dos genes dentro de um evento, i.e., por um efeito linhagem específico, um teste t foi realizado dentro de cada evento usando conjuntos de dados das plantas transgênicas e das plantas nulas correspondentes. "Plantas nulas" ou "segregantes nulos" ou "nulizigotos" se referem a plantas tratadas da mesma maneira que a planta transgênica, mas das quais o transgene foi segregado. Plantas nulas também podem ser descritas como as plantas transformadas negativas homozigotas. O limiar para significância para o teste t foi ajustado em um nível de probabilidade de 10%. Os resultados para alguns eventos podem estar acima ou abaixo deste limiar. Isto é baseado na hipótese de que um gene pode apenas ter um efeito em certas posições no genoma, e que a ocorrência deste efeito dependente de posição não é comum. Este tipo de efeito gênico é também referido neste lugar como um "efeito de linhagem do gene". O valor de p foi obtido comparando o valor de t com a distribuição de t ou alternativamente, comparando o valor de F com a distribuição de F. O valor de p então gera a probabilidade de que a hipótese nula (i.e., que não existe nenhum efeito do transgene) esteja correta.

Os dados obtidos para *FG-GAP* no primeiro experimento foram confirmados em um segundo experimento com plantas T2. Quatro linhagens foram selecionadas para análise adicional. Grupos de sementes das plantas positivas (tanto hetero como homozigotas) em T1, foram triados monitorando expressão de marcador. Para cada evento escolhido, os grupos de sementes heterozigotas foram então retidos para avaliação de T2. Dentro de cada grupo de sementes um número igual de plantas positivas e negativas foi crescido na casa de vegetação para avaliação.

Um número total de 120 plantas transformadas *FG-GAP* foi avaliado na geração T2, isto é 30 plantas por evento das quais 15 foram positivas para o transgene, e 15 negativas.

Devido ao fato de dois experimentos com eventos sobrepostos

terem sido realizados, uma análise combinada foi realizada. Isto é útil para verificar consistência dos efeitos sobre os dois experimentos, e se este é o caso, para acumular evidência de ambos experimentos a fim de aumentar confiança na conclusão. O método usado foi uma abordagem de modelo misturado que leva em consideração a estrutura multivariada dos dados (i.e. experimento - evento - segregantes). Valores de P foram obtidos comparando teste de proporção de probabilidade com distribuições qui quadrado.

Exemplo 20: Avaliação de transformantes FG-GAP: mensuração de parâmetros relacionados a produção

Mediante análise das sementes como descrito acima, os requerentes encontraram que plantas transformadas com a construção gênica *FG-GAP* tiveram uma maior produção de sementes, expressa como número de grãos cheios e peso total de sementes, comparado com plantas carecendo do transgene *FG-GAP*. Os valores de p mostram que os aumentos foram significantes. Também o índice de colheita foi aumentado (+9%).

Os resultados obtidos para plantas na geração T1 estão resumidos Tabela K:

Tabela K:

	% de diferença	**valor de p de teste F**
Número de sementes cheias	+19	0,0051
Peso total de sementes	+17	0,0199

Estes resultados positivos foram obtidos novamente na geração T2. Em Tabela L, dados mostram os aumentos gerais em% para o número de grãos cheios, peso total de sementes e índice de colheita, calculados a partir dos dados das linhagens individuais da geração T2, e os respectivos valores de p. Estes dados T2 foram reavaliados em uma análise combinada com os resultados para a geração T1, e os valores de p obtidos mostram que os efeitos observados foram altamente significantes.

Tabela L:

	Geração T2		Análise combinada
	% de diferença	**valor de p de teste F**	**valor de p**
Número de sementes cheias	+17	0,0247	0,0004
Peso total de sementes	+17	0,0283	0,0014
Índice de colheita	+20	0,0030	0,0007

EXEMPLO C: CYP90B

Exemplo 21: Clonagem de gene de cDNA de CYP90B de Oryza sativa

O cDNA de CYP90B de *Oryza sativa* foi amplificado por PCR usando como modelo uma biblioteca de cDNA de plântula de *Oryza sativa* (Invitrogen, Paisley, UK). Depois de transcrição reversa de RNA extraído de plântulas, os cDNAs foram clonados em pCMV Sport 6.0. Tamanho médio de inserto do banco foi de 1,6 kb e o número original de clones foi da ordem de $1,67x10^7$ cfu. Título original foi determinado como sendo 3,34 $x10^6$ cfu/ml depois de primeira amplificação de $6x10^{10}$ cfu/ml. Depois de extração de plasmídeo, 200 ng de modelo foram usados em uma mistura de PCR de 50 µl. Iniciadores (SEQ ID NO: 107; sentido, códon de início em negrito, sítio AttB1 em itálico: 5' *GGGGACAAGTTTGTACAAAAAAGCAGG*CTTAAACAA**TG**GCCGCCATG ATGGC 3') e (SEQ ID NO: 108; reverso, complementar, sítio AttB2 em itálico: 5' *GGGGACCACTTTGTACAAGAAAGCTGGGT* TTACTCCTGCTCATCATCC 3'), os quais incluem os sítios AttB para recombinação Gateway, foram usados para amplificação por PCR. PCR foi realizado usando DNA polimerase Hifi Taq em condições padrão. Um fragmento de PCR de 1585 bp (incluindo sítios attB; 1521 bp de início a término) foi amplificado e purificado também usando métodos padrão. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante a qual o fragmento de PCR se recombina *in vivo* com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada". Plasmídeo pDONR201 foi adquirido a partir de Invitrogen, como

parte da tecnologia Gateway®.

Exemplo 22: Construção de Vetor

O clone de entrada foi subseqüentemente usado em uma reação LR com vetores de destinação usados para transformação de *Oryza sativa*. Estes vetores contêm como elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; uma gaveta de expressão de marcador triável; e uma gaveta Gateway pretendida para recombinação in vivo LR com a seqüência de interesse já clonada no clone de entrada. Quatro promotores de arroz diferentes localizados antes desta gaveta Gateway foram usados para expressar o CYP90B de *Oryza sativa*: prolamina RP6, oleosina 18 kDa, GOS2 e HMGB1.

Depois da etapa de recombinação LR, os vetores de expressão resultantes (promotor de prolamina RP6, oleosina 18 kDa, GOS2 e HMGB1 – veja Figura 14) foram transformados em linhagem LBA4044 *Agrobacterium* e subseqüentemente para plantas de *Oryza sativa*. Plantas transformadas de arroz foram permitidas crescer e foram então avaliadas para o parâmetros descritos nos exemplos abaixo. Para transformação de outras safras veja Exemplo 40.

Exemplo 23: Descrição do procedimento de avaliação fenotípica

Aproximadamente 15 a 20 transformantes T0 de arroz independentes foram gerados por construção. Os transformantes primários foram transferidos de uma câmera de safra de tecido para uma casa de vegetação para crescimento e colheita de semente T1. Quatro ou cinco eventos, dos quais a progênie T1 segregou 3:1 para presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 plântulas T1 contendo o transgene (hetero e homozigotas) e aproximadamente 10 plântulas T1 carecendo do transgene (nulizigotos) foram selecionadas monitorando expressão de marcador visual. As plantas transgênicas e as plantas de controle adequadas foram crescidas lado a lado em posições

aleatórias. Do estágio de semeadura até o estágio de maturidade as plantas foram passadas diversas vezes através de uma cabine de formação de imagem digital. Em cada momento imagens digitais (2048x1536 pixéis, 16 milhões de cores) foram tiradas de cada planta de pelo menos 6 ângulos diferentes.

Três T1 eventos foram adicionalmente avaliados na geração T2 seguindo o mesmo procedimento de avaliação que para a geração T1 mas com mais indivíduos por evento.

Mensurações de parâmetros relacionados a sementes

As panículas primárias maduras foram coletadas, contadas, ensacadas, etiquetadas com código de barras e então secas por três dias em um forno a 37°C. As panículas foram então trilhadas e todas as sementes foram coletadas e contadas. As cascas cheias foram separadas daquelas vazias usando um dispositivo que sopra ar. As cascas vazias foram descartadas e a fração remanescente foi contada novamente. As cascas cheias foram pesadas em uma balança analítica. O número de grãos cheios foi determinado contando o número de cascas cheias que permaneceram depois da etapa de separação. A produção total de sementes foi medida pesando todas as cascas cheias coletadas de uma planta. Número total de sementes por planta foi medido contando o número de cascas coletadas de uma planta. Peso de Mil Sementes (TKW) é extrapolado a partir do número de grãos cheios contado e seu peso total. O índice de colheita (HI) na presente invenção é definido como a proporção entre a produção total de sementes e a parte aérea (mm^2), multiplicada por um fator 10^6. O número total de flores por panícula como definido na presente invenção é a proporção entre o número total de sementes e o número de panículas primárias maduras. A taxa de enchimento de sementes como definido na presente invenção é a proporção (expressa como uma%) do número de grãos cheios sobre o número total de sementes (ou floretes). Parâmetros individuais de sementes (largura, comprimento e área) foram medidos usando um dispositivo feito sob encomenda consistindo de

dois componentes principais, um dispositivo de pesagem e formação de imagem, acoplado a aplicativo computacional para análise de imagem. Tanto sementes com casca como descascadas foram usadas para estas mensurações.

Análise estatística: teste F

Uma ANOVA de dois fatores (análise de variantes) foi usada como um modelo estatístico para a avaliação geral de características fenotípicas vegetais. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformadas com o gene da presente invenção. O teste F foi realizado para verificar por um efeito do gene sobre todos os eventos de transformação e por um efeito geral do gene, também conhecido como um efeito gênico global. O limiar para significância para um verdadeiro efeito gênico global foi ajustado em um nível de probabilidade de 5% para o teste F. Um valor significante de teste F aponta para um efeito gênico, significando que não é apenas a presença ou posição do gene que está causando as diferenças em fenótipo.

Exemplo 24: Resultados de CYP90B de Oryza sativa sob o controle de promotores não constitutivos

24.1 Plantas transgênicas expressando CYP90B sob o controle do promotor endosperma específico

Os resultados de produção de sementes e mensuração de HI para plantas transgênicas expressando CYP90B sob o controle do promotor endosperma específico (prolamina RP6) são mostrados em Tabela M e N, respectivamente. O número de eventos com um aumento é indicado, assim como os valores de p do teste F para as gerações T1 e T2.

Tabela M: Resultados de mensuração de produção de sementes de plantas transgênicas expressando CYP90B sob o controle do promotor endosperma específico.

	Número de eventos mostrando um aumento	% de Diferença	Valor de P de teste F
Geração T1	3 de 4	11	0,1572
Geração T2	3 de 3	13	0,0103

Tabela N: Resultados de mensuração de HI de plantas transgênicas expressando CYP90B sob o controle do promotor endosperma específico.

	Número de eventos mostrando um aumento	% de Diferença	Valor de P de teste F
Geração T1	4 de 4	11	0,047
Geração T2	3 de 3	10	0,0392

As plantas transgênicas de arroz expressando CYP90B sob o controle do promotor endosperma específico (prolamina RP6) apresentam uma colheita aumentada, devido a um aumento em produção de sementes enquanto que biomassa de parte aérea de planta permanece não modificada (dados não mostrados), quando comparada com plantas de controle.

24.2 Plantas transgênicas expressando CYP90B sob o controle do promotor embrião/aleurona específico

Os resultados de mensuração de TKW para plantas transgênicas expressando CYP90B sob o controle de um promotor de embrião/aleurona (oleosina 18 kDa) são mostrados em Tabela O. O número de eventos com um aumento é indicado assim como os valores de p do teste F para as gerações T1 e T2.

Tabela O: Resultados de mensuração de TKW de plantas transgênicas expressando CYP90B sob o controle do promotor de embrião/aleurona.

	Número de eventos mostrando um aumento	% de Diferença	Valor de P de teste F
Geração T1	5 de 5	4	0,0002
Geração T2	2 de 3	1	0,2428

Os resultados de mensuração de área média de semente para plantas transgênicas expressando CYP90B sob o controle do promotor de oleosina 18 kDa são mostrados em Tabela P. O número de eventos com um

aumento é indicado assim como os valores de p do teste F para as gerações T1 e T2.

Tabela P: Resultados de mensuração de área média de semente de plantas transgênicas expressando CYP90B sob o controle do promotor de embrião/aleurona.

	Número de eventos mostrando um aumento	**% de Diferença**	**Valor de P de teste F**
Geração T1	5 de 5	3	<0,0001
Geração T2	3 de 3	2	0,0272

Os resultados de mensuração de comprimento médio de semente para plantas transgênicas expressando CYP90B sob o controle do promotor de oleosina 18 kDa são mostrados em Tabela Q. O número de eventos com um aumento é indicado assim como os valores de p do teste F para as gerações T1 e T2.

Tabela Q: Resultados de mensuração de comprimento médio de semente de plantas transgênicas expressando CYP90B sob o controle do promotor de embrião/aleurona.

	Número de eventos mostrando um aumento	**% de Diferença**	**Valor de P de teste F**
Geração T1	5 de 5	3	<0,0001
Geração T2	3 de 3	1	0,0086

Plantas transgênicas de arroz expressando CYP90B sob o controle de um promotor de embrião/aleurona (oleosina 18 kDa) têm sementes com TKW, área de semente e comprimento de semente aumentado. Nenhum aumento significante em produção de sementes foi observado.

Exemplo 25: Avaliação e Resultados de* Oryza sativa *CYP90B sob o controle de promotores constitutivos

25.1 Plantas transgênicas expressando CYP90B sob o controle do promotor constitutivo GOS2

Os resultados de mensuração de avaliação para plantas transgênicas expressando CYP90B sob o controle do promotor constitutivo GOS2 são mostrados em Tabela R. O número de eventos com um aumento é

indicado, assim como os valores de p do teste F para a geração T1. Nenhuma avaliação de geração T2 é realizada quando resultados negativos são obtidos na geração T1.

Tabela R: Resultados de mensuração de avaliação de plantas transgênicas expressando CYP90B sob o controle do promotor constitutivo GOS2.

	Número de eventos mostrando um aumento	**% de Diferença**	**Valor de P de teste F**
Biomassa acima do solo	5 de 5	-13	<0,0001
Altura	5 de 5	-7	<0,0001
Número de grãos cheios	5 de 5	-53	<0,0001
Número de sementes	5 de 5	-32	<0,0001
Produção de sementes	5 de 5	-53	<0,0001
HI	5 de 5	-46	<0,0001

25.2 Plantas transgênicas expressando CYP90B sob o controle do promotor constitutivo HMBG1

Os resultados de mensuração de avaliação para plantas transgênicas expressando CYP90B sob o controle do promotor constitutivo HMBG1 são mostrados em Tabela S. O número de eventos com um aumento é indicado, assim como os valores de p do teste F para a geração T1. Nenhuma avaliação de geração T2 é realizada quando resultados negativos são obtidos na geração T1.

Tabela S: Resultados de mensuração de avaliação de plantas transgênicas expressando CYP90B sob o controle do promotor constitutivo HMBG1.

	Número de eventos mostrando um aumento	**% de Diferença**	**Valor de P de teste F**
Biomassa acima do solo	5 de 5	-18	<0,0001
Altura	5 de 5	-6	<0,0001
Número de grãos cheios	5 de 5	-56	<0,0001
Número de sementes	5 de 5	-33	<0,0001
Produção de sementes	5 de 5	-56	<0,0001
HI	5 de 5	-46	<0,0001

Plantas transgênicas de arroz expressando CYP90B sob o controle de dois diferentes promotores constitutivos mostram biomassa de

parte aérea de planta, altura de planta, número de grãos cheios, produção de sementes e HI fortemente reduzidos comparados com plantas de controle.

EXEMPLO D: CDC27

Exemplo 26: Clonagem de um gene de Arabidopsis thaliana codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo

O gene de *Arabidopsis thaliana* codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo (CDS0171_2) foi amplificado por PCR usando como modelo uma biblioteca de cDNA de plântula de *Arabidopsis thaliana* (Invitrogen, Paisley, UK). Depois de transcrição reversa de RNA extraído de plântulas, os cDNAs foram clonados em pCMV Sport 6.0. Tamanho médio de inserto do banco foi de 1,5 kb e o número original de clones foi da ordem de $1,59x10^7$ cfu. Título original foi determinado como sendo $9,6x10^5$ cfu/ml, e depois da primeira amplificação de 10^{10} cfu/ml. Depois de extração de plasmídeo, 200 ng de modelo foram usados em uma mistura de PCR de 50 µl. Iniciadores (SEQ ID NO: 149; sentido, códon de início em negrito, sítio AttB1 em itálico: 5'- *GGGGACAAGTTTGTACAAAAAAGCAGG*CTTCACA**ATG**CAACAA CTGTCAACTTC 3') e (SEQ ID NO: 150; reverso, complementar, sítio AttB2 em itálico: 5' *GGGGACCACTTTGTACAAGAAAGCTGGG*TTGGAG TAGC TATGGTTTCAC-3'), os quais incluem os sítios AttB para recombinação Gateway, foram usados para amplificação por PCR. PCR foi realizado usando DNA polimerase Hifi Taq em condições padrão. Um fragmento de PCR de 1816 bp (incluindo sítios attB; 1737 bp de início a término) foi amplificado e purificado também usando métodos padrão. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante a qual o fragmento de PCR se recombina *in vivo* com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada". Plasmídeo pDONR201 foi adquirido a partir de Invitrogen, como

parte da tecnologia Gateway®.

Exemplo 27: Construção de Vetor

O clone de entrada foi subseqüentemente usado em uma reação LR com um vetor de destinação usado para transformação de *Oryza sativa*. Este vetor contém como elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; uma gaveta de expressão de marcador triável; e uma gaveta Gateway pretendida para recombinação in vivo LR com a seqüência de interesse já clonada no clone de entrada. Um promotor OSH1 de arroz (SEQ ID NO: 151) para expressão em meristema apical de broto foi localizado antes desta gaveta Gateway.

Depois da etapa de recombinação LR, o vetor de expressão resultante mostrado em Figura 18 foi transformado em linhagem LBA4044 de *Agrobacterium* e subseqüentemente para plantas de *Oryza sativa*. Plantas transformadas de arroz foram permitidas crescer e foram então avaliadas para os parâmetros descritos em Exemplos 28 e 29. Para transformação de outras safras veja Exemplo 40.

Exemplo 28: Descrição do procedimento de avaliação fenotípica

Aproximadamente 15 a 20 transformantes T0 de arroz independentes foram gerados. Os transformantes primários foram transferidos de uma câmera de safra de tecido para uma casa de vegetação para crescimento e colheita de semente T1. Cinco eventos, dos quais a progênie T1 segregou 3:1 para presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 plântulas T1 contendo o transgene (hetero e homo-zigotas) e aproximadamente 10 plântulas T1 carecendo do transgene (nulizigotos) foram selecionadas monitorando expressão de marcador visual. As plantas transgênicas e as plantas de controle adequadas foram crescidas lado a lado em posições aleatórias. Do estágio de semeadura até o estágio de maturidade as plantas foram passadas diversas vezes através de uma cabine de formação de imagem digital. Em cada momento imagens

digitais (2048x1536 pixéis, 16 milhões de cores) foram tiradas de cada planta de pelo menos 6 ângulos diferentes.

Três dos eventos avaliados em T1 foram adicionalmente avaliados na geração T2 seguindo o mesmo procedimento de avaliação que para a geração T1 mas com mais indivíduos por evento.

Mensurações de parâmetros relacionados a sementes

As panículas primárias maduras foram coletadas, contadas, ensacadas, etiquetadas com código de barras e então secas por três dias em um forno a 37°C. As panículas foram então trilhadas e todas as sementes foram coletadas e contadas. As cascas cheias foram separadas daquelas vazias usando um dispositivo que sopra ar. As cascas vazias foram descartadas e a fração remanescente foi contada novamente. As cascas cheias foram pesadas em uma balança analítica. O número de grãos cheios foi determinado contando o número de cascas cheias que permaneceram depois da etapa de separação. A produção total de sementes foi medida pesando todas as cascas cheias coletadas de uma planta. Número total de sementes por planta foi medido contando o número de cascas coletadas de uma planta. Peso de Mil Sementes (TKW) é extrapolado a partir do número de grãos cheios contadas e seu peso total. O índice de colheita (HI) na presente invenção é definido como a proporção entre a produção total de sementes e a parte aérea (mm^2), multiplicada por um fator 10^6. O número total de flores por panícula como definido na presente invenção é a proporção entre o número total de sementes e o número de panículas primárias maduras. A taxa de enchimento de sementes como definida na presente invenção é a proporção (expressa como uma%) do número de grãos cheios sobre o número total de sementes (ou floretes).

Análise estatística: teste F

Uma ANOVA de dois fatores (análise de variantes) foi usada como um modelo estatístico para a avaliação geral de características

fenotípicas vegetais. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformadas com o gene da presente invenção. O teste F foi realizado para verificar por um efeito do gene sobre todos os eventos de transformação e por um efeito geral do gene, também conhecido como um efeito gênico global. O limiar para significância para um verdadeiro efeito gênico global foi ajustado em um nível de probabilidade de 5% para o teste F. Um valor significante de teste F aponta para um efeito gênico, significando que não é apenas a presença ou posição do gene que está causando as diferenças em fenótipo.

Exemplo 29: Resultados da avaliação de plantas transgênicas de arroz expressando um* ácido nucleico CDC27 modificado de Arabidopsis thaliana *sob o controle de um promotor de meristema apical de broto

Os resultados de mensuração de avaliação (produção de sementes, número de grãos cheios, e HI) para plantas transgênicas expressando um ácido nucleico CDC27 modificado sob o controle de um promotor de meristema apical de broto (OSH1) são mostrados em Tabelas T a V. O número de eventos com um aumento, a% de diferença com plantas de controle adequadas, assim como os valores de p do teste F para as gerações T1 e T2 são indicados.

Tabela T: Resultados de mensuração de produção de sementes de plantas transgênicas expressando um ácido nucleico CDC27 modificado sob o controle de um promotor de meristema apical de broto.

	Número de eventos mostrando um aumento	**% de Diferença**	**Valor de P de teste F**
Geração T1	3 de 5	35	0,0113
Geração T2	2 de 3	11	0,0083

Tabela U: Resultados de mensuração de número de grãos cheios de plantas transgênicas expressando um ácido nucleico CDC27 modificado sob o controle de um promotor de meristema apical de broto.

	Número de eventos mostrando um aumento	**% de Diferença**	**Valor de P de teste F**
Geração T1	3 de 5	36	0,0083
Geração T2	2 de 3	10	0,0099

Tabela V: Resultados de mensuração de índice de colheita de plantas transgênicas expressando um ácido nucleico CDC27 modificado sob o controle de um promotor de meristema apical de broto.

	Número de eventos mostrando um aumento	**% de Diferença**	**Valor de P de teste F**
Geração T1	3 de 5	34	0,0053
Geração T2	2 de 3	6	0,0188

Plantas transgênicas de arroz expressando um ácido nucleico CDC27 modificado sob o controle de promotor de meristema apical de broto tiveram produção significantemente aumentada de sementes, número aumentado de sementes cheias e índice de colheita aumentado.

EXEMPLO E: AT-hook

Exemplo 30: Clonagem de gene de ácido nucleico codificando AT-hook de Oryza sativa

O gene de *Oryza sativa* codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 (veja SEQ ID NO: 152) foi amplificado por PCR usando como modelo uma biblioteca de cDNA de plântula de *Oryza sativa* (Invitrogen, Paisley, UK). Depois de transcrição reversa de RNA extraído de plântulas, os cDNAs foram clonados em pCMV Sport 6.0. Tamanho médio de inserto do banco foi de 1,6 kb e o número original de clones foi da ordem de $1,67x10^7$ cfu. Título original foi determinado como sendo $3,34 x10^6$ cfu/ml depois de primeira amplificação de $6x10^{10}$ cfu/ml. Depois de extração de plasmídeo, 200 ng de modelo foram usados em uma mistura de PCR de 50 µl. Iniciadores (SEQ ID NO: 196; sentido, iniciador AttB1: 5'- ggggacaagtttgtacaaaaaagcaggcttaaacaatggat

ccggtcacgg -3') e (SEQ ID NO: 197; reverso, complementar, iniciador AttB2: 5'-ggggaccactttgtacaagaaagctgggtggaatcgatccatctcagaa -3'), os quais incluem os sítios AttB para recombinação Gateway, foram usados para amplificação por PCR. PCR foi realizado usando DNA polimerase Hifi Taq em condições padrão. Um fragmento de PCR (incluindo sítios attB; de início a término) foi amplificado e purificado usando métodos padrão. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante a qual o fragmento de PCR recombinado *in vivo* com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada". Plasmídeo pDONR201 foi adquirido a partir de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 31: Construção de Vetor

O clone de entrada foi subseqüentemente usado em uma reação LR com um vetor de destinação contendo o promotor de prolamina usado para transformação de *Oryza sativa*. Este vetor contém como elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; uma gaveta de expressão de marcador triável; e uma gaveta Gateway pretendida para recombinação in vivo LR com a seqüência de interesse já clonada no clone de entrada. Um promotor de prolamina de arroz (SEQ ID NO: 195) para expressão endosperma específica foi localizado antes desta gaveta Gateway.

Depois da etapa de recombinação LR, o vetor de expressão resultante mostrado em Figura 22 foi transformado em linhagem LBA4044 de *Agrobacterium* e subseqüentemente para plantas de *Oryza sativa*. Plantas transformadas de arroz foram permitidas crescer e foram então avaliadas para os parâmetros descritos abaixo. Para transformação de outras safras veja Exemplo 40.

Exemplo 32: Avaliação e Resultados

Aproximadamente 15 a 20 transformantes T0 de arroz

independentes foram gerados. Os transformantes primários foram transferidos de uma câmera de safra de tecido para uma casa de vegetação para crescimento e colheita de semente T1. Sete eventos, dos quais a progênie T1 segregou 3:1 para presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 plântulas T1 contendo o transgene (hetero e homo-zigotas) e aproximadamente 10 plântulas T1 carecendo do transgene (nulizigotos) foram selecionadas monitorando expressão de marcador visual.

32.1 Análise estatística: teste F

Uma ANOVA de dois fatores (análise de variantes) foi usada como um modelo estatístico para a avaliação geral de características fenotípicas vegetais. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformadas com o gene da presente invenção. O teste F foi realizado para verificar por um efeito do gene sobre todos os eventos de transformação e por um efeito geral do gene (também referidas como a efeito gênico global). O limiar para significância para um verdadeiro efeito gênico global foi ajustado em um nível de probabilidade de 5% para o teste F. Um valor significante de teste F aponta para um efeito gênico, significando que não é apenas a presença ou posição do gene que está causando as diferenças em fenótipo.

32.2 Mensurações de parâmetros relacionados a sementes

As panículas primárias maduras foram coletadas, contadas, ensacadas, etiquetadas com código de barras e então secas por três dias em um forno a 37°C. As panículas foram então trilhadas e todas as sementes foram coletadas e contadas. As cascas cheias foram separadas daquelas vazias usando um dispositivo que sopra ar. As cascas vazias foram descartadas e a fração remanescente foi contada novamente. As cascas cheias foram pesadas em uma balança analítica. O número de grãos cheios foi determinado contando p número de cascas cheias que permaneceram depois da etapa de

separação. A produção total de sementes foi medida pesando todas as cascas cheias coletadas de uma planta. Número total de sementes por planta foi medido contando o número de cascas coletadas de uma planta. Peso de Mil Sementes (TKW) foi extrapolado a partir do número de grãos cheios contadas e seu peso total. O índice de colheita (HI) foi expresso como uma proporção entre a produção total de sementes e a parte aérea (mm^2), multiplicada por um fator 10^6. O número total de flores por panícula foi expresso como uma proporção entre o número total de sementes e o número de panículas primárias maduras. A taxa de enchimento de sementes foi expressa como uma% do número de grãos cheios sobre o número total de sementes (ou floretes).

Tabela W: Dados comparativos para mostrar a diferença em produção de sementes obtida usando um promotor endosperma específico (prolamina) comparado com um promotor raiz específico (promotor RCc3)

CDS3129 = SEQ ID NO 1/2			Peso Total	Número de grãos cheios	Flores por panícula	Índice de colheita	Número total de sementes
promotor raiz específico	4 eventos	% de diferença	9,20%	10,00%	0,00%	11,20%	0,00%
		No de p de teste F	0,0451	0,0266	0,7744	0,0027	0,4410
promotor endosperma específico	5 eventos	% de diferença	46,10%	48,50%	16,40%	39,40%	19,50%
		No de p de teste F	< 0,0001	< 0,0001	< 0,0001	< 0,0001	< 0,0001

A Tabela mostra a% de diferença em vários parâmetros para plantas transgênicas comparadas com plantas de controle correspondentes (nulizigotos); também mostrado na Tabela é o valor de p do teste F que indica o efeito geral do gene. Como mostrado na Tabela, vários parâmetros de produção de sementes foram aumentados em plantas expressando um ácido nucleico codificando AT-hook (SEQ ID NO: 152) sob o controle de um promotor endosperma específico, ao passo que nenhum aumento (de fato uma diminuição significante) foi obtido para plantas expressando o mesmo transgene sob o controle de um promotor raiz específico em plantas transgênicas.

EXEMPLO F: fatores de transcrição DOF

Exemplo 33: Clonagem de gene de fator de transcrição DOF de Arabidopsis thaliana (SEQ ID NO: 198)

O gene de fator de transcrição *DOF* de *Arabidopsis thaliana* foi amplificado por PCR usando como modelo uma biblioteca de cDNA de plântula de *Arabidopsis thaliana* (Invitrogen, Paisley, UK). Depois de transcrição reversa de RNA extraído de plântulas, os cDNAs foram clonados em pCMV Sport 6.0. Tamanho médio de inserto do banco foi de 1,5 kb e número original de clones foi de $1,59x10^7$ cfu. Título original foi determinado como sendo $9,6x10^5$ cfu/ml depois de primeira amplificação de $6x10^{11}$ cfu/ml. Depois de extração de plasmídeo, 200 ng de modelo foram usados em uma mistura de PCR de 50 µl. Iniciador (SEQ ID NO: 223) (sentido iniciador AttB1: 5' ggggacaagtttgtacaaaaaa gcaggcttaaacaatgggtggatcgatggc 3') e (SEQ ID NO: 224) (iniciador AttB2 complementar reverso: 5' ggggaccactttgtacaagaaagctgggtcgttaatgatccgacaaaaca 3'), os quais incluem os sítios AttB para recombinação Gateway, foram usados para amplificação por PCR. PCR foi realizado usando DNA polimerase Hifi Taq em condições padrão. Um fragmento de PCR (incluindo sítios attB; de início a término) foi amplificado e purificado também usando métodos padrão. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada durante as quais o fragmento de PCR recombinado *in vivo* com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada". Plasmídeo pDONR201 foi adquirido a partir de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 33a: Construção de Vetor

O clone de entrada foi subseqüentemente usado em uma reação LR com um vetor de destinação contendo GOS2 usado para transformação de *Oryza sativa*. Este vetor continha como elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal;

uma gaveta de expressão de marcador triável; e uma gaveta Gateway pretendida para recombinação in vivo LR com a seqüência de interesse já clonada no clone de entrada. Um promotor GOS2 de arroz (SEQ ID NO: 225) para expressão constitutiva foi localizado antes desta gaveta Gateway.

Depois da etapa de recombinação LR, o vetor de expressão resultante mostrado em Figura 26 foi transformado em linhagem LBA4044 de *Agrobacterium* e subseqüentemente para plantas de *Oryza sativa*. Plantas transformadas de arroz foram permitidas crescer e foram então avaliadas para os parâmetros descritos abaixo. Para transformação de outras safras veja Exemplo 40.

Exemplo 34: Clonagem de gene de fator de transcrição DOF de Arabidopsis thaliana (SEQ ID NO: 226)

O gene de fator de transcrição *DOF* de *Arabidopsis thaliana* foi amplificado por PCR usando como modelo uma biblioteca de cDNA de plântula de *Arabidopsis thaliana* (Invitrogen, Paisley, UK). Depois de transcrição reversa de RNA extraído de plântulas, os cDNAs foram clonados em pCMV Sport 6.0. Tamanho médio de inserto do banco foi de 1,5 kb e número original de clones foi de $1,59x10^7$ cfu. Título original foi determinado como sendo $9,6x10^5$ cfu/ml depois de primeira amplificação de $6x10^{11}$ cfu/ml. Depois de extração de plasmídeo, 200 ng de modelo foram usados em uma mistura de PCR de 50 µl. Iniciador (SEQ ID NO: 256) (iniciador AttB1 sentido: 5' ggggacaagtttgtacaaaaaa gcaggcttaaacaatgatgatggagactagagatc3') and (SEQ ID NO: 257) (iniciador AttB2 complementar reverso: 5' ggggaccactttgtacaagaaagctgggtcatatgtaactctaaatctgttca3'), os quais incluem os sítios AttB para recombinação Gateway, foram usados para amplificação por PCR. PCR foi realizado usando DNA polimerase Hifi Taq em condições padrão. Um fragmento de PCR (incluindo sítios attB; de início a término) foi amplificado e purificado também usando métodos padrão. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada durante a qual o

fragmento de PCR recombinado *in vivo* com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada". Plasmídeo pDONR201 foi adquirido a partir de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 34a: Construção de Vetor

O clone de entrada foi subseqüentemente usado em uma reação LR com um vetor de destinação contendo prolamina usado para transformação de *Oryza sativa*. Este vetor contém como elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; uma gaveta de expressão de marcador triável; e uma gaveta Gateway pretendida para recombinação in vivo LR com a seqüência de interesse já clonada no clone de entrada. Um promotor de prolamina de arroz (SEQ ID NO: 258) para expressão semente específica foi localizado antes desta gaveta Gateway.

Depois da etapa de recombinação LR, o vetor de expressão resultante mostrado em Figura 27 foi transformado em linhagem LBA4044 de *Agrobacterium* e subseqüentemente para plantas de *Oryza sativa*. Plantas transformadas de arroz foram permitidas crescer e foram então avaliadas para os parâmetros descritos abaixo. Para transformação de outras safras veja Exemplo 40.

Exemplo 35: Avaliação e Resultados

Aproximadamente 15 a 20 transformantes T0 de arroz independentes foram gerados. Os transformantes primários foram transferidos de uma câmera de safra de tecido para uma casa de vegetação para crescimento e colheita de semente T1. Sete eventos, dos quais a progênie T1 segregou 3:1 para presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 plântulas T1 contendo o transgene (hetero e homo-zigotas) e aproximadamente 10 plântulas T1 carecendo do transgene (nulizigotos) foram selecionadas monitorando expressão de marcador visual. Aproximadamente 4 eventos T1 foram adicionalmente

avaliados na geração T2 seguindo o mesmo procedimento de avaliação que para a geração T1 mas com mais indivíduos por evento.

Plantas de cinco eventos foram crescidas em condições normais até o estágio de espigamento. Umidade do solo foi monitorada continuamente usando sensores de umidade inseridos nos potes de diversas plantas de controle não transgênicas escolhidas aleatoriamente. Em uma primeira fase, os potes foram saturados para um valor máximo de 60% para reduzir a variabilidade de pote para pote. Uma vez que os potes foram saturados, irrigação foi retida até que um teor de umidade no solo de abaixo de 20% fosse obtido. As plantas foram então re-irrigadas até umidade no solo ter atingido o nível máximo de 60% novamente. As plantas foram então representadas para avaliar os seguintes parâmetros relacionados a raízes e relacionados a sementes.

Parâmetros relacionados a raízes

Plantas foram crescidas em potes especialmente projetados com fundos transparentes para permitir visualização das raízes. Uma câmera digital registrou imagens através do fundo do pote durante crescimento vegetal. Características de raiz tal como área total projetada (a qual pode ser correlacionada a volume total de raiz), diâmetro e comprimento médio de raízes acima de um certo limiar de espessura (comprimento de raízes grossas, ou comprimento de raízes finas) foram deduzidos a partir da imagem gerada usando aplicativo computacional apropriado.

Mensurações de parâmetros relacionados a sementes

As panículas primárias maduras foram coletadas, contadas, ensacadas, etiquetadas com código de barras e então secas por três dias em um forno a 37°C. As panículas foram então trilhadas e todas as sementes foram coletadas e contadas. As cascas cheias foram separadas daquelas vazias usando um dispositivo que sopra ar. As cascas vazias foram descartadas e a fração remanescente foi contada novamente. As cascas cheias foram pesadas

em uma balança analítica. O número de grãos cheios foi determinado contando o número de cascas cheias que permaneceram depois da etapa de separação. A produção total de sementes foi medida pesando todas as cascas cheias coletadas de uma planta. Número total de sementes por planta foi medido contando o número de cascas coletadas de uma planta. Peso de Mil Sementes (TKW) foi extrapolado a partir do número de grãos cheios contadas e seu peso total. O índice de colheita (HI) na presente invenção é definido como a proporção entre a produção total de sementes e a parte aérea (mm^2), multiplicada por um fator 10^6. O número total de flores por panícula como definido na presente invenção é a proporção entre o número total de sementes e o número de panículas primárias maduras. A taxa de enchimento de sementes como definida na presente invenção é a proporção (expressa como uma%) do número de grãos cheios sobre o número total de sementes (ou floretes).

Análise estatística: teste F

Uma ANOVA de dois fatores (análise de variantes) foi usada como um modelo estatístico para a avaliação geral de características fenotípicas vegetais. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformadas com o gene da presente invenção. O teste F foi realizado para verificar por um efeito do gene sobre todos os eventos de transformação e por um efeito geral do gene, também conhecido como um efeito gênico global. O limiar para significância para um verdadeiro efeito gênico global foi ajustado em um nível de probabilidade de 5% para o teste F. Um valor significante de teste F aponta para um efeito gênico, significando que não é apenas a presença ou posição do gene que está causando as diferenças em fenótipo.

Tabela X abaixo mostra os resultados da avaliação de T2 para plantas transgênicas expressando um ácido nucleico codificando um fator de transcrição DOF sob o controle de um promotor GOS2 e os resultados da

avaliação de T2 para plantas transgênicas expressando um ácido nucleico codificando um fator de transcrição DOF sob o controle de um promotor de prolamina. Embora não mostrados, resultados comparáveis foram obtidos para plantas de T1. O valor de p do teste F é mostrado para os parâmetros listados na Tabela, assim como a porcentagem de diferença entre plantas transgênicas versos nulizigotos.

Tabela X: Resultados de Avaliação de T2

	Peso total de semente		Número de Grãos cheios		Índice de colheita		Taxa de enchimento	
	% de Diferença	Valor de P	% de Diferença	Valor de P	% de Diferença	Valor de P	% de Diferença	Valor de P
Exemplo 1 (GOS2) [Referência interna CD12148; CDS3325]	34,0%	0,0070	36,0%	0,0061	32%	0,0032	21%	0,0011
Exemplo 2 (prolamina) [Referência interna CD12225; CDS3319]	18%	0,0348	17%	0,0385	19%	0,035	12%	0,1373

Em adição aos parâmetros relacionados a sementes descritos acima, os seguintes parâmetros de raiz também foram aumentados em plantas transgênicas comparadas com nulizigotos: 14% de aumento em biomassa total de raiz, 7% de aumento em número de raízes grossas (limiar interno), 36% de aumento em número de raízes finas (limiar interno) e uns 8% de aumento em diâmetro médio de raízes.

Os resultados mencionados acima foram obtidos em condições de estresse brando por seca; resultados similares seriam esperados em condições normais ou sem estresse.

EXEMPLO G: CKI

Exemplo 36: Clonagem de um gene de Oryza sativa codificando um polipeptídeo CKI4

O gene de *Oryza sativa* codificando um polipeptídeo CKI4 foi amplificado por PCR usando como modelo uma biblioteca de cDNA de safra celular em suspensão de *Oryza sativa* clonada no vetor pAD-Gal4-2.1 de kit

HybriZAP-2.1 (Stratagene, La Jolla, Califórnia USA), de acordo com as instruções de fabricante. Tamanho médio de inserto do banco foi de 1,5 kb e o número original de clones foi da ordem de $2x10^6$ pfu. Título original foi determinado como sendo $4x10^6$ pfu/ml e depois da primeira amplificação de 10^{10} pfu/ml. Depois de extração de plasmídeo, 200 ng de modelo foram usados em uma mistura de PCR de 50 µl. Iniciadores (SEQ ID NO: 284; sentido, códon de início em negrito, sítio AttB1 em itálico: 5'-*GGGGACAAGTTTGTACAAAAAAGCAGG*CTTCACA**ATG**GGCAAGTACA TGCGCAAGGCC-3') e (SEQ ID NO: 285; reverso, complementar, sítio AttB2 em itálico: 5'- *GGGGACCACTTTGTACAAGAAAGCTGGG T*GGAGCAGAGAGGTCCATGGTGCCC-3'), os quais incluem os sítios AttB para recombinação Gateway, foram usados para amplificação por PCR. PCR foi realizado usando DNA polimerase Hifi Taq em condições padrão. Um fragmento de PCR de 662 bp (incluindo sítios attB; 585 bp de início a término) foi amplificado e purificado também usando métodos padrão. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante a qual o fragmento de PCR se recombina *in vivo* com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada". Plasmídeo pDONR201 foi adquirido a partir de Invitrogen, como parte da tecnologia Gateway®.

Exemplo 37: Construção de Vetor

O clone de entrada foi subseqüentemente usado em uma reação LR com um vetor de destinação usado para transformação de *Oryza sativa*. Este vetor contém como elementos funcionais dentro dos limites de T-DNA: um marcador selecionável vegetal; uma gaveta de expressão de marcador triável; e duas gavetas Gateway em orientação oposta pretendidas para recombinação in vivo LR com a seqüência de interesse já clonada no clone de entrada. As duas gavetas Gateway foram separadas por DNA não codificante (neste caso um fragmento de 315 bp de uma região de

acoplamento de matriz (MAR) de tabaco, referência NCBI U67919, fragmento de 774 a 1088 bp), para promover formação de uma estrutura em forma de grampo do mRNA depois de transcrição. Um promotor de prolamina RP6 de arroz (SEQ ID NO: 281) para expressão endosperma específica foi localizado antes da primeira gaveta Gateway, em orientação oposta com respeito ao promotor.

O clone de entrada também foi usado em uma reação LR com outro vetor de destinação usado para transformação de *Oryza sativa*. Este vetor foi idêntico àquele descrito acima, exceto que o promotor de prolamina RP6 foi substituído por promotor de beta-expansina de arroz de SEQ ID NO: 282.

Depois da etapa de recombinação LR, os dois vetores de expressão resultantes (Figura 32 para ambos vetores) foram transformados em linhagem LBA4044 de *Agrobacterium* e subseqüentemente para plantas de *Oryza sativa*. Plantas transformadas de arroz foram permitidas crescer e foram então avaliadas para os parâmetros descritos em Exemplos 38 e 39. Para transformação de outras safras veja Exemplo 40.

Exemplo 38: Descrição do procedimento de avaliação fenotípica

Aproximadamente 15 a 20 transformantes T0 de arroz independentes foram gerados. Os transformantes primários foram transferidos de uma câmera de safra de tecido para uma casa de vegetação para crescimento e colheita de semente T1. Quatro a cinco eventos, dos quais a progênie T1 segregou 3:1 para presença/ausência do transgene, foram retidos. Para cada um destes eventos, aproximadamente 10 plântulas T1 contendo o transgene (hetero e homozigotas) e aproximadamente 10 plântulas T1 carecendo do transgene (nulizigotos) foram selecionadas monitorando expressão de marcador visual. As plantas transgênicas e as plantas de controle adequadas foram crescidas lado a lado em posições aleatórias. Do estágio de semeadura até o estágio de maturidade as plantas foram passadas diversas

vezes através de uma cabine de formação de imagem digital. Em cada momento imagens digitais (2048x1536 pixéis, 16 milhões de cores) foram tiradas de cada planta de pelo menos 6 ângulos diferentes.

Os mesmos eventos avaliados em T1 foram adicionalmente avaliados na geração T2 seguindo o mesmo procedimento de avaliação que para a geração T1.

Mensurações de parâmetros relacionados a sementes

As panículas primárias maduras foram coletadas, contadas, ensacadas, etiquetadas com código de barras e então secas por três dias em um forno a 37°C. As panículas foram então trilhadas e todas as sementes foram coletadas e contadas. As cascas cheias foram separadas daquelas vazias usando um dispositivo que sopra ar. As cascas vazias foram descartadas e a fração remanescente foi contada novamente. As cascas cheias foram pesadas em uma balança analítica. O número de grãos cheios foi determinado contando o número de cascas cheias que permaneceram depois da etapa de separação. A produção total de sementes foi medida pesando todas as cascas cheias coletadas de uma planta. Número total de sementes por planta foi medido contando o número de cascas coletadas de uma planta. O índice de colheita (HI) na presente invenção é definido como a proporção entre a produção total de sementes e a parte aérea (mm^2), multiplicada por um fator 10^6. O número total de flores por panícula como definido na presente invenção é a proporção entre o número total de sementes e o número de panículas primárias maduras. A taxa de enchimento de sementes como definida na presente invenção é a proporção (expressa como uma%) do número de grãos cheios sobre o número total de sementes (ou floretes).

Análise estatística: teste F

Uma ANOVA de dois fatores (análise de variantes) foi usada como um modelo estatístico para a avaliação geral de características fenotípicas vegetais. Um teste F foi realizado em todos os parâmetros

medidos de todas as plantas de todos os eventos transformadas com o gene da presente invenção. O teste F foi realizado para verificar por um efeito do gene sobre todos os eventos de transformação e por um efeito geral do gene, também conhecido como um efeito gênico global. O limiar para significância para um verdadeiro efeito gênico global foi ajustado em um nível de probabilidade de 5% para o teste F. Um valor significante de teste F aponta para um efeito gênico, significando que não é apenas a presença ou posição do gene que está causando as diferenças em fenótipo.

Exemplo 39: Resultados da avaliação de planta transgênica de arroz com expressão de CKI4 reduzida no endosperma

Os resultados de mensuração de avaliação (produção de sementes, número de grãos cheios, número total de sementes e flores por panícula) para plantas transgênicas com expressão de CKI4 reduzida no endosperma são apresentados em Tabela Y abaixo. O número de plantas com um aumento em um parâmetro, o aumento percentual médio assim como o valor de P da geração T2 são mostrados, e comparados com resultados obtidos com plantas transgênicas com expressão de CKI4 reduzida usando um promotor de beta expansina para expressão preferencial em tecido de broto.

Os resultados mostram que expressão reduzida de CKI4 no endosperma gera plantas com peso de semente, número de grãos cheios, total número de sementes e flores por panícula significantemente aumentados, comparadas com nulizigotos e comparadas com plantas transgênicas com expressão preferencialmente reduzida de CKI4 em tecido de broto (usando um promotor de beta expansina).

Tabela Y: Resultados de mensuração de avaliação para plantas transgênicas com reduzida expressão de CKI4 no endosperma

	Promotor endosperma específico			Promotor broto específico
	Número de plantas com parâmetro aumentado	**% de aumento**	**Valor de P**	**% de aumento**
Produção de sementes	4 de 4	17%	0,006	<17%
Número de grãos cheios	4 de 4	19%	0,0018	<19%
Número total de sementes	4 de 4	16%	0,0014	<16%
Flores por panícula	3 de 4	9%	0,00 6	<9%

Exemplo 40: Transformação de Milho, Trigo, Soja, Colza e Alfafa

Transformação de milho

Transformação de milho (*Zea mays*) é realizada com uma modificação do método descrito por Ishida et al. (1996) Nature Biotech 14(6): 745-50. Transformação é dependente de genótipo em milho e apenas genótipos específicos são receptivos para transformação e regeneração. A linhagem natural A188 (University of Minnesota) ou híbridos com A188 como um ancestral são boas fontes de material doador para transformação, mas outros genótipos também podem ser usados com sucesso. Espigas são coletadas a partir de planta de milho aproximadamente 11 dias depois de polinização (DAP) quando o comprimento do embrião imaturo é de cerca de 1 a 1,2 mm. Embriões imaturos são co-cultivados com *Agrobacterium tumefaciens* contendo o vetor de expressão, e plantas transgênicas são recuperadas através de organogênese. Embriões excisados são crescidos em meio de indução de calo, então meio de regeneração de milho, 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 2-3 semanas, ou até que brotos se desenvolvam. Os brotos verdes são transferidos de cada embrião para meio de enraizamento de milho e incubados a 25 °C por 2-3 semanas, até que raízes se desenvolvam. Os brotos com raiz são transplantados para solo na casa de vegetação. Sementes T1 são

produzidas a partir de plantas que exibem tolerância ao agente de seleção e que contêm uma única cópia do inserto de T-DNA.

Transformação de trigo

Transformação de trigo é realizada com o método descrito por Ishida et al. (1996) Nature Biotech 14(6): 745-50. A cultivar Bobwhite (disponível a partir de CIMMYT, Mexico) é comumente usada em transformação. Embriões imaturos são co-cultivados com *Agrobacterium tumefaciens* contendo o vetor de expressão, e plantas transgênicas são recuperadas através de organogênese. Depois de incubação com *Agrobacterium*, os embriões são crescidos *in vitro* em meio de indução de calo, então 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 2-3 semanas, ou até que brotos se desenvolvam. Os brotos verdes são transferidos de cada embrião para meio de enraizamento e incubados a 25 °C por 2-3 semanas, até que raízes se desenvolvam. Os brotos com raiz são transplantados para solo na casa de vegetação. Sementes T1 são produzidas a partir de plantas que exibem tolerância ao agente de seleção e que contêm uma única cópia do inserto de T-DNA.

Transformação de soja

Soja é transformada de acordo com uma modificação do método descrito na Patente Norte-Americana de Texas A&M 5.164.310. Diversas variedades comerciais de soja são receptivas para transformação por este método. A cultivar Jack (disponível a partir da Illinois Seed foundation) é comumente usada para transformação. Sementes de soja são esterilizadas para semeadura *in vitro*. O hipocótilo, a radícula e um cotilédone são excisados a partir de plântulas jovens de sete dias de idade. O epicótilo e o cotilédone remanescente são adicionalmente crescidos para desenvolver nós axilares. Estes nós axilares são excisados e incubados com *Agrobacterium tumefaciens*

contendo o vetor de expressão. Depois do tratamento de co-cultivo, os explantes são lavados e transferidos para meio de seleção. Brotos regenerados são excisados e colocados em um meio de prolongamento de broto. Brotos menores do que 1 cm são colocados em meio de enraizamento até que raízes se desenvolvam. Os brotos com raiz são transplantados para solo na casa de vegetação. Sementes T1 são produzidas a partir de plantas que exibem tolerância ao agente de seleção e que contêm uma única cópia do inserto de T-DNA.

Transformação de colza/canola

Pecíolos cotiledonares e hipocótilos de plântulas jovens de 5-6 dias de idade são usados como explantes para safra de tecido e transformados de acordo com Babic et al. (1998, Plant Cell Rep 17: 183-188). A cultivar comercial Westar (Agriculture Canada) é a variedade padrão usada para transformação, mas outras variedades também podem ser usadas. Sementes de canola são superfície-esterilizadas para semeadura *in vitro*. Os explantes de pecíolo de cotilédone com o cotilédone acoplado são excisados das plântulas *in vitro*, e inoculados com *Agrobacterium* (contendo o vetor de expressão) banhando a extremidade de corte do explante de pecíolo na suspensão bacteriana. Os explantes são então cultivados por 2 dias em meio MSBAP-3 contendo 3 mg/l de BAP, 3% de sacarose, 0,7% de pH ytagar a 23°C, 16 h de luz. Depois de dois dias de co-cultivo com *Agrobacterium*, os explantes de pecíolo são transferidos para meio MSBAP-3 contendo 3 mg/l de BAP, cefotaxima, carbenicilina, ou timentina (300 mg/l) por 7 dias, e então cultivados em meio MSBAP-3 com cefotaxima, carbenicilina, ou timentina e agente de seleção até regeneração de broto. Quando os brotos estão com 5 – 10 mm de comprimento, eles são cortados e transferidos para meio de prolongamento de broto (MSBAP-0,5, contendo 0,5 mg/l de BAP). Brotos de cerca de 2 cm de comprimento são transferidos para o meio de enraizamento (MS0) para indução de raiz. Os brotos com raiz são transplantados para solo

na casa de vegetação. Sementes T1 são produzidas a partir de plantas que exibem tolerância ao agente de seleção e que contêm uma única cópia do inserto de T-DNA.

Transformação de alfafa

Um clone regenerante de alfafa (*Medicago sativa*) é transformado usando o método de (McKersie et al., 1999 Plant physiol 119: 839–847). Regeneração e transformação de alfafa é dependente de genótipo e portanto uma planta regenerante é requerida. Métodos para obter plantas regenerantes foram descritos. Por exemplo, estas podem ser selecionadas a partir da cultivar Rangelander (Agriculture Canada) ou qualquer outra variedade comercial de alfafa como descrito por Brown DCW and A Atanassov (1985. Plant Cell Tissue Organ Culture 4: 111-112). Alternativamente, a variedade RA3 (University of Wisconsin) foi selecionada para uso em safra de tecido (Walker et al., 1978 Am J Bot 65:654-659). Explantes de pecíolo são co-cultivados com uma safra noturna de *Agrobacterium tumefaciens* C58C1 pMP90 (McKersie et al., 1999 Plant physiol 119: 839–847) ou LBA4404 contendo o vetor de expressão. Os explantes são co-cultivados por 3 d no escuro em meio de indução SH contendo 288 mg/ L de Pro, 53 mg/ L de tioprolina, 4,35 g/ L de K_2SO_4, e 100 µm de acetoseringona. Os explantes são lavados em meio Murashige-Skoog com meia força (Murashige and Skoog, 1962) e colocados no mesmo meio de indução SH sem acetoseringona mas com um agente de seleção adequado e antibiótico adequado para inibir crescimento de *Agrobacterium*. Depois de diversas semanas, embriões somáticos são transferidos para meio de desenvolvimento BOi2Y não contendo nenhum regulador de crescimento, nenhum antibiótico, e 50 g/L de sacarose. Embriões somáticos são subseqüentemente germinados em meio Murashige-Skoog com meia força. Plântulas enraizadas foram transplantadas em potes e crescidas em uma casa de vegetação. Sementes T1 são produzidas a partir de plantas que exibem

tolerância ao agente de seleção e que contêm uma única cópia do inserto de T-DNA.

LISTAGEM DE SEQÜÊNCIAS

<110> CropDesign N.V.
<120> "PLANTAS TENDO CARACTERÍSTICAS DE CRESCIMENTO MELHORADAS E MÉTODOS PARA FAZER AS MESMAS"
<130> PF57958
<150> EP 05111597.0
<151> 2005-12-01
<150> US 60/742,352
<151> 2005-12-05
<150> EP 05111691.1
<151> 2005-12-05
<150> EP 05111786.9
<151> 2005-12-07
<150> US 60/748,903
<151> 2005-12-08
<150> US 60/749,219
<151> 2005-12-09
<150> EP 05111996.4
<151> 2005-12-12
<150> US 60/750,143
<151> 2005-12-14
<150> EP 05112562.3
<151> 2005-12-21
<150> US 60/753,650
<151> 2005-12-23
<150> EP 05113110.0
<151> 2005-12-30
<150> EP 05113111.8
<151> 2005-12-30
<150> US 60/756,086
<151> 2006-01-04
<150> US 60/756,042
<151> 2006-01-04
<160> 285
<170> PatentIn versão 3.3
<210> 1
<211> 353
<212> DNA
<213> Oryza sativa
<400> 1
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> 2
<211> 105
<212> PRT
<213> Oryza sativa
<400> 2
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> 3
<211> 56
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm08170
<400> 3
ggggacaagt ttgtacaaaa aagcaggctt aaacaatgga aggtgtaggt gctagg 56

<210> 4
<211> 51
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm08171
<400> 4
ggggaccact ttgtacaaga aagctgggtc aaaaacaaaa ataaattccc c 51

<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 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 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> 6

<211> 3
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo conservado 1a
<400> 6
Tyr Phe Ser
1
<210> 7
<211> 3
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo conservado 1b
<400> 7
Tyr Phe Thr
1
<210> 8
<211> 3
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo conservado 1c
<400> 8
Tyr Phe Gly
1
<210> 9
<211> 3
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo conservado 1d
<400> 9
Tyr Leu Gly
1
<210> 10
<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> 10
Val Leu Ala Phe Met Pro Thr
1 5
<210> 11
<211> 3
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo conservado 3
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir="Pro"
<400> 11

Ser Tyr Leu
1

<210> 12
<211> 178
<212> PRT
<213> Oryza sativa
<400> 12

```
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> 13
<211> 126
<212> PRT
<213> Oryza sativa
<400> 13

```
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> 14
<211> 105
<212> PRT
<213> Zea mays
<220>
<221> INSEGURO
<222> (65)..(75)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 14

```
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> 15
<211> 110
<212> PRT
<213> Triticum aestivum
<400> 15
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> 16
<211> 110
<212> PRT
<213> Hordeum vulgare
<400> 16
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> 17
<211> 99
<212> PRT
<213> Saccharum officinarum
<220>
<221> INSEGURO
<222> (62)..(62)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 17
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> 18
<211> 99
<212> PRT
<213> Saccharum officinarum
<400> 18

```
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> 19
<211> 107
<212> PRT
<213> Sorghum bicolor
<400> 19

```
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> 20
<211> 130
<212> PRT
<213> Arabidopsis thaliana
<400> 20

```
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> 21
<211> 135
<212> PRT
<213> Arabidopsis thaliana
<400> 21
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> 22
<211> 139
<212> PRT
<213> Vitis vinifera
<400> 22
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> 23
<211> 126
<212> PRT
<213> Citrus sp.
<220>
<221> INSEGURO
<222> (103)..(103)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 23

```
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> 24
<211> 118
<212> PRT
<213> Lycorpersicon esculentum
<400> 24
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> 25
<211> 87
<212> PRT
<213> Lycorpersicon esculentum
<400> 25
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> 26
<211> 106
<212> PRT
<213> Arabidopsis thaliana
<400> 26
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> 27
<211> 537
<212> DNA
<213> Oryza sativa
<400> 27
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> 28
<211> 538
<212> DNA
<213> Zea mays
<220>
<221> misc_feature
<222> (177)..(208)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (493)..(493)
<223> n is a, c, g, ou t
<400> 28
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> 29
<211> 578
<212> DNA
<213> Vitis vinifera
<400> 29
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> 30
<211> 704

<212> DNA
<213> Citrus reticulata
<220>
<221> misc_feature
<222> (619)..(619)
<223> n is a, c, g, ou t
<400> 30
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> 31
<211> 382
<212> DNA
<213> Lycorpersicon esculentum
<400> 31
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> 32
<211> 624
<212> DNA
<213> Lycorpersicon esculentum
<400> 32
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> 33
<211> 1130
<212> DNA
<213> Oryza sativa
<400> 33
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 ttcccccctc tcacaagaga gagagcgcca 1020
gactctcccc aggtgaggtg agaccagtct ttttgctcga ttcgacgcgc ctttcacgcc 1080
gcctcgcgcg gatctgaccg cttccctcgc ccttctcgca ggattcagcc 1130

<210> 34
<211> 77
<212> PRT
<213> Medicago sativa
<400> 34

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> 35
<211> 78
<212> PRT
<213> Medicago sativa
<400> 35

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> 36
<211> 613
<212> DNA
<213> Hordeum vulgare
<400> 36

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> 37
<211> 617
<212> DNA
<213> Saccharum officinarum
<220>
<221> misc_feature
<222> (451)..(451)
<223> n is a, c, g, ou t
<400> 37

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> 38
<211> 638
<212> DNA
<213> Saccharum officinarum
<400> 38
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> 39
<211> 778
<212> DNA
<213> Sorghum bicolor
<400> 39
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> 40
<211> 732
<212> DNA
<213> Arabidopsis thaliana
<400> 40
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> 41
<211> 891
<212> DNA
<213> Arabidopsis thaliana
<400> 41
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> 42
<211> 732
<212> DNA
<213> Arabidopsis thaliana
<400> 42
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> 43
<211> 519
<212> DNA
<213> Zea mays
<400> 43
atgcgaggtt cgggatttaa gatgttcggc tttaggggca gttcttctga agcaggggac 60
gggcgattcg accaccggag ctcagatctg attacaaaac gttcagaaaa cacaaggcgt 120
tctcacaccg cctttcactt cttgcttact ttggcaacca ctcactgcga ctggtctcca 180
cctccaccta caccaagaac acatggcaag ccgatctacg cgatggaagg aggggcggca 240
atacaaagga ggaatgccgt gaagcggcat ctgcagcagc gtcagcagga ggcggatttc 300
ctcgacaaga aggtcatcgc gtccacctac ttcagcatcg gggcgttcct cgtgctcgcc 360
tgcctcaccg tctcgctgct gatactgccg ctggtgctgc ctcccctgcc gccgccgccg 420
tcgctgctgt tgtggctgcc ggtctgcctg ctcgtcttgc tggttgtact ggccttcatg 480
ccgacagatg tgcgcagcat ggcctcctct tacctgtaa 519
<210> 44
<211> 98
<212> PRT
<213> Zea mays
<400> 44
Met Glu Gly Gly Ala Ala Ile Gln Arg Arg Asn Ala Val Lys Arg His
1 5 10 15
Leu Gln Gln Arg Gln Gln Glu Ala Asp Phe Leu Asp Lys Lys Val Ile
20 25 30
Ala Ser Thr Tyr Phe Ser Ile Gly Ala Phe Leu Val Leu Ala Cys Leu
35 40 45

Thr Val Ser Leu Leu Ile Leu Pro Leu Val Leu Pro Pro Leu Pro Pro
50 55 60
Pro Pro Ser Leu Leu Leu Trp Leu Pro Val Cys Leu Leu Val Leu Leu
65 70 75 80
Val Val Leu Ala Phe Met Pro Thr Asp Val Arg Ser Met Ala Ser Ser
85 90 95
Tyr Leu

<210> 45
<211> 2745
<212> DNA
<213> Arabidopsis thaliana
<400> 45
atgagcaatg ctttttata atgccaactt tgtacaaaaa agcaggctta aacaatgaaa 60
tctcgagcga ggcagtgtct gctggtttgt ttgctctgtc tttccttaac gaatctctcc 120
tatggagaaa ataagttcag agagcgtaaa gccaccgatg acgagctggg ctaccccgat 180
attgatgaag atgctttatt gaatactcag tgcccgaaaa aattggagct gcgatggcaa 240
actgaagtca cttctagcgt ttatgctaca cccttgattg ctgatattaa cagtgatgga 300
aagcttgaca ttgttgttcc atcttttgtt cattacctcg aagttcttga aggagctgat 360
ggagacaaga tgccaggttg gcctgctttt caccagtcaa atgtgcactc gagtcctctt 420
ctatttgata tcgacaaaga tggtgttaga gaaattgctc tggctaccta caatgccgaa 480
gtgctctttt tcagggtatc aggctttttg atgtcagata agctagaagt gccacgtaga 540
aaagtgcaca agaactggca tgtgggactt aatcctgatc ctgttgaccg ttcacatcct 600
gatgttcatg atgatgtgct tgaggaggaa gctatggcaa tgaagtcatc gaccactcaa 660
acgaatgcaa ctaccacaac accaaatgtt acagtctcga tgaccaaaga agttcatggc 720
gctaattcat atgtgtcaac tcaagaggat caaaagagac cagagaataa tcaaacagaa 780
gctattgtaa agcctactcc agagctacat aattcctcca tggatgctgg agcaaataat 840
ttggcagcaa atgctactac agctggctca agagaaaacc tcaatagaaa tgtaaccacc 900
aatgaggtgg atcaaagcaa aattagtgga gataagaatg aaactgttat taaattaaat 960
actagtacgg gtaattcctc agaaactctg gggacatctg gaaacagtag tacggcagag 1020
acagtaacca aaagtgggag gcgacttctg gaagaggatg gttcgaaaga atctgtggac 1080
agccattcgg acagtaaaga caacagtgag ggtgtccgca tggcgacagt agaaaatgat 1140
ggaggcttag aagctgacgc agattcatcg tttgagttgt tgcgtgagaa tgatgagtta 1200
gctgatgaat acagttatga ttatgacgat tatgttgatg agaaaatgtg gggtgatgag 1260
gaatgggttg aggggcagca cgagaactca gaagattatg tgaatattga cgcccatata 1320
ctatgcactc ctgtaattgc tgacatagac aaagatggag tacaggagat gattgttgct 1380
gtttcctatt tcttcgaccc cgagtactat gataatccag aacatctgaa agagcttggt 1440
ggtatcgaca ttaaaaatta tattgctagt tcaattgtgg ttttcaatct tgatactaaa 1500
caagtcaagt ggatcaaaga gctagatttg agtacggata aagcaaactt ccgtgcttat 1560
atttattctt caccaacggt tgttgatttg gatggcgatg gttatttgga tatccttgtc 1620
ggaacttcct ttggcttatt ctacgccatg gatcatcgtg gaaacatcag agaaaaattc 1680
ccactggaaa tggctgaaat tcaaggagca gtggttgcgg ccgacataaa tgatgatgga 1740
aagattgaac ttgtaactac tgattcacac ggaaatatag cagcatggac cacccaagga 1800
gtggaaattt gggaagcaca tcttaagagc cttgttcccc agggtccttc tataggcgat 1860
gttgatggtg acggacacac ggaggttgtg gttcctacat catcaggaaa catatacgtt 1920
cttagtggca aggatggttc tattgtccgt ccttacccat acaggactca tggaagagtg 1980
atgaaccaac ttcttcttgt tgatctgaac aagcgaggtg agaaaaagaa gggtctcacc 2040
atcgttacta catcctttga cggttacctg tatctcatag atggacccac ctcgtgcact 2100
gacgttgttg acattggcga aacttcatac agcatggtct tggctgataa tgttgacggt 2160
ggagatgatc tcgatcttat tgtctcaact atgaatggaa acgtcttttg cttctcaacg 2220
ccttcttctc accatcccct taaggcttgg agatctagtg atcaaggcag gaacaataag 2280
gccaatcgtt atgatcgtga aggcgttttt gtcacgcatt cgaccagagg tttccgtgat 2340
gaggaaggca aaaacttctg ggctgagatc gagatcgttg ataaatatag atatccatct 2400
ggttcacaag caccctacaa cgttactacg acgctgttgg ttccaggcaa ttaccaggga 2460
gagaggagga taacgcagag ccagatctat gaccgccctg gaaaataccg gataaaacta 2520
ccaactgtgg gagtgagaac aacaggaact gtaatggtgg agatggcaga taagaatgga 2580
ctccatttct cagacgaatt ctcactaact ttccatatgt attactacaa gcttctgaaa 2640
tggcttcttg ttctcccgat gctcgggatg ttcggtctgc tcgtgatact acggcctcaa 2700
gaagccgtgc ctctcccgtc tttttcccgc aacacagatt tatga 2745
<210> 46
<211> 896
<212> PRT
<213> Arabidopsis thaliana

<400> 46

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

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

<210> 47
<211> 53
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm06643
<400> 47
ggggacaagt ttgtacaaaa aagcaggctt aaacaatgaa atctcgagcg agg 53

<210> 48
<211> 52
<212> DNA
<213> Seqüência artificial
<220>

<223> iniciador: prm06644
<400> 48
ggggaccact ttgtacaaga aagctgggtc ctgtttacag atggtaccta gt 52
<210> 49
<211> 4938
<212> DNA
<213> Seqüência artificial
<220>
<223> cassete de expressão
<400> 49
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 acttattttt 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 taataccccct 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 ttcatttaaa tcaactaggg atatcacaag 2220
tttgtacaaa aaagcaggct taaacaatga aatctcgagc gaggcagtgt ctgctggttt 2280
gtttgctctg tctttcctta acgaatctct cctatggaga aaataagttc agagagcgta 2340
aagccaccga tgacgagctg ggctaccccg atattgatga agatgcttta ttgaatactc 2400
agtgcccgaa aaaattggag ctgcgatggc aaactgaagt cacttctagc gtttatgcta 2460
caccccttgat tgctgatatt aacagtgatg gaaagcttga cattgttgtt ccatcttttg 2520
ttcattacct cgaagttctt gaaggagctg atggagacaa gatgccaggt tggcctgctt 2580
ttcaccagtc aaatgtgcac tcgagtcctc ttctatttga tatcgacaaa gatggtgtta 2640
gagaaattgc tctggctacc tacaatgccg aagtgctctt tttcagggta tcaggctttt 2700
tgatgtcaga taagctagaa gtgccacgta gaaaagtgca caagaactgg catgtgggac 2760
ttaatcctga tcctgttgac cgttcacatc ctgatgttca tgatgatgtg cttgaggagg 2820
aagctatggc aatgaagtca tcgaccactc aaacgaatgc aactaccaca acaccaaatg 2880
ttacagtctc gatgaccaaa gaagttcatg gcgctaattc atatgtgtca actcaagagg 2940
atcaaaagag accagagaat aatcaaacag aagctattgt aaagcctact ccagagctac 3000
ataattcctc catggatgct ggagcaaata atttggcagc aaatgctact acagctggct 3060
caagagaaaa cctcaataga aatgtaacca ccaatgaggt ggatcaaagc aaaattagtg 3120
gagataagaa tgaaactgtt attaaattaa atactagtac gggtaattcc tcagaaactc 3180

tggggacatc tggaaacagt agtacggcag agacagtaac caaaagtggg aggcgacttc 3240
tggaagagga tggttcgaaa gaatctgtgg acagccattc ggacagtaaa gacaacagtg 3300
agggtgtccg catggcgaca gtagaaaatg atggaggctt agaagctgac gcagattcat 3360
cgtttgagtt gttgcgtgag aatgatgagt tagctgatga atacagttat gattatgacg 3420
attatgttga tgagaaaatg tggggtgatg aggaatgggt tgaggggcag cacgagaact 3480
cagaagatta tgtgaatatt gacgcccata tactatgcac tcctgtaatt gctgacatag 3540
acaaagatgg agtacaggag atgattgttg ctgtttccta tttcttcgac cccgagtact 3600
atgataatcc agaacatctg aaagagcttg gtggtatcga cattaaaaat tatattgcta 3660
gttcaattgt ggttttcaat cttgatacta aacaagtcaa gtggatcaaa gagctagatt 3720
tgagtacgga taaagcaaac ttccgtgctt atatttattc ttcaccaacg gttgttgatt 3780
tggatggcga tggttatttg gatatccttg tcggaacttc ctttggctta ttctacgcca 3840
tggatcatcg tggaaacatc agagaaaaat tcccactgga aatggctgaa attcaaggag 3900
cagtggttgc ggccgacata aatgatgatg gaaagattga acttgtaact actgattcac 3960
acggaaatat agcagcatgg accacccaag gagtggaaat ttgggaagca catcttaaga 4020
gccttgttcc ccagggtcct tctataggcg atgttgatgg tgacggacac acggaggttg 4080
tggttcctac atcatcagga aacatatacg ttcttagtgg caaggatggt tctattgtcc 4140
gtccttaccc atacaggact catggaagag tgatgaacca acttcttctt gttgatctga 4200
acaagcgagg tgagaaaaag aagggtctca ccatcgttac tacatccttt gacggttacc 4260
tgtatctcat agatggaccc acctcgtgca ctgacgttgt tgacattggc gaaacttcat 4320
acagcatggt cttggctgat aatgttgacg gtggagatga tctcgatctt attgtctcaa 4380
ctatgaatgg aaacgtcttt tgcttctcaa cgccttctcc tcaccatccc cttaaggctt 4440
ggagatctag tgatcaaggc aggaacaata aggccaatcg ttatgatcgt gaaggcgttt 4500
ttgtcacgca ttcgaccaga ggtttccgtg atgaggaagg caaaaacttc tgggctgaga 4560
tcgagatcgt tgataaatat agatatccat ctggttcaca agcaccctac aacgttacta 4620
cgacgctgtt ggttccaggc aattaccagg gagagaggag gataacgcag agccagatct 4680
atgaccgccc tggaaaatac cggataaaac taccaactgt gggagtgaga acaacaggaa 4740
ctgtaatggt ggagatggca gataagaatg gactccattt ctcagacgaa ttctcactaa 4800
ctttccatat gtattactac aagcttctga aatggcttct tgttctcccg atgctcggga 4860
tgttcggtct gctcgtgata ctacggcctc aagaagccgt gcctctcccg tctttttccc 4920
gcaacacaga tttatgat 4938

<210> 50
<211> 10
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo conservado 1
<220>
<221> VARIANTE
<222> (9)..(9)
<223> /substituir="Glu"
<400> 50
Phe Asp Gly Tyr Leu Tyr Leu Ile Asp Gly
1 5 10

<210> 51
<211> 5
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo conservado 2
<220>
<221> INSEGURO
<222> (3)..(4)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> VARIANTE
<222> (5)..(5)
<223> /substituir="Glu"
<400> 51
Asp Gly Xaa Xaa Asp
1 5

<210> 52
<211> 9
<212> PRT

<213> Seqüência artificial
<220>
<223> motivo conservado 3
<220>
<221> INSEGURO
<222> (2)..(2)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> INSEGURO
<222> (4)..(4)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> INSEGURO
<222> (7)..(8)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> VARIANTE
<222> (9)..(9)
<223> /substituir="Glu"
<400> 52
Asp Xaa Asp Xaa Asp Gly Xaa Xaa Asp
1 5
<210> 53
<211> 36
<212> PRT
<213> Seqüência artificial
<220>
<223> seqüência de domínio conservado FG-GAP
<400> 53
Gly Gly Ser Ser Val Ala Ala Gly Asp Leu Asn Gly Asp Gly Arg Pro
1 5 10 15
Asp Leu Val Val Gly Ala Pro Gly Ala Asp Gly Gly Thr Asp Gly Ser
20 25 30
Val Tyr Leu Leu
35
<210> 54
<211> 2449
<212> DNA
<213> Arabidopsis thaliana
<400> 54
attcttcatc gttctctctg ctctggtaag gagattgctg ctacctactg atgcaaatct 60
gccattgata catctgcttg cggttagggc aaagatgagg aaacgcgatt tggctatttt 120
gatgctctct ggatttgcta tattcttcac tcttcagcac gagggtgatt ttgcgttcaa 180
agaagcatgg tttcatttgt atgatgaata cccagtcaaa tacgaagctg atcgtctccc 240
accacctatt gtagctgatc ttaatggtga tggaaagaag gaggttctcg ttgctactaa 300
tgatgccaaa attcaggttc tggagcctca ttccaggcgt gtggatgaag gttttagtga 360
agcacgtgtt cttgcggaaa tcactctttt gcctgacaag atccgtgttg cgtcagggag 420
acgtgctgtg gccatggcca caggtgttat tgataggtac tataaaaatg gaactcccca 480
gaagcaggtt gtggtcgttg ttacctcagg ttggtctgtg ctctgctttg atcacaacct 540
gaaaaagctg tgggaaacga atctgcagga ggatttccca cataatgcac accatagaga 600
gatagcaatt tcgataagca attacacatt gaagcatggt gatacgggtt tggttattgt 660
tggtggacgg atggagatgc agccatataa tcacatggac ccatttgagg aacttggcat 720
gacagcacaa aatgctgatc aacacagaag aagtgctact gagaatcagg cctctgaaga 780
ttctggagcc ataaacttgc gtcacttttc tgtctatgca tttgctggca agactggcct 840
tcttcgatgg agtaaaaaga ctgatgatgt tgaagctcac acctcagatg catcacaatt 900
aattccacaa cacaattaca agcttgatgt gcatgctcta aatagccgtc acccaggaga 960
gtttgagtgc agggaattta gagaatcaat tcttagtgtc atgccccatc gctgggaccg 1020
acgtgaagat acattattga agcttgctca tttcaggcga cacaagagga aaacattgaa 1080
gaagcaggct ggtagtaagt ctacagctta tccgtttcac aaacctgagg aacacacacc 1140
cgctggaaag gacttgtcaa gaaagattcc aaaattgatt ggaaaggctg cacgctatgc 1200
tggctcggca aaacccaaga agggtatgca atacattccg acgataacta attacacgaa 1260
gctttggtgg gttcctaatg ttgttgtggc tcatcaaaag gaaggaattg aagctattca 1320
tttgcctact ggtcgaacac tttgcaagct ttctctactt gaaggtggac ttcacgccga 1380

cataaacgga gatggtgttc tcgatcatgt ccagactgtc ggaggcaatg ttggagagag 1440
aactgtagtg agcgggtcaa tggaagtgtt gaagccttgc tgggcagtgg caacctcagg 1500
cgttcccatc cgggaacagc tctttaacgt ctcgatctgc catcactccc cttttaactt 1560
cttgcactat ggaggagatt actcacgaca cttcgcccag gcaagagaca cctctactct 1620
ggagatcgca actcccattc tcatccccag agatgatgga cacaaacacc gcaaaggcag 1680
tcacggagat gtaatcttct tgacaaaccg tggagaggtg acatcataca cgcctgatgt 1740
gcacggccac gacgcagtct ggcaatggca gcttcagaca gaagccacat ggtcgaatct 1800
cccgtctcca tcgggtttaa ctgaatcagg gacggtggtc ccaaccctga aaccattctc 1860
gttgcgcatt catgataacc agcctatgat ccttgccggg ggagatcaag cagcggtcat 1920
catctctccg ggaggaagca tattggcttc catcgaatta ccgtctcaac cgactcatgc 1980
acttatcact gatgacttct cgaacgatgg tctaacggat gtgattgtga tgacctcaaa 2040
tggggtttac gggtttgttc aaaccagaca gccaggggct ctgttcttca gttcgttggt 2100
gggttgtctc ttagtagtta tggcagttat tttcgttact cagcacttaa actccattca 2160
aggtaagcct cgaccatcat ctagctttta atagaaatct cgtagagttt tcttcctctc 2220
tacggaagtt aaaccggcag gtccgcttta gcagcttcac ccacatcggt tttggtaaac 2280
cggaaatatt actgatatac cagtgtagat tcagtgcctt atttcctggt ttcagctgta 2340
taactttata ctttgtagaa ttcaagtaaa aaatggtaga ggcaaataga ggactttatg 2400
tttttgtcaa ttttgatgtt ttaatggaat tatgtgacat ttaatttat 2449

<210> 55
<211> 698
<212> PRT
<213> Arabidopsis thaliana
<400> 55

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

```
Arg Arg His Lys Arg Lys Thr Leu Lys Lys Gln Ala Gly Ser Lys Ser
                325                 330                 335
Thr Ala Tyr Pro Phe His Lys Pro Glu Glu His Thr Pro Ala Gly Lys
            340                 345                 350
Asp Leu Ser Arg Lys Ile Pro Lys Leu Ile Gly Lys Ala Ala Arg Tyr
        355                 360                 365
Ala Gly Ser Ala Lys Pro Lys Lys Gly Met Gln Tyr Ile Pro Thr Ile
    370                 375                 380
Thr Asn Tyr Thr Lys Leu Trp Trp Val Pro Asn Val Val Val Ala His
385                 390                 395                 400
Gln Lys Glu Gly Ile Glu Ala Ile His Leu Pro Thr Gly Arg Thr Leu
                405                 410                 415
Cys Lys Leu Ser Leu Leu Glu Gly Gly Leu His Ala Asp Ile Asn Gly
            420                 425                 430
Asp Gly Val Leu Asp His Val Gln Thr Val Gly Gly Asn Val Gly Glu
        435                 440                 445
Arg Thr Val Val Ser Gly Ser Met Glu Val Leu Lys Pro Cys Trp Ala
    450                 455                 460
Val Ala Thr Ser Gly Val Pro Ile Arg Glu Gln Leu Phe Asn Val Ser
465                 470                 475                 480
Ile Cys His His Ser Pro Phe Asn Phe Leu His Tyr Gly Gly Asp Tyr
                485                 490                 495
Ser Arg His Phe Ala Gln Ala Arg Asp Thr Ser Thr Leu Glu Ile Ala
            500                 505                 510
Thr Pro Ile Leu Ile Pro Arg Asp Asp Gly His Lys His Arg Lys Gly
        515                 520                 525
Ser His Gly Asp Val Ile Phe Leu Thr Asn Arg Gly Glu Val Thr Ser
    530                 535                 540
Tyr Thr Pro Asp Val His Gly His Asp Ala Val Trp Gln Trp Gln Leu
545                 550                 555                 560
Gln Thr Glu Ala Thr Trp Ser Asn Leu Pro Ser Pro Ser Gly Leu Thr
                565                 570                 575
Glu Ser Gly Thr Val Val Pro Thr Leu Lys Pro Phe Ser Leu Arg Ile
            580                 585                 590
His Asp Asn Gln Pro Met Ile Leu Ala Gly Gly Asp Gln Ala Ala Val
        595                 600                 605
Ile Ile Ser Pro Gly Gly Ser Ile Leu Ala Ser Ile Glu Leu Pro Ser
    610                 615                 620
Gln Pro Thr His Ala Leu Ile Thr Asp Asp Phe Ser Asn Asp Gly Leu
625                 630                 635                 640
Thr Asp Val Ile Val Met Thr Ser Asn Gly Val Tyr Gly Phe Val Gln
                645                 650                 655
Thr Arg Gln Pro Gly Ala Leu Phe Phe Ser Ser Leu Val Gly Cys Leu
            660                 665                 670
Leu Val Val Met Ala Val Ile Phe Val Thr Gln His Leu Asn Ser Ile
        675                 680                 685
Gln Gly Lys Pro Arg Pro Ser Ser Ser Phe
    690                 695
<210> 56
<211> 2834
<212> DNA
<213> Oryza sativa
<400> 56
atgcgtcccc tcctcgcctt cgcggcggta tgcgcccttc tcgtggctgc cgcagcgccg     60
gcggccgcgg aggaggagaa ggcgaacaag ttccggcagc gcgaggccac cgacgacatg    120
cttggatacc cccaccttga tgaagatgct ttattgaaga ccaagtgtcc aaaacatgta    180
gagctgagat ggcagactga agttagttcc agcatttatg caactccgtt gatcgctgat    240
atcaacagcg atggaaagtt ggaagtagtg gtgccatcat ttgttcatta cctggaagtt    300
cttgaaggct ctgatgggga caaattgcca ggatggcctg catttcacca gtcaaatgtt    360
cattcaagtc cacttctata tgatattgac aaggacggga cacgggaaat agttttggca    420
acttacaatg gtgtagtgaa tttcttcagg gtatcaggtt atatgatgat ggacaagcta    480
gaagtacctc gtaggaaggt acacaaagac tggtatgttg ggctgaatac agatcccgtt    540
gaccgctccc atccagatgt tcatgacagc tcaattgcaa agaaagctgc ttctgaagaa    600
```

```
tctcacccaa atattcagga caagccagtt gtgaatgaat cttctaagga atcccagtca   660
cggagcacaa atgactcaac aacacgagga gttgattcca tgaaacatgc gtctaaggaa   720
gagccagttg aaagtaaacc taattctacc cgaggacagg agaatatgga tgtgttaaac   780
aatctaaaca gcacagatgc tgggaacaac tctagtttaa gtactacaac agagaatgca   840
tcacatgttc agagaaggtt gcttcaaaca gatgagaaaa gtaatcaagc aggaagctca   900
gaaactgatg caagtgatac cggaacagca aaagcagcta ctgttgaaaa tagcgagcct   960
ctagaggctg atgctgatgc atcatttaat ttgtttcggg atgtagaaga tctgcctgat  1020
gagtacaatt acgattatga tgactacgtt gatgaaacca tgtggggaga cgaggactgg  1080
aaagaacaac aacatgaaaa ggcagaagat tatgtgagca tagatgctca catcttgtcc  1140
accccagtga ttgcagatat tgacagagat ggcatacaag aaatggtgat ttctgtatct  1200
tacttctttg accacgagta ttatgataaa ccagaacatc taaaggagtt aggagggatt  1260
gacattggca aatatattgc aagcagtata gttgttttta accttgacac aagacaagtc  1320
aaatggactg cagaacttga tttgagtaca gatagtggaa attttactgc ccatgcatat  1380
tcttcgccga ccgtggttga tttggatggt gatggaaatt tggatattct tgttggaact  1440
tcctttggct tgttttatgt tatcgatcat cgtggtaagg ttaggaacaa gttcccactt  1500
gagatggcgg agattcatgc accagtcatt gcagcagata tcaatgatga tgggaaaatc  1560
gagatggtca ctgctgatgt ccatggcaat gtagcagcat ggactgcaga gggagaagaa  1620
atctgggagg ttcatcttaa gagtcttatc ccacagcgcc ctactgttgg tgatgttaat  1680
ggagatggtc gcactgaagt tgtggtccca accgtgtcag gaaacatata tgttcttagt  1740
ggaaaggatg gctcaaaaat tcagcctttc ccatatagaa cgcatggaag gatcatgagt  1800
cctgttctat tgcttgacat gagcaaacat gatgaaaagt ctaaaggcct cacccttgct  1860
actacatcct ttgatggtta cttgtatttg attgagggct caagtggctg tgcagatgtt  1920
gttgacattg gagagacctc gtacagtatg gttttggctg ataatgttga tggtggagat  1980
gacctcgatc ttattgttac taccatgaat ggcaatgtct tctgcttctc cactccctct  2040
ccgcaccatc cacttaagga atggagatca tcaaaccagg gaagaaacaa tgctgcatat  2100
cgttacaacc gtgaaggtat ttatgttaaa cacggttcca gaacattccg tgacgaagag  2160
ggcaagcatt tctgggtaga gtttgagatt gtggacaagt acagggttcc ttatgggaac  2220
caagctcctt ataacgtgac ggttactcta cttgtccctg ggaattatca aggagaaagg  2280
cgcattgtgg ttaatgcagc ttataatgaa ccaggcaagc agcggatgaa gcttcccaca  2340
gttcctgtga gaaccacagg aactgtgctt gtggaaatgg ttgacaaaaa cgggttctac  2400
ttctctgacg agttctcgct caccttccac atgcattatt acaagcttct gaaatggctc  2460
gtgcttcttc caatgcttgg gatgtttagc gttcttgtca tcctgcggcc acaagaaggc  2520
gctccgttgc catcattttc aaggaatatt gattagtgat ttgcacagtg aaacgtcagg  2580
tcctcataga agcagataag cagaacaaga tattagcagg atatccaatc acccatgaag  2640
gagcttgtcc acatcatact gcatcgggga gctcaatctg cacagaaaaa acccctcacg  2700
atgctgcatc ccatcagtta gctagttcgg gaagtacagc ttttagtcat gcaaaaagaa  2760
ttttgtcaca cttgagattg ttttctaact tatgatattt acacaaggaa ctggcagctg  2820
tctgatcctg gttt                                                    2834
<210> 57
<211> 851
<212> PRT
<213> Oryza sativa
<400> 57
Met Arg Pro Leu Leu Ala Phe Ala Ala Val Cys Ala Leu Leu Val Ala
1               5                  10                  15
Ala Ala Ala Pro Ala Ala Ala Glu Glu Glu Lys Ala Asn Lys Phe Arg
            20                  25                  30
Gln Arg Glu Ala Thr Asp Asp Met Leu Gly Tyr Pro His Leu Asp Glu
        35                  40                  45
Asp Ala Leu Leu Lys Thr Lys Cys Pro Lys His Val Glu Leu Arg Trp
    50                  55                  60
Gln Thr Glu Val Ser Ser Ser Ile Tyr Ala Thr Pro Leu Ile Ala Asp
65                  70                  75                  80
Ile Asn Ser Asp Gly Lys Leu Glu Val Val Val Pro Ser Phe Val His
                85                  90                  95
Tyr Leu Glu Val Leu Glu Gly Ser Asp Gly Asp Lys Leu Pro Gly Trp
            100                 105                 110
Pro Ala Phe His Gln Ser Asn Val His Ser Ser Pro Leu Leu Tyr Asp
        115                 120                 125
Ile Asp Lys Asp Gly Thr Arg Glu Ile Val Leu Ala Thr Tyr Asn Gly
    130                 135                 140
Val Val Asn Phe Phe Arg Val Ser Gly Tyr Met Met Met Asp Lys Leu
145                 150                 155                 160
```

```
Glu Val Pro Arg Arg Lys Val His Lys Asp Trp Tyr Val Gly Leu Asn
                165                 170                 175
Thr Asp Pro Val Asp Arg Ser His Pro Asp Val His Asp Ser Ser Ile
            180                 185                 190
Ala Lys Lys Ala Ala Ser Glu Glu Ser His Pro Asn Ile Gln Asp Lys
        195                 200                 205
Pro Val Val Asn Glu Ser Ser Lys Glu Ser Gln Ser Arg Ser Thr Asn
    210                 215                 220
Asp Ser Thr Thr Arg Gly Val Asp Ser Met Lys His Ala Ser Lys Glu
225                 230                 235                 240
Glu Pro Val Glu Ser Lys Pro Asn Ser Thr Arg Gly Gln Glu Asn Met
                245                 250                 255
Asp Val Leu Asn Asn Leu Asn Ser Thr Asp Ala Gly Asn Asn Ser Ser
            260                 265                 270
Leu Ser Thr Thr Thr Glu Asn Ala Ser His Val Gln Arg Arg Leu Leu
        275                 280                 285
Gln Thr Asp Glu Lys Ser Asn Gln Ala Gly Ser Ser Glu Thr Asp Ala
    290                 295                 300
Ser Asp Thr Gly Thr Ala Lys Ala Ala Thr Val Glu Asn Ser Glu Pro
305                 310                 315                 320
Leu Glu Ala Asp Ala Asp Ala Ser Phe Asn Leu Phe Arg Asp Val Glu
                325                 330                 335
Asp Leu Pro Asp Glu Tyr Asn Tyr Asp Tyr Asp Asp Tyr Val Asp Glu
            340                 345                 350
Thr Met Trp Gly Asp Glu Asp Trp Lys Glu Gln Gln His Glu Lys Ala
        355                 360                 365
Glu Asp Tyr Val Ser Ile Asp Ala His Ile Leu Ser Thr Pro Val Ile
    370                 375                 380
Ala Asp Ile Asp Arg Asp Gly Ile Gln Glu Met Val Ile Ser Val Ser
385                 390                 395                 400
Tyr Phe Phe Asp His Glu Tyr Tyr Asp Lys Pro Glu His Leu Lys Glu
                405                 410                 415
Leu Gly Gly Ile Asp Ile Gly Lys Tyr Ile Ala Ser Ser Ile Val Val
            420                 425                 430
Phe Asn Leu Asp Thr Arg Gln Val Lys Trp Thr Ala Glu Leu Asp Leu
        435                 440                 445
Ser Thr Asp Ser Gly Asn Phe Thr Ala His Ala Tyr Ser Ser Pro Thr
    450                 455                 460
Val Val Asp Leu Asp Gly Asp Gly Asn Leu Asp Ile Leu Val Gly Thr
465                 470                 475                 480
Ser Phe Gly Leu Phe Tyr Val Ile Asp His Arg Gly Lys Val Arg Asn
                485                 490                 495
Lys Phe Pro Leu Glu Met Ala Glu Ile His Ala Pro Val Ile Ala Ala
            500                 505                 510
Asp Ile Asn Asp Asp Gly Lys Ile Glu Met Val Thr Ala Asp Val His
        515                 520                 525
Gly Asn Val Ala Ala Trp Thr Ala Glu Gly Glu Glu Ile Trp Glu Val
    530                 535                 540
His Leu Lys Ser Leu Ile Pro Gln Arg Pro Thr Val Gly Asp Val Asn
545                 550                 555                 560
Gly Asp Gly Arg Thr Glu Val Val Val Pro Thr Val Ser Gly Asn Ile
                565                 570                 575
Tyr Val Leu Ser Gly Lys Asp Gly Ser Lys Ile Gln Pro Phe Pro Tyr
            580                 585                 590
Arg Thr His Gly Arg Ile Met Ser Pro Val Leu Leu Leu Asp Met Ser
        595                 600                 605
Lys His Asp Glu Lys Ser Lys Gly Leu Thr Leu Ala Thr Thr Ser Phe
    610                 615                 620
Asp Gly Tyr Leu Tyr Leu Ile Glu Gly Ser Ser Gly Cys Ala Asp Val
625                 630                 635                 640
Val Asp Ile Gly Glu Thr Ser Tyr Ser Met Val Leu Ala Asp Asn Val
                645                 650                 655
Asp Gly Gly Asp Asp Leu Asp Leu Ile Val Thr Thr Met Asn Gly Asn
```

660 665 670
Val Phe Cys Phe Ser Thr Pro Ser Pro His His Pro Leu Lys Glu Trp
675 680 685
Arg Ser Ser Asn Gln Gly Arg Asn Asn Ala Ala Tyr Arg Tyr Asn Arg
690 695 700
Glu Gly Ile Tyr Val Lys His Gly Ser Arg Thr Phe Arg Asp Glu Glu
705 710 715 720
Gly Lys His Phe Trp Val Glu Phe Glu Ile Val Asp Lys Tyr Arg Val
725 730 735
Pro Tyr Gly Asn Gln Ala Pro Tyr Asn Val Thr Val Thr Leu Leu Val
740 745 750
Pro Gly Asn Tyr Gln Gly Glu Arg Arg Ile Val Val Asn Ala Ala Tyr
755 760 765
Asn Glu Pro Gly Lys Gln Arg Met Lys Leu Pro Thr Val Pro Val Arg
770 775 780
Thr Thr Gly Thr Val Leu Val Glu Met Val Asp Lys Asn Gly Phe Tyr
785 790 795 800
Phe Ser Asp Glu Phe Ser Leu Thr Phe His Met His Tyr Tyr Lys Leu
805 810 815
Leu Lys Trp Leu Val Leu Leu Pro Met Leu Gly Met Phe Ser Val Leu
820 825 830
Val Ile Leu Arg Pro Gln Glu Gly Ala Pro Leu Pro Ser Phe Ser Arg
835 840 845
Asn Ile Asp
850

<210> 58
<211> 2411
<212> DNA
<213> Oryza sativa
<400> 58

gagtgtctag gccagtccag tccagtccag tccaggacgg tccagtccgg tcccgtagag 60
cgccggcgcg tgtagcttcc tcctccgctc caggctccag cgagggaggg agggagacca 120
ccgtctccgc cggctttgag agagaaagag aaagagagag agagagagcg gcgagatgcg 180
gaagcgggat ctgggcatcc tcctcctcgc cgccttcgcc gtcttcttct cgctccagca 240
cgacggcgac ctctccttcc gcgaggcctg gtaccacctc tccgatgccg actaccccat 300
caagcacgac gccgaccgcc tcccctctcc cctagtcgcc gacctcaacg gcgacggcaa 360
acccgaggtc ctcatcccca cccacgacgc caagatccag gtcctccagc cccaccccag 420
gccctccccc gacgatgcct ccttccacga cgcccgcctc atggctgatg tctccctcct 480
cccctccaac gttcgcctct cctccgggag gcgccccgtc gccatggccg ttggcaccgt 540
cgaccgccac tacgcccacg ccccctcccc ctccaagcag ctcctcgtcg tcgtcacctc 600
cggatggtcc gtcatgtgct tcgaccacaa cctcaaaaag ctctgggagg ctaatctcca 660
ggacgatttc ccccacgccg cccaccatcg cgaggttgcc atttccatta ccaactacac 720
tctcaagcat ggggatgcag gtttggtcat cgtcggagga aggatggaaa tgcaacatca 780
ttcagcagag cttttcgatg aatttatggt gtcagaacac aacagggaag agcaccgtag 840
aagcgccagt gagaagcagg cttctgagac aggcaacaca gacctgcgtc attttgccct 900
ttatgctttt gctggccgca ctggtgaatt aagatggagc cgaaagaatg agaatatccc 960
atcgcaacca tccgatgctt cagtgctgat accacaacac aattacaagc ttgatgccca 1020
tgcccttaat agtcgtcatc ctggtcagtt cgaatgtcgg gaatttagag aatcagttct 1080
tggggtcatg cctcatcatt gggataggag agaggatact tttctgcaac ttgcccattt 1140
taggaggcat aaaaggaaag cactgaagaa aacacctgga aaagctgttg taaataacgt 1200
gcacaagccc agtgaacata atccacctgg aaaggatgtt tccaataggt tagcaaatgt 1260
gattgggaaa gctgcagata tggctaattc aaataaaatc aagaagtcac aaaggacgct 1320
ttatgttccg acaatcacca actatactca agtttggtgg gttcctaatg ttgttgttgc 1380
ccacgaaaaa gaagggatag aggctgttca tctagcttct ggacgtacaa tctgcaagct 1440
tcatttaaca gaaggaggcc ttcatgcaga tattaatgga gatggggttc tagaccatgt 1500
tcaggttgtt ggtgcaaatg gcatcgagca aacagttgtt agtggttcaa tggaagtgct 1560
gaaaccttgt tgggcagtag ctacatctgg tgtgccagtg cgggagcaac tttttaatgt 1620
ttctatctgc cattacaaca atttcaattt gtttcatcat ggtgactttt caagaagttt 1680
tgggaggaca tttgatacaa ctggcttaga ggtggcgact cctattctgc tccagagaga 1740
tgatggtcat aaacacagga gaggaagcca cggggatatc atctttctta cgagtcgtgg 1800
ggaggtgacc tcgtactctc caggtctact tggtcatgat gcaatatgga gatggcaact 1860
gtcaacaggt gcaacatggt ccaaccttcc gtctccatca gggatgatgg aaaacattgt 1920
agttccaact ttgaaggctt tctctctgcg agcctatgac ccaaaacagg taatcatcgc 1980

```
gggtggtgat ctggaggctg tggtgatttc tccttctggt ggtttattgg catccattga   2040
actccctgca cctccaaccc atgcgctggt acttgaagat ttcaatggtg atggtttaac   2100
tgatattatt ctggtaacat cgggggggagt ctatgggttt gtgcagacaa gacatcctgg   2160
ggctcttttc ttcagcacgc ttgtgggttg cttgatagtt gttatcggag tgatatttgt   2220
ttcactgcac ctcaactcct cgaacagcgg taaacctagg gcttcaaccg actataggtg   2280
acgacattat gtcagtcgac agaattcctt gtcaatattg aggtgttgtt atagcatttg   2340
atttagattt ttcatataat caactgccaa ctgggttatg taaacttgtg cggcgaatgc   2400
atttttggcg t                                                        2411
<210> 59
<211> 701
<212> PRT
<213> Oryza sativa
<400> 59
Met Arg Lys Arg Asp Leu Gly Ile Leu Leu Leu Ala Ala Phe Ala Val
1               5                  10                  15
Phe Phe Ser Leu Gln His Asp Gly Asp Leu Ser Phe Arg Glu Ala Trp
            20                  25                  30
Tyr His Leu Ser Asp Ala Asp Tyr Pro Ile Lys His Asp Ala Asp Arg
        35                  40                  45
Leu Pro Ser Pro Leu Val Ala Asp Leu Asn Gly Asp Gly Lys Pro Glu
    50                  55                  60
Val Leu Ile Pro Thr His Asp Ala Lys Ile Gln Val Leu Gln Pro His
65                  70                  75                  80
Pro Arg Pro Ser Pro Asp Asp Ala Ser Phe His Asp Ala Arg Leu Met
                85                  90                  95
Ala Asp Val Ser Leu Leu Pro Ser Asn Val Arg Leu Ser Ser Gly Arg
            100                 105                 110
Arg Pro Val Ala Met Ala Val Gly Thr Val Asp Arg His Tyr Ala His
        115                 120                 125
Ala Pro Ser Pro Ser Lys Gln Leu Leu Val Val Val Thr Ser Gly Trp
    130                 135                 140
Ser Val Met Cys Phe Asp His Asn Leu Lys Lys Leu Trp Glu Ala Asn
145                 150                 155                 160
Leu Gln Asp Asp Phe Pro His Ala Ala His His Arg Glu Val Ala Ile
                165                 170                 175
Ser Ile Thr Asn Tyr Thr Leu Lys His Gly Asp Ala Gly Leu Val Ile
            180                 185                 190
Val Gly Gly Arg Met Glu Met Gln His His Ser Ala Glu Leu Phe Asp
        195                 200                 205
Glu Phe Met Val Ser Glu His Asn Arg Glu Glu His Arg Arg Ser Ala
    210                 215                 220
Ser Glu Lys Gln Ala Ser Glu Thr Gly Asn Thr Asp Leu Arg His Phe
225                 230                 235                 240
Ala Leu Tyr Ala Phe Ala Gly Arg Thr Gly Glu Leu Arg Trp Ser Arg
                245                 250                 255
Lys Asn Glu Asn Ile Pro Ser Gln Pro Ser Asp Ala Ser Val Leu Ile
            260                 265                 270
Pro Gln His Asn Tyr Lys Leu Asp Ala His Ala Leu Asn Ser Arg His
        275                 280                 285
Pro Gly Gln Phe Glu Cys Arg Glu Phe Arg Glu Ser Val Leu Gly Val
    290                 295                 300
Met Pro His His Trp Asp Arg Arg Glu Asp Thr Phe Leu Gln Leu Ala
305                 310                 315                 320
His Phe Arg Arg His Lys Arg Lys Ala Leu Lys Lys Thr Pro Gly Lys
                325                 330                 335
Ala Val Val Asn Asn Val His Lys Pro Ser Glu His Asn Pro Pro Gly
            340                 345                 350
Lys Asp Val Ser Asn Arg Leu Ala Asn Val Ile Gly Lys Ala Ala Asp
        355                 360                 365
Met Ala Asn Ser Asn Lys Ile Lys Lys Ser Gln Arg Thr Leu Tyr Val
    370                 375                 380
Pro Thr Ile Thr Asn Tyr Thr Gln Val Trp Trp Val Pro Asn Val Val
385                 390                 395                 400
```

```
Val Ala His Glu Lys Glu Gly Ile Glu Ala Val His Leu Ala Ser Gly
                405                 410                 415
Arg Thr Ile Cys Lys Leu His Leu Thr Glu Gly Gly Leu His Ala Asp
            420                 425                 430
Ile Asn Gly Asp Gly Val Leu Asp His Val Gln Val Val Gly Ala Asn
        435                 440                 445
Gly Ile Glu Gln Thr Val Val Ser Gly Ser Met Glu Val Leu Lys Pro
    450                 455                 460
Cys Trp Ala Val Ala Thr Ser Gly Val Pro Val Arg Glu Gln Leu Phe
465                 470                 475                 480
Asn Val Ser Ile Cys His Tyr Asn Asn Phe Asn Leu Phe His His Gly
                485                 490                 495
Asp Phe Ser Arg Ser Phe Gly Arg Thr Phe Asp Thr Thr Gly Leu Glu
            500                 505                 510
Val Ala Thr Pro Ile Leu Leu Gln Arg Asp Asp Gly His Lys His Arg
        515                 520                 525
Arg Gly Ser His Gly Asp Ile Ile Phe Leu Thr Ser Arg Gly Glu Val
    530                 535                 540
Thr Ser Tyr Ser Pro Gly Leu Leu Gly His Asp Ala Ile Trp Arg Trp
545                 550                 555                 560
Gln Leu Ser Thr Gly Ala Thr Trp Ser Asn Leu Pro Ser Pro Ser Gly
                565                 570                 575
Met Met Glu Asn Ile Val Val Pro Thr Leu Lys Ala Phe Ser Leu Arg
            580                 585                 590
Ala Tyr Asp Pro Lys Gln Val Ile Ile Ala Gly Gly Asp Leu Glu Ala
        595                 600                 605
Val Val Ile Ser Pro Ser Gly Gly Leu Leu Ala Ser Ile Glu Leu Pro
    610                 615                 620
Ala Pro Pro Thr His Ala Leu Val Leu Glu Asp Phe Asn Gly Asp Gly
625                 630                 635                 640
Leu Thr Asp Ile Ile Leu Val Thr Ser Gly Gly Val Tyr Gly Phe Val
                645                 650                 655
Gln Thr Arg His Pro Gly Ala Leu Phe Phe Ser Thr Leu Val Gly Cys
            660                 665                 670
Leu Ile Val Val Ile Gly Val Ile Phe Val Ser Leu His Leu Asn Ser
        675                 680                 685
Ser Asn Ser Gly Lys Pro Arg Ala Ser Thr Asp Tyr Arg
    690                 695                 700
<210> 60
<211> 1103
<212> DNA
<213> Triticum aestivum
<220>
<221> misc_feature
<222> (16)..(16)
<223> n is a, c, g, ou t
<400> 60
gcgttccgga gctttnaaaa gagtcaggag ggattgacat aggaccatat attgcatgca     60
gtatagttgt gggtaacctt gacacaaaac aagttaaatg gacagcagaa ctcgatttga    120
gtaccgaaag cgggaaattc cttgcccatg catattcgtc tccaactgtg gttgatttgg    180
atggtgatgg aaatttggat atccttgtcg gaacttccta tggcttgttt tatgttcttg    240
atcatcacgg taagactagg aaaaatttcc cccttgagat ggctgagatc catgcaccag    300
tcattgcagc agacatcaat gatgatggta agatcgagat ggtcactgct gatgtgcatg    360
gtaatgtagc agcttggact gcagagggag acgaaatctg ggaggtgcat ctgaagagcc    420
ttgttccaca gcgacctact gtcggggacg tcaatggaga tggccacact gatgttgtgg    480
tcccaactgt atcaggaaac atctacgttc ttagtggaaa ggatggctca aaagttcagc    540
ctttcccata tagaacacat ggaaggatca tgagtccggt cttgttagtt gacatgagca    600
aacgtggaga aaagacgcaa ggactaaccc ttgctactac gtcctttgat ggttatttgt    660
atttgatcga gggctctagt ggctgtgcag atgttgtcga cattggagag acctcgtaca    720
ctatgggttt ggctgacaat gtggatggcg gagatgacct tgatctgatt gttactacca    780
tgattggcaa tgtcttttgc tttttcactc cctcaccgga acatcctctc aaggaatgga    840
gatcatcaaa ccaggaaagg aataatgctg catataggca ccacccgtca aggtatttat    900
gtaaggcatg gttcaggaca acccgggaat gaaaagggta aacatttcgg ggtcgatttg    960
```

aaaattggag acaagtacag ggttccctat gggaattaac tccttaaaat gtcagggtac 1020
cttactcgtc ctgggaatta ccagggaaac aggggtattg gggtttgcca attttaaaat 1080
gaaccgggca caaggaattg ggt 1103
<210> 61
<211> 1152
<212> DNA
<213> Zea mays
<220>
<221> misc_feature
<222> (232)..(256)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (416)..(474)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (525)..(529)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (1025)..(1047)
<223> n is a, c, g, ou t
<400> 61
acaggcttat tataggtttg ttttctacta atagttatga agagcgtttt gacacagtat 60
ccgagagttt tgcttctcct ctccgagggt cgcttctctt agaatctgac aattaccgtt 120
tcaaaggcct actgtaaact ctcaatggtg acaatatgct tgttctttgg gtgagccaat 180
attgaaaatg gctgtgtaca aaattgttct gcagaactaa atgggtacat tnnnnnnnnn 240
nnnnnnnnnn nnnnnnaaat ggttaactgt tgggatacta cactgaacaa gacatttagg 300
tcatggagat gttggggagt cacttcgtca gcttcaatgg gtcaccggca tttcaccact 360
agtcgatatt ccgtgaaaac gatggcagcg gcgcaccttc ttggggcctc aagatnnnnn 420
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnagtaat 480
gcatgtggaa ggtgagtgag aactcgtcgg agaagtacag cccannnnna tcgatcatct 540
ccacgaccac agttccagta gttctaacag gaacagtggg aagcatcatc cgttgcttcc 600
ctggttgatg atacacacca ctgacaacaa tgcgcctatc tccttgataa ttccccggaa 660
caagtaaagt taccgtcacg ttataaggag cttgcttccc atagggcacc ctgtacttgt 720
caacgatttc aaactctatc cagaaattct tgccctcttc gtcacggaag gtcctggaac 780
catgcttaac atagatacct tcacggttgt actgatatgc agcattattc cttccttggt 840
ttgatgacct ccattcctta agcggatggt gtggtgattg agtggagaag cagaagacgt 900
tgccattcat agtggtaaca ataagatcaa ggtcatctcc accatccacg ttatcagcca 960
aaaccatagt gtacgaggtc tctccaatgt caacaacatc tgcacagcca ctcgatccct 1020
caatnnnnnn nnnnnnnnnn nnnnnnnacg ttgttgcaag agttagccct tttgcattgt 1080
ctccatgttt gctcatgtcg agaagtagaa caggactcat gattcttcca tgtgttctat 1140
atgggaaagg tt 1152
<210> 62
<211> 720
<212> DNA
<213> Solanum tuberosum
<400> 62
gtttcaatct agataccaag caagttaaat ggactgcaca gttggactta agtactgacg 60
acgggaaatt ccgtgcctat atatactctt ctcctacggt agttgatttg gatggtgatg 120
gaaacatgga cattctagtg gggacctcct atggcttctt ttatgtgttg gatcacaacg 180
gcaaagtgag ggaaaagttc cctctcgaaa tggctgaaat ccaaggagca gtagttgcag 240
ctgatatcaa tgatgatgga aagattgaac tagttacaac agattcacat ggaaatgttg 300
ctgcttggac cgcacaaggc acagaaattt gggaaacgca tctcaagagc cttgttcctc 360
agggaccggt tattggcgat gtggatggag atggccatac agatgtcgtt gtcccaacac 420
tttctgggaa tatatatgtt ctgaatggca aggacggctc atttgtacgt ccatatcctt 480
ataggactca tggtagggtg atgaatcgag cacttcttgt cgacttgagc aaacgtgggg 540
agaagaaaaa agggcttaca attgtcacaa tgtcatttga tggttatttg tatctcatag 600
atggaccaac atcatgtgct gatgttgtag atattggtga aacttcatac agcatggtct 660
tggctgataa tgttgatggt ggcgatgatc ttgatcttat tgtaacaacc atgaatggta 720
<210> 63
<211> 904

<212> DNA
<213> Aquilegia
<400> 63
gtgaatattg attcacacat cctttgcact cctgtgattg cagacatcga caatgacgga 60
gtatcagaaa tggttgtcgc tgtatcctat ttctttgatc atgagtatta tgacaaccca 120
gagcatctct ctgagttggg tggcatagat attgggaaat atgtagcggg aggaattgta 180
gtatttgatc ttgatacaag acaagttaag tggacaacag agctagatct tagtacagac 240
acaggggact tccgtgctta tatatattct tctccaacgg tggtcgattt ggatggagat 300
ggaaatttgg acattcttgt tgggacatct tttggcttgt tttatgtttt ggaccataat 360
ggcaagataa gaaacaagtt ccctctcgaa atggctgaga ttcagggctc tgtcattgcg 420
gcggatataa atgatgatgg aaaaattgaa ttggttacaa ctgatactca tggaaatgtt 480
gctgcatgga ccccagaagg agaagaaatt tgggaagtac atcttaagag tcttgttcca 540
caacgtccaa cagttggtga tgttgatggc gatggtcata ctgatgtggt ggttcctaca 600
ctatcaggga acatatacgt tctcagtggc aaggatggct cttttgttca cccataccca 660
tatcgtactc atgggagagt catgaatcaa gttcttttag tagatctaag taaacgtgaa 720
gagaaacaga agggactcac tcttgtcaca acatcatttg atggctattt gtaccttatt 780
gacggaccat catcttgtgc tgatgttgtt gatattggcg agacttcata tagtatggtc 840
ttggcagaca atgttgatgg tggggatgat cttgatctaa ttgttacaac tatgaatggg 900
aatg 904
<210> 64
<211> 708
<212> DNA
<213> Brassica napus
<400> 64
atttgggaag tgcatctaaa gagtcttgtt ccccagggtc cttcaatagg cgatgttgat 60
ggtgatggac atactgatgt cgtggttcct acaacgtcag gaaacattta tgttcttagt 120
ggcaaggatg gttcgattat acgtccgtac ccgtacagaa ctcatggaag agtgatgaac 180
caacttcttc ttgttgatct gaacaagcga ggtgagaaaa agaaggggct caccattgtt 240
accacatcct ttgacggtta catgtacctc atagatggac ccacctcatg cacggacgct 300
gttgatattg gtgaaacttc atacagcatg gtcttggctg ataatgttga cggtggagat 360
gatcttgatc taatcgtctc aactatgaat ggaaacgtct tttgcttctc tacaccttct 420
cctcaccatc ccctcaaggc gtggagatcg actgatcaag ggaggaacaa taaggcaatc 480
cgttacggtc gtgaaggtgt ctttgtcact cattcaacca gaggcttccg tgacgaggaa 540
ggcaaaaact tctgggctga gatcgagatt gttgataact acagataccc atctggttca 600
caagcaccct acaacgttac tacgacgttg ttggtaccag gcaactacca aggagatagg 660
aggataacac atagccagat ctatgaccga ccaggaaaat acagaatt 708
<210> 65
<211> 834
<212> DNA
<213> Citrus sinensis
<400> 65
gttgatgggg atggccatac tgatgttgta gttccaacac tatcggggaa catttacgtt 60
cttagtggca aggatggcag taaagtccgt ccttatcctt atagaactca tggaagggtg 120
atgaatcaag tccttcttgt tgatttaact aaacgcgggg agaaaagcaa gggactcaca 180
attgttacaa catcttttga tggctatttg taccttatag atggcccaac atcatgtgct 240
gatgtagtcg acattggcga aacttcatat agcatggtct tagctgacaa tgttgacggt 300
ggagatgatc ttgatcttat tgttaccaca atgaacggca atgtttttg cttctcaact 360
cctgctccac accatcccct caaggcatgg agatcaatta accaaggaag aaacaatgtt 420
gcaatccgtt acaaccgtga aggaatctat gttacacatc catcaagagc tttccgtgat 480
gaggaaggca gaaacttctg ggtggagatt gagattgtag atgaatatag attcccatct 540
gggtcccaag ctccatataa tgtcactaca acattgttgg tccccggcaa ttaccaaggt 600
gagaggagga ttaagcaaag ccaaatcttt gcacgtcgtg gaaaatatag aatcaaattg 660
ccgacagtcg gggtaaggac gacagggact gttttggtgg agatggttga caagaacgga 720
ctttatttct cagacgaatt ctcgcttaca ttccatatgt attactataa actactaaag 780
tggctcctag tcctcccgat gctcgggatg tttggtgtgc ttgtcatcct tcgt 834
<210> 66
<211> 906
<212> DNA
<213> Asparagus officinalis
<400> 66
gggtggaaag gatggatcgt ttgttcgacc cttcccctat aggacacgag ggagattgat 60
gagtccagtt cttctggctg atttaagcaa acgtgatgaa aagttaaagg gcctaactct 120
tgttacaact gcatttgatg gctatttgta cctaattgat gggtccactg gttgtacgga 180

tgttgttgac attggcgaga catcatacac catggtcttg gcagataatg ttgatggtgg 240
ggacgatctt gaccttattg ttactactat gaatggcaat gtcttttgct tttccacacc 300
ctcaccgcat catcctctga aggaatggag atcatccaac cagggaagaa acaatgcagc 360
aagtagatat aaccgtgaag gaatatacat atcacatgga tccagagctt tccgcgatga 420
ggaaggtaaa catttctggg tggagctgga gataatagac aagtacaggt ttccctcagg 480
gcatcaaggg ccttataatg tcacaacaac attattagtt cctggaaatt accaaggaga 540
aaggcgaatt gtcgtcaaca atgtatataa tcaacctgga aagcaacgga taaagctgcc 600
aacagttcct gttcgaacaa cagggactgt gctggtggag atggttgaca agaatggatt 660
gtatttctct gatgaatttt ctcttacatt tcacatgcat tactacaagc tactgaagtg 720
gcttctgttt ctaccgatgc tcggaatgtt tggagttctt gtcattctac gtccccagga 780
aggtgcaccg ttaccatcgt tttcacgaaa tttggattga tattaatgat atcatccaat 840
gaagatgaga atactccaaa gtgcgatgct atctgacatg tatgcctttg tacatctgaa 900
tgtgtg 906

<210> 67
<211> 779
<212> DNA
<213> Populus
<220>
<221> misc_feature
<222> (34)..(39)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (663)..(663)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (686)..(688)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (739)..(744)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (746)..(748)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (760)..(761)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (766)..(769)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (773)..(777)
<223> n is a, c, g, ou t
<400> 67
tgaatcaagt tcttcttctt gacttaagta aacnnnnnng agaaaaacaa gggactcaca 60
cttgttacaa catcattcga tggttatctt taccttatag atggaccaac ttcttgtgct 120
gatgttgttg atattggtga aacttcatat agcatggtct tggcagacaa tgttgatggt 180
ggagatgatc ttgatctcat agtttcaaca atgaatggaa atgtcttttg cttttcaact 240
cctgttccac atcatcccct gaaggcttgg agatctaata atcaaggaag aaacaatgta 300
gcaaaccgct acaaccgtga aggggtgtat attaaacctt catcaagaag tttccgtgat 360
gaggagggga agagcttctg ggtggaattt gagattgtgg acaagtatag aatcccgtct 420
gggtctcaag caccttataa tgtcactaca accctgttag ttcctggcaa ttatcaaggt 480
gaacgacgga taaagcaaaa tcaaatcttt gaccgtccag gaaaatatcg gataaaactt 540
ccaacagttg gagtgagaac tactggaact gttttggtgg agatggttga taagaatgga 600
ctctatttct cagatgactt ctcgcttaca ttccacatgc attactataa actgctgaag 660
tgnctccttg tcctcccaat gcttgnnntg tttggtgtgc ttgtcatcct tcgtccacaa 720
gaggccgtgc ccttgccann nnnntnnngg aatactgacn ngtgannnnc atnnnnnac 779

<210> 68
<211> 723
<212> DNA
<213> Populus
<400> 68
ggcacgagca cacttgttac aacatcattt gatggttatt tgtaccttat agacggacca 60
acttcttgtg ctgatgttgt tgatattggt gaaacttcat acagcatggt cttggcagat 120
aatgttgatg gtggagatga tctagatctc atagtctcaa caatgaatgg gaatgtcttt 180
tgcttttcaa ctcctgttcc acatcaccct ctcaaggctt ggagatcttc taatcaagga 240
agaaacaacg tggtgaaccg ctacaaccgt gaaggggttt atgttacacc ttcatcaaga 300
agttttcgtg atgaggaggg aaagagcttc tgggtggaat ttgagattgt agacaagtat 360
agattcccat ctgggtctca agcaccttat aatgtcacta caacccttt agttcctggc 420
aattatcaag gtgagagacg aataaagcaa agccaaatct ttgaccgtcc aggaaattat 480
cgggtaaaac ttccaacagt tggagtgagg actactggaa ctgttttggt ggagatggtt 540
gataagaacg ggctctattt ctcagatgac ttctcccttа cgtttcacat gcattactat 600
aaactgctga agtggctcct agtcctccca atgcttggaa tgttttgtgt gcttgtcatc 660
cttcgtccac agaggccatg ccccttacat cattttcagg aatactgact tgtgatatca 720
tct 723
<210> 69
<211> 704
<212> DNA
<213> Euphorbia esula
<220>
<221> misc_feature
<222> (10)..(11)
<223> n is a, c, g, ou t
<400> 69
ctgatacccn ntgaaatgtt gctgcttgga cttcacaagg aaaggaaatt tgggagaggc 60
atctcaaaag tcttatttcc cagggcccat cagttggtga tgtggatggg gatggccata 120
ccgatgtggt agtccctact atatctggga atatatatgt tttgagcggc aaggatgggt 180
caaatgttcg cccataccca tatagaaccc atgggagagt aatgaatcaa gttctccttg 240
ttgatttaag taaacgtggg gagaaaagca ccggactcag ccttgtaaca acttcatttg 300
atggatactt gtaccttata gatgggccca catcctgcgc tgatgtcgtt gatattggtg 360
aaacctcata tagcatggtc ttagcggaca atgtagacgg aggagatgac cttgacctag 420
tagtcacaac aatgaatggg aacgtctttt gcttcgcgac tcctgttcca catcatcccc 480
ttaaggcatg gagatcggct aatcaaggta gaaataatgt ggcaaacaga tttaaccgtg 540
aaggggtcta tgttacacct tcatcaagag cttttcgaga cgaggaagga aagaatttct 600
gggtggaaat cgaaatcgtg gataaatata gataccttc tggttctcaa gcaccttata 660
aagtcactac aaccctgtta gttccgggta attatcaagg tgag 704
<210> 70
<211> 1088
<212> DNA
<213> Ceratopteris richardii
<400> 70
ccacgcgtcc gggttacaac agatgcccgg ggaaatgttg ctgcatggac aaataaagga 60
aaagaaattt gggaagttca cttaaaaagc ttgattgctc agggtccgac tgtcggagat 120
attgatggag acggaacatt ggatattgtt gtgccaactg tgtcaggaaa tatatatgtt 180
ttgcatggaa agactggtgt tactatgaag ccatatccat ttcgtaccca tgggcgtgtg 240
atggctcctg ttcttctggt ggacttggcc aagctaggca ccgagagaag aggtctggtc 300
ctggtggttt catcatttga tgggtacctc tacctcatag atggtcccac agcctgtgct 360
gatgtggttg atgtaggaga gacatcctac acaatggtgc ttgcagacaa cgttgatgga 420
ggagatgatt tggatttaat agtgacgacc atgaatggga atgttttctg cttctctaca 480
cctgctcctt atcatccttt acggtcatgg acctcacaga atcagggtag gagcaatctt 540
gcatctcgta taagacatga agggatttat gctacatctg gctcgagaag tttcagggat 600
gaaggtggtg aaagtttttg ggtgcatttc aatattattg atgagaaccg cttaccttcg 660
gggccttaca atattacggc tacccttctt gtcccccaca actacatggg tccaaggcgc 720
attacagaga gtcagttaat tcaacaacca ggaacgcaca agttcaaact gccatgtgta 780
tctgtacgat catctgggac tgtagttcta gagatggtgg ataaacatgg gttctatttc 840
actgatgagt attcattgac attccacatg cactattatc ggctactgaa atggatggta 900
gtgcttccaa tgcttggaat gctttttgtc ttgttgtggc ttcatcctga ggatgccata 960
tcagttcttc cttccttttc aagggatcat taaaatgaca tatgagctca atttttggta 1020
ctcctaaatc caaaccttat tgtaaatgta tacattttga catttatcca gtctactaaa 1080
gaaatgca 1088

<210> 71
<211> 830
<212> DNA
<213> Welwitschia mirabilis
<400> 71
atgacaatca tgaacacttg aaggagctgg gtgatattga tattagcaaa tatgtttcag 60
gtgctatagt ggtatttaat cttgatacaa agcaagtgaa gtggagtact cagttggatc 120
ttagcacaac ctctgggact tttaatgcat acatctattc ttctcctact gtggttgacc 180
tggatggtga cggcaatttg gatattattg ttggcacgtc atttggtttc ttttatgtat 240
tggatcacca tggaaaaaat agagaaggct ttccattaca gatgggtgag attcaaggac 300
aagtaattgc tgctgatata aatgatgatg ggaagattga aatggttaca acagataccc 360
gtggaaatgt ggctgcttgg acttctcagg ggaaagaaat atgggagata catttaaaaa 420
gtcttattgc tcaggggcct acagtaggtg atattgatgg agatggtcat acagatttgg 480
tggttcctac agtatcagga aatatttatg tattaaatgg gaaagatgga tcattggtga 540
agccatttcc ttatcgtact catggaagag tcatgagtcc ggtacttttg gttgacctta 600
gcaaacgtgg ggaaaagcaa aagggcttaa ctcttgcagc attatcattt gatggttact 660
tctacttaat tgatggtcaa acagcttgtg cagatgttgt tgatattgga gagacttcct 720
actctatggc tttggcggac aatgtagatg gcggcgatga tcttgatttt attatcacaa 780
ctatgaatgg caatgtgttc tgcttttcaa cccctgctcc acatcatcca 830
<210> 72
<211> 8730
<212> DNA
<213> Medicago sativa
<400> 72
atgaggaagc gtgatttggc gattcttatg ctctgcgctt tcgctatttt cttctctctt 60
caggtattgt acccttttaac cccctaattc actttctcac tcttaaaccc taaccctcat 120
tttcaattca attgattcag caagatggtg gtgtttcatt caaagacgcg tggatgcatc 180
taacagatga atacccaatc aaatacgaag ccgaacgcct tcctcctcct gttgtcgccg 240
atctcaacgg cgatggtaag aaagaagttc tcgtagctac ccacgatgcc aaaattcagg 300
ttctttcatt tcccaatttt cctcaatttt ttattcactt ataaaattcc aggttttatt 360
tggtttcctt cacaactgga cattgtcaaa aaaacatttt ttttctttca gtaaaccatc 420
aatgtagttg atgaactata actctcttca atttagtttg taagtagtct ctaaaattca 480
tcaattaatc cctgctatta atgtttggtt ctacaccgaa gaaagttgac tttgaatgag 540
ttgattgtgt aaactctatt tgggctaaaa ttgatttgaa ggtaaagtga tttatgtttg 600
gactttggat acattcatgt aaaagtgaat tgaacaggtg aaaaatcaac tctagaatca 660
gaagctacaa tttctaactt taagtagaat gtagaatcaa ttctggaggt aaaatcaatt 720
ttactctaga agaaccaaat atgtcaatca attttgggtc ctccaaaatt agaaccaaac 780
atacagaaag tagagactaa aattgatgga ttttcatata ctttagggac ttaattgacc 840
ggttttatat actttaaaga ctacttatca tatttatata gtttagggac tagattgatg 900
gtacaagtaa tatagtagta actggatttt gtatttatgg tcggtgtctt agattttgga 960
gccacatagt aggcgtgttg atgaaggatt cagtgaggca cgtgtgttag cggaggtgtc 1020
tctgttgcct gacaaagtgc gtgtcatgtc tgggagacga cctgttgcta tggccactgg 1080
ctttattgac cgccatagaa ttggacaacc acataagcag gttttggttg tagtaacatc 1140
gggttggttt gtaatgtgtt ttgattccaa cctccaaaag ttgtgggaaa ataatttgca 1200
ggtatgatgt ttatgttgtg ccatgtttca gttgatagaa ttttcagtta aacattttga 1260
ttatcttgtt atgttttagt ttaccaactg ttttagtggt tgttaacagg aggatttccc 1320
ccataatgct caccacaggg aagtttcaat atctataagc aattatactc tcaagcatgg 1380
agatacagga ttgatcattg ttggtgggag aatggaaatg cagcctcatg tatgtgtcac 1440
tttccttctt gttttctggt ttgttattct ttttagcaga gtatgtgtgt ctgtgttcta 1500
tctgctgatg gaaaaactta gaatttaatg tggatggtgt ttaccatctt ttatgaaagc 1560
gaaaactctc aacttcatga gtagcataag attctggcta tcttgttttc ttggaagtga 1620
tgaaaatgaa agttcctgag taattttatt tgataaagta tgatggtaat gtgagggttt 1680
gctgctgctt ctcgagtttc agaatttttt atttgttatt tgattctttc tatggtaatg 1740
attttggaat tgttcatctt tggaattatt ggcattatct tctaatctga tggttctgtt 1800
ggcttatgtt gattgctttc cattttcaga tttttatgga ccctttttgaa gaaatgggaa 1860
tgggagctag atttgctgag cagcatcgaa gaagtgctac tgaaaaggag gtatagatac 1920
gtttctagtt gatgctaaat tttatgccca ttttttttact atctgtagca tgctaatttt 1980
ggtttaaact ttgagacact atttgcttta tttatttcac tttaattttg ggtcttcaag 2040
gttgacctta ttttatgctc ctaaacagtc aattcaaatt cctattttga tataaataga 2100
taatataaaa gttcaccatg tcttggcatg ataaagccaa taaataaaat ctgtgcacct 2160
tgcagtaagt agtctataaa tctgaacctt catgatcagg tagctaaata aacaggagac 2220
agtttcactc gttttgttta gatttttgtt tttacaaaaa aatgtatcag gtgattttag 2280
ttttcttccc attttaaggt tgaagacttt aacttgttta ttgtgttcct ttttctggat 2340

aggcttctga aaacaccgga actgtggatt tacgccattt tgcattttat gcatttgctg 2400
gtcgatcagg tgtagaacga tggagcagaa aaacagaggt ggtgttcttt ttttttgtt 2460
ttcagtgcgc tctattattt gggtgttgtc aacatgcaaa tatgaaacaa ttatttattt 2520
ttaatttaaa aatgcatttt cttgtttttt aaaagagctg tcttttctat ctggaaattc 2580
attaggtgat taagtgaact ttctctgtcc tccctctctg tacataactt tattcagggg 2640
gtagcaactt agttctagac agtatgtttt gtcctgcaac agacatgaaa actgaagaaa 2700
agttatatag ttaagctcac aaagttcaga acataatgag ttagagccta aatgtcaata 2760
atagattata aattatacct ttgaagatat tggagtcttg ttgcaagcat ctgcacttgg 2820
attttaacgc actggtgcac aaatatgtat tttttgaaaa gaaaatctat caaacaaatg 2880
gaaccacctc ccaaccagca taaactttag tcttttctgt ctttgggtgg tagtaccaga 2940
tagatgtttt acaatcaaca tagcggtcca gtttgattga atatgtgatt tacgctagtc 3000
caacatttat tttgttaaat gatcttgtgt atccttgtcc ttttttttttt cttttctttt 3060
ctaattgcta gtattgtgat ttaattgttt agaacattga agcagcagct tcttcagacg 3120
catcacagtt aattccacag cataactaca aacttgatgt tcatgctctg aatagacgtc 3180
aacctggaga ggtacttagc tatagctgtt caatttaatc ctaaattcct aatggtatat 3240
tgaattgatg gaatatggta aattggagtg gcataaaata ccattccatc ctttttaagc 3300
aggggaatca ttgcttttct ttttcataca tccattttat tctaatctag aaatatgcat 3360
gttgattttt atgttctccc ctttagctca acgtactgat tttgttatat gtaagtaata 3420
ctgtaatttc agatttatta ctttttccat tcctttcctc atcaatctta agtaataccg 3480
tagaatagat tcctagtcat atctttatgt tcttcacctt ttctcttcaa atgatcaaac 3540
actaaagcaa atatttgttt gttaatatgg agtttactat gagaactcta gtacaatagt 3600
actcatccaa atctacttga gtaacttggt catctccttc ctttgtgtgt ggaaatagca 3660
gcatgtgcac ataaaggagt gtttcaatca aatttaatat tattgtgaaa cctttgtata 3720
tgtttatcca tatctaccca atagatagct tcctgtgaca tgtgaaatta tagtttgaat 3780
gcagggaatt cagagaatca atcctgggag ttatgccaca tcaatgggta cactctattt 3840
tcaaccttga ttaacgtttg gtactaaagc ataatctcaa atgtcatagt taactgatat 3900
tagatacatg cactcatcag gataggagag aagatacttt attgaagttg gtccacttca 3960
atcggcataa gaggaaaaca ttgaagaaaa cacctggaaa gactatcaat tacccttcg 4020
acaagcctga ggaaaaccat ccaccaggga aggactcaac caaaaaaatt tcaaacataa 4080
ttgggaaagc agcaaatttt gctggttcag caaaatccaa gaaggtaatt agaaaccatg 4140
tcatgcaatg caactggtaa caattcattc aattgttgac ataatgataa ttagttttaa 4200
gatggctaca actggacaac tcaaaaattc ttttccaaac atgatttcag ctacaagctt 4260
ttgttgtctt tatccagtaa cagagctgtc tgtttacttt ctaactactg catgtccaat 4320
tttttgttat gtgatctaaa tctttgtgtt ggtaatattc tcagcatctg ctgattttaa 4380
ttcattttgg ctaaacctat tttgaatttg ttatgataac cattgcttag catttttaat 4440
tttttttttа aactaagcac tggtcagata aaccatctca cttgttatat ttgtatttgt 4500
tcttaaagat ttaagtctag ctacattttt tatttctcat ttaaattttt tgaatcctga 4560
atttagggag ccaagatatt tagttataaa aataatgatt aataataaat atgcagttct 4620
ccccattgtc gtgttaattt ggtagatagc atagtagtca atcttggata gcggtgcgta 4680
gtggccaacc ccaaaaatgg gatagcgtat agcggtatga caaatagcgg tcaccgatgt 4740
ttgattattt tttggacaat tatatgaata atagcaaagt ataaacaaac ttatatattt 4800
ctcaacaaaa aaaacttata ttaaatataa cactcaaatc tcacgagaca ttcataaatc 4860
aaaacaccga aaacaaaaga caggaacaaa taaaaaaaat cataaaacag gagaaaccat 4920
aagatagaag atagaagagt gtgtttcgcg gcacaggtga gaagagagaa aaagttgaaa 4980
aatgttaggg tttacgaaat cctcaaaatg ccctaaaaaa ggttttaaaa atagggtcaa 5040
tttggaaatt tcctactaaa ataaatagcg gctgctatcc aacccgcacc gctatagcag 5100
ctgaagcggc cgctatttga ttccgcgcta tttgataccg ctatgctaca cgcggccgct 5160
atagcgccac tatcggctgc tattgactac tatggtagat agaaactgtt aggttaccta 5220
gtagggctga tgatgtgtta gtgggctcaa gtaatgagtg agaggcccat tatttggagg 5280
gtgtatgaac gtgggaagaa gagaggagag aagggactaa gggtgcttcc tagctggcag 5340
ttagttagta tgaggcagag ggagtgaata tatttcttct gaggagctgt tgctctagtt 5400
tatccgtcat agtgatttca ttgttattct agtttagcca ggctttggga gtatttggta 5460
tatctacctc gaagatgcta ggtcctaaaa tgatgcagtt ggaaggttcg attgcgggag 5520
tctctctgtt ggtgatgatt taaacacaat ttcatatact cctggtgtgg ttgaagcgtt 5580
gggggttgcc tttggattca tcagatactg ctggtgtgca tttgggtggg ggtgggggt 5640
catatggtgc agatattggg caaatgtgtg gttggcttct gcatatcggc ccttatgcag 5700
tcttgctgga tacatatgtg ttggggttag aagaggtcat gccaaagtac caattactct 5760
cttaatacct gtctaataag actctaatgc tttaaattga cgaggatttg cattttcac 5820
ctttcactgg caaacatatg ggttaaactg gtagatgttt agcgtgttca agtacaagct 5880
gtgttcccat agatgaagta cttcatgttt ggggagtttg atcgttaaac tacctaattg 5940
tttgtttaaa tgaattttga attccctctt ctaataattt tagtgcttcc aaagattata 6000
atgcatataa aacaatttga ggattatctt gttcttaggt gaaaaacagt ttttttccca 6060
tttgaaatta gataaacaga gtagttgtag atttctactg tggaaagcat aacatgtcta 6120

agctggttaa atatacttaa ttgatggttt ttgtcattaa agtaatttct ttgcatacat 6180
ttttatgttc tcatgttatt gcttaatcgg aaaggggttc tataaaatgg atggcatttt 6240
atgggattcc gagaagcctt tttctttttt ttgccagggt ggggggtttg cattatttgt 6300
cataatgttt tttcttttgg aatgcagtat ccaccatatg ttcccaccat aaccaactat 6360
actaaggttt ggtgggttcc taatgttgtt gtggctcatc taaaggaggg gatagaagtc 6420
cttcatctgg catctggtcg aacactatgt aaggtaaata taaagccact gatcagcttg 6480
aaaatcaaaa tacaaaaaaa aaaaaaaaaa gtttacaagt agttgtacaa catttgctcc 6540
tgtggttgta gatttgtagt attaaactgc catggtttca gccaaaatct tatttagcat 6600
tcacgttact tttgaaattt ttcgtattct gtgaaatgac acttctattt atctcagctt 6660
caccttcagg aaggtggtct acatgctgat attaatggtg atggagttct ggatcatgtg 6720
caggttcact ttttttcctt cacatttttt gataaacaaa caagccaaat ctttattaat 6780
ttagcatgat aattgttaac ttaacggggg attttcctct gcctgatgtc tgaaagattt 6840
atctaatcag gctgttggag gaaatggtgc tgagcaaact gtagttagtg ggtccatgga 6900
tgttctacgt ccttgttggg ctgttgcaac gtctggcgta ccagtacggg aacaactctt 6960
caatgtatct atttgtcatt atacccattt taacttattc caacatggag aactttatag 7020
aggcttcaac cgaggttcag atatgtcttc tttggaggta gcaacaccca ttctcattcc 7080
tagaagcgat ggtcacaagc accggaaggg aagccacggg gatgttatct tcttgacaaa 7140
ccggggagag gtacgctctt ttttgttcga ctaaatctgg ttatgtaaca tttgatagtt 7200
tgtttgagca ccaatcaaac aaacctagtt cttaacttct gttaacaaat tatttcaagt 7260
cacgcacgac taatatattt tggtggtatt tgtgagataa tccgtacctc cgtactgact 7320
ttatgaaaat tgctccttgc agataacttc gcacactcct ggtttgcatg gtcatgatgc 7380
tgtttggcag tggcaacaat caactggtgt cacatggtca aacctacctt ccccagcagg 7440
gatgatggaa ggtggtttgg tgattcccac actaaagcct tttcctttgc ggttgcatga 7500
caatcatgag atgatccttg cagctggaga acaagaggct gtggtaatat cacccggagg 7560
tagcatattg gctacaattg aactccctgg ttcacctaca catgtattga tccgtgaaga 7620
cttctcaaat gatgggctca ctgaccttat tctcgtcacc tcatccggag tgtacggctt 7680
tgtccagact cgacaacctg gtgctctctt cttcagtgtg ctgatcggct gtctcatagt 7740
cgtgatggga attatatttg ttacccagca cataaattcc atgaagggga agcctcgtcc 7800
atcatctggt cctaggtgaa atgcaggacc tgaaatttat acattaactt tgcaggcgct 7860
tacagatgaa gcagaagaca aagcggtgtg cgcatcgtat ggagcttcaa gattgttgga 7920
aatgcataga atgtaaggag attttgagga ggtgttgtac tgtctaggat ggatgattgt 7980
aaaatttaga cagaagtgac agaatgtcct tcatgagaga aacatgatcc attaatataa 8040
ggaaactata gtctttgcct acgtatctgg ctgttgttta gaagttatac tgcccacact 8100
tgacaattgt cacccgttcc gaattagttg atacttgata gttgatgcat aacagaaact 8160
caccttggtc tttacacttt tcgctaaaag gattttctct tgtcattaca caattcttgt 8220
aaatacgacc atcgtttatt tgcttcagat ttagcagtgt aaaacattaa atctagcaac 8280
caacttgctc caatctgtct tggtcttctt cagaatacac ttcttaaaag tcaattaaga 8340
aatgacttcc tgattccatc aaattaccaa tgtggttgat cccatcatac atcaagtgct 8400
atgtacgtaa tccctgttag tctttggcat caccttaagt gttttagcta ttgcttttcc 8460
tgtgcgtcga gaaatgtttc aaataaccat ctgaatgatt caatattgca tcatataata 8520
ttgcaaaccc aaaattagtg tgttcctata atggataaag cttgaaaata atgcgagtgg 8580
tcgattagtg cgattggtgc aatgttggat tcattaacaa ctggtattga aaatattgca 8640
taaagctatc aacttctcca tataccatgc tacattgtct tttgcctgat gatgattttt 8700
gacatccaac cgtgtaaata caatattagt 8730

<210> 73
<211> 57
<212> PRT
<213> Medicago sativa
<400> 73

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

<210> 74
<211> 25
<212> PRT
<213> Medicago sativa
<400> 74

```
Leu Ser Gln Leu His Leu Gln Glu Gly Gly Leu His Ala Asp Ile Asn
```

1 5 10 15
Gly Asp Gly Val Leu Asp His Val Gln
20 25

<210> 75
<211> 107
<212> PRT
<213> Medicago sativa
<400> 75
Gln Ala Val Gly Gly Asn Gly Ala Glu Gln Thr Val Val Ser Gly Ser
1 5 10 15
Met Asp Val Leu Arg Pro Cys Trp Ala Val Ala Thr Ser Gly Val Pro
20 25 30
Val Arg Glu Gln Leu Phe Asn Val Ser Ile Cys His Tyr Thr His Phe
35 40 45
Asn Leu Phe Gln His Gly Glu Leu Tyr Arg Gly Phe Asn Arg Gly Ser
50 55 60
Asp Met Ser Ser Leu Glu Val Ala Thr Pro Ile Leu Ile Pro Arg Ser
65 70 75 80
Asp Gly His Lys His Arg Lys Gly Ser His Gly Asp Val Ile Phe Leu
85 90 95
Thr Asn Arg Gly Glu Val Arg Ser Phe Leu Phe
100 105

<210> 76
<211> 156
<212> PRT
<213> Medicago sativa
<400> 76
Gln Ile Thr Ser His Thr Pro Gly Leu His Gly His Asp Ala Val Trp
1 5 10 15
Gln Trp Gln Gln Ser Thr Gly Val Thr Trp Ser Asn Leu Pro Ser Pro
20 25 30
Ala Gly Met Met Glu Gly Gly Leu Val Ile Pro Thr Leu Lys Pro Phe
35 40 45
Pro Leu Arg Leu His Asp Asn His Glu Met Ile Leu Ala Ala Gly Glu
50 55 60
Gln Glu Ala Val Val Ile Ser Pro Gly Gly Ser Ile Leu Ala Thr Ile
65 70 75 80
Glu Leu Pro Gly Ser Pro Thr His Val Leu Ile Arg Glu Asp Phe Ser
85 90 95
Asn Asp Gly Leu Thr Asp Leu Ile Leu Val Thr Ser Ser Gly Val Tyr
100 105 110
Gly Phe Val Gln Thr Arg Gln Pro Gly Ala Leu Phe Phe Ser Val Leu
115 120 125
Ile Gly Cys Leu Ile Val Val Met Gly Ile Ile Phe Val Thr Gln His
130 135 140
Ile Asn Ser Met Lys Gly Lys Pro Arg Pro Ser Ser
145 150 155

<210> 77
<211> 1521
<212> DNA
<213> Oryza sativa
<400> 77
atggccgcca tgatggcgtc cataaccagc gagctgctct tctttctccc cttcatcctc 60
cttgccctgc tcacgttcta caccaccacc gtggccaaat gccacggcgg gcactggtgg 120
cgaggtggga cgacgccggc gaagaggaag cggatgaacc tgccgcccgg cgccgccggg 180
tggccgctcg tcggcgagac gttcggctac ctccgcgccc accccgccac ctccgtcggc 240
cgcttcatgg agcagcacat cgcacggtac gggaagatat accggtcgag cctgttcggg 300
gagcggacgg tggtgtcggc ggacgcgggg ctcaaccggt acatcctgca gaacgagggg 360
aggctgttcg agtgcagcta cccgcgcagc atcggcggca tcctgggcaa gtggtccatg 420
ctggtcctcg tcggggaccc gcaccgcgag atgcgcgcca tctccctcaa cttcctctcc 480
tccgtccgcc tccgcgccgt cctcctcccc gaggtcgagc gccacaccct cctcgtcctc 540
cgcgcctggc ccccttcctc caccttctcc gctcagcacc aagccaagaa gttcacgttc 600
aacctgatgg cgaagaacat aatgagcatg gacccggggg aggaagagac ggagcggctg 660

cggcgggagt acatcacctt catgaagggc gtggtctccg cgccgctcaa cctgcccggg 720
acgccctact ggaaggctct caagtcgcgt gctgccattc tcggagtaat agagaggaaa 780
atggaagagc gggttgagaa gctgagcaag gaggatgcaa gcgtagagca agacgatctt 840
ctcggatggg ctctgaaaca atctaacctt tcaaaagagc aaatcctgga cctcttgctg 900
agcttgctct tcgccgggca cgagacgtcg tccatggcgc tcgccctcgc catcttcttc 960
cttgaaggct gccccaaggc tgtccaagaa ctgagggagg agcatcttgg gattgcaagg 1020
agacaaaggc taagagggga gtgcaaattg agctgggaag actacaaaga gatggttttc 1080
acgcaatgtg tcataaacga gacgttgcgg ctaggaaacg tggtcaggtt cctgcaccgg 1140
aaggtcatca aggacgtgca ctacaagggt tatgacattc caagcggatg gaagatcctg 1200
ccggtgttag ccgcggtgca tctggactcg tccctgtacg aggaccccca gcgcttcaat 1260
ccctggagat ggaagagtag cggatcatcc ggcggcttgg ctcagagcag cagcttcatg 1320
ccgtacggcg gcgggacgcg gctgtgcgcc gggtcggagc tcgcgaagct ggagatggcc 1380
gtgttcttgc accacctggt gctcaacttc aggtgggagc tcgccgagcc ggaccaagcc 1440
ttcgtcttcc ccttcgtcga cttccccaag ggccttccca ttagggttca tagaattgca 1500
caggatgatg agcaggagta a 1521

<210> 78
<211> 505
<212> PRT
<213> Oryza sativa
<400> 78

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

```
            340                 345                 350
Glu Asp Tyr Lys Glu Met Val Phe Thr Gln Cys Val Ile Asn Glu Thr
        355                 360                 365
Leu Arg Leu Gly Asn Val Val Arg Phe Leu His Arg Lys Val Ile Lys
    370                 375                 380
Asp Val His Tyr Lys Gly Tyr Asp Ile Pro Ser Gly Trp Lys Ile Leu
385                 390                 395                 400
Pro Val Leu Ala Ala Val His Leu Asp Ser Ser Leu Tyr Glu Asp Pro
                405                 410                 415
Gln Arg Phe Asn Pro Trp Arg Trp Lys Ser Ser Gly Ser Ser Gly Gly
            420                 425                 430
Leu Ala Gln Ser Ser Ser Phe Met Pro Tyr Gly Gly Gly Thr Arg Leu
        435                 440                 445
Cys Ala Gly Ser Glu Leu Ala Lys Leu Glu Met Ala Val Phe Leu His
    450                 455                 460
His Leu Val Leu Asn Phe Arg Trp Glu Leu Ala Glu Pro Asp Gln Ala
465                 470                 475                 480
Phe Val Phe Pro Phe Val Asp Phe Pro Lys Gly Leu Pro Ile Arg Val
                485                 490                 495
His Arg Ile Ala Gln Asp Asp Glu Gln
            500                 505
<210> 79
<211> 1542
<212> DNA
<213> Arabidopsis thaliana
<400> 79
atgttcgaaa cagagcatca tactctctta cctcttcttc ttctcccatc gcttttgtct     60
cttcttctct tcttgattct cttgaagaga agaaatagaa aaaccagatt caatctacct    120
ccgggtaaat ccggttggcc atttcttggt gaaaccatcg gttatcttaa accgtacacc    180
gccacaacac tcggtgactt catgcaacaa catgtctcca agtatggtaa gatatataga    240
tcgaacttgt ttggagaacc aacgatcgta tcagctgatg ctggacttaa tagattcata    300
ttacaaaacg aaggaaggct ctttgaatgt agttatccta gaagtatagg tgggattctt    360
gggaaatggt cgatgcttgt tcttgttggt gacatgcata gagatatgag aagtatctcg    420
cttaacttct taagtcacgc acgtcttaga actattctac ttaaagatgt tgagagacat    480
actttgtttg ttcttgattc ttggcaacaa aactctattt tctctgctca agacgaggcc    540
aaaaagttta cgtttaatct aatggcgaag catataatga gtatggatcc tggagaagaa    600
gaaacagagc aattaaagaa agagtatgta actttcatga aaggagttgt ctctgctcct    660
ctaaatctac caggaactgc ttatcataaa gctcttcagt cacgagcaac gatattgaag    720
ttcattgaga ggaaaatgga agagagaaaa ttggatatca aggaagaaga tcaagaagaa    780
gaagaagtga aaacagagga tgaagcagag atgagtaaga gtgatcatgt taggaaacaa    840
agaacagacg atgatctttt gggatgggtt ttgaaacatt cgaatttatc gacggagcaa    900
attctcgatc tcattcttag tttgttattt gccggacatg agacttcttc tgtagccatt    960
gctctcgcta tcttcttctt gcaagcttgc cctaaagccg ttgaagagct tagggaagag   1020
catcttgaga tcgcgagggc caagaaggaa ctaggagagt cagaattaaa ttgggatgat   1080
tacaagaaaa tggactttac tcaatgtgtt ataaatgaaa ctcttcgatt gggaaatgta   1140
gttaggtttt tgcatcgcaa agcactcaaa gatgttcggt acaaaggata cgatatccct   1200
agtgggtgga aagtgttacc ggtgatctca gccgtacatt tggataattc tcgttatgac   1260
caacctaatc tctttaatcc ttggagatgg caacagcaaa acaacggagc gtcatcgtca   1320
ggaagtggta gtttttcgac gtggggaaac aactacatgc cgtttggagg agggccaagg   1380
ctatgtgctg gttcagagct agccaagtta gaaatggcag tgtttattca tcatctagtt   1440
cttaaattca attgggaatt agcagaagat gatcaaccat ttgcttttcc ttttgttgat   1500
tttcctaacg gtttgcctat tagggtttct cgtattctgt aa                      1542
<210> 80
<211> 513
<212> PRT
<213> Arabidopsis thaliana
<400> 80
Met Phe Glu Thr Glu His His Thr Leu Leu Pro Leu Leu Leu Leu Pro
1               5                   10                  15
Ser Leu Leu Ser Leu Leu Leu Phe Leu Ile Leu Leu Lys Arg Arg Asn
            20                  25                  30
Arg Lys Thr Arg Phe Asn Leu Pro Pro Gly Lys Ser Gly Trp Pro Phe
        35                  40                  45
```

```
Leu Gly Glu Thr Ile Gly Tyr Leu Lys Pro Tyr Thr Ala Thr Thr Leu
    50                  55                  60
Gly Asp Phe Met Gln Gln His Val Ser Lys Tyr Gly Lys Ile Tyr Arg
65                  70                  75                  80
Ser Asn Leu Phe Gly Glu Pro Thr Ile Val Ser Ala Asp Ala Gly Leu
                85                  90                  95
Asn Arg Phe Ile Leu Gln Asn Glu Gly Arg Leu Phe Glu Cys Ser Tyr
            100                 105                 110
Pro Arg Ser Ile Gly Gly Ile Leu Gly Lys Trp Ser Met Leu Val Leu
        115                 120                 125
Val Gly Asp Met His Arg Asp Met Arg Ser Ile Ser Leu Asn Phe Leu
    130                 135                 140
Ser His Ala Arg Leu Arg Thr Ile Leu Leu Lys Asp Val Glu Arg His
145                 150                 155                 160
Thr Leu Phe Val Leu Asp Ser Trp Gln Gln Asn Ser Ile Phe Ser Ala
                165                 170                 175
Gln Asp Glu Ala Lys Lys Phe Thr Phe Asn Leu Met Ala Lys His Ile
            180                 185                 190
Met Ser Met Asp Pro Gly Glu Glu Glu Thr Glu Gln Leu Lys Lys Glu
        195                 200                 205
Tyr Val Thr Phe Met Lys Gly Val Val Ser Ala Pro Leu Asn Leu Pro
    210                 215                 220
Gly Thr Ala Tyr His Lys Ala Leu Gln Ser Arg Ala Thr Ile Leu Lys
225                 230                 235                 240
Phe Ile Glu Arg Lys Met Glu Glu Arg Lys Leu Asp Ile Lys Glu Glu
                245                 250                 255
Asp Gln Glu Glu Glu Glu Val Lys Thr Glu Asp Glu Ala Glu Met Ser
            260                 265                 270
Lys Ser Asp His Val Arg Lys Gln Arg Thr Asp Asp Asp Leu Leu Gly
        275                 280                 285
Trp Val Leu Lys His Ser Asn Leu Ser Thr Glu Gln Ile Leu Asp Leu
    290                 295                 300
Ile Leu Ser Leu Leu Phe Ala Gly His Glu Thr Ser Ser Val Ala Ile
305                 310                 315                 320
Ala Leu Ala Ile Phe Phe Leu Gln Ala Cys Pro Lys Ala Val Glu Glu
                325                 330                 335
Leu Arg Glu Glu His Leu Glu Ile Ala Arg Ala Lys Lys Glu Leu Gly
            340                 345                 350
Glu Ser Glu Leu Asn Trp Asp Asp Tyr Lys Lys Met Asp Phe Thr Gln
        355                 360                 365
Cys Val Ile Asn Glu Thr Leu Arg Leu Gly Asn Val Val Arg Phe Leu
    370                 375                 380
His Arg Lys Ala Leu Lys Asp Val Arg Tyr Lys Gly Tyr Asp Ile Pro
385                 390                 395                 400
Ser Gly Trp Lys Val Leu Pro Val Ile Ser Ala Val His Leu Asp Asn
                405                 410                 415
Ser Arg Tyr Asp Gln Pro Asn Leu Phe Asn Pro Trp Arg Trp Gln Gln
            420                 425                 430
Gln Asn Asn Gly Ala Ser Ser Ser Gly Ser Gly Ser Phe Ser Thr Trp
        435                 440                 445
Gly Asn Asn Tyr Met Pro Phe Gly Gly Gly Pro Arg Leu Cys Ala Gly
    450                 455                 460
Ser Glu Leu Ala Lys Leu Glu Met Ala Val Phe Ile His His Leu Val
465                 470                 475                 480
Leu Lys Phe Asn Trp Glu Leu Ala Glu Asp Asp Gln Pro Phe Ala Phe
                485                 490                 495
Pro Phe Val Asp Phe Pro Asn Gly Leu Pro Ile Arg Val Ser Arg Ile
            500                 505                 510
Leu
<210> 81
<211> 1590
<212> DNA
<213> Saccharum officinarum
```

<220>
<221> misc_feature
<222> (394)..(394)
<223> n is a, c, g, ou t
<400> 81
atgggcgcca tgatggcctc cataaccagc gagctcctct tcttccttcc cttcatcctg 60
ctggccctcc tcgccctgta caccaccact gtcgccaaat gcgacggcac ccaccagtgg 120
cgccggccga agaagaagcg gccgaacctg cccccgggcg ccctcggatg gcctttcgtc 180
ggcgagacct tcggctacct ccgcgcccac ccggccacct ccgtgggcct cttcatggag 240
cagcatgtcg cacggtacgg caagatatac cggtcgagcc tgttcgggga gcggacggtg 300
gtgtcggcgg acgcggggct caaccgctac atcctgcaga acgaggggcg gctgttcgag 360
tgcagctacc cgcgcagcat cggcggcatc ctgngcaagt ggtccatgct ggtgctggtg 420
ggcgacgcgc accgcgagat gcgcgccatc tcgctcaact ttctcagctc cgtccgcctc 480
cgcgccgtgc tgctcccgga ggtggagcgc cacaccctgc tggtgctccg ctcatggccg 540
ccctccgacg gcacggtctc cgcgcagcac caagccaaga agttcacgtt taacctgatg 600
gcgaagaaca taatgagcat ggaccccggc gaggaggaga cggagcggct gcggctggag 660
tacatcacct tcatgaaggg cgtcgtgtca gcgccgctca acttcccggg cacggcctac 720
tggaaggcgc tcaagtcacg cgcgtccata cttggagtaa tagagaggaa gatggaggac 780
aggcttcaga aaatgagtaa ggagaactca agtgtggagg aagacgatct tcttggatgg 840
gccctgaagc agtccaatct gtcaaaggaa cagatcctgg acctcttgct gagcctgctc 900
ttcgccgggc acgagacttc gtcaatggcg ctagccctcg ccatcttctt ccttgaagga 960
tgccccaagg ctgttcaaga actccgggag gagcatctcg agattgctag gagacaaagg 1020
ctaagagggg cgttcaaatt gagctgggaa gactacaagg aaatggtttt cacgccatgg 1080
tgtataaacg agacattgcg ggttggcaac gtggtcaggt tcctgcaccg gaaggtcatc 1140
caagatgtgc actacaatgg gtacgacata ccacgcgggt ggaaaatcct gccggtgtta 1200
gcggcggtgc atctggattc gtcgctgtac aaggacccct accggttcaa cccttggaga 1260
tggaagagca acgcgccgag cagcttcatg ccgtacggcg gcgggccgcg gctgtgcgcc 1320
gggtcggagc tggccaagct ggagatcgcc atcttcctgc accacctggt gctcaacttc 1380
cggtgggagc tggcggagcc ggaccaagcc ttcgtctacc ccttcgtcga cttccccaag 1440
ggcctcccga tcagggtcca gcggatcgcc gacgaccagg gggcatcgca gcgttttgac 1500
cgagaagcac gatgtagcgg tacgggtaca caagcaagta caaaattttc gttgcatttt 1560
gagggaattg gggctgggtg gttatggtaa 1590
<210> 82
<211> 529
<212> PRT
<213> Saccharum officinarum
<220>
<221> INSEGURO
<222> (132)..(132)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 82
Met Gly Ala Met Met Ala Ser Ile Thr Ser Glu Leu Leu Phe Phe Leu
1 5 10 15
Pro Phe Ile Leu Leu Ala Leu Leu Ala Leu Tyr Thr Thr Thr Val Ala
20 25 30
Lys Cys Asp Gly Thr His Gln Trp Arg Arg Pro Lys Lys Lys Arg Pro
35 40 45
Asn Leu Pro Pro Gly Ala Leu Gly Trp Pro Phe Val Gly Glu Thr Phe
50 55 60
Gly Tyr Leu Arg Ala His Pro Ala Thr Ser Val Gly Leu Phe Met Glu
65 70 75 80
Gln His Val Ala Arg Tyr Gly Lys Ile Tyr Arg Ser Ser Leu Phe Gly
85 90 95
Glu Arg Thr Val Val Ser Ala Asp Ala Gly Leu Asn Arg Tyr Ile Leu
100 105 110
Gln Asn Glu Gly Arg Leu Phe Glu Cys Ser Tyr Pro Arg Ser Ile Gly
115 120 125
Gly Ile Leu Xaa Lys Trp Ser Met Leu Val Leu Val Gly Asp Ala His
130 135 140
Arg Glu Met Arg Ala Ile Ser Leu Asn Phe Leu Ser Ser Val Arg Leu
145 150 155 160
Arg Ala Val Leu Leu Pro Glu Val Glu Arg His Thr Leu Leu Val Leu
165 170 175

```
Arg Ser Trp Pro Pro Ser Asp Gly Thr Val Ser Ala Gln His Gln Ala
            180                 185                 190
Lys Lys Phe Thr Phe Asn Leu Met Ala Lys Asn Ile Met Ser Met Asp
        195                 200                 205
Pro Gly Glu Glu Glu Thr Glu Arg Leu Arg Leu Glu Tyr Ile Thr Phe
    210                 215                 220
Met Lys Gly Val Val Ser Ala Pro Leu Asn Phe Pro Gly Thr Ala Tyr
225                 230                 235                 240
Trp Lys Ala Leu Lys Ser Arg Ala Ser Ile Leu Gly Val Ile Glu Arg
                245                 250                 255
Lys Met Glu Asp Arg Leu Gln Lys Met Ser Lys Glu Asn Ser Ser Val
            260                 265                 270
Glu Glu Asp Asp Leu Leu Gly Trp Ala Leu Lys Gln Ser Asn Leu Ser
        275                 280                 285
Lys Glu Gln Ile Leu Asp Leu Leu Leu Ser Leu Leu Phe Ala Gly His
    290                 295                 300
Glu Thr Ser Ser Met Ala Leu Ala Leu Ala Ile Phe Phe Leu Glu Gly
305                 310                 315                 320
Cys Pro Lys Ala Val Gln Glu Leu Arg Glu Glu His Leu Glu Ile Ala
                325                 330                 335
Arg Arg Gln Arg Leu Arg Gly Ala Phe Lys Leu Ser Trp Glu Asp Tyr
            340                 345                 350
Lys Glu Met Val Phe Thr Pro Trp Cys Ile Asn Glu Thr Leu Arg Val
        355                 360                 365
Gly Asn Val Val Arg Phe Leu His Arg Lys Val Ile Gln Asp Val His
    370                 375                 380
Tyr Asn Gly Tyr Asp Ile Pro Arg Gly Trp Lys Ile Leu Pro Val Leu
385                 390                 395                 400
Ala Ala Val His Leu Asp Ser Ser Leu Tyr Lys Asp Pro Tyr Arg Phe
                405                 410                 415
Asn Pro Trp Arg Trp Lys Ser Asn Ala Pro Ser Ser Phe Met Pro Tyr
            420                 425                 430
Gly Gly Gly Pro Arg Leu Cys Ala Gly Ser Glu Leu Ala Lys Leu Glu
        435                 440                 445
Ile Ala Ile Phe Leu His His Leu Val Leu Asn Phe Arg Trp Glu Leu
    450                 455                 460
Ala Glu Pro Asp Gln Ala Phe Val Tyr Pro Phe Val Asp Phe Pro Lys
465                 470                 475                 480
Gly Leu Pro Ile Arg Val Gln Arg Ile Ala Asp Asp Gln Gly Ala Ser
                485                 490                 495
Gln Arg Phe Asp Arg Glu Ala Arg Cys Ser Gly Thr Gly Thr Gln Ala
            500                 505                 510
Ser Thr Lys Phe Ser Leu His Phe Glu Gly Ile Gly Ala Gly Trp Leu
        515                 520                 525
Trp
<210> 83
<211> 1437
<212> DNA
<213> Allium cepa
<400> 83
atggagatca tattagtgtc tacattgatc atatcgctac taatatttct tggatttaga     60
agcaatggga aaacggagag aaaattgctg cctacactac caccaggcaa tcttggaggt    120
tggccattca tcggtgacac cattccgttc atgacacctc attcttctgc tttgttgggc    180
acttacatcg atcaaaatat ttccaaatat gggaggatat ttcgaatgaa cttgttagga    240
aaggcaacga tcgtgtctgt agaccctgat ttcaacagat atattctaca gaatgaagga    300
agattgtttg aaaatagctg cccaacgagc attaaagaga ttttgggaaa atggtctatg    360
cttgcattag ctggggatat acacagagaa atgagatcca ttgctgtgaa tttcatgaac    420
agtgttaagc ttagaactta ttttttaaag gatattgata ttcaggctgt taatattctt    480
gatgcttgga aggtcaactc tactttctct gcacaggatg aaggaaagaa gtttgcattt    540
aacctcatgg tgaagcatct aatgaatatg gatcctggaa tgccagagac agaagaaatc    600
agaaaagagt acatttttctt catggagggg atggcttcca ttcctttaaa ctttcctgga    660
acagcctaca gaagagcttt acagtcaagg tccaggattc tggcaataat ggggcaaaag    720
cttgacgaaa ggatgcagaa aataaaagaa ggctgtaaag gactggaaga agaggatctt    780
```

cttgcctcag ttgcaagaaa tactaacata acaagagacc agattcttga tttgatgatc 840
agcatgcttt ttgctggtca tgaaacttct tctgccgcta tttctcttgc catttatttc 900
ctgcaagctt cccccgatgt tcttaaaaag cttcgagagg agcacataaa gattgcaaaa 960
caaaagaaag aaaggggcga aactgaattg aactgggatg attacaagca aatggaattc 1020
acaaactgcg ttattcatga aaccctaaga ttaggcaaca tcgttaagtt tttgcatcgg 1080
aaaaccatca aagatgttca atacaaaggt tatgaaattc catgtgggtg ggaagtagtg 1140
ccaatcatct cagcagcaca tttggattct tctatctttg acaacccaaa agttatgaat 1200
ccttcgaggt gggaggcgat attttcagca ggagcaaaga gcaatataat gtcattcagc 1260
ggtggacctc ggttatgtcc aggagcagag ttggcgaaac tggagatggc tatttttctt 1320
catcatcttg tacagaggtt cgactgggaa ttggtggaga aggataaccc tgtatcattc 1380
cccttccttg gatttcccaa gaaattgcct atcaaaatca cagctcttaa acattga 1437

<210> 84
<211> 478
<212> PRT
<213> Allium cepa
<400> 84

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

```
    370                 375                 380
Ala Ala His Leu Asp Ser Ser Ile Phe Asp Asn Pro Lys Val Met Asn
385                 390                 395                 400
Pro Ser Arg Trp Glu Ala Ile Phe Ser Ala Gly Ala Lys Ser Asn Ile
                405                 410                 415
Met Ser Phe Ser Gly Gly Pro Arg Leu Cys Pro Gly Ala Glu Leu Ala
            420                 425                 430
Lys Leu Glu Met Ala Ile Phe Leu His His Leu Val Gln Arg Phe Asp
        435                 440                 445
Trp Glu Leu Val Glu Lys Asp Asn Pro Val Ser Phe Pro Phe Leu Gly
    450                 455                 460
Phe Pro Lys Lys Leu Pro Ile Lys Ile Thr Ala Leu Lys His
465                 470                 475
<210> 85
<211> 1491
<212> DNA
<213> Zinnia elegans
<400> 85
atgtgttcaa caaccctaaa tatgtgtgac cttgagttct tcatccttgc atcttgtctt     60
gtcttggctc tttttctcat cttgaagctt gtcaaaagaa gaacaaacaa tggttcgact    120
cgaaatcttc caccgggcaa catgggctgg ccgttcatcg gtgaaaccat cggttacctc    180
caaccgtatt cggctacaac catcggcaag tttatggaac aacacatatc caagtatggg    240
aagatataca aatctagttt gtttggtgag ccaacaatag tttctgctga tccagggttg    300
aataagtaca tattgcaaaa tgaagggagg ttatttgaat gtagttatcc aagaagcata    360
gggggtattc ttggcaaatg gtccatgttg gttttggttg gtgacatgca tagagacatg    420
aggcaaattt cactcaactt tttgtccaat gcaaggctta aaactcaact agtaaatgaa    480
gttgagaaaa atactttgtg ggtattagat tcttggaaag aaaactcacc tttttgtgcc    540
caagaagaag ccaagaagtt tacttttaat ctaatggcaa cacatatcat gagtttagac    600
ccgggtgaac cggagaccga gcgattgaag aaagagtatg taactttcat gaaaggtgtg    660
gtttctcccc ctttaaactt ccctggaact gcatactgga aagctttaaa gtctcgagcg    720
acgattctta aattcatcga aacaaaaatg gaggagcgga ttaggatgga cgaaggaaac    780
ggattaggga aactagacaa tgatcttctt ggatggtcta tgaagaactc aaatctcact    840
aaagagcaaa tactcgattt ggtattgagt ctactctttg ctggtcacga aacgtcttcg    900
gtttcgatat cgttagccgt ttacttcctt gaagcttgtc ctaccgcggt tcgtcaactt    960
agagaagaac atgaagaaat tgtgatgaaa aaaaagctat tgggtgagaa gtatctcact   1020
tgggatgact acaaaaagat ggagtttact cagtgtgtga tcaatgagac gctaagattc   1080
gggaatgtgg tgagattcct ccacagaaag gctattaaag atgtgaggta taaaggatat   1140
gacattccat gtggttggaa agtgctgcca gtgattgcag ccgtgcattt ggatcctaca   1200
cattttgacc aaccttacct ttttgatcca tggagatggc agaacgcaag tgtcacgtca   1260
tctacttgtt caaccccgcc atcagcaagt aacttcatgc catttggtgg agggccccgc   1320
ttatgcacag ggtcagagct agcgaaacta gagatggcga tatttatcca ccatttggtc   1380
ctaaaatacg agtgggaatt ggttgactca gatgaagcat tcgcttatcc atatctcgac   1440
tttccaaaag gtctgccaat caaaatccgt caccgaaaac aatcatgtta g             1491
<210> 86
<211> 496
<212> PRT
<213> Zinnia elegans
<400> 86
Met Cys Ser Thr Thr Leu Asn Met Cys Asp Leu Glu Phe Phe Ile Leu
1                   5                  10                  15
Ala Ser Cys Leu Val Leu Ala Leu Phe Leu Ile Leu Lys Leu Val Lys
            20                  25                  30
Arg Arg Thr Asn Asn Gly Ser Thr Arg Asn Leu Pro Pro Gly Asn Met
        35                  40                  45
Gly Trp Pro Phe Ile Gly Glu Thr Ile Gly Tyr Leu Gln Pro Tyr Ser
    50                  55                  60
Ala Thr Thr Ile Gly Lys Phe Met Glu Gln His Ile Ser Lys Tyr Gly
65                  70                  75                  80
Lys Ile Tyr Lys Ser Ser Leu Phe Gly Glu Pro Thr Ile Val Ser Ala
                85                  90                  95
Asp Pro Gly Leu Asn Lys Tyr Ile Leu Gln Asn Glu Gly Arg Leu Phe
            100                 105                 110
Glu Cys Ser Tyr Pro Arg Ser Ile Gly Gly Ile Leu Gly Lys Trp Ser
```

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

<210> 87
<211> 1422
<212> DNA
<213> Medicago trunculata
<400> 87

```
atgtctaact catacttaac ttgcagtttt ctttcttcca tctttgttct ttctttgatt     60
ttcattttca tcaaaagaaa gaaaacaagg tataatcttc cacctggaaa aatgggatgg    120
ccctttatag gagaaaccat tggttatttg aagccttaca ctgccaccac aatgggagaa    180
tttatggaaa atcacatagc aaggtatggg acaatttaca agtcaaattt gtttggaggg    240
ccagctattg tatcagcaga tgcagaattg aataggttca tattacaaaa tgatggaaaa    300
ttgtttgagt gtagctatcc aaaaagcatt ggtggaatac ttggaaaatg gtcaatgttg    360
gttttagtag gtgacatgca tagggaaatg aggaatatat cactaaactt tatgagctat    420
gctaggctta aaacacattt tttgaaagat atggagaagc atacccttt tgttctaagc    480
tcttggaaag aaaattgtac attttcagct caagatgaag caaaaaagtt caccttcaat    540
ttgatggcca aacaaatcat gagtttggat ccagggaatc ttgagacaga acagttgaaa    600
aaagagtatg tctgtttcat gaaaggtgtt gtttctgctc ctttgaattt gccaggaact    660
```

gcatacagaa aagcattaaa gtctaggaac aatatattga agttcataga ggggaaaatg 720
gaagaaaggg tgaagagaaa ccaagaagga aaaaaaggga tggaggaaaa tgatcttcta 780
aattgggttt taaagcattc aaatctttcc actgagcaaa ttcttgactt gattctaagt 840
ttacttttttg ctggccatga aacttcatct gtggctatag ctctagctat ttacttttttg 900
cctagttgtc ctcaagctat acaacaatta agggaagagc atagagaaat agctagatcc 960
aagaagaaag caggggaggt tgaattaact tgggatgatt ataaaagaat ggaatttact 1020
cattgtgttg tgaatgaaac actaaggttg ggtaatgttg tgagattcct tcacaggaag 1080
gctatcaaag atgttcatta caaaggttat gacattccat gtggatggaa agtccttccg 1140
gtgatttcag cggtacattt ggatccttca aattttgacc aacctcaaca cttcaatcct 1200
tggagatggc agatgagcaa taacaacttc atgccatttg gaggaggacc aaggctatgt 1260
gcaggattag aattagccaa acttgaaatg gctgttttca ttcaccatat catcctcaaa 1320
tacaactggg acatggtcga tgttgatcaa cctattgtat acccttttgt tgattttccc 1380
aaaggtttgc caattagagt ccaaagccaa gccactcttt aa 1422

<210> 88
<211> 480
<212> PRT
<213> Medicago trunculata
<400> 88

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

```
         355                 360                 365
Gly Tyr Asp Ile Pro Cys Gly Trp Lys Val Leu Pro Val Ile Ser Ala
    370                 375                 380
Val His Leu Asp Pro Ser Asn Phe Asp Gln Pro Gln His Phe Asn Pro
385                 390                 395                 400
Trp Arg Trp Gln Gln Asn Asn Asp Gly Ala Ser Gly Asn Ser Asn Ile
                405                 410                 415
Phe Leu Pro Phe Gly Gly Gly Pro Arg Leu Cys Ala Gly Leu Glu Leu
            420                 425                 430
Ala Lys Leu Glu Met Ala Val Phe Ile His His Ile Ile Leu Lys Tyr
        435                 440                 445
Asn Trp Asp Met Val Asp Val Asp Gln Pro Ile Val Tyr Pro Phe Val
    450                 455                 460
Asp Phe Pro Lys Gly Leu Pro Ile Arg Val Gln Ser Gln Ala Thr Leu
465                 470                 475                 480
<210> 89
<211> 1485
<212> DNA
<213> Populus trichocarpa
<220>
<221> misc_feature
<222> (1110)..(1110)
<223> n is a, c, g, ou t
<400> 89
atgtctcact cagagcttgt tgtctttctc cttccatcga ttttatcact actcttgctc     60
ttcattctcg tgcaaagaaa gcaagtaaga tttaatctcc caccgggcaa catggggtgg    120
ccatttcttg gagaaaccat tggctacctg aagccttact ctgctacttc aataggagaa    180
ttcatggaac agcacatatc aaggtatgga aagatttaca agtccaattt gtttggggag    240
ccaacaatag tatccgcaga tgctggactt agcagattta tactacagaa tgagggaaga    300
ttatttgaat gcagctatcc aaaaagtatt ggtggaattc ttggaaaatg gtccatgatg    360
gttcttgttg gagacatgca tagagacatg aggattatat ctctcaactt tttgagccat    420
gccaggttaa gaactcatct attgaaagaa gtggagaagc aaaccctgct tgttcttagc    480
tcttggaagg agaattgtac attttcagct caagatgaag caaacaagtt taccttcaat    540
tggatggcaa aacatatcat gagcttggat cctggaaaga cagagactga gcagctgaaa    600
aaagagtatg ttactttcat gaaaggagta gtttcaggtc ctataaattt tcctggaacc    660
ccatatagaa aagccttgaa gtctcgatca atcatcttga aatttataga gcgaaagatg    720
gaggagagaa ttggagaaac gaagggtgga gtagaaaact tggaagacga tgatcttctt    780
ggatgggtct tgaagcattc aaatctttat acggagcaaa tccttgaatt aatcttaagc    840
ttgctctttg ctggccacga aacttcttct gtgtccatag ctctagccat atccttcttg    900
caagcttgtc ctggttctat tcaacagtta aaagaagaac atattcaaat ctccagagcc    960
aagaaacggt caggagagac ggaattgacc tgggatgatt acaaaaaaat ggaattcact   1020
caatgcgtta taagcgagac agtgaggctt ggaaacgtag tcaggtttgt tcacagaaaa   1080
gctctaaaag atgttcggta caaagggtan gacattccat gtggatggaa agtgttacca   1140
gtaatctcat ccgtccattt agattcaact cttttcgacc aacctcaaca cttcaatcca   1200
tggagatggc agcagcacaa caatgctcgt ggatcttcta cttgttcgag tgcggcggcg   1260
gcggcggcgg cggcggtgag tagtaatcac ttcatgccat ttgggggagg accgcgactc   1320
tgtgcaggat cggaattggc aaaacttgaa atggcagttt tcattcacca tttggttctg   1380
aacttccatt gggaattggt cggtgccgat caagcctttg cctttccttt tgttgatttt   1440
cctaaaggct tgccaataag agtcaagcac cacacagtca tataa                   1485
<210> 90
<211> 494
<212> PRT
<213> Populus trichocarpa
<220>
<221> INSEGURO
<222> (370)..(370)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 90
Met Ser His Ser Glu Leu Val Val Phe Leu Leu Pro Ser Ile Leu Ser
1               5                   10                  15
Leu Leu Leu Leu Phe Ile Leu Val Gln Arg Lys Gln Val Arg Phe Asn
            20                  25                  30
Leu Pro Pro Gly Asn Met Gly Trp Pro Phe Leu Gly Glu Thr Ile Gly
```

```
          35                  40                  45
Tyr Leu Lys Pro Tyr Ser Ala Thr Ser Ile Gly Glu Phe Met Glu Gln
    50                  55                  60
His Ile Ser Arg Tyr Gly Lys Ile Tyr Lys Ser Asn Leu Phe Gly Glu
65                  70                  75                  80
Pro Thr Ile Val Ser Ala Asp Ala Gly Leu Ser Arg Phe Ile Leu Gln
                85                  90                  95
Asn Glu Gly Arg Leu Phe Glu Cys Ser Tyr Pro Lys Ser Ile Gly Gly
            100                 105                 110
Ile Leu Gly Lys Trp Ser Met Met Val Leu Val Gly Asp Met His Arg
        115                 120                 125
Asp Met Arg Ile Ile Ser Leu Asn Phe Leu Ser His Ala Arg Leu Arg
    130                 135                 140
Thr His Leu Leu Lys Glu Val Glu Lys Gln Thr Leu Leu Val Leu Ser
145                 150                 155                 160
Ser Trp Lys Glu Asn Cys Thr Phe Ser Ala Gln Asp Glu Ala Asn Lys
                165                 170                 175
Phe Thr Phe Asn Trp Met Ala Lys His Ile Met Ser Leu Asp Pro Gly
            180                 185                 190
Lys Thr Glu Thr Glu Gln Leu Lys Lys Glu Tyr Val Thr Phe Met Lys
        195                 200                 205
Gly Val Val Ser Gly Pro Ile Asn Phe Pro Gly Thr Pro Tyr Arg Lys
    210                 215                 220
Ala Leu Lys Ser Arg Ser Ile Ile Leu Lys Phe Ile Glu Arg Lys Met
225                 230                 235                 240
Glu Glu Arg Ile Gly Glu Thr Lys Gly Gly Val Glu Asn Leu Glu Asp
                245                 250                 255
Asp Asp Leu Leu Gly Trp Val Leu Lys His Ser Asn Leu Tyr Thr Glu
            260                 265                 270
Gln Ile Leu Glu Leu Ile Leu Ser Leu Leu Phe Ala Gly His Glu Thr
        275                 280                 285
Ser Ser Val Ser Ile Ala Leu Ala Ile Ser Phe Leu Gln Ala Cys Pro
    290                 295                 300
Gly Ser Ile Gln Gln Leu Lys Glu Glu His Ile Gln Ile Ser Arg Ala
305                 310                 315                 320
Lys Lys Arg Ser Gly Glu Thr Glu Leu Thr Trp Asp Asp Tyr Lys Lys
                325                 330                 335
Met Glu Phe Thr Gln Cys Val Ile Ser Glu Thr Val Arg Leu Gly Asn
            340                 345                 350
Val Val Arg Phe Val His Arg Lys Ala Leu Lys Asp Val Arg Tyr Lys
        355                 360                 365
Gly Xaa Asp Ile Pro Cys Gly Trp Lys Val Leu Pro Val Ile Ser Ser
    370                 375                 380
Val His Leu Asp Ser Thr Leu Phe Asp Gln Pro Gln His Phe Asn Pro
385                 390                 395                 400
Trp Arg Trp Gln Gln His Asn Asn Ala Arg Gly Ser Ser Thr Cys Ser
                405                 410                 415
Ser Ala Ala Ala Ala Ala Ala Ala Ala Val Ser Ser Asn His Phe Met
            420                 425                 430
Pro Phe Gly Gly Gly Pro Arg Leu Cys Ala Gly Ser Glu Leu Ala Lys
        435                 440                 445
Leu Glu Met Ala Val Phe Ile His His Leu Val Leu Asn Phe His Trp
    450                 455                 460
Glu Leu Val Gly Ala Asp Gln Ala Phe Ala Phe Pro Phe Val Asp Phe
465                 470                 475                 480
Pro Lys Gly Leu Pro Ile Arg Val Lys His His Thr Val Ile
                485                 490
<210> 91
<211> 1407
<212> DNA
<213> Aquilegia formosa x Aquilegia pubescens
<220>
<221> misc_feature
```

<222> (1031)..(1032)
<223> n is a, c, g, ou t
<400> 91
atgcctgagc ttgtgttttt cttttcatta gctccagcaa ttttagcact aatacttctc 60
ttgaaactct tcaaaaggaa gaaaaagtca tataatcttc caccaggaaa catgggttgg 120
ccatatctag gcgaaactct tggttatttg aagccttatt gtgctatcac tactggagat 180
ttcatggagc aacatatatc aaggtatggg aagatctaca agtcaaattt atttggttat 240
cctacaatag tttcagttga tcctgaatta aatcgatatg tattacaaaa cgaaggaaga 300
ctgtttgaat gtagttatcc aagtagttta ggtgggattc ttggcaaatg gtcaatgtta 360
gttttggttg gagacatgca taaaaacatg aggatgatct ctgtcaactt catgagcagc 420
gcaagacttc gaacacatct aattcaagat gtggagactc aagccttatt ggtgctgaaa 480
tcttggcagg ttgataaaaa gattttggcc caagatgaag caaagaagtt taccttcaat 540
ttaattgtaa aaaatataat gagcatggaa cctggaacac ctgaaagtga gaagcttagg 600
agggaataca ttacattcat gaaaggaatc atttctgcgc ctttgaattt gcctggaact 660
gcatatagaa gagccttaaa gtctcgatcg aacattctgc aacttatcga gcataacatg 720
aatgagagac tccaaaagac taacggagat ggtaagaaag tggaagatga tgatctactt 780
ggatgggtct tgaagcattc aaatcttacc actgaacaaa ttcttgattt gatactaagt 840
atgcttttcg cgggtcatga aacttcctca gtatctatat ctctagccat atacctttg 900
caaggatgcc taaaagcagt tgaagagtta agggaagagc atattagaat tgctaaagca 960
aaagaacagg caggacagag atctggatta aattgggaag attacaaaca catggaattc 1020
actcaatgtg nntttcttca tagaaaaact ctcaaagatg ttcagtacaa agggtatgac 1080
attccatgtg gttggaaagt tcttccagta tttgcagcag ttcatttgga ccctttaaat 1140
ttcgaccaac ctcatgcttt caatccatgg agatggcaga atgggaagac aagcacgacg 1200
actaacaact tcatgccgtt tggtggtgga ttacggttat gtgctggttc agagctagcc 1260
aagttggaaa tggctatttt cattcaccat ttggtcttga attatgattg ggatatagca 1320
gaaccagatc aaccatttgc ctacccattt gttgaatttc caaaaggtct accaattaag 1380
gtctatgacc atcagtgctt aacatga 1407
<210> 92
<211> 468
<212> PRT
<213> Aquilegia formosa x Aquilegia pubescens
<220>
<221> INSEGURO
<222> (344)..(344)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 92
Met Pro Glu Leu Val Phe Phe Phe Ser Leu Ala Pro Ala Ile Leu Ala
1 5 10 15
Leu Ile Leu Leu Leu Lys Leu Phe Lys Arg Lys Lys Lys Ser Tyr Asn
20 25 30
Leu Pro Pro Gly Asn Met Gly Trp Pro Tyr Leu Gly Glu Thr Leu Gly
35 40 45
Tyr Leu Lys Pro Tyr Cys Ala Ile Thr Thr Gly Asp Phe Met Glu Gln
50 55 60
His Ile Ser Arg Tyr Gly Lys Ile Tyr Lys Ser Asn Leu Phe Gly Tyr
65 70 75 80
Pro Thr Ile Val Ser Val Asp Pro Glu Leu Asn Arg Tyr Val Leu Gln
85 90 95
Asn Glu Gly Arg Leu Phe Glu Cys Ser Tyr Pro Ser Ser Leu Gly Gly
100 105 110
Ile Leu Gly Lys Trp Ser Met Leu Val Leu Val Gly Asp Met His Lys
115 120 125
Asn Met Arg Met Ile Ser Val Asn Phe Met Ser Ser Ala Arg Leu Arg
130 135 140
Thr His Leu Ile Gln Asp Val Glu Thr Gln Ala Leu Leu Val Leu Lys
145 150 155 160
Ser Trp Gln Val Asp Lys Lys Ile Leu Ala Gln Asp Glu Ala Lys Lys
165 170 175
Phe Thr Phe Asn Leu Ile Val Lys Asn Ile Met Ser Met Glu Pro Gly
180 185 190
Thr Pro Glu Ser Glu Lys Leu Arg Arg Glu Tyr Ile Thr Phe Met Lys
195 200 205
Gly Ile Ile Ser Ala Pro Leu Asn Leu Pro Gly Thr Ala Tyr Arg Arg

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

<210> 93
<211> 641
<212> DNA
<213> Triticum aestivum
<400> 93
atggccgcca tcatggcctc cataaccagc gagctccttt tcttcctccc cttcatcctc 60
ctggccctgc tcaccttcta caccagcgcc gtggctaaat gccatggcct ccactggtgg 120
agcggccgga cgaagaagag gcggccgaac ctgccgcccg gcgccgtcgg ctggcccttc 180
atcggcgaga ccttcgggta cctccgcgcc cacccggcca cctccatcgg ccagttcatg 240
gaccagcaca tcgcacggta cgggaagata taccggtcga gcctgttcgg ggaccggacg 300
gtggtgtcgg cggacgcggg gctgaaccgg tacatcctgc agaacgaggg gcggctgttc 360
gagtgtagct acccgcggag catcggcggc atccttggca aatggtcgat gctggtgctc 420
gtcggcgacc cccaccgcga gatgcgcttc atctccctca acttcctcag ctccgtccgc 480
ctccgcgccg tgctcctccc ggaggtggag cgccacaccc tcctcgtcct ccgcgactgg 540
ctgccttact cctcctcctc cgtcttctcc gcgcagcacg aagccaagaa gttcacgttt 600
aacctgatgg cgaagaacat catgagcatg gaccccggcg a 641

<210> 94
<211> 213
<212> PRT
<213> Triticum aestivum
<400> 94
Met Ala Ala Ile Met Ala Ser Ile Thr Ser Glu Leu Leu Phe Phe Leu
1 5 10 15
Pro Phe Ile Leu Leu Ala Leu Leu Thr Phe Tyr Thr Ser Ala Val Ala
20 25 30
Lys Cys His Gly Leu His Trp Trp Ser Gly Arg Thr Lys Lys Arg Arg
35 40 45
Pro Asn Leu Pro Pro Gly Ala Val Gly Trp Pro Phe Ile Gly Glu Thr
50 55 60
Phe Gly Tyr Leu Arg Ala His Pro Ala Thr Ser Ile Gly Gln Phe Met

65 70 75 80
Asp Gln His Ile Ala Arg Tyr Gly Lys Ile Tyr Arg Ser Ser Leu Phe
85 90 95
Gly Asp Arg Thr Val Val Ser Ala Asp Ala Gly Leu Asn Arg Tyr Ile
100 105 110
Leu Gln Asn Glu Gly Arg Leu Phe Glu Cys Ser Tyr Pro Arg Ser Ile
115 120 125
Gly Gly Ile Leu Gly Lys Trp Ser Met Leu Val Leu Val Gly Asp Pro
130 135 140
His Arg Glu Met Arg Phe Ile Ser Leu Asn Phe Leu Ser Ser Val Arg
145 150 155 160
Leu Arg Ala Val Leu Leu Pro Glu Val Glu Arg His Thr Leu Leu Val
165 170 175
Leu Arg Asp Trp Leu Pro Tyr Ser Ser Ser Ser Val Phe Ser Ala Gln
180 185 190
His Glu Ala Lys Lys Phe Thr Phe Asn Leu Met Ala Lys Asn Ile Met
195 200 205
Ser Met Asp Pro Gly
210

<210> 95
<211> 791
<212> DNA
<213> Triticum aestivum
<400> 95
ggaaggccct caagtcgcga gccaccatac ttggagtaat agagaggaaa atggaggaaa 60
ggcttgagaa aatgaataag gaggcctcga gcatggaaga agatgatctt ctcgggtggg 120
cgatgaagca gtccaatctt tcgaaagaac aaatattgga cctcttgctg agcttgctgt 180
ttgccgggca tgagacatcg tcaatggcgc tcgccctcgc catcttcttc ctcgaaggtt 240
gccctaaagc cgttgaagaa ctgcgggagg agcatcttga gattgctagg agacagaagc 300
tgagagggga gtgcaaactg agctgggaag actacaaaga gatggttttc acgcaatgcg 360
ttataaacga gaccttgcgg ctaggcaacg tggtcaggtt cctgcaccgg aaggtcattc 420
gagatgtgca ctacaatggg tatgacatcc cgagcggatg gaaaattctg ccggtgttag 480
cggcggtgca tctcgactcg tcgctgtacg aggaccccag cagcttcaac ccttggagat 540
ggaagggcaa cgcgtccggc gtggcgcaga acagtaactt catgccctac ggcggcggca 600
cgaggctctg cgccgggtcg gagctcgcca agctcgagat ggccatcttc ctgcaccacc 660
tggtgctcaa cttccggtgg gagctcgccg agcccgacca ggcgttcgtg tacccgttcg 720
tcgacttccc caagggcctg cccatcaggg tccataggat tgcacaggaa gaagaaggag 780
aagaggagta a 791

<210> 96
<211> 261
<212> PRT
<213> Triticum aestivum
<400> 96
Lys Ala Leu Lys Ser Arg Ala Thr Ile Leu Gly Val Ile Glu Arg Lys
1 5 10 15
Met Glu Glu Arg Leu Glu Lys Met Asn Lys Glu Ala Ser Ser Met Glu
20 25 30
Glu Asp Asp Leu Leu Gly Trp Ala Met Lys Gln Ser Asn Leu Ser Lys
35 40 45
Glu Gln Ile Leu Asp Leu Leu Leu Ser Leu Leu Phe Ala Gly His Glu
50 55 60
Thr Ser Ser Met Ala Leu Ala Leu Ala Ile Phe Phe Leu Glu Gly Cys
65 70 75 80
Pro Lys Ala Val Glu Glu Leu Arg Glu Glu His Leu Glu Ile Ala Arg
85 90 95
Arg Gln Lys Leu Arg Gly Glu Cys Lys Leu Ser Trp Glu Asp Tyr Lys
100 105 110
Glu Met Val Phe Thr Gln Cys Val Ile Asn Glu Thr Leu Arg Leu Gly
115 120 125
Asn Val Val Arg Phe Leu His Arg Lys Val Ile Arg Asp Val His Tyr
130 135 140
Asn Gly Tyr Asp Ile Pro Ser Gly Trp Lys Ile Leu Pro Val Leu Ala
145 150 155 160

Ala Val His Leu Asp Ser Ser Leu Tyr Glu Asp Pro Ser Ser Phe Asn
165 170 175
Pro Trp Arg Trp Lys Gly Asn Ala Ser Gly Val Ala Gln Asn Ser Asn
180 185 190
Phe Met Pro Tyr Gly Gly Gly Thr Arg Leu Cys Ala Gly Ser Glu Leu
195 200 205
Ala Lys Leu Glu Met Ala Ile Phe Leu His His Leu Val Leu Asn Phe
210 215 220
Arg Trp Glu Leu Ala Glu Pro Asp Gln Ala Phe Val Tyr Pro Phe Val
225 230 235 240
Asp Phe Pro Lys Gly Leu Pro Ile Arg Val His Arg Ile Ala Gln Glu
245 250 255
Glu Glu Gly Glu Glu
260

<210> 97
<211> 789
<212> DNA
<213> Euphorbia esula
<220>
<221> misc_feature
<222> (325)..(326)
<223> n is a, c, g, ou t
<400> 97
agcatggacc ctggaaaacc agagactgaa aagctcaaga aagaatatgt tactttcatg 60
aaaggagttg tttctgctcc tattaatttg cctggaactg cttatagaag ggccttacag 120
tctcgatcga caattttgaa gttcatagag gagaaaatgg aggaaagaaa tgaaaaatta 180
aaagaaggaa aagcagagga agaagaagaa gatgatcttc taggatgggt tttaaagcat 240
tcaaatcttt caactgagca aattctagat ttggtattaa gtttaatgtt tgctggccat 300
gaaacttctt cagtagcaat tcttnnagct atctattttt tgcaagattc tcctgctgct 360
cttcaacagc taagggaaga acataaggaa attgaaaaag ccaagaagca gtcaggagag 420
aagggattga actgggatgt ttacaaaaat atggaattca ctcagtgtgt tatcaatgaa 480
acactaagac ttggaaatgt agtcaggttc cttcatagga agactattaa acacgttcaa 540
tacaaaggat atgacattcc acgtggatgg aaagtgctgc cagtgattgc agccgtgcat 600
ttggacagta gccattttga gaagcctcaa cacttcaatc catggagatg gttgcaccag 660
aataatggga tacaaaatat gaataataat ttcatgccat ttgggggagg accaagatta 720
tgtgcaggat cagaattagc aaaactagaa atggctattt ttattcatca tttggtcctt 780
aattaccag 789

<210> 98
<211> 263
<212> PRT
<213> Euphorbia esula
<220>
<221> INSEGURO
<222> (109)..(109)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 98
Ser Met Asp Pro Gly Lys Pro Glu Thr Glu Lys Leu Lys Lys Glu Tyr
1 5 10 15
Val Thr Phe Met Lys Gly Val Val Ser Ala Pro Ile Asn Leu Pro Gly
20 25 30
Thr Ala Tyr Arg Arg Ala Leu Gln Ser Arg Ser Thr Ile Leu Lys Phe
35 40 45
Ile Glu Glu Lys Met Glu Glu Arg Asn Glu Lys Leu Lys Glu Gly Lys
50 55 60
Ala Glu Glu Glu Glu Glu Asp Asp Leu Leu Gly Trp Val Leu Lys His
65 70 75 80
Ser Asn Leu Ser Thr Glu Gln Ile Leu Asp Leu Val Leu Ser Leu Met
85 90 95
Phe Ala Gly His Glu Thr Ser Ser Val Ala Ile Leu Xaa Ala Ile Tyr
100 105 110
Phe Leu Gln Asp Ser Pro Ala Ala Leu Gln Gln Leu Arg Glu Glu His
115 120 125
Lys Glu Ile Glu Lys Ala Lys Lys Gln Ser Gly Glu Lys Gly Leu Asn

```
    130                 135                 140
Trp Asp Val Tyr Lys Asn Met Glu Phe Thr Gln Cys Val Ile Asn Glu
145                 150                 155                 160
Thr Leu Arg Leu Gly Asn Val Val Arg Phe Leu His Arg Lys Thr Ile
                165                 170                 175
Lys His Val Gln Tyr Lys Gly Tyr Asp Ile Pro Arg Gly Trp Lys Val
            180                 185                 190
Leu Pro Val Ile Ala Ala Val His Leu Asp Ser Ser His Phe Glu Lys
        195                 200                 205
Pro Gln His Phe Asn Pro Trp Arg Trp Leu His Gln Asn Asn Gly Ile
    210                 215                 220
Gln Asn Met Asn Asn Asn Phe Met Pro Phe Gly Gly Gly Pro Arg Leu
225                 230                 235                 240
Cys Ala Gly Ser Glu Leu Ala Lys Leu Glu Met Ala Ile Phe Ile His
                245                 250                 255
His Leu Val Leu Asn Tyr Gln
            260
<210> 99
<211> 955
<212> DNA
<213> Gossypium hirsutum
<400> 99
atgcctgact cagagcctat tttcttgctt cttccatcta ttttatcttt gattctgttc     60
ttcattctca tcaagagaaa gcaaagaagg tataatcttc caccagggaa catggggtgg    120
ccttttctcg gcgaaacaat cggttacttg aggccttact ctgctacttc agtaggtgaa    180
ttcatgcacc agcatatatc aaggtatggg aatatctaca aatcgaattt gtttggtgag    240
aagacaatag tgtctgcaga tgctgggttg aacaagttca tattacaaaa cgaagggaga    300
ttattcgagt gcagttaccc gagaagcatt ggtggcattc ttgggaaatg gtcgatgctg    360
gttttggttg gggatatgca tagagacatg aggattatat cactcaactt cttgagcaac    420
gccaggctga ggactcatct tttgagagaa gtggagaaac atactttgct tgttctaaat    480
acttggaaag agaagtgcat attttcagct caggatgaag caaaaaagtt cactttcaat    540
ttgatggcaa aaaatatcat gagcatggac cctggacatc cagagacgga gcagctaaag    600
aaagaatatg ttactttcat gaaaggagtc gtttctgctc ctttaaattt acctggaact    660
gcatacagaa aagccttaca atctcgatca acaatcctga aatttattga aaagaaaatg    720
gaagtaagga taaggaaaat gaaggaagga aaagaaaact cagaggaaga tgatcttctt    780
gaatgggtct taaagcattc taatctttcc acagagcaaa tccttgactt gattttgagc    840
ttgctttttg ctggacatga gacttcctcg gtagccataa ccttagccat ctacttcttg    900
ccaggttgtc ctttggccat tcaacagttg agagaggaac accttgaagt tgcag          955
<210> 100
<211> 333
<212> PRT
<213> Gossypium hirsutum
<220>
<221> INSEGURO
<222> (326)..(326)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 100
Met Pro Asp Ser Glu Pro Ile Phe Leu Leu Leu Pro Ser Ile Leu Ser
1               5                   10                  15
Leu Ile Leu Phe Phe Ile Leu Ile Lys Arg Lys Gln Arg Arg Tyr Asn
            20                  25                  30
Leu Pro Pro Gly Asn Met Gly Trp Pro Phe Leu Gly Glu Thr Ile Gly
        35                  40                  45
Tyr Leu Arg Pro Tyr Ser Ala Thr Ser Val Gly Glu Phe Met His Gln
    50                  55                  60
His Ile Ser Arg Tyr Gly Asn Ile Tyr Lys Ser Asn Leu Phe Gly Glu
65                  70                  75                  80
Lys Thr Ile Val Ser Ala Asp Ala Gly Leu Asn Lys Phe Ile Leu Gln
                85                  90                  95
Asn Glu Gly Arg Leu Phe Glu Cys Ser Tyr Pro Arg Ser Ile Gly Gly
            100                 105                 110
Ile Leu Gly Lys Trp Ser Met Leu Val Leu Val Gly Asp Met His Arg
        115                 120                 125
```

```
Asp Met Arg Ile Ile Ser Leu Asn Phe Leu Ser Asn Ala Arg Leu Arg
    130                 135                 140
Thr His Leu Leu Arg Glu Val Glu Lys His Thr Leu Leu Val Leu Asn
145                 150                 155                 160
Thr Trp Lys Glu Lys Cys Ile Phe Ser Ala Gln Asp Glu Ala Lys Lys
                165                 170                 175
Phe Thr Phe Asn Leu Met Ala Lys Asn Ile Met Ser Met Asp Pro Gly
            180                 185                 190
His Pro Glu Thr Glu Gln Leu Lys Lys Glu Tyr Val Thr Phe Met Lys
        195                 200                 205
Gly Val Val Ser Ala Pro Leu Asn Leu Pro Gly Thr Ala Tyr Arg Lys
    210                 215                 220
Ala Leu Gln Ser Arg Ser Thr Ile Leu Lys Phe Ile Glu Lys Lys Met
225                 230                 235                 240
Glu Val Arg Ile Arg Lys Met Lys Glu Gly Lys Glu Asn Ser Glu Glu
                245                 250                 255
Asp Asp Leu Leu Glu Trp Val Leu Lys His Ser Asn Leu Ser Thr Glu
            260                 265                 270
Gln Ile Leu Asp Leu Ile Leu Ser Leu Leu Phe Ala Gly His Glu Thr
        275                 280                 285
Ser Ser Val Ala Ile Thr Leu Ala Ile Tyr Phe Leu Pro Gly Cys Pro
    290                 295                 300
Leu Ala Ile Gln Gln Leu Arg Glu Asn Thr Leu Lys Leu Gln Ser Gln
305                 310                 315                 320
Thr Asn Gln Arg Asp Xaa Leu Asn Trp Asp Asp Thr Glu
                325                 330
<210> 101
<211> 832
<212> DNA
<213> Lycopersicon esculentum
<400> 101
atgggttggc cttttcttgg tgaaactatt ggctatttga gaccttattc agctactact     60
attggagatt tcatgcaaga tcatatctct aggtatggga aaattttcaa gtcaaatttg    120
tttggagagc caacaatagt ttcagcagat gcagggctaa acagatacat tctgcagaat    180
gaagggagat tatttgagtg taattatcca agaagtatag gtgggatact tggtaaatgg    240
tctatgttag ttcaagttgg acaaatgcat agagatatga ggatgatttc tctgaatttt    300
ttgagcaatg ctaggctaag gaatcaactt ttaagtgaag ttgaaaagca tactttgctt    360
gttcttggct cttggaaaca ggattctgtt gtttgtgcac aagatgaagc aaagaagtta    420
acattcaact ttatggcaga gcatatcatg agtctacaac ctggaaatcc agagacagag    480
aagctgaaaa aagagtacat cacatttatg aaaggagtgg tttctgctcc attgaatttt    540
ccaggaacag cttacagaaa ggccttacag tctcgatcaa caattcttgg atttattgag    600
agaaaaatgg aggagaggct taaggaaatg aacagaaacg aaaacgacct tctaggttgg    660
gttctgaaga attcaaatct ctcaaaagag caaattcttg atttgctact gagtttgctc    720
tttgctggcc atgaaacttc atcagtagca atagctctgt ctattttctt actcgaaagc    780
tgtcctgctg ctgttcaaca attaacagaa gagcacttgg agatttcccg gg             832
<210> 102
<211> 277
<212> PRT
<213> Lycopersicon esculentum
<400> 102
Met Gly Trp Pro Phe Leu Gly Glu Thr Ile Gly Tyr Leu Arg Pro Tyr
1               5                   10                  15
Ser Ala Thr Thr Ile Gly Asp Phe Met Gln Asp His Ile Ser Arg Tyr
            20                  25                  30
Gly Lys Ile Phe Lys Ser Asn Leu Phe Gly Glu Pro Thr Ile Val Ser
        35                  40                  45
Ala Asp Ala Gly Leu Asn Arg Tyr Ile Leu Gln Asn Glu Gly Arg Leu
    50                  55                  60
Phe Glu Cys Asn Tyr Pro Arg Ser Ile Gly Gly Ile Leu Gly Lys Trp
65                  70                  75                  80
Ser Met Leu Val Gln Val Gly Gln Met His Arg Asp Met Arg Met Ile
                85                  90                  95
Ser Leu Asn Phe Leu Ser Asn Ala Arg Leu Arg Asn Gln Leu Leu Ser
```

```
            100                 105                 110
Glu Val Glu Lys His Thr Leu Leu Val Leu Gly Ser Trp Lys Gln Asp
        115                 120                 125
Ser Val Val Cys Ala Gln Asp Glu Ala Lys Lys Leu Thr Phe Asn Phe
    130                 135                 140
Met Ala Glu His Ile Met Ser Leu Gln Pro Gly Asn Pro Glu Thr Glu
145                 150                 155                 160
Lys Leu Lys Lys Glu Tyr Ile Thr Phe Met Lys Gly Val Val Ser Ala
                165                 170                 175
Pro Leu Asn Phe Pro Gly Thr Ala Tyr Arg Lys Ala Leu Gln Ser Arg
            180                 185                 190
Ser Thr Ile Leu Gly Phe Ile Glu Arg Lys Met Glu Glu Arg Leu Lys
        195                 200                 205
Glu Met Asn Arg Asn Glu Asn Asp Leu Leu Gly Trp Val Leu Lys Asn
    210                 215                 220
Ser Asn Leu Ser Lys Glu Gln Ile Leu Asp Leu Leu Leu Ser Leu Leu
225                 230                 235                 240
Phe Ala Gly His Glu Thr Ser Ser Val Ala Ile Ala Leu Ser Ile Phe
                245                 250                 255
Leu Leu Glu Ser Cys Pro Ala Ala Val Gln Gln Leu Thr Glu Glu His
            260                 265                 270
Leu Glu Ile Ser Arg
        275
<210> 103
<211> 711
<212> DNA
<213> Solanum tuberosum
<400> 103
atgtctgact tagagttttt tctttttctt gttcctccaa tcttggcagt ccttattatt     60
cttaatctat tcaaaagaaa acacaaattt caaaatcttc caccagggga tatgggttgg    120
ccttttcttg gtgaaactat tggttatttg agaccttact cagttactac tattggagat    180
ttcatgcaag atcatatctc aaggtatggg aaaattttca agtcaaattt gtttggagag    240
ccaacaattg tttcagcaga tgcagggctt aacagataca ttctgcagaa tgaagggaga    300
ttatttgagt gtaattatcc aagaagtata ggtgggatac ttggtaaatg gtctatgttg    360
gttcaagttg gacaaatgca tagagatatg aggatgattt ctctgaattt tttgagcaat    420
gctagactca ggaatcaact tttaagtgaa gttgaaaagc atactgtgct tgttcttggc    480
tcttggaaac aggattctgt tgtttgtgca caagatgaag caaagaagtt tacattcaac    540
tttatggcag agcatatcat gagtctacaa cctggaaatc cagagacaga gaagctgaag    600
aaagagtaca tcacatttat gaaaggagtg gtttctgctc cattgaattt tccaggaaca    660
gcttacagaa aagccctaca gtctcgatca acaattcttg gaattattga g              711
<210> 104
<211> 237
<212> PRT
<213> Solanum tuberosum
<400> 104
Met Ser Asp Leu Glu Phe Phe Leu Phe Leu Val Pro Pro Ile Leu Ala
1               5                   10                  15
Val Leu Ile Ile Leu Asn Leu Phe Lys Arg Lys His Lys Phe Gln Asn
            20                  25                  30
Leu Pro Pro Gly Asp Met Gly Trp Pro Phe Leu Gly Glu Thr Ile Gly
        35                  40                  45
Tyr Leu Arg Pro Tyr Ser Val Thr Thr Ile Gly Asp Phe Met Gln Asp
    50                  55                  60
His Ile Ser Arg Tyr Gly Lys Ile Phe Lys Ser Asn Leu Phe Gly Glu
65                  70                  75                  80
Pro Thr Ile Val Ser Ala Asp Ala Gly Leu Asn Arg Tyr Ile Leu Gln
                85                  90                  95
Asn Glu Gly Arg Leu Phe Glu Cys Asn Tyr Pro Arg Ser Ile Gly Gly
            100                 105                 110
Ile Leu Gly Lys Trp Ser Met Leu Val Gln Val Gly Gln Met His Arg
        115                 120                 125
Asp Met Arg Met Ile Ser Leu Asn Phe Leu Ser Asn Ala Arg Leu Arg
    130                 135                 140
```

Asn Gln Leu Leu Ser Glu Val Glu Lys His Thr Val Leu Val Leu Gly
145 150 155 160
Ser Trp Lys Gln Asp Ser Val Val Cys Ala Gln Asp Glu Ala Lys Lys
165 170 175
Phe Thr Phe Asn Phe Met Ala Glu His Ile Met Ser Leu Gln Pro Gly
180 185 190
Asn Pro Glu Thr Glu Lys Leu Lys Lys Glu Tyr Ile Thr Phe Met Lys
195 200 205
Gly Val Val Ser Ala Pro Leu Asn Phe Pro Gly Thr Ala Tyr Arg Lys
210 215 220
Ala Leu Gln Ser Arg Ser Thr Ile Leu Gly Ile Ile Glu
225 230 235

<210> 105
<211> 431
<212> DNA
<213> Solanum tuberosum
<400> 105
ggaaagctgt gaaagatgtt cgatacaaag gttatgacat tccatgtgga tggaaagtat 60
tgccggtgat ttcagcagcg catttagatc cttcactttt tgaccgacct cacgactttg 120
atccttggag atggcagaat gcagaagagt cgccttcagg taaaggagga agcacaggca 180
caagcagcac aacaaaaagt agtaataatt tcatgccatt tgggggaggt ccacgtctat 240
gtgcaggatc tgaactggcc aaacttgaga tggccatttt cattcactat cttgttctta 300
attttcactg gaaattagct gcacctgatc aggcttttgc ctatccttac gtagattttc 360
ccaatgccct accaatcact atccaacatc gatcgtcaaa taaattaaac caccacactt 420
actacctcta a 431

<210> 106
<211> 142
<212> PRT
<213> Solanum tuberosum
<400> 106
Lys Ala Val Lys Asp Val Arg Tyr Lys Gly Tyr Asp Ile Pro Cys Gly
1 5 10 15
Trp Lys Val Leu Pro Val Ile Ser Ala Ala His Leu Asp Pro Ser Leu
20 25 30
Phe Asp Arg Pro His Asp Phe Asp Pro Trp Arg Trp Gln Asn Ala Glu
35 40 45
Glu Ser Pro Ser Gly Lys Gly Gly Ser Thr Gly Thr Ser Ser Thr Thr
50 55 60
Lys Ser Ser Asn Asn Phe Met Pro Phe Gly Gly Gly Pro Arg Leu Cys
65 70 75 80
Ala Gly Ser Glu Leu Ala Lys Leu Glu Met Ala Ile Phe Ile His Tyr
85 90 95
Leu Val Leu Asn Phe His Trp Lys Leu Ala Ala Pro Asp Gln Ala Phe
100 105 110
Ala Tyr Pro Tyr Val Asp Phe Pro Asn Ala Leu Pro Ile Thr Ile Gln
115 120 125
His Arg Ser Ser Asn Lys Leu Asn His His Thr Tyr Tyr Leu
130 135 140

<210> 107
<211> 52
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm06540
<400> 107
ggggacaagt ttgtacaaaa aagcaggctt aaacaatggc cgccatgatg gc 52

<210> 108
<211> 48
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm06520
<400> 108

ggggaccact ttgtacaaga aagctgggtt tactcctgct catcatcc 48

<210> 109
<211> 654
<212> DNA
<213> Oryza sativa
<400> 109
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> 110
<211> 1236
<212> DNA
<213> Oryza sativa
<400> 110
ggtcagccaa tacattgatc cgttgccaat catgcaaagt attttggctg tggccgagtg 60
ccggaattga taattgtgtt ctgactaaat taaatgacca gaagtcgcta tcttccaatg 120
tatccgaaac ctggattaaa caatcctgtt ctgttctcta gcccctcctg catggccgga 180
ttgtttttt gacatgtttt cttgactgag gcctgtttgt tctaaacttt ttcttcaaac 240
ttttaacttt ttcatcacat cagaactttt ctacacatat aaacttttaa cttttccgtc 300
acatcgttcc aatttcaatc aaactttcaa ttttggcgtg aactaaacac accctgagtc 360
ttttattgct cctccgtacg ggttggctgg ttgagaatag gtattttcag agagaaaatc 420
tagatattgg gaggaacttg gcatgaatgg ccactatatt tagagcaatt ctacggtcct 480
tgaggaggta ccatgaggta ccaaaatttt agtgtaaatt ttagtatctc attataacta 540
ggtattatga ggtaccaaat ttacaataga aaaaatagta cttcatggta ctttcttaag 600
taccgtaaaa ttgctcctat atttaagggg atgtttatat ctatccatat ccataatttg 660
attttgataa gaaaaaatgt gagcacacca agcatgtcca tgaccttgca ctcttggctc 720
actcgtcaac tgtgaagaac ctcaaaaatg ctcaatatag ctacaggtgc ctgaaaaaat 780
aactttaaag ttttgaacat cgatttcact aaacaacaat tattatctcc ctctgaaaga 840
tgatagttta gaactctaga atcattgtcg gcggagaaag taaattattt tccccaaatt 900
tccagctatg aaaaaaccct caccaaacac catcaaacaa gagttcacca aaccgcccat 960
gcggccatgc tgtcacgcaa cgcaccgcat tgcctgatgg ccgctcgatg catgcatgct 1020
tccccgtgca catatccgac agacgcgccg tgtcagcgag ctcctcgacc gacctgtgta 1080
gcccatgcaa gcatccaccc ccgccacgta cacccccctcc tcctccctac gtgtcaccgc 1140
tctctccacc tatatatgcc cacctggccc ctctcctccc atctccactt cacccgatcg 1200
cttcttcttc ttcttcgttg cattcatctt gctagc 1236

<210> 111
<211> 1455
<212> DNA
<213> Glycine max
<220>
<221> misc_feature
<222> (1043)..(1044)
<223> n is a, c, g, ou t
<400> 111
atgtctgact cactcttaac tttctattct ctttcagcca ttcttgctct tctcccaatc 60
ttcatcttca ttctcatcaa aagaaagcaa agcaaaccca ggctcaacct tcccccaggt 120
aacatgggtt ggccatttct tggtgaaacc attggctatt tgaagcctta ttctgccacc 180
acagtagggg aattcatgga gcaacacata gcaaggtatg gtacaattta caagtcaaaa 240
ctgtttgcgg ggccagcaat agtgtcagca gatgcaggac tcaacaggtt cattctacaa 300
aacgaaggga aattgttcga gtgcagctat cctagaagca tcggtggaat actaggaaaa 360
tggtccatgt tggtcttagt tggtgacatg catagagaca tgcgggttat atcactcaac 420
tttctaagcc acgccaggct cagaacacac ctcttgaaag aggtggagaa gcaatccctc 480
ttggttctga actcttggag ccaaaattcc atattctcag cccaagatga agctaagaag 540
ttcaccttca atttaatggc taagcatatc atgagcatgg atcctgggga tatcgagaca 600
gagcaactaa agaaagagta cgtcactttc atgaaagggg tggtttccgc gccattgaat 660

ttacctggaa ctgcataccg aaaggcattg aagtctcggt ccattatact aaagttcata 720
gaggggaaaa tggaagagag agttagaaga atccaagagg ggaatgagag tttggaggaa 780
gatgatcttc taaattgggt tttgaagaat tcaaatcttt caaccgagca aattcttgac 840
ttgattctca gcttgctctt cgctggccat gaaacttcgt cggtagccat agctctagcc 900
atttacttct tacccggttg tcctcaagct atacaacagt taaaggaaga acacagagaa 960
attgccagag ccaaaaagca agcaggggaa gttgaactca cttgggatga ctacaaacga 1020
atggaattta ctcattgtgt aannttggga aatgttgtga ggtttctcca caggaaggct 1080
gtgaaagatg ttaactataa aggttatgac attccatgtg ggtggaaagt cctcccggtg 1140
attgcagccg tgcatctgga tccttcactt tttgaccaac ctcaacactt caatccatgg 1200
agatggcaga acaatggcag tcatggaagt tgtccaagca agaacacagc aaacaacaac 1260
tttcttccgt tcggaggagg accacgatta tgtgcaggat cagagttagc taagcttgaa 1320
atggctgttt tcattcacca tctcattctc aactaccatt gggaattggc tgataccgat 1380
caagcttttg cctacccttt tgtcgacttc cccaaaggcc tacccgttag agtccaagcc 1440
cattctttac tttga 1455

<210> 112
<211> 530
<212> PRT
<213> Glycine max
<220>
<221> INSEGURO
<222> (391)..(392)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 112

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

Ile Ala Arg Ala Lys Lys Gln Ala Gly Glu Val Glu Leu Thr Trp Asp
325 330 335
Asp Tyr Lys Arg Met Glu Phe Thr His Cys Val Ser Gly Met Ile Ile
340 345 350
Asn Glu Leu Trp Arg Ser Thr Cys Leu Phe Phe Tyr Phe Met Tyr Ser
355 360 365
Asn Cys Tyr Tyr Tyr Tyr Leu Met Asn Met Leu Val Val Ser Asn Ile
370 375 380
Ile Trp Asp Asp Glu Ala Xaa Xaa Val Gly Leu Gly Asn Val Val Arg
385 390 395 400
Phe Leu His Arg Lys Ala Val Lys Asp Val Asn Tyr Lys Gly Tyr Asp
405 410 415
Ile Pro Cys Gly Trp Lys Val Leu Pro Val Ile Ala Ala Val His Leu
420 425 430
Asp Pro Ser Leu Phe Asp Gln Pro Gln His Phe Asn Pro Trp Arg Trp
435 440 445
Gln Asn Asn Gly Ser His Gly Ser Cys Pro Ser Lys Asn Thr Ala Asn
450 455 460
Asn Asn Phe Leu Pro Phe Gly Gly Gly Pro Arg Leu Cys Ala Gly Ser
465 470 475 480
Glu Leu Ala Lys Leu Glu Met Ala Val Phe Ile His His Leu Ile Leu
485 490 495
Asn Tyr His Trp Glu Leu Ala Asp Thr Asp Gln Ala Phe Ala Tyr Pro
500 505 510
Phe Val Asp Phe Pro Lys Gly Leu Pro Val Arg Val Gln Ala His Ser
515 520 525
Leu Leu
530

<210> 113
<211> 1458
<212> DNA
<213> Gossypium hirsutum
<400> 113

atgcctgact cagagcctat tttcttgctt ctaccatcta ttttatcttt gattctgttc 60
ttcattctca tcaagagaaa gcaaagaagg tataatcttc caccagggaa catggggtgg 120
ccttttctcg gcgaaacaat cggttacttg aggccttact ctgctacttc agtaggtgaa 180
ttcatgcacc agcatatatc aaggtatggg aatatctaca aatcgaattt gtttggtgag 240
aagacaatag tgtctgcaga tgctgggttg aacaagttca tattacaaaa cgaagggaga 300
ttattcgagt gcagttaccc gagaagcatt ggtggcattc ttgggaaatg gtcgatgctg 360
gttttggttg gggatatgca tagagacatg aggattatat cactcaactt cttgagcaac 420
gccaggctga ggactcatct tttgagagaa gtggagaaac atactttgct tgttctaaat 480
acttggaaag agaagtgcat attttcagct caggatgaag caaaaaagtt cactttcaat 540
ttggtggcaa aaaatatcat gagcatggac cctggacatc cagagacgga gcagctaaag 600
aaagaatatg ttactttcat gaaaggagtc gtttctgctc ctttaaattt acctggaact 660
gcatacagaa aagccttaca atctcgatca acaatcctga aatttattga aaagaaaatg 720
gaagtaagga taaggaaaat gaaggaagga aaagaaaact cagaggaaga tgatcttctt 780
gaatgggtct taaagcattc taatctttcc acagagcaaa tccttgactt gattttgagc 840
ttgctttttg ctggacatga gacctcctcg gtagccataa ccttagccat ctacttcttg 900
ccaggttgtc ctttggccat tcaacagttg agagaagaac accttgaagt tgccagagcc 960
aagaaccaat caggagagac tgaactcaac tgggatgatt acaagaaaat ggagttcact 1020
caatgtgtta ttaacgagac acttaggctt ggtaatgtcg tcagatttct ccacagaaaa 1080
gctctcaaag atattagata taaaggttat gatattccat gtgggtggaa agtgcttcca 1140
gtgattgcag cagtgcactt ggatccctgt ctttttgacc accctcaact cttcaatcca 1200
tggcgatggc agcaaaataa tgggagtcga ggggcgggga cggcaacgtc atcagcgagc 1260
agcagcaatt acttcatgcc attcgggggga ggaccacggc tatgtgcagg aacagagctg 1320
gctaaactgg aaatggcggt gttcatccac catttggtcc tcaactacca gtgggagtta 1380
gccgatacgg atgaagcctt tgccttccct tttgtcgact tccctaaagg cctacccatc 1440
agagtcttca aatcttaa 1458

<210> 114
<211> 485
<212> PRT
<213> Gossypium hirsutum
<400> 114

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

<210> 115

<211> 1521
<212> DNA
<213> Zea mays
<400> 115

```
atgggcgcca tgatggcctc cataaccagc gagctcctct tcttccttcc cttcatcctg     60
ctggccctcc tcgccttgta caccaccacc gtcgccaaat gccacggcac ccacccgtgg    120
cgccgtcaga agaagaagcg gcccaacctg cccccgggcg cccgcggatg gcccttggtc    180
ggcgaaactt tcggctacct ccgcgcccac ccggccacct ccgtgggccg cttcatggag    240
cggcatgtcg cacggtacgg gaagatatac cggtcgagcc tgttcgggga gcggacggtg    300
gtgtcggcgg acgcggggct gaaccgctac atcctgcaga acgaggggcg gctgttcgag    360
tgcagctacc cgcgcagcat cggcggcatc ctgggcaagt ggtccatgct ggtgctcgtg    420
ggcgacgcgc accgcgagat gcgcgctatc tcgctcaact tcctcagctc cgtccgcctc    480
cgcgccgtgc tgctccccga ggtggagcgc cacaccctgc tggtcctccg ctcgtggccg    540
ccctccgacg gcaccttctc cgcccagcac gaagccaaga agttcacgtt taacctgatg    600
gcgaagaaca taatgagcat ggaccccggc gaggaggaga cggagcggct gcggctggag    660
tacatcacct tcatgaaggg cgtcgtgtca gcgccgctca acttcccggg cacggcctac    720
tggaaggcgc tcaagtcacg cgcgtccata cttggagtga tagagaggaa gatggaggac    780
aggcttgaga agatgagcag ggagaagtca agcgtggagg aggacgacct tcttggatgg    840
gccctgaagc aatccaacct gtccaaggaa cagatcctgg acctcttgct gagcctgctc    900
ttcgcggggc acgagacttc gtccatggcg ctcgccctcg ccatcttctt cctcgaaggg    960
tgccctaagg ccgtgcaaga actccgggag gagcatctcc tgattgctag gagacaaagg   1020
ctaagagggg cgtctaaatt gagctgggaa gactacaagg aaatggtttt cacgcagtgt   1080
gttataaacg agacattgcg gctcggcaac gtggtcaggt tcctgcaccg gaaggtcatc   1140
cgagatgtac actacaatgg gtacgacata ccgcgggggt ggaaaatcct gccggttcta   1200
gcggcggtgc acctggactc gtcgctgtac gaggacccca gccggttcaa cccttggaga   1260
tggaagagca acaacgcgcc aagcagcttc atgccgtacg gcggcgggcc gcggctgtgc   1320
gccgggtcgg agctggccaa gctggagatg gccatcttcc tgcaccacct ggtgctcaac   1380
ttccggtggg agctggcgga gccggaccag gccttcgttt acccttttcgt cgacttcccc   1440
aagggcctcc cgatcagggt ccagcgggtc gccgacgacc aaggccatcg tagcgttttg   1500
accgagagca caagaggctg a                                             1521
```

<210> 116
<211> 506
<212> PRT
<213> Zea mays
<400> 116

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

```
Met Lys Gly Val Val Ser Ala Pro Leu Asn Phe Pro Gly Thr Ala Tyr
225                 230                 235                 240
Trp Lys Ala Leu Lys Ser Arg Ala Ser Ile Leu Gly Val Ile Glu Arg
                245                 250                 255
Lys Met Glu Asp Arg Leu Glu Lys Met Ser Arg Glu Lys Ser Ser Val
            260                 265                 270
Glu Glu Asp Asp Leu Leu Gly Trp Ala Leu Lys Gln Ser Asn Leu Ser
        275                 280                 285
Lys Glu Gln Ile Leu Asp Leu Leu Leu Ser Leu Leu Phe Ala Gly His
    290                 295                 300
Glu Thr Ser Ser Met Ala Leu Ala Leu Ala Ile Phe Phe Leu Glu Gly
305                 310                 315                 320
Cys Pro Lys Ala Val Gln Glu Leu Arg Glu Glu His Leu Leu Ile Ala
                325                 330                 335
Arg Arg Gln Arg Leu Arg Gly Ala Ser Lys Leu Ser Trp Glu Asp Tyr
            340                 345                 350
Lys Glu Met Val Phe Thr Gln Cys Val Ile Asn Glu Thr Leu Arg Leu
        355                 360                 365
Gly Asn Val Val Arg Phe Leu His Arg Lys Val Ile Arg Asp Val His
    370                 375                 380
Tyr Asn Gly Tyr Asp Ile Pro Arg Gly Trp Lys Ile Leu Pro Val Leu
385                 390                 395                 400
Ala Ala Val His Leu Asp Ser Ser Leu Tyr Glu Asp Pro Ser Arg Phe
                405                 410                 415
Asn Pro Trp Arg Trp Lys Ser Asn Asn Ala Pro Ser Ser Phe Met Pro
            420                 425                 430
Tyr Gly Gly Gly Pro Arg Leu Cys Ala Gly Ser Glu Leu Ala Lys Leu
        435                 440                 445
Glu Met Ala Ile Phe Leu His His Leu Val Leu Asn Phe Arg Trp Glu
    450                 455                 460
Leu Ala Glu Pro Asp Gln Ala Phe Val Tyr Pro Phe Val Asp Phe Pro
465                 470                 475                 480
Lys Gly Leu Pro Ile Arg Val Gln Arg Val Ala Asp Asp Gln Gly His
                485                 490                 495
Arg Ser Val Leu Thr Glu Ser Thr Arg Gly
            500                 505
<210> 117
<211> 1566
<212> DNA
<213> Zea mays
<400> 117
atgggcgcca tgatggcctc cataaccagc gagctcctct tcttccttcc cttcatcctg     60
ctggccctcc tcgccttgta caccaccacc gtcgccaaat gccacggcac ccaccagtgg    120
cgccgtcaga agaagaagcg gcccaacctg cccccgggcg cccgcggatg gcccttggtc    180
ggcgagactt tcggctacct ccgcgcccac ccggccacct ccgtgggccg cttcatggag    240
cggcatgtcg cacggtacgg gaagatatac cggtcgagcc tgttcgggga gcggacggtg    300
gtgtcggcgg acgcggggct gaaccggtac atcctgcaga acgaggggcg gctgttcgag    360
tgcagctacc cgcgcagcat cggcggcatc ctgggcaagt ggtccatgct ggtgctcgtg    420
ggcgacgcgc accgcgagat gcgcgctatc tcgctcaact tcctcagctc cgtccgcctc    480
cgcgccgtgc tgctccccga ggtggagcgc cacaccctgc tggtcctccg ctcctggccg    540
cctccgaccg gcaccttctc cgcccagcac gaagccaaga agttcacgtt taacctgatg    600
gcgaagaaca taatgagcat ggaccccggc gaggaggaga cggagcggct gcggctggag    660
tacatcacct tcatgaaggg cgtcgtgtca gcgccgctca acttcccggg cacggcctac    720
tggaaggcgc tcaagtcgcg cgcgtccata cttggagtga tagagaggaa gatggaggac    780
aggcttgaga agatgagcag ggagaagtca agcgtggagg aggacgacct tcttggatgg    840
gccctgaagc aatccaacct gtccaaggaa cagatcctgg acctcttgct gagcctgctc    900
ttcgcggggc acgagacttc gtccatggcg ctcgccctcg ccatcttctt cctcgaaggg    960
tgccctaagg ccgtgcaaga actccgggag gagcatctcc tgattgctag gagacaaagg   1020
ctaaggggcg cgtccaaatt gagctgggaa gactacaagg aaatggtttt cacgcagtgt   1080
gttataaacg agacattgcg gctcggcaac gtggtcaggt tcctgcaccg gaaggtcatc   1140
cgagatgtac actacaatgg gtacgacata ccgcgggggt ggaaaatcct gccggttcta   1200
gcggcggtgc acctggactc gtcgctgtac gaggatccca gccggttcaa cccttggaga   1260
tggaaggtca gtgccctgcc ccctcccctc gcaaaaggca caagccaagg gacaagcaag   1320
```

aggtttacac cgttcggtgg tggcccccgg ctctgcccag gatcagagct cgctaaagtg 1380
gagactgctt tcttcctcca tcaccttgtc ctcaattata gatggagaat tgatggcgat 1440
gacattccaa tggcataccc gtatgtggag tttcagagag gtctgccaat agaaatcgag 1500
ccaacgtccc ctgaatttga ctgtcctgga gctacagcca tcagttatca caccagagag 1560
aaatga 1566

<210> 118
<211> 521
<212> PRT
<213> Zea mays
<400> 118

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

```
            420                 425                 430
Gly Thr Ser Gln Gly Thr Ser Lys Arg Phe Thr Pro Phe Gly Gly Gly
        435                 440                 445
Pro Arg Leu Cys Pro Gly Ser Glu Leu Ala Lys Val Glu Thr Ala Phe
    450                 455                 460
Phe Leu His His Leu Val Leu Asn Tyr Arg Trp Arg Ile Asp Gly Asp
465                 470                 475                 480
Asp Ile Pro Met Ala Tyr Pro Tyr Val Glu Phe Gln Arg Gly Leu Pro
                485                 490                 495
Ile Glu Ile Glu Pro Thr Ser Pro Glu Phe Asp Cys Pro Gly Ala Thr
            500                 505                 510
Ala Ile Ser Tyr His Thr Arg Glu Lys
        515                 520
<210> 119
<211> 1257
<212> DNA
<213> Brassica rapa
<220>
<221> misc_feature
<222> (967)..(969)
<223> n is a, c, g, ou t
<400> 119
atgttcgaaa cagagcatac tctcgtgcct cttcttcttc tcccatcact tctatctctt     60
ctcctcttct tgattctctt gaagagacga agtcgacaca gtttcaatct ccctcctgga    120
aaatctggat ggccatttct aggcgaaacc atcggatatc tcaaacctta ctctgccaaa    180
actctcggtt acttcatgca acaacatatc tccaagtatg ggaagatata tagatcgaat    240
ttgtttggag aaccaacgat cgtatcagct gatgcaggac tcaacaggtt catattacaa    300
aacgaaggaa gactctttga atgtagttat cctcgaagta ttggtgggat tcttgggaaa    360
tggtcgatgc ttgttcttgt tggagacatg catagagaca tgagaagtat ctcgctaaac    420
tttctaagtc acgctcgtct cagaacgatt cttcttaagg acgttgagag gcatactttg    480
ttcgttctta attcttggca acaacattct gttttctctg ctcaagatga ggccaaaaag    540
tttacgttta atctaatggc gaagcatata atgagtatgg atcctggaga agaagagaca    600
gagcagttaa agaaagagta tgtgactttc atgaaagggg ttgtttctgc tcctctcaat    660
ctcccaggaa ctgcttatcg taaagctctc cagcagtcac gagggacgat attgaagttt    720
attgagaaga aaatggaaga gagaaaatca gagattcaag aagaagacga agaagatgaa    780
gcagagatta gtagaagtga tcattatgag agaaaacata gagcagatga tgatcttttg    840
ggatgggttc taaaacattc caatctttcg actgagcaaa ttctcgatct tattctcagt    900
ttattatttg ccggacatga gacatcatct gtagccatcg ctctcgctat ctacttttta    960
gcgggcnnnt tgaaggaaga gcatcttgaa atcgcgaggg tgaagaagga acttggagag   1020
tcagaattga attgggatga ttacaagaca atggacttta ctcatagtgt tataaatgag   1080
actcttcgac taggaaatgt agtaaggttt ttgcatcgta aagcactcaa aaacgttcgg   1140
tataaaggat acgatatccc aagtgggtgg aaagtgttac cagtgatctc agccgtacat   1200
ttggataact cccgttacga cgaacctaat ctctttaatc cttggagatg gcaacag      1257
<210> 120
<211> 419
<212> PRT
<213> Brassica rapa
<220>
<221> INSEGURO
<222> (323)..(323)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 120
Met Phe Glu Thr Glu His Thr Leu Val Pro Leu Leu Leu Leu Pro Ser
1               5                   10                  15
Leu Leu Ser Leu Leu Leu Phe Leu Ile Leu Leu Lys Arg Arg Ser Arg
            20                  25                  30
His Ser Phe Asn Leu Pro Pro Gly Lys Ser Gly Trp Pro Phe Leu Gly
        35                  40                  45
Glu Thr Ile Gly Tyr Leu Lys Pro Tyr Ser Ala Lys Thr Leu Gly Tyr
    50                  55                  60
Phe Met Gln Gln His Ile Ser Lys Tyr Gly Lys Ile Tyr Arg Ser Asn
65                  70                  75                  80
Leu Phe Gly Glu Pro Thr Ile Val Ser Ala Asp Ala Gly Leu Asn Arg
```

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

<210> 121
<211> 848
<212> DNA
<213> Hordeum vulgare
<400> 121

```
atggccgcca tgatggcatc cataaccagc gagctccttt tcttcctccc cttcatcctc     60
ctggccctgc tcaccttcta caccagcagc gtggccaaat gccatggcct ccaccggtgg    120
agcggccgga cgaagaagaa gcggccgaat ctgccgcccg gcgccgccgg ctggcctttc    180
gtcggcgaga ccttcgggta cctccgcgcc cacccggcca cctccatcgg ccagttcatg    240
aaccagcaca tcgcacggta cgggaagata taccggtcga gcctgttcgg ggagcggacg    300
gtggtgtcgg cggacgcggg gctgaaccgg tacatcctgc agaacgaggg gcggctgttc    360
gagtgcagct acccgcggag catcggcggc atcctgggca aatggtccat gctggtgctc    420
gtgggcgacc cccaccgcga gatgcgctcc atctccctca acttcctcag ctccctccgc    480
ctccgcgccg tgctcctccc ggaggtggag cgccacaccc tcctcgtcct ccgcgactgg    540
ctgccttcct cctcctccgc cgtcttctcc gcccagcacg aagccaagaa gttcacgttt    600
aacctgatgg cgaagaacat catgagcatg gaccccggcg aggaggagac ggagcggctg    660
aggctcgagt acatcacctt catgaagggg gtggtgtccg cgcccctcaa cttccccggg    720
acggcctact ggaaggccct caagtctcga gccaccatac ttggggtaat agagaggaaa    780
atggaggata ggctcgagaa aatgaacaag gaggcctcga gcatgaagaa gatatctctc    840
gggggggcg                                                            848
```

<210> 122
<211> 282
<212> PRT
<213> Hordeum vulgare
<400> 122

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

<210> 123
<211> 1200
<212> DNA
<213> Lotus japonicus
<400> 123

```
atgtctgact catacctcac tttctgtttt ctttcttcca ttcttgctct cactgtaatc     60
atcattttca tgaaaagaaa gaaggcaagg cttaaccttc cccctggaaa aatgggatgg    120
ccctttcttg gggaaaccat tgggtacttg aagccatact ctgctaccac agtaggagat    180
tttatggaaa agcacatagc aaggtatggt acaatttaca agtcaaaatt gtttggtgag    240
cctgcaatag tttctgcaga tgcagagttg aacaggttca tattacagaa cgaagggaag    300
ctgtttgagt gcagctatcc aagaagcatt ggaggaatac ttggaaaatg gtccatgttg    360
gtcttagttg gggacatgca tagggacatg agaaacattt cactgaactt tctgagcctt    420
gctaggctca aaacacatct attgaaagaa gtggagaagc actctcttct agttctaggc    480
tcttggaaag aaaattgtac attctcagct caagatgaag caaagaagtt cacattcaat    540
ttgatggcga aacatatcat gagcttggat cctgggaatc tagagacaga acagctgaag    600
aaagagtatg tctctttcat gaatggtgtg gtgtctgcac ctttgaattt tcccggaact    660
gcatacagaa aagcattaaa gtctaggtcc accatactga agttcataga gggaaaaatg    720
gaagaaagga tcaaaagaaa ccaaaatttg gaggaagatc ttctaaactg ggttgtgatg    780
cattcaaatc tttcaactga gcaaattctt gacctggttc tgagcttgct ctttgcaggc    840
catgaaactt catctgtggc tatagcttta gctatttact ttttgccagg ttgtcctaaa    900
gctatacaac aattaaggga agaacataga gaaatagcca ggtccaagaa gcaagcaggg    960
gaggttgaat taacttggga tgattacaaa agaatggagt ttactcaatg tgtagttgtg   1020
```

aatgaaacac tgaggttggg aaatgttgtg aggttccttc acaggaaggc tctgaaagat 1080
gttcggtaca aaggttatga cattccacgt gggtggaaag tcctccctgt gatttcagct 1140
atgcatctgg atcctgcact tttttaccaa cctcaacact tcaatccatg gagatggaag 1200

<210> 124
<211> 400
<212> PRT
<213> Lotus japonicus
<400> 124

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

<210> 125
<211> 1092
<212> DNA
<213> Malus domestica
<400> 125

agagacatga ggatgatatc tctcaacttc ttgagccatg ccagactcag aacccatctg 60
atgagagaag ttgaaaagca cactcttctt gttttgggga gttggaagga aaactctgtt 120
ttttcagctc aggatgaagc taagaagttc acgttcaact tgatggccaa acatatcatg 180
agcttggatc ctggaaaacc agagactgag cagctgaaga aattgtatgt tactttcatg 240
aaaggtgtgg tttctcctcc tctgaattta ccaggaacag cttacagaag agccttacag 300
tcaagatcaa cgattctgaa gtttatagag tgcaaaatga aagaaagatt gatggaggga 360
accgaaaaca ttggggaaga tgatctgctt ggatgggttc tgaagaattc aaatctttca 420
aaggagcaaa ttcttgactt aatattgagc ttgctctttg ctggccatga aacttcatca 480
gtgtctatag ccttagcaat ttacttttta ccaagctgcc ctaatgcaat tctgcagtta 540
agggaagaac acagtgaaat tgccaaagcc aagaaactgg caggcgagac agagttgaat 600
tgggaggact acaagaaaat ggagttcacc gaatgtgtaa tctgtgagac acttcggctt 660
gggaatgtgg tgagattttt acacagaaag gctctgaagg atgttcggta caaagggtat 720
gacattccat ctgggtggaa agtgcttccg gtgattgcag ccgtgcattt ggatccttta 780
ctttttgacc acccgcaaca cttcaatcca tggagatggc agcagaataa caaccaccac 840
caccgtggat catcgtcgtc atcatcatgc tacacgagca tgacaagtca caactttatg 900
ccatttgggg gaggaccacg actttgcgcc ggttcagaat tggccaaact tgaaatggcc 960
gtgttcatcc accaccttgt cctcaacttc cactgggagt tagccgatcc tttcgacaaa 1020
ccttttgctt tccccttcgt cgatttccaa aatggcctac caatcacagc ccaccgctac 1080
gtaccaaact aa 1092

<210> 126
<211> 363
<212> PRT
<213> Malus domestica
<400> 126

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

305 310 315 320
Val Phe Ile His His Leu Val Leu Asn Phe His Trp Glu Leu Ala Asp
325 330 335
Pro Phe Asp Lys Pro Phe Ala Phe Pro Phe Val Asp Phe Gln Asn Gly
340 345 350
Leu Pro Ile Thr Ala His Arg Tyr Val Pro Asn
355 360

<210> 127
<211> 748
<212> DNA
<213> Vitis vinifera
<400> 127

agacttcatg gagcaacaca tatcaaggtt cggagaaatc tacaagtcca atctgtttgg 60
cgagccaacc atagtctcag cggattctgg gctgaacaga ttcatactac agaacgaagg 120
aaaattgttt gaatgcagct atcccagaag cataggtgga attcttggga aatggtccat 180
gctggtttta gttggagaca tgcatagaga catgagaacc atctccctca acttcttgag 240
ccatggcagg cttaggactc atctcctacc agaggtggtg aagcacactt tgcttgttct 300
aagctcttgg aaggagaatt gtacattttc tgctcaagat gaagctaaga agtttacctt 360
caatctgatg gcaaaacata tcatgagctt ggatcccgga aagccggaga ctgagcagct 420
taagaaagag tatgttactt tcatgaaagg agtggtatct gctcctttga atttccctgg 480
aactgcatac agaaaagccc tacagtctcg gtccaccatc ttaaaattta tcgagctgaa 540
gatggaagag aggattcaga aactgagggg aggagggttt gagaatatgg aggatgacga 600
tcttcttgga tgggtgctga agcattccaa cctttcaact gagcaaatcc tggacttggt 660
actgagcttg ctctttgctg gccatgaaac ttcatcagtg gcaatagctt tagctatcta 720
cttcttggaa ggctgtccca acgccgtt 748

<210> 128
<211> 248
<212> PRT
<213> Vitis vinifera
<400> 128

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

<210> 129

<211> 1737
<212> DNA
<213> Arabidopsis thaliana
<400> 129

```
atgcaacaac tgtcaacttc cctcggctta aacacttaca acgaggaacg taattcaact     60
tctactaaaa acacgagttc tgaagattat agtccaaggc agtctaaaca cacacaaagc    120
catggcctta aagatatctc cggaaatttc cattctcatg gagttaatgg aggtgtttcg    180
aacatgtcat tctataatac gccttcgcca gtggctgcac agctatccgg tatagctcca    240
ccaccacttt tccggaattt tcagccagct gttgcaaacc caaactccct tattactgac    300
agttctccaa agtccactgt taactctact cttcaagcac ctagaagaaa gtttgtagat    360
gaaggaaagt tacgtaagat ttctggcaga ctattttctg attctggtcc acgacggagt    420
tcaagactgt ctgctgattc aggggcaaac attaattcaa gtgttgcaac agtaagcgga    480
aatgtgaaca acgcttccaa gtatttggga ggttctaaat tgagttcttt ggcacttcgt    540
tctgtaacac ttcggaaggg acactcctgg gcaaatgaaa acatggatga aggggtccgt    600
ggggaacctt ttgatgattc aaggcctaat actgcctcaa cgactggttc tatggcttcc    660
aatgatcaag aagacgaaac aatgtcgatt ggtggcatag caatgagttc tcaaacaatc    720
acaattggtg tttcggaaat tttaaacctc cttaggacac tcggagaagg gtgtagactt    780
tcatacatgt acaggtgtca ggaggcactg gatacgtata tgaaacttcc acataagcat    840
tataatacag gatgggttct ttcccaggtc gggaaagcat actttgaact aattgactat    900
ttagaggctg aaaaggcatt ccgtcttgcc cgtctggctt ctccttattg cttagaagga    960
atggatatat actctacggt cctctatcat ttgaaggaag acatgaagct gagttacttg   1020
gctcaggaac taatatcaac cgatcgctta gctcctcaat cttggtgtgc tatgggaaat   1080
tgctatagct tgcaaaagga ccatgagacc gcactgaaga atttcctacg agctgttcaa   1140
ctgaatccaa gatttgcata tgcacatacc ttatgtggcc acgaatacac aactcttgag   1200
gattttgaga acggaatgaa aagttaccaa aacgcacttc gtgtagatac aagacactac   1260
aacgcatggt acgggcttgg aatgatatat ctacgccaag agaagttaga gttctcagag   1320
catcacttca gaatggcttt cctaataaac ccgagttcct ctgttataat gtcttattta   1380
gggacatctt tgcatgcctt gaagagaagt gaggaagcac tagagataat ggagcaagcc   1440
atagtagcag atagaaaaaa ccctcttcca atgtaccaga aagctaacat acttgtctgc   1500
ttagaaagat tagatgaagc tctagaagtt cttgaggagc tcaaagagta tgcgccttca   1560
gagagcagcg tttacgcttt aatgggcagg atctataagc ggcgaaacat gcacgataaa   1620
gccatgcttc atttcggtct agctttagat atgaaaccgc ctgcaactga cgttgctgca   1680
ataaaggctg caatggagaa attgcatgtt ccagatgaga tcgatgagag cccgtga      1737
```

<210> 130
<211> 578
<212> PRT
<213> Arabidopsis thaliana
<400> 130

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

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

<210> 131
<211> 2235
<212> DNA
<213> Arabidopsis thaliana
<400> 131

```
atggaagcta tgcttgtgga ctgtgtaaac aacagtcttc gtcattttgt ctacaaaaat     60
gctattttca tgtgcgagcg tctctgcgct gagtttcctt ctgaggttaa tttgcagcta    120
ttagccacca gctacctgca gaataatcaa gcttacagtg catatcatct gctaaaggga    180
acacaaatgg ctcagtcccg atacttgttc gcattatcat gcttccagat ggaccttctc    240
aatgaagctg aatctgcact ctgccctgtt aatgaacctg gtgcggagat cccaaatggt    300
gcagcaggcc attaccttct tggacttatt tacaagtata ctgatagaag gaagaatgct    360
gctcaacaat ttaaacagtc cttgacaata gaccctctac tttgggctgc atatgaggaa    420
ttatgtatat taggtgctgc tgaggaagca actgcagttt ttggtgaaac agctgctctc    480
tccattcaaa agcagtatat gcaacaactg tcaacttccc tcggcttaaa cacttacaac    540
```

```
gaggaacgta attcaacttc tactaaaaac acgagttctg aagattatag tccaaggcag    600
tctaaacaca cacaaagcca tggccttaaa gatatctccg gaaatttcca ttctcatgga    660
gttaatggag gtgtttcgaa catgtcattc tataatacgc cttcgccagt ggctgcacag    720
ctatccggta tagctccacc accacttttc cggaattttc agccagctgt tgcaaaccca    780
aactccctta ttactgacag ttctccaaag tccactgtta actctactct tcaagcacct    840
agaagaaagt ttgtagatga aggaaagtta cgtaagattt ctggcagact attttctgat    900
tctggtccac gacggagttc aagactgtct gctgattcag gggcaaacat taattcaagt    960
gttgcaacag taagcggaaa tgtgaacaac gcttccaagt atttgggagg ttctaaattg   1020
agttctttgg cacttcgttc tgtaacactt cggaagggac actcctgggc aaatgaaaac   1080
atggatgaag gggtccgtgg ggaacctttt gatgattcaa ggcctaatac tgcctcaacg   1140
actggttcta tggcttccaa tgatcaagaa gacgaaacaa tgtcgattgg tggcatagca   1200
atgagttctc aaacaatcac aattggtgtt tcggaaattt taaacctcct taggacactc   1260
ggagaagggt gtagactttc atacatgtac aggtgtcagg aggcactgga tacgtatatg   1320
aaacttccac ataagcatta taatacagga tgggttcttt cccaggtcgg gaaagcatac   1380
tttgaactaa ttgactattt agaggctgaa aaggcattcc gtcttgcccg tctggcttct   1440
ccttattgct tagaaggaat ggatatatac tctacggtcc tctatcattt gaaggaagac   1500
atgaagctga gttacttggc tcaggaacta atatcaaccg atcgcttagc tcctcaatct   1560
tggtgtgcta tgggaaattg ctatagcttg caaaaggacc atgagaccgc actgaagaat   1620
ttcctacgag ctgttcaact gaatccaaga tttgcatatg cacatacctt atgtggccac   1680
gaatacacaa ctcttgagga ttttgagaac ggaatgaaaa gttaccaaaa cgcacttcgt   1740
gtagatacaa gacactacaa cgcatggtac gggcttggaa tgatatatct acgccaagag   1800
aagttagagt tctcagagca tcacttcaga atggctttcc taataaaccc gagttcctct   1860
gttataatgt cttatttagg gacatctttg catgccttga agagaagtga ggaagcacta   1920
gagataatgg agcaagccat agtagcagat agaaaaaacc ctcttccaat gtaccagaaa   1980
gctaacatac ttgtctgctt agaaagatta gatgaagctc tagaagttct tgaggagctc   2040
aaagagtatg cgccttcaga gagcagcgtt tacgctttaa tgggcaggat ctataagcgg   2100
cgaaacatgc acgataaagc catgcttcat ttcggtctag ctttagatat gaaaccgcct   2160
gcaactgacg ttgctgcaat aaaggctgca atggagaaat tgcatgttcc agatgagatc   2220
gatgagagcc cgtga                                                    2235
<210> 132
<211> 744
<212> PRT
<213> Arabidopsis thaliana
<400> 132
Met Glu Ala Met Leu Val Asp Cys Val Asn Asn Ser Leu Arg His Phe
1               5                  10                  15
Val Tyr Lys Asn Ala Ile Phe Met Cys Glu Arg Leu Cys Ala Glu Phe
            20                  25                  30
Pro Ser Glu Val Asn Leu Gln Leu Leu Ala Thr Ser Tyr Leu Gln Asn
        35                  40                  45
Asn Gln Ala Tyr Ser Ala Tyr His Leu Leu Lys Gly Thr Gln Met Ala
    50                  55                  60
Gln Ser Arg Tyr Leu Phe Ala Leu Ser Cys Phe Gln Met Asp Leu Leu
65                  70                  75                  80
Asn Glu Ala Glu Ser Ala Leu Cys Pro Val Asn Glu Pro Gly Ala Glu
                85                  90                  95
Ile Pro Asn Gly Ala Ala Gly His Tyr Leu Leu Gly Leu Ile Tyr Lys
            100                 105                 110
Tyr Thr Asp Arg Arg Lys Asn Ala Ala Gln Gln Phe Lys Gln Ser Leu
        115                 120                 125
Thr Ile Asp Pro Leu Leu Trp Ala Ala Tyr Glu Glu Leu Cys Ile Leu
    130                 135                 140
Gly Ala Ala Glu Glu Ala Thr Ala Val Phe Gly Glu Thr Ala Ala Leu
145                 150                 155                 160
Ser Ile Gln Lys Gln Tyr Met Gln Gln Leu Ser Thr Ser Leu Gly Leu
                165                 170                 175
Asn Thr Tyr Asn Glu Glu Arg Asn Ser Thr Ser Thr Lys Asn Thr Ser
            180                 185                 190
Ser Glu Asp Tyr Ser Pro Arg Gln Ser Lys His Thr Gln Ser His Gly
        195                 200                 205
Leu Lys Asp Ile Ser Gly Asn Phe His Ser His Gly Val Asn Gly Gly
    210                 215                 220
Val Ser Asn Met Ser Phe Tyr Asn Thr Pro Ser Pro Val Ala Ala Gln
```

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

Pro Asp Glu Ile Asp Glu Ser Pro
740

<210> 133
<211> 2181
<212> DNA
<213> Arabidopsis thaliana
<400> 133

atgatggaga atctactggc gaattgtgtc cagaaaaacc ttaaccattt tatgttcacc 60
aatgctatct tcctttgcga acttcttctc gcccaatttc catctgaggt gaacctgcaa 120
ttgttagcca ggtgttactt gagtaacagt caagcttata gtgcatatta tatccttaaa 180
ggttcaaaaa cgcctcagtc tcggtattta tttgcattct catgctttaa gttggatctt 240
cttggagagg ctgaagctgc attgttgccc tgtgaagatt atgctgaaga agttcctggt 300
ggtgcagctg ggcattatct tcttggtctt atatatagat attctgggag gaagaactgt 360
tcaatacaac agtttaggat ggcattgtca tttgatccat tgtgttggga agcatatgga 420
gaactttgta gtttaggtgc cgctgaagaa gcctcaacag ttttcgggaa tgttgcttcc 480
cagcgtctta aaacttgtgt agaacaaaga ataagcttct cagaaggagc aaccatagac 540
cagattacag attctgataa ggccttaaaa gatacaggtt tatcgcaaac agaacacatt 600
ccaggagaga accaacaaga tctgaaaatt atgcagcagc ctggagatat tccaccaaat 660
actgacaggc aacttagtac aaacggatgg gacttgaaca caccttctcc agtgctttta 720
caggtaatgg atgctccacc gcctctgctt cttaagaata tgcgtcgtcc agcagtggaa 780
ggatctttga tgtctgtaca tggagtgcgt gtgcgtcgaa gaaacttttt tagtgaagaa 840
ttgtcagcag aggctcaaga agaatctggg cgccgccgta gtgctagaat agcagcaagg 900
aaaaagaatc ctatgtcgca gtcatttgga aaagattccc attggttaca tctttcacct 960
tccgagtcaa actatgcacc ttctctttcc tcgatgattg gaaaatgcag aatccaaagc 1020
agcaaagaag cgattcctga taccgttact ctaaatgatc cagcaacgac gtcaggccag 1080
tctgtaagtg acactggaag ctctgttgat gatgaggaaa agtcaaatcc tagtgaatct 1140
tccccggatc gtttcagcct tatttctgga atttcagaag tgctaggcat tctgaaaatt 1200
cttggagatg gccacaggca tttacatatg tacaagtgtc aggaagcttt gttggcatat 1260
caaaagctat ctcagaaaca atacaataca cactgggttc tcatgcaggt tggaaaagca 1320
tattttgagc tacaagacta cttcaacgct gactcttcct ttactcttgc tcatcaaaag 1380
tatccttatg ctttggaagg aatggataca tactccactg ttctttatca cctgaaagaa 1440
gagatgaggt tgggctatct ggctcaggaa ctgatttcag ttgatcgcct gtctccagaa 1500
tcctggtgtg cagttgggaa ctgttacagt ttgcgtaagg atcatgatac tgctctcaaa 1560
atgtttcaga gagctatcca actgaatgaa agattcacat atgcacatac cctttgtggc 1620
cacgagtttg ccgcattgga agaattcgag gatgcagaga gatgctaccg gaaggctctg 1680
ggcatagata cgagacacta taatgcatgg tacggtcttg gaatgaccta tcttcgtcag 1740
gagaaattcg agtttgcgca gcatcaattt caactggctc tccaaataaa tccaagatct 1800
tcagtcatca tgtgttacta tggaattgct ttgcatgagt caaagagaaa cgatgaggcg 1860
ttgatgatga tggagaaggc tgtactcact gatgcaaaga atccgctccc caagtactac 1920
aaggctcaca tattaaccag cctaggtgat tatcacaaag cacagaaagt tttagaagag 1980
ctcaaagaat gtgctcctca agaaagcagt gtccatgcat cgcttggcaa aatatacaat 2040
cagctaaagc aatacgacaa agccgtgtta catttcggca ttgctttgga tttaagccct 2100
tctccatctg atgctgtcaa gataaaggct tacatggaga ggttgatact accagacgag 2160
ctggtgacgg aggaaaattt g 2181

<210> 134
<211> 728
<212> PRT
<213> Arabidopsis thaliana
<400> 134

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

```
Arg Tyr Ser Gly Arg Lys Asn Cys Ser Ile Gln Gln Phe Arg Met Ala
        115                 120                 125
Leu Ser Phe Asp Pro Leu Cys Trp Glu Ala Tyr Gly Glu Leu Cys Ser
    130                 135                 140
Leu Gly Ala Ala Glu Glu Ala Ser Thr Val Phe Gly Asn Val Ala Ser
145                 150                 155                 160
Gln Arg Leu Gln Lys Thr Cys Val Glu Gln Arg Ile Ser Phe Ser Glu
                165                 170                 175
Gly Ala Thr Ile Asp Gln Ile Thr Asp Ser Asp Lys Ala Leu Lys Asp
            180                 185                 190
Thr Gly Leu Ser Gln Thr Glu His Ile Pro Gly Glu Asn Gln Gln Asp
        195                 200                 205
Leu Lys Ile Met Gln Gln Pro Gly Asp Ile Pro Pro Asn Thr Asp Arg
    210                 215                 220
Gln Leu Ser Thr Asn Gly Trp Asp Leu Asn Thr Pro Ser Pro Val Leu
225                 230                 235                 240
Leu Gln Val Met Asp Ala Leu Pro Pro Leu Leu Leu Lys Asn Met Arg
                245                 250                 255
Arg Pro Ala Val Glu Gly Ser Leu Met Ser Val His Gly Val Arg Val
            260                 265                 270
Arg Arg Arg Asn Phe Phe Ser Glu Glu Leu Ser Ala Glu Ala Gln Glu
        275                 280                 285
Glu Ser Gly Arg Arg Arg Ser Ala Arg Ile Ala Ala Arg Lys Lys Asn
    290                 295                 300
Pro Met Ser Gln Ser Phe Gly Lys Asp Ser His Trp Leu His Leu Ser
305                 310                 315                 320
Pro Ser Glu Ser Asn Tyr Ala Pro Ser Leu Ser Ser Met Ile Gly Lys
                325                 330                 335
Cys Arg Ile Gln Ser Ser Lys Glu Val Ile Pro Asp Thr Val Thr Leu
            340                 345                 350
Asn Asp Pro Ala Thr Thr Ser Gly Gln Ser Val Ser Asp Ile Gly Ser
        355                 360                 365
Ser Val Asp Asp Glu Glu Lys Ser Asn Pro Ser Glu Ser Ser Pro Asp
    370                 375                 380
Arg Phe Ser Leu Ile Ser Gly Ile Ser Glu Val Leu Ser Leu Leu Lys
385                 390                 395                 400
Ile Leu Gly Asp Gly His Arg His Leu His Met Tyr Lys Cys Gln Glu
                405                 410                 415
Ala Leu Leu Ala Tyr Gln Lys Leu Ser Gln Lys Gln Tyr Asn Thr His
            420                 425                 430
Trp Val Leu Met Gln Val Gly Lys Ala Tyr Phe Glu Leu Gln Asp Tyr
        435                 440                 445
Phe Asn Ala Asp Ser Ser Phe Thr Leu Ala His Gln Lys Tyr Pro Tyr
    450                 455                 460
Ala Leu Glu Gly Met Asp Thr Tyr Ser Thr Val Leu Tyr His Leu Lys
465                 470                 475                 480
Glu Glu Met Arg Leu Gly Tyr Leu Ala Gln Glu Leu Ile Ser Val Asp
                485                 490                 495
Arg Leu Ser Pro Glu Ser Trp Cys Ala Val Gly Asn Cys Tyr Ser Leu
            500                 505                 510
Arg Lys Asp His Asp Thr Ala Leu Lys Met Phe Gln Arg Ala Ile Gln
        515                 520                 525
Leu Asn Glu Arg Phe Thr Tyr Ala His Thr Leu Cys Gly His Glu Phe
    530                 535                 540
Ala Ala Leu Glu Glu Phe Glu Asp Ala Glu Arg Cys Tyr Arg Lys Ala
545                 550                 555                 560
Leu Gly Ile Asp Thr Arg His Tyr Asn Ala Trp Tyr Gly Leu Gly Met
                565                 570                 575
Thr Tyr Leu Arg Gln Glu Lys Phe Glu Phe Ala Gln His Gln Phe Gln
            580                 585                 590
Leu Ala Leu Gln Ile Asn Pro Arg Ser Ser Val Ile Met Cys Tyr Tyr
        595                 600                 605
Gly Ile Ala Leu His Glu Ser Lys Arg Asn Asp Glu Ala Leu Met Met
```

```
    610                 615                 620
Met Glu Lys Ala Val Leu Thr Asp Ala Lys Asn Pro Leu Pro Lys Tyr
625                 630                 635                 640
Tyr Lys Ala His Ile Leu Thr Ser Leu Gly Asp Tyr His Lys Ala Gln
                645                 650                 655
Lys Val Leu Glu Glu Leu Lys Glu Cys Ala Pro Gln Glu Ser Ser Val
            660                 665                 670
His Ala Ser Leu Gly Lys Ile Tyr Asn Gln Leu Lys Gln Tyr Asp Lys
        675                 680                 685
Ala Val Leu His Phe Gly Ile Ala Leu Asp Leu Ser Pro Ser Pro Ser
    690                 695                 700
Asp Ala Val Lys Ile Lys Ala Tyr Met Glu Arg Leu Ile Leu Pro Asp
705                 710                 715                 720
Glu Leu Val Thr Glu Glu Asn Leu
                725
<210> 135
<211> 1835
<212> DNA
<213> Oryza sativa
<220>
<221> misc_feature
<222> (729)..(732)
<223> n is a, c, g, ou t
<400> 135
atggaaaccc taatggtgga ccgcgtccac ggcagcctcc gcctcttcat gcaccgcaac     60
gccgtcttcc tctgcgagcg cctctgcgcc cagttccccg ccgagacaaa tgtccagttg    120
tagcaacttg ctaccttcac aacaaccagc catatgctgc ataccacatc ttgaaaggaa    180
agaagctgcc agagtcccgg tacttgtttg ctatgtcatg cttccgaatg aacctcttac    240
gggaagctga agaagccttg tgtcctgtca atgaaccaaa tattgaggtt ccaagtggtg    300
caacagggca ctaccttctt ggagtaattt acaggtacac tggcagagtg gaagctgcag    360
ctgagcaatt tgtacaagct ctgactcttg atcctcttct atgggcagca tacgaggaat    420
tgtgcatact aggtgttgct gaagatgcaa atgaatgttt cagtgaagca acagctctac    480
gtcttcagca ggaactcaca tccacatcaa atgtggaaaa gtcaaacttt gttaatgaaa    540
atcggtttct atcttccaat gtgtcagcaa gttttggtga tagtcctaag caaattaaac    600
agctgcatgc taacaccact gcagaagtat ctggttatcc tcatgtaaag tcaactgcat    660
tgcatatgca gaacggtgca ccatctaatt tatcacagtt tgacactcca tcgccaactt    720
caacgcagnn nnataatgta acttcaactt cgtcttctac aagtatagtt gatggaagat    780
atcccgagca agagaaatct gaacgagttc tgtcacagga ctccaaatta gctattggta    840
tcagggagct aatggcactc ttgcggacac taggggaagg gtataggctt tcttgcttgt    900
ttaagtgtca ggaagcattg gaagtatata gaaagctccc agaggcacaa tttaatactg    960
gatgggttct ttgccaggtt gggaagacat attttgaact cgtcaattat ttagaagccg   1020
atcatttttt tgagttagcg catcgactat caccatgcac gttggaggga atggacattt   1080
actccactgt tctttatcat ttgaatgagg aaatgcggct aagttacctt gctcaagatc   1140
ttgtttctat tgatcgacta tctccccaag catggtgtgc tgtgggaaat tgctttgcct   1200
tgaggaaaga tcatgagact gccttgaaga attttcaacg tgctgtacag cttgactcaa   1260
gagttgcata cgctcacacg ctatgcggtc acgatataaa actataccga tctgcacttc   1320
aggtagatga aagacactac aatgcctggt atggccttgg agtggtgtac cttcgccagg   1380
aaaagtttga gtttgctgag catcatttca gaagggcatt ccagataaat ccttgctctt   1440
ctgttcttat gtgctatctt gggatggcct tgcatgcttt aaagaggaat gaggaagcct   1500
tggaaatgat ggagaaggct atatttgctg ataagaagaa tccactcccc aagtatcaaa   1560
aggctttaat ccttctaggc ctacaaaaat accctgatgc tctggatgag ttggaacggc   1620
taaaggaaat tgcacctcat gaaagtagta tgtatgcact gatgggaaag atttacaagc   1680
aacttaacat tcttgacaag gctgtatttt gctttggcat tgccctggat ttgaaacctc   1740
ctgctgctga cgttgctata atacaatctg caatggagaa agtacacctt ccagatgaac   1800
ttatggatga tgatgatgat gatgatgaga tttaa                              1835
<210> 136
<211> 757
<212> PRT
<213> Oryza sativa
<400> 136
Met Glu Thr Leu Met Val Asp Arg Val His Gly Ser Leu Arg Leu Phe
1               5                   10                  15
Met His Arg Asn Ala Val Phe Leu Cys Glu Arg Leu Cys Ala Gln Phe
```

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

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

<210> 137
<211> 2253
<212> DNA
<213> Solanum tuberosum
<400> 137

```
atggaaaccc tactagctga atctgtgcaa aacagccttg gccaatttat gtaccacaac     60
gccattttca tgtgtgaacg actctgtgcc gagttcccca ctgagacaaa tatgcagctt    120
ttagctggct gctacctgca caaccaacag gcttatgctg catatcatct tctcaagggg    180
acaagtatgg ctcaatcccg ctacttgttt gcactatcat gctttcagat ggatcttctc    240
actgaagctg agacagcact ttgccctcct aatgagccaa ctgcagaggt tccaaatggt    300
gcagctgggc attaccttct tggtcttatt tacaggtata cagatagaag aaatagttcc    360
atccagcatt tcaatcaggc attgttattg gatccattgc tatgggctgc atatgaggag    420
ttgtgtatac taggtgctgc agaagaagca gctgcagttt ttggggaagc atctttgctt    480
tgcattcaga aacaacacat agaccaaggg aaccaatctc aaaatttaca agcatccact    540
gatgatcaga atgtagcttc tacgaatatt gtctcaggcg acatcagccc tatgcaatca    600
aaacatacac acagccataa tcttcgagaa atgtctggaa attataatgg agcagctgct    660
atccagaatt taggtggggt ttctactaac atgtcattct acaacactcc ctcaccaatg    720
gcatcacagt tgtcaggagt ggttccacct ccagtttgta gaaattttca gcaaaatgga    780
actaatgcat ctgtggctgg tgctgataat tctccacgag caactgtcaa ttcaaccatt    840
caggcccctc ggagaaagtt tgttgatgag gggaagttaa gaaagatatc tgggaggtta    900
ttttctgatt ctggccctcg acgaaattca aggcttgctg gagaatctac tggaaacaca    960
aattcaaatg tatctggtgc ttctggaaat ggaacaattc attcttccaa atattatggt   1020
agttcgaagc tgagctcaat gactttacgt tccatgacaa gtcgaaaggc acaatcttgg   1080
gctaccgaaa actatggtga agggactcgc aatgacattt ctaatgattc tcggctaaat   1140
atgactatgt cacacccttc tggagatgct agacctcttg aacaagaaag gccccgaact   1200
tctgcttctg gggttaatgt aagcagcaca tctatccctc tcagtgcttc agagatattg   1260
gccctttca ggtttcttgg ggaaggctat agactttctt gtttatatag atgtcaggat   1320
gcactggatg tttataacaa actcccacac aaacattatc acactggatg ggttctttct   1380
cagattggaa gagcatactt cgaaatggtt gattacctag aagcagatca tgcatttggc   1440
cttgctcgtc tggcctcacc ttatagttta gaaggaatgg acgtgtactc gacagtgttg   1500
tttcatctca aggaggacat gaagttgagc tatctggcgc aggtgctggt atcaactgat   1560
agattagctc ctcaatcttg gtgtgctatg gggaattgct atagtttaca gaaagaccat   1620
gaaactgctc ttaaaaattt tcaacgagct gtacaactaa atcctagatt tgcatacggg   1680
```

```
cacacgcttt gtggtcatga atatgttgct ttagaagatt ttgaaaatgc tattaagagc   1740
tatcagagtg cacttcgtgt ggatgccagg cattacaatg cctggtatgg gcttggaatg   1800
atctatctcc gacaggagaa gtttgaattt tcagagcatc actttcgaat ggctttgggt   1860
ataaatccac agtcttctgt tatcatgtca tatcttggca ctgcattaca cgctctgaag   1920
aaaaatgaag aggcattgga agtgatggag ctggctattg tagcagacaa gaaaaaccct   1980
cttccaatgt atcagaaggc taacatcctt gtgagcacgg aaagttttga tgccgcttta   2040
gaagtcttag aggaacttaa agagcatgct cctcgtgaga gcagtgtcta tgctttgatg   2100
ggtcggatat acaagaggcg taatatgtac gacaaagcca tgcttcattt tggagtggca   2160
ttagatttaa aaccatctgc aactgatgtt gctaccatta aggctgccat tgaaaagctg   2220
catgtaccag atgagatgga agatgaatta taa                               2253
<210> 138
<211> 750
<212> PRT
<213> Solanum tuberosum
<400> 138
Met Glu Thr Leu Leu Ala Glu Ser Val Gln Asn Ser Leu Gly Gln Phe
1               5                  10                  15
Met Tyr His Asn Ala Ile Phe Met Cys Glu Arg Leu Cys Ala Glu Phe
            20                  25                  30
Pro Thr Glu Thr Asn Met Gln Leu Leu Ala Gly Cys Tyr Leu His Asn
        35                  40                  45
Gln Gln Ala Tyr Ala Ala Tyr His Leu Leu Lys Gly Thr Ser Met Ala
    50                  55                  60
Gln Ser Arg Tyr Leu Phe Ala Leu Ser Cys Phe Gln Met Asp Leu Leu
65                  70                  75                  80
Thr Glu Ala Glu Thr Ala Leu Cys Pro Pro Asn Glu Pro Thr Ala Glu
                85                  90                  95
Val Pro Asn Gly Ala Ala Gly His Tyr Leu Leu Gly Leu Ile Tyr Arg
            100                 105                 110
Tyr Thr Asp Arg Arg Asn Ser Ser Ile Gln His Phe Asn Gln Ala Leu
        115                 120                 125
Leu Leu Asp Pro Leu Leu Trp Ala Ala Tyr Glu Glu Leu Cys Ile Leu
    130                 135                 140
Gly Ala Ala Glu Glu Ala Ala Ala Val Phe Gly Glu Ala Ser Leu Leu
145                 150                 155                 160
Cys Ile Gln Lys Gln His Ile Asp Gln Gly Asn Gln Ser Gln Asn Leu
                165                 170                 175
Gln Ala Ser Thr Asp Asp Gln Asn Val Ala Ser Thr Asn Ile Val Ser
            180                 185                 190
Gly Asp Ile Ser Pro Met Gln Ser Lys His Thr His Ser His Asn Leu
        195                 200                 205
Arg Glu Met Ser Gly Asn Tyr Asn Gly Ala Ala Ala Ile Gln Asn Leu
    210                 215                 220
Gly Gly Val Ser Thr Asn Met Ser Phe Tyr Asn Thr Pro Ser Pro Met
225                 230                 235                 240
Ala Ser Gln Leu Ser Gly Val Val Pro Pro Pro Val Cys Arg Asn Phe
                245                 250                 255
Gln Gln Asn Gly Thr Asn Ala Ser Val Ala Gly Ala Asp Asn Ser Pro
            260                 265                 270
Arg Ala Thr Val Asn Ser Thr Ile Gln Ala Pro Arg Arg Lys Phe Val
        275                 280                 285
Asp Glu Gly Lys Leu Arg Lys Ile Ser Gly Arg Leu Phe Ser Asp Ser
    290                 295                 300
Gly Pro Arg Arg Asn Ser Arg Leu Ala Gly Glu Ser Thr Gly Asn Thr
305                 310                 315                 320
Asn Ser Asn Val Ser Gly Ala Ser Gly Asn Gly Thr Ile His Ser Ser
                325                 330                 335
Lys Tyr Tyr Gly Ser Ser Lys Leu Ser Ser Met Thr Leu Arg Ser Met
            340                 345                 350
Thr Ser Arg Lys Ala Gln Ser Trp Ala Thr Glu Asn Tyr Gly Glu Gly
        355                 360                 365
Thr Arg Asn Asp Ile Ser Asn Asp Ser Arg Leu Asn Met Thr Met Ser
    370                 375                 380
```

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

<210> 139
<211> 1998
<212> DNA
<213> Schizosaccharomyces pombe
<400> 139

```
atgacagatc gattgaaatg tttaatatgg tattgcattg ataatcagaa ttatgataat     60
tcaatttttt attcagaacg tttacatgca attgaagatt caaacgagag tttgtatctt    120
ttggcatatt cgcatttcct aaacctcgat tacaatattg tatacgactt attagataga    180
gtaattagtc atgttccttg cacatactta tttgcaagga ccagccttat tttaggcaga    240
tataaacaag gaataagtgc tgtggaggcc tgtcgatcga attggcgctc cattcagcca    300
aacataaatg actcaattag cagtcgtgga catccagatg cctcttgcat gcttgatgtt    360
ttgggtacta tgtataaaaa ggcagggttc ctcaaaaaag ctacagattg ttttgtagaa    420
gctgtctcca ttaacccata taatttctct gctttccaga atttaactgc aataggcgtg    480
ccactcgatg ctaataatgt atttgttatt ccaccctacc ttacggcaat gaagggtttt    540
gaaaaatctc aaacgaatgc tacagcttcg gtaccagaac cgtctttttt gaagaaaagt    600
aaagagtctt cctcatcttc caacaagttt tcggtttctg aatcgatagc aaatagttat    660
tcaaactcat ccatttcagc atttactaag tggtttgata gggttgacgc ttctgagctt    720
```

```
ccaggaagtg agaaggaacg acatcaaagc ttgaaattac aatctcaatc tcagactagc    780
aaaaaccttt tggctttcaa tgatgctcaa aaagctgatt ctaacaatag ggatacgtct    840
ttaaaatccc actttgtgga acctagaacc caagcattaa gaccaggagc tcgtttaaca    900
tataaattac gcgaagcgag aagttctaaa agaggagaga gcacacctca aagcttccgc    960
gaagaggaca ataatttgat ggaattacta aagttattcg gtaagggtgt ttacctgctc   1020
gcccagtata agttacgaga ggctttaaat tgtttccaaa gcttgcccat cgaacagcaa   1080
aatacacctt ttgttcttgc aaagcttgga ataacctact ttgaactggt tgattacgaa   1140
aaatctgaag aagtgtttca aaaattaagg gacttgtcgc cttcacgtgt caaagatatg   1200
gaagtctttt caactgcact ttggcatttg caaaagtctg ttcctttatc ttaccttgcc   1260
catgaaactt tggaaactaa tccttattcc ccagaatcat ggtgcattct tgctaattgc   1320
ttctcacttc aacgtgaaca ctcgcaggca ttaaaatgta ttaatagagc tattcaattg   1380
gatccaactt ttgaatatgc ttatacgctt caagggcacg agcattctgc aaacgaagaa   1440
tacgaaaaat cgaaaacatc tttccgcaaa gcaattagag taaatgttcg acattacaat   1500
gcatggtatg gcctgggaat ggtttattta aaaacagggc gaaatgatca agcagacttt   1560
cattttcaac gtgctgcaga gatcaatcct aacaactctg tactcatcac ttgtattggt   1620
atgatttacg aacgctgcaa agattacaaa aaagcacttg atttttatga tcgggcatgt   1680
aaactggatg aaaagtcttc gcttgccagg ttcaagaaag ccaaagtgct tattttatta   1740
catgatcacg ataaagcact cgttgaattg gaacaattaa aggcaattgc gccagatgaa   1800
gcaaatgttc atttttgct cggcaaaatt ttcaagcaaa tgcggaaaaa aaatttagcc   1860
ttaaagcact tcactatagc atggaactta gacggcaagg ctacgcatat tattaaggaa   1920
tcgattgaaa atctggatat tcccgaagaa aatttgttaa ctgaaacagg tgaaatttat   1980
aggaatctgg aaacttaa                                                 1998
<210> 140
<211> 665
<212> PRT
<213> Schizosaccharomyces pombe
<400> 140
Met Thr Asp Arg Leu Lys Cys Leu Ile Trp Tyr Cys Ile Asp Asn Gln
1               5                  10                  15
Asn Tyr Asp Asn Ser Ile Phe Tyr Ser Glu Arg Leu His Ala Ile Glu
            20                  25                  30
Asp Ser Asn Glu Ser Leu Tyr Leu Leu Ala Tyr Ser His Phe Leu Asn
        35                  40                  45
Leu Asp Tyr Asn Ile Val Tyr Asp Leu Leu Asp Arg Val Ile Ser His
    50                  55                  60
Val Pro Cys Thr Tyr Leu Phe Ala Arg Thr Ser Leu Ile Leu Gly Arg
65                  70                  75                  80
Tyr Lys Gln Gly Ile Ser Ala Val Glu Ala Cys Arg Ser Asn Trp Arg
                85                  90                  95
Ser Ile Gln Pro Asn Ile Asn Asp Ser Ile Ser Ser Arg Gly His Pro
            100                 105                 110
Asp Ala Ser Cys Met Leu Asp Val Leu Gly Thr Met Tyr Lys Lys Ala
        115                 120                 125
Gly Phe Leu Lys Lys Ala Thr Asp Cys Phe Val Glu Ala Val Ser Ile
    130                 135                 140
Asn Pro Tyr Asn Phe Ser Ala Phe Gln Asn Leu Thr Ala Ile Gly Val
145                 150                 155                 160
Pro Leu Asp Ala Asn Asn Val Phe Val Ile Pro Pro Tyr Leu Thr Ala
                165                 170                 175
Met Lys Gly Phe Glu Lys Ser Gln Thr Asn Ala Thr Ala Ser Val Pro
            180                 185                 190
Glu Pro Ser Phe Leu Lys Lys Ser Lys Glu Ser Ser Ser Ser Ser Asn
        195                 200                 205
Lys Phe Ser Val Ser Glu Ser Ile Ala Asn Ser Tyr Ser Asn Ser Ser
    210                 215                 220
Ile Ser Ala Phe Thr Lys Trp Phe Asp Arg Val Asp Ala Ser Glu Leu
225                 230                 235                 240
Pro Gly Ser Glu Lys Glu Arg His Gln Ser Leu Lys Leu Gln Ser Gln
                245                 250                 255
Ser Gln Thr Ser Lys Asn Leu Leu Ala Phe Asn Asp Ala Gln Lys Ala
            260                 265                 270
Asp Ser Asn Asn Arg Asp Thr Ser Leu Lys Ser His Phe Val Glu Pro
        275                 280                 285
```

```
Arg Thr Gln Ala Leu Arg Pro Gly Ala Arg Leu Thr Tyr Lys Leu Arg
    290                 295                 300
Glu Ala Arg Ser Ser Lys Arg Gly Glu Ser Thr Pro Gln Ser Phe Arg
305                 310                 315                 320
Glu Glu Asp Asn Asn Leu Met Glu Leu Leu Lys Leu Phe Gly Lys Gly
                325                 330                 335
Val Tyr Leu Leu Ala Gln Tyr Lys Leu Arg Glu Ala Leu Asn Cys Phe
            340                 345                 350
Gln Ser Leu Pro Ile Glu Gln Gln Asn Thr Pro Phe Val Leu Ala Lys
        355                 360                 365
Leu Gly Ile Thr Tyr Phe Glu Leu Val Asp Tyr Glu Lys Ser Glu Glu
    370                 375                 380
Val Phe Gln Lys Leu Arg Asp Leu Ser Pro Ser Arg Val Lys Asp Met
385                 390                 395                 400
Glu Val Phe Ser Thr Ala Leu Trp His Leu Gln Lys Ser Val Pro Leu
                405                 410                 415
Ser Tyr Leu Ala His Glu Thr Leu Glu Thr Asn Pro Tyr Ser Pro Glu
            420                 425                 430
Ser Trp Cys Ile Leu Ala Asn Cys Phe Ser Leu Gln Arg Glu His Ser
        435                 440                 445
Gln Ala Leu Lys Cys Ile Asn Arg Ala Ile Gln Leu Asp Pro Thr Phe
    450                 455                 460
Glu Tyr Ala Tyr Thr Leu Gln Gly His Glu His Ser Ala Asn Glu Glu
465                 470                 475                 480
Tyr Glu Lys Ser Lys Thr Ser Phe Arg Lys Ala Ile Arg Val Asn Val
                485                 490                 495
Arg His Tyr Asn Ala Trp Tyr Gly Leu Gly Met Val Tyr Leu Lys Thr
            500                 505                 510
Gly Arg Asn Asp Gln Ala Asp Phe His Phe Gln Arg Ala Ala Glu Ile
        515                 520                 525
Asn Pro Asn Asn Ser Val Leu Ile Thr Cys Ile Gly Met Ile Tyr Glu
    530                 535                 540
Arg Cys Lys Asp Tyr Lys Lys Ala Leu Asp Phe Tyr Asp Arg Ala Cys
545                 550                 555                 560
Lys Leu Asp Glu Lys Ser Ser Leu Ala Arg Phe Lys Lys Ala Lys Val
                565                 570                 575
Leu Ile Leu Leu His Asp His Asp Lys Ala Leu Val Glu Leu Glu Gln
            580                 585                 590
Leu Lys Ala Ile Ala Pro Asp Glu Ala Asn Val His Phe Leu Leu Gly
        595                 600                 605
Lys Ile Phe Lys Gln Met Arg Lys Lys Asn Leu Ala Leu Lys His Phe
    610                 615                 620
Thr Ile Ala Trp Asn Leu Asp Gly Lys Ala Thr His Ile Ile Lys Glu
625                 630                 635                 640
Ser Ile Glu Asn Leu Asp Ile Pro Glu Glu Asn Leu Leu Thr Glu Thr
                645                 650                 655
Gly Glu Ile Tyr Arg Asn Leu Glu Thr
            660                 665
<210> 141
<211> 2421
<212> DNA
<213> Aspergillus niger
<400> 141
atgactccgt ccacatcaca tatttcgagc cagctaaggc agctgatata ttaccatctt     60
gacaacaacc ttgctcggaa cgcgctgttc cttgccggtc gtttacacgc ctacgaacct    120
cggacgtcgg aagcttcgta cctattagct ctgtgttacc tacaaaatgg tcaggtgaaa    180
gcagcatggg aaactagcaa gcattttggg tcgaggggtg cgcatcttgg atgttcttac    240
gtctacgcgc aggcttgtct tgacctcggg aaatatacgg acggtattaa cgcgctagag    300
cgaagtaagg gacaatggac ttcgcgaaac cactggaata aacacagtga gacgcgacga    360
caacacttgc cggatgctgc ggcagtttta tgtttgcaag gaaaattatg gcaggcacac    420
aaggaacaca acaaggctgt ggagtgttac gctgcagctt taaagctgaa tcccttcatg    480
tgggatgcat tcttgaatct gtgcgaaact ggtgttgatt tgcgtgtttc aaacatatat    540
aagatgagcc cggaattgta cagcatggta tcgtctgctg cgctcgaaga tgttgaatcc    600
```

caggttctac ctccggacgg tccactccag acacaagtta acccaaatcc gagcttggac 660
ccgttcactg ccggtacgac tcgcagcgat tcgacctcta ctcacgggag ctcagctttg 720
tgggaaaagt tgaatggaag cacagttagt gtggcatctt caggaacagg accgcatctt 780
ccgcgggaag gcatggagac tcctggtggc caaagtagcg aatctgatga tccacgtgta 840
accaacggta atggcacaga tgtttttgag ccgccgctag cccctgcaaa gaagaatcga 900
acgatccaaa caataggcgg cgatcatccg atggatcccc cgccaaagat gcggccaacc 960
ggtattcgac cgaggacccg aaccaaattc gagtcagatg aaggtcatac tgagagagac 1020
gcgggcatgg gtcacaggtt aggcgatcgg aaacgaacgg tctctggaca agttgcgcat 1080
ccgtcggtac cgcattcaac cgaccaaggt gtcgggcaac ggcgaagtgt acgtctcttt 1140
aaccagataa agccatcaac gaacaaaata tctagcacgg cattaggagt gaaagagggc 1200
cgggaggtga aaaaggtgag aactacgggg aataaggcgc gcacaacaac aagctcaaat 1260
gtgggccgtg ttgttagcgg caacaatcga cgacacgctg gggagatcca tgacggagac 1320
agcaaagaat accgtggaac atcttcaaca tcaaatggct ctcagaacgc atcgtctaag 1380
ctcgctatat ctgagcggac caaatcggtt gaagccttgg cttggatttt ggacttgttt 1440
ttcaagatag cctccggcta tttctgtctc agccggtaca aatgctccga tgctatccag 1500
attttcagct ctctatccca gggccaacgg gagacaccgt gggttcttgc tcaaattgga 1560
cgagcttact acgagcaagc aatgtataca gaggccgaga aatactttgt ccgggtgaag 1620
gccatggcgc cttcccggtt agaagatatg gagatctact cgacggttct ttggcatttg 1680
aagaatgatg tcgaacttgc ttacctggcc catgagttaa tggatgtcga tcgcttatca 1740
ccagaagctt ggtgtgctgt cggtaactcg ttctcacacc agcgggacca cgatcaagct 1800
ctgaagtgct tcaagcgtgc cactcaattg gatcctcatt tcgcgtatgg gttcacgcta 1860
cagggccatg agtatgttgc caacgaggaa tatgacaagg cattggatgc ctatagaagc 1920
ggcatcaacg cggacagtcg acattataat gcctggtacg ggctgggaac ggtctacgat 1980
aagatgggca aacttgactt tgctgaacag cacttccgta atgcggccaa gattaatcct 2040
tctaatgctg ttttaatatg ctgcattggc ttggtcttag agaagatgaa caatccgaag 2100
tcagcgttga tccagtacaa cagggcctgc accctcgcgc ctcattctgt tcttgctcga 2160
ttccgaaaag cccgtgcgtt gatgaaattg caggatctca agtcagcgct tacggagctg 2220
aaggttctga aggacatggc gccggatgag gcaaacgttc attacctcct gggcaaactg 2280
tataagatgc ttcgtgacaa gggaaatgcc attaagcact tcacaactgc attaaacctt 2340
gatcccaagg cggcgcagta catcaaagat gcaatggaag cccttgacga cgatgaagag 2400
gatgaagaag atatggcgtg a 2421

<210> 142
<211> 806
<212> PRT
<213> Aspergillus niger
<400> 142

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

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

```
Phe Arg Lys Ala Arg Ala Leu Met Lys Leu Gln Asp Leu Lys Ser Ala
                725                 730                 735
Leu Thr Glu Leu Lys Val Leu Lys Asp Met Ala Pro Asp Glu Ala Asn
            740                 745                 750
Val His Tyr Leu Leu Gly Lys Leu Tyr Lys Met Leu Arg Asp Lys Gly
        755                 760                 765
Asn Ala Ile Lys His Phe Thr Thr Ala Leu Asn Leu Asp Pro Lys Ala
    770                 775                 780
Ala Gln Tyr Ile Lys Asp Ala Met Glu Ala Leu Asp Asp Asp Glu Glu
785                 790                 795                 800
Asp Glu Glu Asp Met Ala
                805
```

<210> 143
<211> 2472
<212> DNA
<213> Homo sapiens
<400> 143

```
atgacggtgc tgcaggaacc cgtccaggct gctatatggc aagcactaaa ccactatgct     60
taccgagatg cggttttcct cgcagaacgc ctttatgcag aagtacactc agaagaagcc    120
ttgtttttac tggcaacctg ttattaccgc tcaggaaagg catataaagc atatagactc    180
ttgaaaggac acagttgtac tacaccgcaa tgcaaatacc tgcttgcaaa atgttgtgtt    240
gatctcagca agcttgcaga aggggaacaa atcttatctg gtggagtgtt taataagcag    300
aaaagccatg atgatattgt tactgagttt ggtgattcag cttgctttac tctttcattg    360
ttgggacatg tatattgcaa gacagatcgg cttgccaaag gatcagaatg ttaccaaaag    420
agccttagtt taaatccttt cctctggtct ccctttgaat cattatgtga aataggtgaa    480
aagccagatc ctgaccaaac atttaaattc acatctttac agaactttag caactgtctg    540
cccaactctt gcacaacaca agtacctaat catagtttat ctcacagaca gcctgagaca    600
gttcttacgg aaacacccca ggacacaatt gaattaaaca gattgaattt agaatcttcc    660
aattcaaagt actccttgaa tacagattcc tcagtgtctt atattgattc agctgtaatt    720
tcacctgata ctgtcccact gggaacagga acttccatat tatctaaaca ggttcaaaat    780
aaaccaaaaa ctggtcgaag tttattagga ggaccagcag ctcttagtcc attaacccca    840
agttttggga ttttgccatt agaaacccca agtcctggag atggatccta tttacaaaac    900
tacactaata cacctcctgt aattgatgtg ccatccaccg gagccccttc aaaaaagtct    960
gttgccagaa tcggccaaac tggaacaaag tctgtcttct cacagagtgg aaatagccga   1020
gaggtaactc caattcttgc acaaacacaa agttctggtc cacaaacaag tacaacacct   1080
caggtattga gccccactat tacatctccc ccaaacgcac tacctcgaag aagttcacga   1140
ctctttacta gtgacagctc cacaaccaag gagaatagca aaaaattaaa aatgaagttt   1200
ccacctaaaa tcccaaacag aaaaacaaaa agtaaaacta ataaaggagg aataactcaa   1260
cctaacataa atgatagcct ggaaattaca aaattggact cttccatcat ttcagaaggg   1320
aaaatatcca caatcacacc tcagattcag gcctttaatc tacaaaaagc agcagcaggt   1380
ttgatgagcc ttcttcgtga aatggggaaa ggttatttag ctttgtgttc atacaactgc   1440
aaagaagcta taaatatttt gagccatcta ccttctcacc actacaatac tggttgggta   1500
ctgtgccaaa ttggaagggc ctattttgaa ctttcagagt acatgcaagc tgaaagaata   1560
ttctcagagg ttagaaggat tgagaattat agagttgaag gcatggagat ctactctaca   1620
acactttggc atcttcaaaa agatgttgct ctttcagttc tgtcaaaaga cttaacagac   1680
atggataaaa attcgccaga ggcctggtgt gctgcaggga actgtttcag tctgcaacgg   1740
gaacatgata ttgcaattaa attcttccag agagctatcc aagttgatcc aaattacgct   1800
tatgcctata ctctattagg gcatgagttt gtcttaactg aagaattgga caaagcatta   1860
gcttgttttc gaaatgctat cagagtcaat cctagacatt ataatgcatg gtatggttta   1920
ggaatgattt attacaagca agaaaaattc agccttgcag aaatgcattt ccaaaaagcg   1980
cttgatatca accctcaaag ttcagtttta ctttgccaca ttggagtagt tcaacatgca   2040
ctgaaaaaat cagagaaggc tttggatacc ctaaacaaag ccattgtcat tgatcccaag   2100
aaccctctat gcaaatttca cagagcctca gttttatttc gaaatgaaaa atataagtct   2160
gctttacaag aacttgaaga attgaaacaa attgttccca aagaatccct cgtttacttc   2220
ttaataggaa aggtttacaa gaagttaggt caaacgcacc tcgccctgat gaatttctct   2280
tgggctatgg atttagatcc taaaggagcc aataaccaga ttaaagaggc aattgataag   2340
cgttatcttc cagatgatga ggagccaata acccaagaag aacagatcat gggaacagat   2400
gaatcccagg agagcagcat gacagatgcg gatgacacac aacttcatgc agctgaaagt   2460
gatgaatttt aa                                                       2472
```

<210> 144
<211> 824
<212> PRT
<213> Homo sapiens

<400> 144

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

```
Asn Thr Gly Trp Val Leu Cys Gln Ile Gly Arg Ala Tyr Phe Glu Leu
            500                 505                 510
Ser Glu Tyr Met Gln Ala Glu Arg Ile Phe Ser Glu Val Arg Arg Ile
        515                 520                 525
Glu Asn Tyr Arg Val Glu Gly Met Glu Ile Tyr Ser Thr Thr Leu Trp
    530                 535                 540
His Leu Gln Lys Asp Val Ala Leu Ser Val Leu Ser Lys Asp Leu Thr
545                 550                 555                 560
Asp Met Asp Lys Asn Ser Pro Glu Ala Trp Cys Ala Ala Gly Asn Cys
                565                 570                 575
Phe Ser Leu Gln Arg Glu His Asp Ile Ala Ile Lys Phe Phe Gln Arg
            580                 585                 590
Ala Ile Gln Val Asp Pro Asn Tyr Ala Tyr Ala Tyr Thr Leu Leu Gly
        595                 600                 605
His Glu Phe Val Leu Thr Glu Glu Leu Asp Lys Ala Leu Ala Cys Phe
    610                 615                 620
Arg Asn Ala Ile Arg Val Asn Pro Arg His Tyr Asn Ala Trp Tyr Gly
625                 630                 635                 640
Leu Gly Met Ile Tyr Tyr Lys Gln Glu Lys Phe Ser Leu Ala Glu Met
                645                 650                 655
His Phe Gln Lys Ala Leu Asp Ile Asn Pro Gln Ser Ser Val Leu Leu
            660                 665                 670
Cys His Ile Gly Val Val Gln His Ala Leu Lys Lys Ser Glu Lys Ala
        675                 680                 685
Leu Asp Thr Leu Asn Lys Ala Ile Val Ile Asp Pro Lys Asn Pro Leu
    690                 695                 700
Cys Lys Phe His Arg Ala Ser Val Leu Phe Ala Asn Glu Lys Tyr Lys
705                 710                 715                 720
Ser Ala Leu Gln Glu Leu Glu Glu Leu Lys Gln Ile Val Pro Lys Glu
                725                 730                 735
Ser Leu Val Tyr Phe Leu Ile Gly Lys Val Tyr Lys Lys Leu Gly Gln
            740                 745                 750
Thr His Leu Ala Leu Met Asn Phe Ser Trp Ala Met Asp Leu Asp Pro
        755                 760                 765
Lys Gly Ala Asn Asn Gln Ile Lys Glu Ala Ile Asp Lys Arg Tyr Leu
    770                 775                 780
Pro Asp Asp Glu Glu Pro Ile Thr Gln Glu Glu Gln Ile Met Gly Thr
785                 790                 795                 800
Asp Glu Ser Gln Glu Ser Ser Met Thr Asp Ala Asp Asp Thr Gln Leu
                805                 810                 815
His Ala Ala Glu Ser Asp Glu Phe
            820
<210> 145
<211> 783
<212> DNA
<213> Saccharum officinarum
<220>
<221> misc_feature
<222> (780)..(783)
<223> n is a, c, g, ou t
<400> 145
atggaaaccc taatggtgga ccgcgtccac agcagcctcc gcctcttcat gcaccgcaac     60
gccgtattcc tctgcgagcg cctctgcgcg cagttcccct ccgagaccaa tgtgcaattg    120
ttagcgacct gctacctcca caacaatcag ccatatgctg cataccacat tttgaaaggg    180
aagaagctgc cggagtcccg gtacttgttt gctacatcat gctttcgaat gaacctcttg    240
cgtgaagcag aagaaactct atgtccagtc aatgaaccaa acatggaggt tccaagtgga    300
gcaacaggac actacctcct tggagtgatt tacaggtgca caggcagaat ttcagctgca    360
gctgaacaat ttacacaagc gttgactcta gatcctcttt tatgggcggc atatgaggaa    420
ttgtgtatat taggtattgc tgaagatact gatgagtgtt ttagtgaatc gactgctcta    480
cgtctccagc aggaacacac atccacggcc actctggtga agtcgaactt cgccaatgaa    540
aatcgagttc tatcatccag ggtctctgca aatcttgggg atattagtcc taagcaaatc    600
aaacagcttc atgctaacaa catagcagaa gtatctggct atcctcatgt aagaccaact    660
gcattgcatg tgcagaacag ttcaacctct aatgtagcac agtttgacac cccatcacca    720
```

actgcagcac agacttctag tatcatgcca ccaccactct ttaggaatgt ccatgcttan 780
nnn 783

<210> 146
<211> 259
<212> PRT
<213> Saccharum officinarum
<400> 146

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

<210> 147
<211> 1314
<212> DNA
<213> Saccharum officinarum
<220>
<221> misc_feature
<222> (416)..(416)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (881)..(881)
<223> n is a, c, g, ou t
<400> 147

attcaaattc aaatacctgg ggtttggagg gaatggtaca ggttattcgt cagggaaatt 60
gcgagtaaac tcgtccacac catcaaaatg gtgttaacca ccatacgttc cgtgcaagtt 120
aggaaaggaa aaccacgggc tacagaaaat tttgatgaag gaagtagata tgaagtcatt 180
gatgaaatgt ggacagacaa tatatcagga acttcatctt ctgtaagtac agctgatgga 240
agatcctttg agcaagataa agctgaacga attctgttgc aagactccaa attggcactt 300
ggtattaggg agatattggg acttgtccga acactcggtg aaggttgtag gctttcttgc 360
ttgtttaagt gccatgaagc cttggaagtc tacagaagac tccctgagac ccattntagc 420
actggatgga gcatatgcca ggttggtaag gcatatttcg aattagttga ttatttggaa 480
gctgatcgtt actttgaatt ggcacaccga ctgtcgcctt gtacgcttga tggaatggac 540

atctattcta ctgttcttta tcatctgaat gaggaaatga gactaagcta ccttgctcaa 600
gagcttattt ccattgatcg actatctcct caagcatggt gtgcagtggg caattgcttt 660
gccttgagga aagatcatga gactgctttg aagaattttc aacgttcggt acagcttgac 720
tcaagatttg catatgctca cactctatgt ggtcatgagt attctgcatt ggaggattat 780
gagaatagta tcaaattcta ccggtgtgca ctgcaggtag atgaaaggca ctacaatgcc 840
tggtatggcc ttggggtggt gtatcttcgc caggaaaagt ntgagtttgc tgagcatcat 900
ttcagaaggg catttcagat aaatcctcgc tcttctgttc tcatgtgcta tcttgggatg 960
gcgttgcatt ctcttaagag gaaggaggag gcattggaaa tgatggagaa agctatagca 1020
gctgataaga agaatccact gcccaagtat cagaaggcct taatccttct aggtcttcag 1080
aagtatcaag aagctctgga tgagttggag cggctaaagg agattgcacc tcatgagagc 1140
agtatgtatg cactgatggg aaagatttac aagcaactca atatccttga caaagctgtt 1200
ttctgctttg gcattgccct ggatttgaaa cctcctgctg ctgatcttgc tataattaag 1260
tccgcaatgg agaaagtaca tctccctgat gaactgatgg aggatgacct gtaa 1314

<210> 148
<211> 437
<212> PRT
<213> Saccharum officinarum
<220>
<221> INSEGURO
<222> (139)..(139)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> INSEGURO
<222> (294)..(294)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 148

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

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

<210> 149
<211> 54
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm00778
<400> 149
ggggacaagt ttgtacaaaa aagcaggctt cacaatgcaa caactgtcaa cttc 54

<210> 150
<211> 49
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm00779
<400> 150
ggggaccact ttgtacaaga aagctgggtt ggagtagcta tggtttcac 49

<210> 151
<211> 1042
<212> DNA
<213> Oryza sativa
<400> 151
ggggagttag gaaccttgac atacaaccaa tgataccatt ttttttcaag cagctagagg 60
tacagaggtt catatttttt agatggctca ttagtgatat tttagtcaaa aatttcagaa 120
gtacggccac gggtgatggc ctgaactaat attttattcg aggtgccgct acatcatcgt 180
ctaaagtaca cgcaagattc aacggaaaaa agaaaccgat cgatcgagat cagttagtta 240
atgaaataac tagatcaact catgtcgtca aaaacaaaag atgctcatct atggacaaca 300
cacgctgatg attcgatcat caaacaaagg tggtagtagt agtaaagcgt atcgtgtttc 360
tcatcagaaa gaagaattaa agaaaaaact aatcccgtct cgcgagccag agaaaattcc 420
ctacaaaagc cactcctttg atttgacatg caaaagcaag gctccactcc tctactaccc 480
tacaactaca caacactgtc tctctatctc caaaggcagt agctgtattg gcttccagct 540
tttcctctct acctctaatg atagcttgga gcaagttcaa tagtatagct aactactagc 600
tctaattcat ctataatcaa tctaatagct tattcataca atagttatat actacattat 660
taatatctgg tcccatctat catacacact gagtctgtgc tatagctgac tacaaatctg 720
tagcccgctg ctcttctctc ttcatttatc ttcttaaaat atatttgcag ctggcttatg 780
gcttatagct tgatattgag agggagagga gtgagagcta gctcagctca gctcaaggta 840
aacaacaagg cacactcttc ctacctcttc tccggttctt ccttttttcct ctttctcttc 900
gtccaagaac ttcacctcaa tagctcgagc tacggcctaa cttttgcgtt gcgcaggaga 960
gctcgatcgc tgcaccaata cttcactgga gatcgcctag ctgcagctag ctagcttatc 1020
ctgcgtgtct tgagcttctc tc 1042

<210> 152
<211> 951
<212> DNA
<213> Oryza sativa
<400> 152

```
atggatccgg tcacggcatc aatacacggt caccatcttc ctccaccgtt caacacccgc     60
gacttccatc accatctcca gcagcagcag caccagctgc atctcaagac cgaggatgac    120
caaggcggcg gcactccggg tgtcttcggc agccgcggca ccaagcgcga ccacgacgac    180
gacgagaaca gtggcaacgg ccatggaagc ggtggtgacg gcggtgacct cgcgctggta    240
ccccccctcgg gtggcgggcc ggacggcgcc gggagcgaga gcgccacgcg ccgcccgagg    300
ggacgcccgg cggggtccaa gaacaagccg aagccaccga tcatcatcac cagggacagc    360
gccaacacgc tccggacgca tgtcatggag gtggccggcg gctgcgacat ctccgagagc    420
atcaccacgt tcgcgcgacg ccggcagcgc ggggtttgcg tgctcagcgg cgccggcacc    480
gtcactaacg tcacgctgcg gcagcccgca tcgcagggag cggtcgttgc gctccacggc    540
cggttcgaga tactctccct ctccggctcc ttcctcccgc cgcccgcccc gccggaggcc    600
acggggctca ccgtctacct ggccggaggc cagggccagg tcgtgggcgg cagcgtcgtc    660
ggcgcgctga ccgcggctgg gcctgtggtg ataatggcgg cgtcttttgc gaacgcggtg    720
tacgagcggc tgccgttgga ggacgacgag ctactggcgg ctcaagggca agccgacagc    780
gctgggttgc tcgccgcggg gcagcaagcg gcgcagctcg ccggcggggc cgtcgatcca    840
agcctcttcc aaggactacc accaaaccta ctcggaaacg tgcagctgcc gccggaagcc    900
gcctacggat ggaaccctgg cgccggcggt ggccgcccgg cgccgttctg a              951
<210> 153
<211> 316
<212> PRT
<213> Oryza sativa
<400> 153
Met Asp Pro Val Thr Ala Ser Ile His Gly His His Leu Pro Pro Pro
1               5                  10                  15
Phe Asn Thr Arg Asp Phe His His His Leu Gln Gln Gln Gln His Gln
            20                  25                  30
Leu His Leu Lys Thr Glu Asp Asp Gln Gly Gly Gly Thr Pro Gly Val
        35                  40                  45
Phe Gly Ser Arg Gly Thr Lys Arg Asp His Asp Asp Asp Glu Asn Ser
    50                  55                  60
Gly Asn Gly His Gly Ser Gly Gly Asp Gly Gly Asp Leu Ala Leu Val
65                  70                  75                  80
Pro Pro Ser Gly Gly Gly Pro Asp Gly Ala Gly Ser Glu Ser Ala Thr
                85                  90                  95
Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys Pro
            100                 105                 110
Pro Ile Ile Ile Thr Arg Asp Ser Ala Asn Thr Leu Arg Thr His Val
        115                 120                 125
Met Glu Val Ala Gly Gly Cys Asp Ile Ser Glu Ser Ile Thr Thr Phe
    130                 135                 140
Ala Arg Arg Arg Gln Arg Gly Val Cys Val Leu Ser Gly Ala Gly Thr
145                 150                 155                 160
Val Thr Asn Val Thr Leu Arg Gln Pro Ala Ser Gln Gly Ala Val Val
                165                 170                 175
Ala Leu His Gly Arg Phe Glu Ile Leu Ser Leu Ser Gly Ser Phe Leu
            180                 185                 190
Pro Pro Pro Ala Pro Pro Glu Ala Thr Gly Leu Thr Val Tyr Leu Ala
        195                 200                 205
Gly Gly Gln Gly Gln Val Val Gly Gly Ser Val Val Gly Ala Leu Thr
    210                 215                 220
Ala Ala Gly Pro Val Val Ile Met Ala Ala Ser Phe Ala Asn Ala Val
225                 230                 235                 240
Tyr Glu Arg Leu Pro Leu Glu Asp Asp Glu Leu Leu Ala Ala Gln Gly
                245                 250                 255
Gln Ala Asp Ser Ala Gly Leu Leu Ala Ala Gly Gln Gln Ala Ala Gln
            260                 265                 270
Leu Ala Gly Gly Ala Val Asp Pro Ser Leu Phe Gln Gly Leu Pro Pro
        275                 280                 285
Asn Leu Leu Gly Asn Val Gln Leu Pro Pro Glu Ala Ala Tyr Gly Trp
    290                 295                 300
Asn Pro Gly Ala Gly Gly Gly Arg Pro Ala Pro Phe
305                 310                 315
<210> 154
<211> 918
```

<212> DNA
<213> Oryza sativa
<400> 154
atggcaggtc tcgacctcgg caccgccgcg acgcgctacg tccaccagct ccaccacctc 60
caccccgacc tccagctgca gcacagctac gccaagcagc acgagccgtc cgacgacgac 120
cccaacggca gcggcggcgg cggcaacagc aacggcgggc cgtacgggga ccatgacggc 180
gggtcctcgt cgtcaggccc tgccaccgac ggcgcggtcg gcgggcccgg cgacgtggtg 240
gcgcgccggc cgcgggggcg cccgcctggc tccaagaaca agccgaagcc gccggtgatc 300
atcacgcggg agagcgccaa cacgctgcgc gcccacatcc tggaggtcgg gagcggctgc 360
gacgtgttcg agtgcgtctc cacgtacgcg cgccggcggc agcgcggcgt gtgcgtgctg 420
agcggcagcg gcgtggtcac caacgtgacg ctgcgtcagc cgtcggcgcc cgcgggcgcc 480
gtcgtgtcgc tgcacgggag gttcgagatc ctgtcgctct cgggctcctt cctcccgccg 540
ccggctcccc ccggcgccac cagcctcacc atcttcctcg ccgggggcca gggacaggtc 600
gtcggcggca acgtcgtcgg cgcgctctac gccgcgggcc cggtcatcgt catcgcggcg 660
tccttcgcca acgtcgccta cgagcgcctc ccactggagg aggaggaggc gccgccgccg 720
caggccggcc tgcagatgca gcagcccggc ggcggcgccg atgctggtgg catgggtggc 780
gcgttcccgc cggacccgtc tgccgccggc ctcccgttct tcaacctgcc gctcaacaac 840
atgcccggtg gcggcggctc acagctccct cccggcgccg acggccatgg ctgggccggc 900
gcacggccac cgttctga 918
<210> 155
<211> 305
<212> PRT
<213> Oryza sativa
<400> 155
Met Ala Gly Leu Asp Leu Gly Thr Ala Ala Thr Arg Tyr Val His Gln
1 5 10 15
Leu His His Leu His Pro Asp Leu Gln Leu Gln His Ser Tyr Ala Lys
20 25 30
Gln His Glu Pro Ser Asp Asp Asp Pro Asn Gly Ser Gly Gly Gly Gly
35 40 45
Asn Ser Asn Gly Gly Pro Tyr Gly Asp His Asp Gly Gly Ser Ser Ser
50 55 60
Ser Gly Pro Ala Thr Asp Gly Ala Val Gly Gly Pro Gly Asp Val Val
65 70 75 80
Ala Arg Arg Pro Arg Gly Arg Pro Pro Gly Ser Lys Asn Lys Pro Lys
85 90 95
Pro Pro Val Ile Ile Thr Arg Glu Ser Ala Asn Thr Leu Arg Ala His
100 105 110
Ile Leu Glu Val Gly Ser Gly Cys Asp Val Phe Glu Cys Val Ser Thr
115 120 125
Tyr Ala Arg Arg Arg Gln Arg Gly Val Cys Val Leu Ser Gly Ser Gly
130 135 140
Val Val Thr Asn Val Thr Leu Arg Gln Pro Ser Ala Pro Ala Gly Ala
145 150 155 160
Val Val Ser Leu His Gly Arg Phe Glu Ile Leu Ser Leu Ser Gly Ser
165 170 175
Phe Leu Pro Pro Pro Ala Pro Pro Gly Ala Thr Ser Leu Thr Ile Phe
180 185 190
Leu Ala Gly Gly Gln Gly Gln Val Val Gly Gly Asn Val Val Gly Ala
195 200 205
Leu Tyr Ala Ala Gly Pro Val Ile Val Ile Ala Ala Ser Phe Ala Asn
210 215 220
Val Ala Tyr Glu Arg Leu Pro Leu Glu Glu Glu Glu Ala Pro Pro Pro
225 230 235 240
Gln Ala Gly Leu Gln Met Gln Gln Pro Gly Gly Gly Ala Asp Ala Gly
245 250 255
Gly Met Gly Gly Ala Phe Pro Pro Asp Pro Ser Ala Ala Gly Leu Pro
260 265 270
Phe Phe Asn Leu Pro Leu Asn Asn Met Pro Gly Gly Gly Gly Ser Gln
275 280 285
Leu Pro Pro Gly Ala Asp Gly His Gly Trp Ala Gly Ala Arg Pro Pro
290 295 300
Phe

305
<210> 156
<211> 906
<212> DNA
<213> Lotus corniculatus
<400> 156
atggatccat tatcagcaca cggccactct cttcctcctc cttttctcca tctgcaccac 60
caacaccaac accagcagca gcatcagcag catcagttcc actctttaca gcagcagcaa 120
accccagcag aagatgaaca gagtggaagc agcggcggca tcaaaaggga acgcgatgaa 180
aacaacaaca gccatgacgg caaagaaggc tccggaggcg gaggcgaaag cgagaattca 240
agaagacctc gaggaaggcc cgccggatcg aagaacaagc ctaagccgcc catcatcatc 300
acccgcgaca gcgccaacgc acttaagacc cacgtcatgg aggtcgccga cggctgcgac 360
atcgtcgaca gcgtctccaa ctttgcaaga cgccgccagc gcggcgtctg catcatgagc 420
ggcactggaa ccgtcaccaa cgtcactctc aggcagccag cttcttccgg cgctgttgtc 480
accctccacg gaaggtttga gattctctcc ctggcaggat cgttcctgcc gccgcctgct 540
ccaccggcag catcaggttt gaccatttac ttggctggtg gacaagggca ggttgttgga 600
ggaagtgttg tgggagctct cattgcttcg ggacctgtgg ttatcatggc agcttcgttc 660
agcaacgctg cgtatgagag gcttcctttg gaagatgagg acccttcatt ggcaatgcaa 720
ggaggttcaa tgggttcacc acccggtggt agtggtggtg gtggtggagt tggtcagcaa 780
cagcagcaac agcttttagg ggatgcaact gccccacttt ttcatggttt gcctccgaat 840
cttctcaatt ctgttcagat gccaaactcc gataacttct ggccatctgg ccgctctcct 900
tactga 906
<210> 157
<211> 301
<212> PRT
<213> Lotus corniculatus
<400> 157
Met Asp Pro Leu Ser Ala His Gly His Ser Leu Pro Pro Pro Phe Leu
1 5 10 15
His Leu His His Gln His Gln His Gln Gln Gln His Gln Gln His Gln
20 25 30
Phe His Ser Leu Gln Gln Gln Gln Thr Pro Ala Glu Asp Glu Gln Ser
35 40 45
Gly Ser Ser Gly Gly Ile Lys Arg Glu Arg Asp Glu Asn Asn Asn Ser
50 55 60
His Asp Gly Lys Glu Gly Ser Gly Gly Gly Gly Glu Ser Glu Asn Ser
65 70 75 80
Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys Pro
85 90 95
Pro Ile Ile Ile Thr Arg Asp Ser Ala Asn Ala Leu Lys Thr His Val
100 105 110
Met Glu Val Ala Asp Gly Cys Asp Ile Val Asp Ser Val Ser Asn Phe
115 120 125
Ala Arg Arg Arg Gln Arg Gly Val Cys Ile Met Ser Gly Thr Gly Thr
130 135 140
Val Thr Asn Val Thr Leu Arg Gln Pro Ala Ser Ser Gly Ala Val Val
145 150 155 160
Thr Leu His Gly Arg Phe Glu Ile Leu Ser Leu Ala Gly Ser Phe Leu
165 170 175
Pro Pro Pro Ala Pro Pro Ala Ala Ser Gly Leu Thr Ile Tyr Leu Ala
180 185 190
Gly Gly Gln Gly Gln Val Val Gly Gly Ser Val Val Gly Ala Leu Ile
195 200 205
Ala Ser Gly Pro Val Val Ile Met Ala Ala Ser Phe Ser Asn Ala Ala
210 215 220
Tyr Glu Arg Leu Pro Leu Glu Asp Glu Asp Pro Ser Leu Ala Met Gln
225 230 235 240
Gly Gly Ser Met Gly Ser Pro Pro Gly Gly Ser Gly Gly Gly Gly Gly
245 250 255
Val Gly Gln Gln Gln Gln Gln Leu Leu Gly Asp Ala Thr Ala Pro
260 265 270
Leu Phe His Gly Leu Pro Pro Asn Leu Leu Asn Ser Val Gln Met Pro
275 280 285

Asn Ser Asp Asn Phe Trp Pro Ser Gly Arg Ser Pro Tyr
290 295 300

<210> 158
<211> 1020
<212> DNA
<213> Arabidopsis thaliana
<400> 158

atggatccag ttcaatctca tggatcacaa agctctcttc ctcctccttt ccatgctaga 60
gatttccaat tacatcttca acaacaacaa caacatcaac aacaacatca acaacaacaa 120
caacaacagt tctttctcca ccatcatcag caaccacaaa gaaaccttga tcaagatcac 180
gagcagcaag gagggtcaat attgaataga tctatcaaga tggatcgcga agagacaagc 240
gataacatgg acaacatcgc taataccaac agcggtagcg aaggtaaaga gatgagttta 300
cacggaggag aaggaggaag cggtggtgga ggaagtggag aacagatgac aagaaggcca 360
agaggaagac cagcaggatc caagaacaaa cctaaagctc caataatcat aacaagagac 420
agcgcaaacg cgcttcgaac tcacgtcatg gagataggag acggatgtga catagttgac 480
tgtatggcta cgttcgctag acgccgccaa agaggcgttt gcgttatgag cggtacagga 540
agcgttacta acgtcactat acgtcagcct ggatcgccac ctggctcggt ggttagcctt 600
cacggccggt ttgaaatcct ctctctttcg ggatctttct tgcctccgcc tgcgccgcct 660
gcagccaccg gactaagcgt ttacctagcc ggaggacaag ggcaggtcgt tggaggtagt 720
gtggtgggac ctttgttgtg ttcgggtcct gtggtggtta tggcggcttc ttttagcaat 780
gcggcgtacg aaaggctgcc tttggaagaa gatgagatgc agacgccagt tcaaggaggc 840
ggtggaggag gaggaggtgg tggtggaatg ggatctcccc cgatgatggg acagcaacaa 900
gctatggcag ctatggcggc ggctcaagga ctaccaccga atcttcttgg ttcggttcag 960
ttgccaccgc cacaacagaa tgatcagcag tattggtcta cgggtcggcc accgtattga 1020

<210> 159
<211> 339
<212> PRT
<213> Arabidopsis thaliana
<400> 159

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

```
Met Gln Thr Pro Val Gln Gly Gly Gly Gly Gly Gly Gly Gly Gly
        275                 280                 285
Gly Met Gly Ser Pro Pro Met Met Gly Gln Gln Gln Ala Met Ala Ala
    290                 295                 300
Met Ala Ala Ala Gln Gly Leu Pro Pro Asn Leu Leu Gly Ser Val Gln
305                 310                 315                 320
Leu Pro Pro Pro Gln Gln Asn Asp Gln Gln Tyr Trp Ser Thr Gly Arg
                325                 330                 335
Pro Pro Tyr
```

<210> 160
<211> 975
<212> DNA
<213> Arabidopsis thaliana
<400> 160

```
atggatccag tacaatctca tggatcacaa agctctctac ctcctccttt ccacgcaaga     60
gactttcaat tacatcttca acaacagcaa caagagttct tcctccacca tcaccagcaa    120
caaagaaacc aaaccgatgg tgaccaacaa ggaggatcag gaggaaaccg acaaatcaag    180
atggatcgtg aagagacaag cgacaacata gacaacatag ctaacaacag cggtagtgaa    240
ggtaaagaca tagatataca cggtggttca ggagaaggag gtggtggctc cggaggagat    300
catcagatga caagaagacc aagaggaaga ccagcgggat ccaagaacaa accaaaacca    360
ccgattatca tcacacggga cagcgcaaac gcgcttagaa cccacgtgat ggagatcgga    420
gatggctgcg acttagtcga aagcgttgcc acttttgcac gaagacgcca acgcggcgtt    480
tgcgttatga gcggtactgg aaatgttact aacgtcacta tacgtcagcc tggatctcat    540
ccttctcctg gctcggtagt tagtcttcac ggaaggttcg agattctatc tctctcagga    600
tcttttctcc ctcctccggc tcctcctaca gccaccggat tgagtgttta cctcgctgga    660
ggacaaggac aggtggttgg aggaagcgta gttggtccgt tgttatgtgc tggtcctgtc    720
gttgtcatgg ctgcgtcttt tagcaatgcg gcgtacgaaa ggttgccttt agaggaagat    780
gagatgcaga cgccggttca tggcggagga ggaggaggat cattggagtc gccgccaatg    840
atgggacaac aactgcaaca tcagcaacaa gctatgtcag gtcatcaagg gttaccacct    900
aatcttcttg gttcggttca gttgcagcag caacatgatc agtcttattg gtcaacggga    960
cgaccaccgt attga                                                     975
```

<210> 161
<211> 324
<212> PRT
<213> Arabidopsis thaliana
<400> 161

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

```
Pro Thr Ala Thr Gly Leu Ser Val Tyr Leu Ala Gly Gly Gln Gly Gln
    210                 215                 220
Val Val Gly Gly Ser Val Val Gly Pro Leu Leu Cys Ala Gly Pro Val
225                 230                 235                 240
Val Val Met Ala Ala Ser Phe Ser Asn Ala Ala Tyr Glu Arg Leu Pro
                245                 250                 255
Leu Glu Glu Asp Glu Met Gln Thr Pro Val His Gly Gly Gly Gly Gly
            260                 265                 270
Gly Ser Leu Glu Ser Pro Pro Met Met Gly Gln Gln Leu Gln His Gln
        275                 280                 285
Gln Gln Ala Met Ser Gly His Gln Gly Leu Pro Pro Asn Leu Leu Gly
    290                 295                 300
Ser Val Gln Leu Gln Gln Gln His Asp Gln Ser Tyr Trp Ser Thr Gly
305                 310                 315                 320
Arg Pro Pro Tyr
```

<210> 162
<211> 954
<212> DNA
<213> Arabidopsis thaliana
<400> 162

```
atggatcagg tctctcgctc tcttcctcca ccttttctct caagagatct ccatcttcac     60
ccacaccatc aattccagca tcagcagcag cagcagcaac agaatcacgg ccacgatata    120
gaccagcacc gaatcggtgg gctaaaacgt gaccgagatg ctgatatcga tcccaacgag    180
cactcttcag ccggaaaaga tcaaagtact cctggctccg gtggagaaag cggcggcgga    240
ggaggaggag ataatcacat cacgagaagg ccacgtggca gaccagcggg atctaagaac    300
aaaccaaaac cgccaatcat catcactcga gacagcgcaa acgctctcaa atctcatgtc    360
atggaagtag caaacggatg tgacgtcatg gaaagtgtca ccgtcttcgc tcgccgtcgc    420
caacgtggca tctgcgtttt gagcggaaac ggcgccgtta ccaacgttac cataagacaa    480
ccagcttcag tacctggtgg tggctcatct gtcgttaact tacacggacg tttcgagatt    540
ctttctctct cgggatcatt ccttcctcct ccggctccac cagctgcgtc aggtctaacg    600
atttacttag ccggtggtca gggacaggtt gttggaggaa gcgtggttgg tccactcatg    660
gcttcaggac ctgtagtgat tatggcagct tcgtttggaa acgctgcgta tgagagactg    720
ccgttggagg aagacgatca agaagagcaa acagctggag cggttgctaa taatatcgat    780
ggaaacgcaa caatgggtgg tggaacgcaa acgcaaactc agacgcagca gcaacagcaa    840
caacagttga tgcaagatcc gacgtcgttt atacaagggt tgcctccgaa tcttatgaat    900
tctgttcaat tgccagctga agcttattgg ggaactccga gaccatcttt ctaa          954
```

<210> 163
<211> 317
<212> PRT
<213> Arabidopsis thaliana
<400> 163

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

```
                165                 170                 175
Arg Phe Glu Ile Leu Ser Leu Ser Gly Ser Phe Leu Pro Pro Pro Ala
            180                 185                 190
Pro Pro Ala Ala Ser Gly Leu Thr Ile Tyr Leu Ala Gly Gly Gln Gly
        195                 200                 205
Gln Val Val Gly Gly Ser Val Val Gly Pro Leu Met Ala Ser Gly Pro
    210                 215                 220
Val Val Ile Met Ala Ala Ser Phe Gly Asn Ala Ala Tyr Glu Arg Leu
225                 230                 235                 240
Pro Leu Glu Glu Asp Asp Gln Glu Glu Gln Thr Ala Gly Ala Val Ala
                245                 250                 255
Asn Asn Ile Asp Gly Asn Ala Thr Met Gly Gly Gly Thr Gln Thr Gln
            260                 265                 270
Thr Gln Thr Gln Gln Gln Gln Gln Gln Gln Leu Met Gln Asp Pro Thr
        275                 280                 285
Ser Phe Ile Gln Gly Leu Pro Pro Asn Leu Met Asn Ser Val Gln Leu
    290                 295                 300
Pro Ala Glu Ala Tyr Trp Gly Thr Pro Arg Pro Ser Phe
305                 310                 315
<210> 164
<211> 798
<212> DNA
<213> Arabidopsis thaliana
<400> 164
atggatgagg tatctcgttc tcatacaccg caatttctat caagtgatca tcagcactat     60
caccatcaaa acgctggacg acaaaaacgc ggcagagaag aagaaggagt tgaacccaac    120
aatatagggg aagacctagc cacctttcct tccggagaag agaatatcaa gaagagaagg    180
ccacgtggca gacctgctgg ttccaagaac aaacccaaag caccaatcat agtcactcgc    240
gactccgcga acgccttcag atgtcacgtc atggagataa ccaacgcctg cgatgtaatg    300
gaaagcctag ccgtcttcgc tagacgccgt cagcgtggcg tttgcgtctt gaccggaaac    360
ggggccgtta caaacgtcac cgttagacaa cctggcggag gcgtcgtcag tttacacgga    420
cggtttgaga ttctttctct ctcgggttcg tttcttcctc caccggcacc accagctgcg    480
tctggtttaa aggtttactt agccggtggt caaggtcaag tgatcggagg cagtgtggtg    540
ggaccgctta cggcatcaag tccggtggtc gttatggcag cttcatttgg aaacgcatct    600
tacgagaggc tgccactaga ggaggaggag gaaactgaaa gagaaataga tggaaacgcg    660
gctagggcga ttggaacgca aacgcagaaa cagttaatgc aagatgcgac atcgtttatt    720
gggtcgccgt cgaatttaat taactctgtt tcgttgccag gtgaagctta ttggggaacg    780
caacgaccgt ctttctaa                                                  798
<210> 165
<211> 265
<212> PRT
<213> Arabidopsis thaliana
<400> 165
Met Asp Glu Val Ser Arg Ser His Thr Pro Gln Phe Leu Ser Ser Asp
1               5                   10                  15
His Gln His Tyr His His Gln Asn Ala Gly Arg Gln Lys Arg Gly Arg
            20                  25                  30
Glu Glu Glu Gly Val Glu Pro Asn Asn Ile Gly Glu Asp Leu Ala Thr
        35                  40                  45
Phe Pro Ser Gly Glu Glu Asn Ile Lys Lys Arg Arg Pro Arg Gly Arg
    50                  55                  60
Pro Ala Gly Ser Lys Asn Lys Pro Lys Ala Pro Ile Ile Val Thr Arg
65                  70                  75                  80
Asp Ser Ala Asn Ala Phe Arg Cys His Val Met Glu Ile Thr Asn Ala
                85                  90                  95
Cys Asp Val Met Glu Ser Leu Ala Val Phe Ala Arg Arg Arg Gln Arg
            100                 105                 110
Gly Val Cys Val Leu Thr Gly Asn Gly Ala Val Thr Asn Val Thr Val
        115                 120                 125
Arg Gln Pro Gly Gly Gly Val Val Ser Leu His Gly Arg Phe Glu Ile
    130                 135                 140
Leu Ser Leu Ser Gly Ser Phe Leu Pro Pro Pro Ala Pro Pro Ala Ala
145                 150                 155                 160
```

Ser Gly Leu Lys Val Tyr Leu Ala Gly Gly Gln Gly Gln Val Ile Gly
165 170 175
Gly Ser Val Val Gly Pro Leu Thr Ala Ser Ser Pro Val Val Val Met
180 185 190
Ala Ala Ser Phe Gly Asn Ala Ser Tyr Glu Arg Leu Pro Leu Glu Glu
195 200 205
Glu Glu Glu Thr Glu Arg Glu Ile Asp Gly Asn Ala Ala Arg Ala Ile
210 215 220
Gly Thr Gln Thr Gln Lys Gln Leu Met Gln Asp Ala Thr Ser Phe Ile
225 230 235 240
Gly Ser Pro Ser Asn Leu Ile Asn Ser Val Ser Leu Pro Gly Glu Ala
245 250 255
Tyr Trp Gly Thr Gln Arg Pro Ser Phe
260 265

<210> 166
<211> 933
<212> DNA
<213> Arabidopsis thaliana
<400> 166

atggcgaatc cttggtgggt agggaatgtt gcgatcggtg gagttgagag tccagtgacg 60
tcatcagctc cttctttgca ccacagaaac agtaacaaca acaacccacc gactatgact 120
cgttcggatc caagattgga ccatgacttc accaccaaca acagtggaag ccctaatacc 180
cagactcaga gccaagaaga acagaacagc agagacgagc aaccagctgt tgaacccgga 240
tccggatccg ggtctacggg tcgtcgtcct agaggtagac ctcctggttc caagaacaaa 300
ccaaagagtc cagttgttgt taccaaagaa agccctaact ctctccagag ccatgttctt 360
gagattgcta cgggagctga cgtggcggaa agcttaaacg cctttgctcg tagacgcggc 420
cggggcgttt cggtgctgag cggtagtggt ttggttacta atgttactct gcgtcagcct 480
gctgcatccg gtggagttgt tagtttacgt ggtcagtttg agatcttgtc tatgtgtggg 540
gcttttcttc ctacgtctgg ctctcctgct gcagccgctg gtttaaccat ttacttagct 600
ggagctcaag gtcaagttgt gggaggtgga gttgctggcc cgcttattgc ctctggaccc 660
gttattgtga tagctgctac gttttgcaat gccacttatg agaggttacc gattgaggaa 720
gaacaacagc aagagcagcc gcttcaacta gaagatggga agaagcagaa agaagagaat 780
gatgataacg agagtgggaa taacggaaac gaaggatcga tgcagccgcc gatgtataat 840
atgcctccta attttatccc aaatggtcat caaatggctc aacacgacgt gtattggggt 900
ggtcctccgc ctcgtgctcc tccttcgtat tga 933

<210> 167
<211> 310
<212> PRT
<213> Arabidopsis thaliana
<400> 167

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

```
            180                 185                 190
Ala Gly Leu Thr Ile Tyr Leu Ala Gly Ala Gln Gly Gln Val Val Gly
        195                 200                 205
Gly Gly Val Ala Gly Pro Leu Ile Ala Ser Gly Pro Val Ile Val Ile
    210                 215                 220
Ala Ala Thr Phe Cys Asn Ala Thr Tyr Glu Arg Leu Pro Ile Glu Glu
225                 230                 235                 240
Glu Gln Gln Gln Glu Gln Pro Leu Gln Leu Glu Asp Gly Lys Lys Gln
                245                 250                 255
Lys Glu Glu Asn Asp Asp Asn Glu Ser Gly Asn Asn Gly Asn Glu Gly
            260                 265                 270
Ser Met Gln Pro Pro Met Tyr Asn Met Pro Pro Asn Phe Ile Pro Asn
        275                 280                 285
Gly His Gln Met Ala Gln His Asp Val Tyr Trp Gly Gly Pro Pro Pro
    290                 295                 300
Arg Ala Pro Pro Ser Tyr
305                 310
<210> 168
<211> 987
<212> DNA
<213> Oryza sativa
<400> 168
atggatccgg tgacggcggc ggcggcgcat gggggtgggc accaccacca ccaccacttc     60
ggagcgccac cggtggcggc gttccaccac cacccgttcc accacggcgg cggggcgcac    120
tacccggcgg cgttccagca gtttcaggag gagcagcagc agcttgtggc ggcggcggcg    180
gcggctggtg ggatggcgaa gcaggagctg gtggatgaga gcaacaacac catcaacagc    240
ggcgggagca acgggagcgg cggggaggag cagaggcagc agtccgggga ggagcagcac    300
cagcaagggg cggcggcgcc ggtggtgatc cggcgtccca ggggccgccc cgccggctcc    360
aagaacaagc ccaagcctcc ggtcatcatc acgcgcgaca gcgccagcgc gctgcgggcg    420
cacgtcctcg aggtcgcctc cgggtgcgac ctcgtcgaca gcgtcgccac gttcgcgcgc    480
cgccgccagg tcggtgtctg cgtgctcagc gccaccggcg ccgtcaccaa cgtctccgtc    540
cggcagcccg gcgcgggccc cggcgccgtc gtcaacctca ccggccgctt cgacatcctc    600
tcgctgtccg gctccttcct cccgccgccg gcgcctccct ccgccaccgg cctcaccgtc    660
tacgtctccg gcggccaggg gcaggtcgtg ggcggcacgg tcgccggacc gctcatcgcc    720
gtcggccccg tcgtcatcat ggccgcctcg ttcgggaacg ccgcctacga gcgcctcccg    780
ctcgaggacg acgagccgcc gcagcacatg gcgggcggcg gccagtcctc gccgccgccg    840
ccgccgctgc cattaccacc acaccagcag ccgattcttc aagaccatct gccacacaac    900
ctgatgaacg gaatccacct ccccggcgac gccgcctacg gctggaccag cggcggcggc    960
ggcggcggcc gcgcggcgcc gtactga                                        987
<210> 169
<211> 328
<212> PRT
<213> Oryza sativa
<400> 169
Met Asp Pro Val Thr Ala Ala Ala Ala His Gly Gly Gly His His His
1               5                   10                  15
His His His Phe Gly Ala Pro Pro Val Ala Ala Phe His His His Pro
            20                  25                  30
Phe His His Gly Gly Gly Ala His Tyr Pro Ala Ala Phe Gln Gln Phe
        35                  40                  45
Gln Glu Glu Gln Gln Gln Leu Val Ala Ala Ala Ala Ala Ala Gly Gly
    50                  55                  60
Met Ala Lys Gln Glu Leu Val Asp Glu Ser Asn Asn Thr Ile Asn Ser
65                  70                  75                  80
Gly Gly Ser Asn Gly Ser Gly Gly Glu Glu Gln Arg Gln Gln Ser Gly
                85                  90                  95
Glu Glu Gln His Gln Gln Gly Ala Ala Ala Pro Val Val Ile Arg Arg
            100                 105                 110
Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys Pro Pro Val
        115                 120                 125
Ile Ile Thr Arg Asp Ser Ala Ser Ala Leu Arg Ala His Val Leu Glu
    130                 135                 140
Val Ala Ser Gly Cys Asp Leu Val Asp Ser Val Ala Thr Phe Ala Arg
```

```
145                 150                 155                 160
Arg Arg Gln Val Gly Val Cys Val Leu Ser Ala Thr Gly Ala Val Thr
                165                 170                 175
Asn Val Ser Val Arg Gln Pro Gly Ala Gly Pro Gly Ala Val Val Asn
            180                 185                 190
Leu Thr Gly Arg Phe Asp Ile Leu Ser Leu Ser Gly Ser Phe Leu Pro
        195                 200                 205
Pro Pro Ala Pro Pro Ser Ala Thr Gly Leu Thr Val Tyr Val Ser Gly
    210                 215                 220
Gly Gln Gly Gln Val Val Gly Gly Thr Val Ala Gly Pro Leu Ile Ala
225                 230                 235                 240
Val Gly Pro Val Val Ile Met Ala Ala Ser Phe Gly Asn Ala Ala Tyr
                245                 250                 255
Glu Arg Leu Pro Leu Glu Asp Asp Glu Pro Pro Gln His Met Ala Gly
            260                 265                 270
Gly Gly Gln Ser Ser Pro Pro Pro Pro Pro Leu Pro Leu Pro Pro His
        275                 280                 285
Gln Gln Pro Ile Leu Gln Asp His Leu Pro His Asn Leu Met Asn Gly
    290                 295                 300
Ile His Leu Pro Gly Asp Ala Ala Tyr Gly Trp Thr Ser Gly Gly Gly
305                 310                 315                 320
Gly Gly Gly Arg Ala Ala Pro Tyr
                325
<210> 170
<211> 777
<212> DNA
<213> Oryza sativa
<400> 170
atggggagca tcgacggcca ctcgctgcag cagcatcagg ggtactccca cggcggcggc     60
gcgggaggga gcaacgagga ggaggaggcg tcgccgccgc ccggcggtgg ctcggctacg    120
gggtcggcgg gccgccggcc gaggggggagg ccgccgggct ccaagaacaa gccgaagccg    180
cccgtcgtgg tgacgcggga gagccccaac gcgatgcgtt cccacgtgct ggagatcgcc    240
agcggcgccg acatcgtcga ggccatcgcg ggcttctccc gccgcaggca gcgcggcgtc    300
tccgtgctca gcgggagcgg cgccgtcacc aacgtcacgc tccggcagcc cgcggggact    360
ggggccgccg ccgtcgcgct gcgggggagg ttcgagatat tgtccatgtc tggcgccttc    420
ctcccggcgc cggcgccgcc aggggccacg gggctcgccg tgtacctcgc cggcgggcag    480
gggcaggtgg tgggtgggag cgtcatgggg gagctgatcg cgtcgggccc cgtcatggtg    540
atcgcggcca cgttcggcaa cgccacgtac gagaggctgc cgctggacca ggaaggcgag    600
gagggcgccg tgctgtccgg gtcggagggc gccgccgcgc agatggagca gcagagcagc    660
ggaggcgccg tcgtgccccc gccgatgtac gccgccgtcc agcagacgcc gccgcacgac    720
atgttcgggc agtgggggca tgcagcggtg gctcggccgc cgccgacatc gttctag       777
<210> 171
<211> 258
<212> PRT
<213> Oryza sativa
<400> 171
Met Gly Ser Ile Asp Gly His Ser Leu Gln Gln His Gln Gly Tyr Ser
1                   5                   10                  15
His Gly Gly Gly Ala Gly Gly Ser Asn Glu Glu Glu Glu Ala Ser Pro
            20                  25                  30
Pro Pro Gly Gly Gly Ser Ala Thr Gly Ser Ala Gly Arg Arg Pro Arg
        35                  40                  45
Gly Arg Pro Pro Gly Ser Lys Asn Lys Pro Lys Pro Pro Val Val Val
    50                  55                  60
Thr Arg Glu Ser Pro Asn Ala Met Arg Ser His Val Leu Glu Ile Ala
65                  70                  75                  80
Ser Gly Ala Asp Ile Val Glu Ala Ile Ala Gly Phe Ser Arg Arg Arg
                85                  90                  95
Gln Arg Gly Val Ser Val Leu Ser Gly Ser Gly Ala Val Thr Asn Val
            100                 105                 110
Thr Leu Arg Gln Pro Ala Gly Thr Gly Ala Ala Ala Val Ala Leu Arg
        115                 120                 125
Gly Arg Phe Glu Ile Leu Ser Met Ser Gly Ala Phe Leu Pro Ala Pro
```

```
    130                 135                 140
Ala Pro Pro Gly Ala Thr Gly Leu Ala Val Tyr Leu Ala Gly Gly Gln
145                 150                 155                 160
Gly Gln Val Val Gly Gly Ser Val Met Gly Glu Leu Ile Ala Ser Gly
                165                 170                 175
Pro Val Met Val Ile Ala Ala Thr Phe Gly Asn Ala Thr Tyr Glu Arg
            180                 185                 190
Leu Pro Leu Asp Gln Glu Gly Glu Glu Gly Ala Val Leu Ser Gly Ser
        195                 200                 205
Glu Gly Ala Ala Ala Gln Met Glu Gln Gln Ser Ser Gly Gly Ala Val
    210                 215                 220
Val Pro Pro Pro Met Tyr Ala Ala Val Gln Gln Thr Pro Pro His Asp
225                 230                 235                 240
Met Phe Gly Gln Trp Gly His Ala Ala Val Ala Arg Pro Pro Pro Thr
                245                 250                 255
Ser Phe
```

<210> 172
<211> 798
<212> DNA
<213> Lycopersicon esculentum
<400> 172

```
atggctggtt tggacttagg ttccgcttct catcatcgtt ttctccctcg tcatctccac     60
gatcctcaag acgatgaaat caatcgcaac aacactcaat tttccgatga tgacaacaac    120
aacaacgata acaataacaa taatagcccc ggggtagcac gtagacctag aggtcgtccc    180
gcggggtcca aaaataagcc taagcctccc gtgatcataa cacgggagag cgctaacgcg    240
ctccgtgcgc atattttaga agtgagtagc ggacatgatg tctttgaatc agttgctact    300
tatgctagaa aaagacaaag aggaatttgc atactgagcg ggagcggtac ggtgaataac    360
gtcaccatac ggcagccaca ggctgccggt tctgtggtga cgttacacgg aagattcgag    420
atattatctt tatccggatc tttcctacca ccacccgctc cccctggggc caccagctta    480
acgatttatt tagcgggtgg tcaaggtcaa gttgttggtg gaaacgttgt gggtgcgcta    540
attgcatcag gaccggttat tgttattgct tcgtcattta ctaatgttgc ttatgagaga    600
ttgcctttgg atgaagaaaa tgagtcaatt cagatgcagc aacaaggaca aagtggtaat    660
tttgctgatc catctaatat tggattacct tttcttaatt tgccattaaa catgccaaat    720
ggtggtggtc aactacaatt ggaaagtgga ggaggtgaag gttggaatgg aaacacaaca    780
aacaggccac aatattag                                                  798
```

<210> 173
<211> 265
<212> PRT
<213> Lycopersicon esculentum
<400> 173

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

```
Val Gly Ala Leu Ile Ala Ser Gly Pro Val Ile Val Ile Ala Ser Ser
            180                 185                 190
Phe Thr Asn Val Ala Tyr Glu Arg Leu Pro Leu Asp Glu Glu Asn Glu
        195                 200                 205
Ser Ile Gln Met Gln Gln Gln Gly Gln Ser Gly Asn Phe Ala Asp Pro
    210                 215                 220
Ser Asn Ile Gly Leu Pro Phe Leu Asn Leu Pro Leu Asn Met Pro Asn
225                 230                 235                 240
Gly Gly Gly Gln Leu Gln Leu Glu Ser Gly Gly Gly Glu Gly Trp Asn
                245                 250                 255
Gly Asn Thr Thr Asn Arg Pro Gln Tyr
            260                 265
```

<210> 174
<211> 813
<212> DNA
<213> Arabidopsis thaliana
<400> 174

```
atggaactta acagatctga agcagacgaa gcaaaggccg agaccactcc caccggtgga     60
gccaccagct cagccacagc ctctggctct tcctccggac gtcgtccacg tggtcgtcct    120
gcaggttcca aaaacaaacc caaacctccg acgattataa ctagagatag tcctaacgtc    180
cttagatcac acgttcttga agtcacctcc ggttcggaca tatccgaggc agtctccacc    240
tacgccactc gtcgcggctg cggcgtttgc attataagcg gcacgggtgc ggtcactaac    300
gtcacgatac ggcaacctgc ggctccggct ggtggaggtg tgattaccct gcatggtcgg    360
tttgacattt tgtctttgac cggtactgcg cttccaccgc ctgcaccacc gggagcagga    420
ggtttgacgg tgtatctagc cggaggtcaa ggacaagttg taggagggaa tgtggctggt    480
tcgttaattg cttcgggacc ggtagtgttg atggctgctt cttttgcaaa cgcagtttat    540
gataggttac cgattgaaga ggaagaaacc ccaccgccga gaaccaccgg ggtgcagcag    600
cagcagccgg aggcgtctca gtcgtcggag gttacgggga gtggggccca ggcgtgtgag    660
tcaaacctcc aaggtggaaa tggtggagga ggtgttgctt tctacaatct tggaatgaat    720
atgaacaatt ttcaattctc cgggggagat atttacggta tgagcggcgg tagcggagga    780
ggtggtggcg gtgcgactag acccgcgttt tag                                 813
```

<210> 175
<211> 270
<212> PRT
<213> Arabidopsis thaliana
<400> 175

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

```
    210                 215                 220
Gly Gly Asn Gly Gly Gly Gly Val Ala Phe Tyr Asn Leu Gly Met Asn
225                 230                 235                 240
Met Asn Asn Phe Gln Phe Ser Gly Gly Asp Ile Tyr Gly Met Ser Gly
                245                 250                 255
Gly Ser Gly Gly Gly Gly Gly Gly Ala Thr Arg Pro Ala Phe
            260                 265                 270
<210> 176
<211> 948
<212> DNA
<213> Arabidopsis thaliana
<400> 176
atggcgaatc catggtggac aggacaagtg aacctatccg gcctcgaaac gacgccgcct     60
ggttcctctc agttaaagaa accagatctc cacatctcca tgaacatggc catggactca    120
ggtcacaata atcatcacca tcaccaagaa gtcgataaca acaacaacga cgacgataga    180
gacaacttga gtggagacga ccacgagcca cgtgaaggag ccgtagaagc ccccacgcgc    240
cgtccacgtg gacgtcctgc tggttccaag aacaaaccaa agccaccgat cttcgtcact    300
cgcgattctc caaatgctct caagagccat gtcatggaga tcgctagtgg gactgacgtc    360
atcgaaaccc tagctacttt tgctaggcgg cgtcaacgtg gcatctgcat cttgagcgga    420
aatggcacag tggctaacgt caccctccgt caaccctcga ccgctgccgt tgcggcggct    480
cctggtggtg cggctgtttt ggctttacaa gggaggtttg agattctttc tttaaccggt    540
tctttcttgc caggaccggc tccacctggt tccaccggtt taacgattta cttagccggt    600
ggtcaaggtc aggttgttgg aggaagcgtg gtgggcccat tgatggcagc aggtccggtg    660
atgctgatcg ccgccacgtt ctctaacgcg acttacgaga gattgccatt ggaggaggaa    720
gaggcagcag agagaggcgg tggtggaggc agcggaggag tggttccggg gcagctcgga    780
ggcggaggtt cgccactaag cagcggtgct ggtggaggcg acggtaacca aggacttccg    840
gtgtataata tgccgggaaa tcttgtttct aatggtggca gtggtggagg aggacagatg    900
agcggccaag aagcttatgg ttgggctcaa gctaggtcag gattttaa                 948
<210> 177
<211> 315
<212> PRT
<213> Arabidopsis thaliana
<400> 177
Met Ala Asn Pro Trp Trp Thr Gly Gln Val Asn Leu Ser Gly Leu Glu
1               5                   10                  15
Thr Thr Pro Pro Gly Ser Ser Gln Leu Lys Lys Pro Asp Leu His Ile
            20                  25                  30
Ser Met Asn Met Ala Met Asp Ser Gly His Asn Asn His His His His
        35                  40                  45
Gln Glu Val Asp Asn Asn Asn Asn Asp Asp Asp Arg Asp Asn Leu Ser
    50                  55                  60
Gly Asp Asp His Glu Pro Arg Glu Gly Ala Val Glu Ala Pro Thr Arg
65                  70                  75                  80
Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys Pro Pro
                85                  90                  95
Ile Phe Val Thr Arg Asp Ser Pro Asn Ala Leu Lys Ser His Val Met
            100                 105                 110
Glu Ile Ala Ser Gly Thr Asp Val Ile Glu Thr Leu Ala Thr Phe Ala
        115                 120                 125
Arg Arg Arg Gln Arg Gly Ile Cys Ile Leu Ser Gly Asn Gly Thr Val
    130                 135                 140
Ala Asn Val Thr Leu Arg Gln Pro Ser Thr Ala Ala Val Ala Ala Ala
145                 150                 155                 160
Pro Gly Gly Ala Ala Val Leu Ala Leu Gln Gly Arg Phe Glu Ile Leu
                165                 170                 175
Ser Leu Thr Gly Ser Phe Leu Pro Gly Pro Ala Pro Pro Gly Ser Thr
            180                 185                 190
Gly Leu Thr Ile Tyr Leu Ala Gly Gly Gln Gly Gln Val Val Gly Gly
        195                 200                 205
Ser Val Val Gly Pro Leu Met Ala Ala Gly Pro Val Met Leu Ile Ala
    210                 215                 220
Ala Thr Phe Ser Asn Ala Thr Tyr Glu Arg Leu Pro Leu Glu Glu Glu
225                 230                 235                 240
```

Glu Ala Ala Glu Arg Gly Gly Gly Gly Gly Ser Gly Gly Val Val Pro
245 250 255
Gly Gln Leu Gly Gly Gly Gly Ser Pro Leu Ser Ser Gly Ala Gly Gly
260 265 270
Gly Asp Gly Asn Gln Gly Leu Pro Val Tyr Asn Met Pro Gly Asn Leu
275 280 285
Val Ser Asn Gly Gly Ser Gly Gly Gly Gly Gln Met Ser Gly Gln Glu
290 295 300
Ala Tyr Gly Trp Ala Gln Ala Arg Ser Gly Phe
305 310 315

<210> 178
<211> 708
<212> DNA
<213> Oryza sativa
<400> 178

atggcgtcca aggagccaag cggcgaccac gaccacgaga tgaacgggac cagcgccggg 60
ggcggcgagc ccaaggacgg cgcggtggtg accggccgca accggcgccc ccgcggacgg 120
ccgccgggct ccaagaacaa gcccaagccg cccatcttcg tgacgcggga cagcccgaac 180
gcgctgcgca gccacgtcat ggaggtggcc ggcggcgccg atgtcgccga gtccatcgcg 240
cacttcgcgc ggcggcggca gcgcggcgtc tgcgtgctca gcggggccgg caccgtgacc 300
gacgtggccc tgcgccagcc ggccgcgccg agcgccgtgg tggcgctccg tgggcggttc 360
gagatcctgt ccctgacggg gacgttcctg ccggggccgg cgccgccggg ctccaccggg 420
ctgaccgtgt acctcgccgg cgggcagggg caggtggtgg gcggcagcgt ggtggggacg 480
ctcaccgcgg cggggccggt catggtgatc gcctccacct tcgccaacgc cacctacgag 540
aggctgccgc tggatcagga ggaggaggaa gcagcggcag gcggcatgat ggcgccgccg 600
ccactcatgg ccggcgccgc cgatccacta cttttcggcg ggggaatgca cgacgccggg 660
cttgctgcat ggcaccatgc ccgccctccg ccgccgccgc cctactag 708

<210> 179
<211> 235
<212> PRT
<213> Oryza sativa
<400> 179

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

<210> 180

<211> 846
<212> DNA
<213> Arabidopsis thaliana
<400> 180
atggcaaacc cttggtggac gaaccagagt ggtttagcgg gcatggtgga ccattcggtc 60
tcctcaggcc atcaccaaaa ccatcaccac caaagtcttc ttaccaaagg agatcttgga 120
atagccatga atcagagcca agacaacgac caagacgaag aagatgatcc tagagaagga 180
gccgttgagg tggtcaaccg tagaccaaga ggtagaccac caggatccaa aaacaaaccc 240
aaagctccaa tctttgtgac aagagacagc cccaacgcac tccgtagcca tgtcttggag 300
atctccgacg gcagtgacgt cgccgacaca atcgctcact tctcaagacg caggcaacgc 360
ggcgtttgcg ttctcagcgg gacaggctca gtcgctaacg tcaccctccg ccaagccgcc 420
gcaccaggag gtgtggtctc tctccaaggc aggtttgaaa tcttatcttt aaccggtgct 480
ttcctccctg gaccttcccc acccgggtca accggtttaa cggtttactt agccggggtc 540
cagggtcagg tcgttggagg tagcgttgta ggcccactct tagccatagg gtcggtcatg 600
gtgattgctg ctactttctc taacgctact tatgagagat tgcccatgga agaagaggaa 660
gacggtggcg gctcaagaca gattcacgga ggcggtgact caccgcccag aatcggtagt 720
aacctgcctg atctatcagg gatggccggg ccaggctaca atatgccgcc gcatctgatt 780
ccaaatgggg ctggtcagct agggcacgaa ccatatacat gggtccacgc aagaccacct 840
tactga 846
<210> 181
<211> 281
<212> PRT
<213> Arabidopsis thaliana
<400> 181
Met Ala Asn Pro Trp Trp Thr Asn Gln Ser Gly Leu Ala Gly Met Val
1 5 10 15
Asp His Ser Val Ser Ser Gly His His Gln Asn His His His Gln Ser
20 25 30
Leu Leu Thr Lys Gly Asp Leu Gly Ile Ala Met Asn Gln Ser Gln Asp
35 40 45
Asn Asp Gln Asp Glu Glu Asp Asp Pro Arg Glu Gly Ala Val Glu Val
50 55 60
Val Asn Arg Arg Pro Arg Gly Arg Pro Pro Gly Ser Lys Asn Lys Pro
65 70 75 80
Lys Ala Pro Ile Phe Val Thr Arg Asp Ser Pro Asn Ala Leu Arg Ser
85 90 95
His Val Leu Glu Ile Ser Asp Gly Ser Asp Val Ala Asp Thr Ile Ala
100 105 110
His Phe Ser Arg Arg Arg Gln Arg Gly Val Cys Val Leu Ser Gly Thr
115 120 125
Gly Ser Val Ala Asn Val Thr Leu Arg Gln Ala Ala Ala Pro Gly Gly
130 135 140
Val Val Ser Leu Gln Gly Arg Phe Glu Ile Leu Ser Leu Thr Gly Ala
145 150 155 160
Phe Leu Pro Gly Pro Ser Pro Pro Gly Ser Thr Gly Leu Thr Val Tyr
165 170 175
Leu Ala Gly Val Gln Gly Gln Val Val Gly Gly Ser Val Val Gly Pro
180 185 190
Leu Leu Ala Ile Gly Ser Val Met Val Ile Ala Ala Thr Phe Ser Asn
195 200 205
Ala Thr Tyr Glu Arg Leu Pro Met Glu Glu Glu Glu Asp Gly Gly Gly
210 215 220
Ser Arg Gln Ile His Gly Gly Gly Asp Ser Pro Pro Arg Ile Gly Ser
225 230 235 240
Asn Leu Pro Asp Leu Ser Gly Met Ala Gly Pro Gly Tyr Asn Met Pro
245 250 255
Pro His Leu Ile Pro Asn Gly Ala Gly Gln Leu Gly His Glu Pro Tyr
260 265 270
Thr Trp Val His Ala Arg Pro Pro Tyr
275 280
<210> 182
<211> 879
<212> DNA

<213> Arabidopsis thaliana
<400> 182
atggctggtc ttgatctagg cacagctttt cgttacgtta atcaccagct ccatcgtccc 60
gatctccacc ttcaccacaa ttcctcctcc gatgacgtca ctcccggagc cgggatgggt 120
catttcaccg tcgacgacga agacaacaac aacaaccatc aaggtcttga cttagcctct 180
ggtggaggat caggaagctc tggaggagga ggaggtcacg gcgggggagg agacgtcgtt 240
ggtcgtcgtc cacgtggcag accaccggga tccaagaaca aaccgaaacc tccggtaatt 300
atcacgcgcg agagcgcaaa cactctaaga gctcacattc ttgaagtaac aaacggctgc 360
gatgttttcg actgcgttgc gacttatgct cgtcggagac agcgagggat ctgcgttctg 420
agcggtagcg gaacggtcac gaacgtcagc atacgtcagc catctgcggc tggagcggtt 480
gtgacgctac aaggaacgtt cgagattctt tctctctccg gatcgtttct tcctcctccg 540
gcacctcccg gagcaacgag tttgacaatt ttcttagccg gaggacaagg tcaggtggtt 600
ggaggaagcg ttgtgggtga gcttacggcg gctggaccgg tgattgtgat tgcagcttcg 660
tttactaatg ttgcttatga gagacttcct ttagaagaag atgagcagca gcaacagctt 720
ggaggaggat ctaacggcgg aggtaatttg tttccggagg tggcagctgg aggaggagga 780
ggacttccgt tctttaattt accgatgaat atgcaaccaa atgtgcaact tccggtggaa 840
ggttggccgg ggaattccgg tggaagaggt cctttctga 879

<210> 183
<211> 292
<212> PRT
<213> Arabidopsis thaliana
<400> 183

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

<210> 184
<211> 918
<212> DNA

<213> Oryza sativa
<400> 184
atggcaggtc tcgacctcgg caccgccgcg acgcgctacg tccaccagct ccaccacctc 60
caccccgacc tccagctgca gcacagctac gccaagcagc acgagccgtc cgacgacgac 120
cccaacggca gcggcggcgg cggcaacagc aacggcgggc cgtacgggga ccatgacggc 180
gggtcctcgt cgtcaggtcc tgccaccgac ggcgcggtcg gcgggcccgg cgacgtggtg 240
gcgcgccggc cgcgggggcg cccgcctggc tccaagaaca agccgaagcc gccggtgatc 300
atcacgcggg agagcgccaa cacgctgcgc gcccacatcc tggaggtcgg gagcggctgc 360
gacgtgttcg agtgcgtctc cacgtacgcg cgccggcggc agcgcggcgt gtgcgtgctg 420
agcggcagcg gcgtggtcac caacgtgacg ctgcgtcagc cgtcggcgcc cgcgggcgcc 480
gtcgtgtcgc tgcacgggag gttcgagatc ctgtcgctct cgggctcctt cctcccgccg 540
ccggctcccc ccggcgccac cagcctcacc atcttcctcg ccgggggcca gggacaggtc 600
gtcggcggca acgtcgtcgg cgcgctctac gccgcgggcc cggtcatcgt catcgcggcg 660
tccttcgcca acgtcgccta cgagcgcctc ccactggagg aggaggaggc gccgccgccg 720
caggccggcc tgcagatgca gcagcccggc ggcggcgccg atgctggtgg catgggtggc 780
gcgttcccgc cggacccgtc tgccgccggc ctcccgttct tcaacctgcc gctcaacaac 840
atgcccggtg gcggcggctc acagctccct cccggcgccg acggccatgg ctgggccggc 900
gcacggccac cgttctga 918
<210> 185
<211> 305
<212> PRT
<213> Oryza sativa
<400> 185
Met Ala Gly Leu Asp Leu Gly Thr Ala Ala Thr Arg Tyr Val His Gln
1 5 10 15
Leu His His Leu His Pro Asp Leu Gln Leu Gln His Ser Tyr Ala Lys
20 25 30
Gln His Glu Pro Ser Asp Asp Asp Pro Asn Gly Ser Gly Gly Gly Gly
35 40 45
Asn Ser Asn Gly Gly Pro Tyr Gly Asp His Asp Gly Gly Ser Ser Ser
50 55 60
Ser Gly Pro Ala Thr Asp Gly Ala Val Gly Gly Pro Gly Asp Val Val
65 70 75 80
Ala Arg Arg Pro Arg Gly Arg Pro Pro Gly Ser Lys Asn Lys Pro Lys
85 90 95
Pro Pro Val Ile Ile Thr Arg Glu Ser Ala Asn Thr Leu Arg Ala His
100 105 110
Ile Leu Glu Val Gly Ser Gly Cys Asp Val Phe Glu Cys Val Ser Thr
115 120 125
Tyr Ala Arg Arg Arg Gln Arg Gly Val Cys Val Leu Ser Gly Ser Gly
130 135 140
Val Val Thr Asn Val Thr Leu Arg Gln Pro Ser Ala Pro Ala Gly Ala
145 150 155 160
Val Val Ser Leu His Gly Arg Phe Glu Ile Leu Ser Leu Ser Gly Ser
165 170 175
Phe Leu Pro Pro Pro Ala Pro Pro Gly Ala Thr Ser Leu Thr Ile Phe
180 185 190
Leu Ala Gly Gly Gln Gly Gln Val Val Gly Gly Asn Val Val Gly Ala
195 200 205
Leu Tyr Ala Ala Gly Pro Val Ile Val Ile Ala Ala Ser Phe Ala Asn
210 215 220
Val Ala Tyr Glu Arg Leu Pro Leu Glu Glu Glu Glu Ala Pro Pro Pro
225 230 235 240
Gln Ala Gly Leu Gln Met Gln Gln Pro Gly Gly Gly Ala Asp Ala Gly
245 250 255
Gly Met Gly Gly Ala Phe Pro Pro Asp Pro Ser Ala Ala Gly Leu Pro
260 265 270
Phe Phe Asn Leu Pro Leu Asn Asn Met Pro Gly Gly Gly Gly Ser Gln
275 280 285
Leu Pro Pro Gly Ala Asp Gly His Gly Trp Ala Gly Ala Arg Pro Pro
290 295 300
Phe
305

<210> 186
<211> 858
<212> DNA
<213> Arabidopsis thaliana
<400> 186
atggctggtc tcgatctagg cacaacttct cgctacgtcc acaacgtcga tggtggcggc 60
ggcggacagt tcaccaccga caaccaccac gaagatgacg gtggcgctgg aggaaaccac 120
catcatcacc atcataatca taatcaccat caaggtttag atttaatagc ttctaatgat 180
aactctggac taggcggcgg tggaggagga gggagcggtg acctcgtcat gcgtcggcca 240
cgtggccgtc cagctggatc gaagaacaaa ccgaagccgc cggtgattgt cacgcgcgag 300
agcgcaaaca ctcttagggc tcacattctt gaagttggaa gtggctgcga cgttttcgaa 360
tgtatctcca cttacgctcg tcggagacag cgcgggattt gcgttttatc cgggacggga 420
accgtcacta acgtcagcat ccgtcagcct acggcggccg gagctgttgt gactctgcgg 480
ggtacttttg agattctttc cctctccgga tcttttcttc cgccacctgc tcctccaggg 540
gcgactagct tgacgatatt cctcgctgga gctcaaggac aggtcgtcgg aggtaacgta 600
gttggtgagt taatggcggc ggggccggta atggtcatgg cagcgtcttt tacaaacgtg 660
gcttacgaaa ggttgccttt ggacgagcat gaggagcact tgcaaagtgg cggcggcgga 720
ggtggaggga atatgtactc ggaagccact ggcggtggcg gagggttgcc tttctttaat 780
ttgccgatga gtatgcctca gattggagtt gaaagttggc aggggaatca cgccggcgcc 840
ggtagggctc cgttttag 858
<210> 187
<211> 285
<212> PRT
<213> Arabidopsis thaliana
<400> 187
Met Ala Gly Leu Asp Leu Gly Thr Thr Ser Arg Tyr Val His Asn Val
1 5 10 15
Asp Gly Gly Gly Gly Gly Gln Phe Thr Thr Asp Asn His His Glu Asp
20 25 30
Asp Gly Gly Ala Gly Gly Asn His His His His His His Asn His Asn
35 40 45
His His Gln Gly Leu Asp Leu Ile Ala Ser Asn Asp Asn Ser Gly Leu
50 55 60
Gly Gly Gly Gly Gly Gly Gly Ser Gly Asp Leu Val Met Arg Arg Pro
65 70 75 80
Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys Pro Pro Val Ile
85 90 95
Val Thr Arg Glu Ser Ala Asn Thr Leu Arg Ala His Ile Leu Glu Val
100 105 110
Gly Ser Gly Cys Asp Val Phe Glu Cys Ile Ser Thr Tyr Ala Arg Arg
115 120 125
Arg Gln Arg Gly Ile Cys Val Leu Ser Gly Thr Gly Thr Val Thr Asn
130 135 140
Val Ser Ile Arg Gln Pro Thr Ala Ala Gly Ala Val Val Thr Leu Arg
145 150 155 160
Gly Thr Phe Glu Ile Leu Ser Leu Ser Gly Ser Phe Leu Pro Pro Pro
165 170 175
Ala Pro Pro Gly Ala Thr Ser Leu Thr Ile Phe Leu Ala Gly Ala Gln
180 185 190
Gly Gln Val Val Gly Gly Asn Val Val Gly Glu Leu Met Ala Ala Gly
195 200 205
Pro Val Met Val Met Ala Ala Ser Phe Thr Asn Val Ala Tyr Glu Arg
210 215 220
Leu Pro Leu Asp Glu His Glu Glu His Leu Gln Ser Gly Gly Gly Gly
225 230 235 240
Gly Gly Gly Asn Met Tyr Ser Glu Ala Thr Gly Gly Gly Gly Gly Leu
245 250 255
Pro Phe Phe Asn Leu Pro Met Ser Met Pro Gln Ile Gly Val Glu Ser
260 265 270
Trp Gln Gly Asn His Ala Gly Ala Gly Arg Ala Pro Phe
275 280 285
<210> 188
<211> 978

<212> DNA
<213> Medicago truncatula
<400> 188
atggatcaag tagcacaagg tcgtcctctt ccacctccat ttctcactag agaccttcat 60
cttcatcctc accatcaatt tcacaccaac caccaaacca atgaagagga acaacaaagt 120
ggcaatggga gcttaagccg aggccaaaaa agagaacgaa acaacgaaga cggcaacaac 180
actcccaccg gaggagaagg aaaagacgac ggtggtagcg gaagtgctgg tggaggaagt 240
ggtggtgaga tgggaagaag accaagagga agaccagcag gttcgaaaaa caaaccaaaa 300
ccacctatca tcatcacgag ggacagcgcg aacgcactcc gatcccacgt gatggaagtt 360
gcaaatggat gtgacatcat ggaaagtgtg acggtctttg cgcgaaggag gcagcgtggt 420
gtctgcatcc ttagcggaag tgggaccgtc acaaacgtga ctctccgtca accagcatcg 480
cctggtgcgg tagtcacact tcatggaaga tttgagatat tatcattatc tggctctttc 540
ctgccgccgc ctgctccacc agcggcgtca ggattagcca tatatctagc tggtggacaa 600
ggacaggtcg tcggtggtag cgtggtggga ccgttgttag cttccggtcc ggttgttatc 660
atggcagctt cctttggaaa tgctgcttat gaaaggctac ctttagaaga tgaagaaaca 720
ccagtgaatg tgccaggaaa tggagggtta gggtcaccgg gaaccatggg aagtcaacaa 780
caacagcagc agaaccagca acagcaacaa cttgtagcag atcctaatgc ttcatcactt 840
ttccatggag ttcctcaaaa tcttctcaat tcatgccaat taccagctga aggttattgg 900
ggtggaagtg ctcgtcctcc ttttttaacc aaaaatgtta ttcacttaat cacttttctc 960
atcatgtttt tcgtttaa 978
<210> 189
<211> 325
<212> PRT
<213> Medicago truncatula
<400> 189
Met Asp Gln Val Ala Gln Gly Arg Pro Leu Pro Pro Pro Phe Leu Thr
1 5 10 15
Arg Asp Leu His Leu His Pro His His Gln Phe His Thr Asn His Gln
20 25 30
Thr Asn Glu Glu Glu Gln Gln Ser Gly Asn Gly Ser Leu Ser Arg Gly
35 40 45
Gln Lys Arg Glu Arg Asn Asn Glu Asp Gly Asn Asn Thr Pro Thr Gly
50 55 60
Gly Glu Gly Lys Asp Asp Gly Gly Ser Gly Ser Ala Gly Gly Gly Ser
65 70 75 80
Gly Gly Glu Met Gly Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys
85 90 95
Asn Lys Pro Lys Pro Pro Ile Ile Ile Thr Arg Asp Ser Ala Asn Ala
100 105 110
Leu Arg Ser His Val Met Glu Val Ala Asn Gly Cys Asp Ile Met Glu
115 120 125
Ser Val Thr Val Phe Ala Arg Arg Arg Gln Arg Gly Val Cys Ile Leu
130 135 140
Ser Gly Ser Gly Thr Val Thr Asn Val Thr Leu Arg Gln Pro Ala Ser
145 150 155 160
Pro Gly Ala Val Val Thr Leu His Gly Arg Phe Glu Ile Leu Ser Leu
165 170 175
Ser Gly Ser Phe Leu Pro Pro Pro Ala Pro Pro Ala Ala Ser Gly Leu
180 185 190
Ala Ile Tyr Leu Ala Gly Gly Gln Gly Gln Val Val Gly Gly Ser Val
195 200 205
Val Gly Pro Leu Leu Ala Ser Gly Pro Val Val Ile Met Ala Ala Ser
210 215 220
Phe Gly Asn Ala Ala Tyr Glu Arg Leu Pro Leu Glu Asp Glu Glu Thr
225 230 235 240
Pro Val Asn Val Pro Gly Asn Gly Gly Leu Gly Ser Pro Gly Thr Met
245 250 255
Gly Ser Gln Gln Gln Gln Gln Gln Asn Gln Gln Gln Gln Gln Leu Val
260 265 270
Ala Asp Pro Asn Ala Ser Ser Leu Phe His Gly Val Pro Gln Asn Leu
275 280 285
Leu Asn Ser Cys Gln Leu Pro Ala Glu Gly Tyr Trp Gly Gly Ser Ala
290 295 300

Arg Pro Pro Phe Leu Thr Lys Asn Val Ile His Leu Ile Thr Phe Leu
305 310 315 320
Ile Met Phe Phe Val
325
<210> 190
<211> 6
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 1
<220>
<221> VARIANTE
<222> (4)..(4)
<223> /substituir="Ile"
<400> 190
Gln Gly Gln Val Gly Gly
1 5
<210> 191
<211> 15
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 2
<400> 191
Ile Leu Ser Leu Ser Gly Ser Phe Leu Pro Pro Pro Ala Pro Pro
1 5 10 15
<210> 192
<211> 8
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 3
<400> 192
Asn Ala Thr Tyr Glu Arg Leu Pro
1 5
<210> 193
<211> 12
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 4
<400> 193
Ser Phe Thr Asn Val Ala Tyr Glu Arg Leu Pro Leu
1 5 10
<210> 194
<211> 42
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 5
<400> 194
Gly Arg Phe Glu Ile Leu Ser Leu Thr Gly Ser Phe Leu Pro Gly Pro
1 5 10 15
Ala Pro Pro Gly Ser Thr Gly Leu Thr Ile Tyr Leu Ala Gly Gly Gln
20 25 30
Gly Gln Val Val Gly Gly Ser Val Val Gly
35 40
<210> 195
<211> 654
<212> DNA
<213> Oryza sativa
<400> 195
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> 196
<211> 51
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador
<400> 196
ggggacaagt ttgtacaaaa aagcaggctt aaacaatgga tccggtcacg g 51

<210> 197
<211> 49
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador
<400> 197
ggggaccact ttgtacaaga aagctgggtg gaatcgatcc atctcagaa 49

<210> 198
<211> 828
<212> DNA
<213> Arabidopsis thaliana
<400> 198
atgggtggat cgatgtcgga gagagcaagg caggccaaca ttcctccact agcgggaccc 60
ctaaagtgtc ctcgatgcga ctccagcaac actaagttct gttactacaa caactataac 120
ctcactcagc ctcgtcactt ctgcaaaggt tgccgtcgct actggacaca agggggcgcc 180
ctgagaaacg tccctgtagg tggaggctgc cggaggaata acaagaaggg caaaaatgga 240
aatttaaaat cttcttcttc ttcgtccaaa cagtcttcct cggtcaacgc tcaaagtcct 300
agctcaggac agctaaggac aaatcatcag ttccctttt caccaactct ttacaatctc 360
actcaactcg gaggtattgg tttgaactta gccgctacta atggcaacaa ccaagctcac 420
cagatcggtt ccagtttgat gatgagcgat ctagggtttc tccatggacg aaatacttca 480
actccgatga cgggaaacat tcatgaaaac aacaacaata ataacaatga aaacaaccta 540
atggcatccg ttggatcttt gagccccttt gctctcttcg atccaacgac ggggctatac 600
gctttccaga acgacggtaa tatcgggaac aacgttggga tatctggctc ttctacttcc 660
atggttgatt ctagggttta tcagacgcct ccggtgaaga tggaagaaca acctaatttg 720
gctaacttgt ctagaccggt ctccggtttg acgtctcctg ggaatcaaac aaatcagtac 780
ttttggcctg gttcggattt ctcgggtcct tctaatgatc tcttgtga 828

<210> 199
<211> 275
<212> PRT
<213> Arabidopsis thaliana
<400> 199

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

100 105 110
Phe Ser Pro Thr Leu Tyr Asn Leu Thr Gln Leu Gly Gly Ile Gly Leu
115 120 125
Asn Leu Ala Ala Thr Asn Gly Asn Asn Gln Ala His Gln Ile Gly Ser
130 135 140
Ser Leu Met Met Ser Asp Leu Gly Phe Leu His Gly Arg Asn Thr Ser
145 150 155 160
Thr Pro Met Thr Gly Asn Ile His Glu Asn Asn Asn Asn Asn Asn Asn
165 170 175
Glu Asn Asn Leu Met Ala Ser Val Gly Ser Leu Ser Pro Phe Ala Leu
180 185 190
Phe Asp Pro Thr Thr Gly Leu Tyr Ala Phe Gln Asn Asp Gly Asn Ile
195 200 205
Gly Asn Asn Val Gly Ile Ser Gly Ser Ser Thr Ser Met Val Asp Ser
210 215 220
Arg Val Tyr Gln Thr Pro Pro Val Lys Met Glu Glu Gln Pro Asn Leu
225 230 235 240
Ala Asn Leu Ser Arg Pro Val Ser Gly Leu Thr Ser Pro Gly Asn Gln
245 250 255
Thr Asn Gln Tyr Phe Trp Pro Gly Ser Asp Phe Ser Gly Pro Ser Asn
260 265 270
Asp Leu Leu
275

<210> 200
<211> 63
<212> PRT
<213> Seqüência artificial
<220>
<223> Domínio DOF
<400> 200
Pro Leu Ala Gly Pro Leu Lys Cys Pro Arg Cys Asp Ser Ser Asn Thr
1 5 10 15
Lys Phe Cys Tyr Tyr Asn Asn Tyr Asn Leu Thr Gln Pro Arg His Phe
20 25 30
Cys Lys Gly Cys Arg Arg Tyr Trp Thr Gln Gly Gly Ala Leu Arg Asn
35 40 45
Val Pro Val Gly Gly Gly Cys Arg Arg Asn Asn Lys Lys Gly Lys
50 55 60
<210> 201
<211> 684
<212> DNA
<213> Arabidopsis thaliana
<400> 201
atggcggaga gagcaaggca ggccaacatt cctccactag cgggacccct aaagtgtcct 60
cgatgcgact ccagcaacac taagttctgt tactacaaca actataacct cactcagcct 120
cgtcacttct gcaaaggttg ccgtcgctac tggacacaag ggggcgccct gagaaacgtc 180
cctgtaggtg gaggctgccg gaggaataac aagaagggca aaaatggaaa tttaaaatct 240
tcttcttctt cgtccaaaca gtcttcctcg gtcaacgctc aaagtcctag ctcaggacag 300
ctaaggacaa atcatcagtt ccctttttca ccaactcttt acaatctcac tcaactcgga 360
ggtattggtt tgaacttagc cgctactaat ggcaacaacc aagctcacca gatcggttcc 420
agtttgatga tgagcgatct agggtttctc catggacgaa atacttcaac tccgatgacg 480
ggaaacattc atgaaaacaa caacaataat aacaatgaaa acaacctaat ggcatccgtt 540
ggatctttga gccccttgc tctcttcgat ccaacgacgg ggctatacgc tttccagaac 600
gacggtaata tcgggaacaa cgttgggata tctggttctt ctacttccat ggttgattct 660
agggtttatc agacgctccg gtga 684
<210> 202
<211> 227
<212> PRT
<213> Arabidopsis thaliana
<400> 202
Met Ala Glu Arg Ala Arg Gln Ala Asn Ile Pro Pro Leu Ala Gly Pro
1 5 10 15
Leu Lys Cys Pro Arg Cys Asp Ser Ser Asn Thr Lys Phe Cys Tyr Tyr

```
            20                  25                  30
Asn Asn Tyr Asn Leu Thr Gln Pro Arg His Phe Cys Lys Gly Cys Arg
        35                  40                  45
Arg Tyr Trp Thr Gln Gly Gly Ala Leu Arg Asn Val Pro Val Gly Gly
    50                  55                  60
Gly Cys Arg Arg Asn Asn Lys Lys Gly Lys Asn Gly Asn Leu Lys Ser
65                  70                  75                  80
Ser Ser Ser Ser Ser Lys Gln Ser Ser Ser Val Asn Ala Gln Ser Pro
                85                  90                  95
Ser Ser Gly Gln Leu Arg Thr Asn His Gln Phe Pro Phe Ser Pro Thr
            100                 105                 110
Leu Tyr Asn Leu Thr Gln Leu Gly Gly Ile Gly Leu Asn Leu Ala Ala
        115                 120                 125
Thr Asn Gly Asn Asn Gln Ala His Gln Ile Gly Ser Ser Leu Met Met
    130                 135                 140
Ser Asp Leu Gly Phe Leu His Gly Arg Asn Thr Ser Thr Pro Met Thr
145                 150                 155                 160
Gly Asn Ile His Glu Asn Asn Asn Asn Asn Asn Asn Glu Asn Asn Leu
                165                 170                 175
Met Ala Ser Val Gly Ser Leu Ser Pro Phe Ala Leu Phe Asp Pro Thr
            180                 185                 190
Thr Gly Leu Tyr Ala Phe Gln Asn Asp Gly Asn Ile Gly Asn Asn Val
        195                 200                 205
Gly Ile Ser Gly Ser Ser Thr Ser Met Val Asp Ser Arg Val Tyr Gln
    210                 215                 220
Thr Leu Arg
225
<210> 203
<211> 993
<212> DNA
<213> Arabidopsis thaliana
<400> 203
atgcctacga attcgaatca tcagcatcat cttcaacacc agcttaacga aaatggaagt     60
ataataagtg gccacggact agtactctct caccaacttc cacctctcca agcaaaccct    120
aaccctaacc accaccatgt cgctacctct gctggtcttc cgtcaaggat gggtggatcg    180
atggcggaga gagcaaggca ggccaacatt cctccactag cgggacccct aaagtgtcct    240
cgatgcgact ccagcaacac taagttctgt tactacaaca actataacct cactcagcct    300
cgttacttct gcaaaggttg ccgtcgctac tggacacaag ggggcgccct gagaaacgtc    360
cctgtaggtg gaggctgccg gaggaataac aagaagggca aaaatggaaa tttaaaatct    420
tcttcttctt cgtccaaaca gtcttcctcg gtcaacgctc aaagtcctag ctcaggacag    480
ctaaggacaa atcatcagtt ccctttttca ccaactcttt acaatctcac tcaactcgga    540
ggtattggtt tgaacttagc cgctactaat ggcaacaacc aagctcacca gatcggttcc    600
agtttgatga tgagcgatct agggtttctc catggacgaa atacttcaac tccgatgacg    660
ggaaacattc atgaaaacaa caacaataat aacaatgaaa acaacctaat ggcatccgtt    720
ggatctttga gccccttttgc tctcttcgat ccaacgacgg ggctatacgc tttccagaac    780
gacggtaata tcgggaacaa cgttgggata tctggttctt ctacttccat ggttgattct    840
agggtttatc agacgcctcc ggtgaagatg gaagaacaac ctaatttggc taacttgtct    900
agaccggtct ccggtttgac gtctcctggg aatcaaacaa atcaatactt ttggcctggt    960
tcggatttct cgggtccttc taatgatatc ttg                                 993
<210> 204
<211> 331
<212> PRT
<213> Arabidopsis thaliana
<400> 204
Met Pro Thr Asn Ser Asn His Gln His His Leu Gln His Gln Leu Asn
1               5                   10                  15
Glu Asn Gly Ser Ile Ile Ser Gly His Gly Leu Val Leu Ser His Gln
            20                  25                  30
Leu Pro Pro Leu Gln Ala Asn Pro Asn Pro Asn His His His Val Ala
        35                  40                  45
Thr Ser Ala Gly Leu Pro Ser Arg Met Gly Gly Ser Met Ala Glu Arg
    50                  55                  60
Ala Arg Gln Ala Asn Ile Pro Pro Leu Ala Gly Pro Leu Lys Cys Pro
```

```
65                  70                  75                  80
Arg Cys Asp Ser Ser Asn Thr Lys Phe Cys Tyr Tyr Asn Asn Tyr Asn
                85                  90                  95
Leu Thr Gln Pro Arg Tyr Phe Cys Lys Gly Cys Arg Arg Tyr Trp Thr
            100                 105                 110
Gln Gly Gly Ala Leu Arg Asn Val Pro Val Gly Gly Gly Cys Arg Arg
        115                 120                 125
Asn Asn Lys Lys Gly Lys Asn Gly Asn Leu Lys Ser Ser Ser Ser Ser
    130                 135                 140
Ser Lys Gln Ser Ser Ser Val Asn Ala Gln Ser Pro Ser Ser Gly Gln
145                 150                 155                 160
Leu Arg Thr Asn His Gln Phe Pro Phe Ser Pro Thr Leu Tyr Asn Leu
                165                 170                 175
Thr Gln Leu Gly Gly Ile Gly Leu Asn Leu Ala Ala Thr Asn Gly Asn
            180                 185                 190
Asn Gln Ala His Gln Ile Gly Ser Ser Leu Met Met Ser Asp Leu Gly
        195                 200                 205
Phe Leu His Gly Arg Asn Thr Ser Thr Pro Met Thr Gly Asn Ile His
    210                 215                 220
Glu Asn Asn Asn Asn Asn Asn Asn Glu Asn Asn Leu Met Ala Ser Val
225                 230                 235                 240
Gly Ser Leu Ser Pro Phe Ala Leu Phe Asp Pro Thr Thr Gly Leu Tyr
                245                 250                 255
Ala Phe Gln Asn Asp Gly Asn Ile Gly Asn Asn Val Gly Ile Ser Gly
            260                 265                 270
Ser Ser Thr Ser Met Val Asp Ser Arg Val Tyr Gln Thr Pro Pro Val
        275                 280                 285
Lys Met Glu Glu Gln Pro Asn Leu Ala Asn Leu Ser Arg Pro Val Ser
    290                 295                 300
Gly Leu Thr Ser Pro Gly Asn Gln Thr Asn Gln Tyr Phe Trp Pro Gly
305                 310                 315                 320
Ser Asp Phe Ser Gly Pro Ser Asn Asp Ile Leu
                325                 330
```

<210> 205
<211> 873
<212> DNA
<213> Oryza sativa
<400> 205

```
atgattccgg gcacgcttgc cgatggcggc ggcggaggcg gcgcggtggg gccggcgaag     60
ccgatgtcga tgtcggagag ggcgcggctg gcgaggatcc cgctgccgga gccggggctc    120
aagtgcccgc gctgcgactc caccaacacc aagttctgct acttcaacaa ctactccctc    180
tcccagcccc gccacttctg ccgcgcctgc cgccgctact ggacccgcgg cggcgcgctc    240
cgcaacgtcc ccgtcggcgg cggctaccgc cgccacgcca agcgcgccaa gcccaagccg    300
gcgtccgcgg cgggctccgc ctcagccgcc accaccaccg ccggctcgac gccagcgggg    360
tcgacgacga cgacgacgac gtcctccacc tgcgccacgc ccaacgcgcc cgccctcccg    420
gcgatgctgg gcggcaacct ctccatcctg ccgccgctgc tccgcctcgc cgacttcgac    480
gccatgagcc tcggctccac cttctctggc atggcggcgg ccgccggcaa gccaccgccc    540
gtcgacgcgg ccggctgcta ctccgtcggc gccgccaccg gcctcgagca atggagacta    600
cagcagatgc agagcttccc gttcttccac gccatggatc accaggcggc gatggcggcg    660
ccaccgccgg caatggcaat gccggggatg ttccagctag gcctagacgg cgacggccat    720
ggcagcggcg gcggcgaaga cggtggagag ctccaccatg cgatgccatc atcgaagaga    780
gaaggctacc caaggggcat gtatggcgat catcacctcg ctggaggata cacctcctac    840
tccagtgcaa ccacaggtaa ccatctcttg taa                                 873
```

<210> 206
<211> 290
<212> PRT
<213> Oryza sativa
<400> 206

```
Met Ile Pro Gly Thr Leu Ala Asp Gly Gly Gly Gly Gly Gly Ala Val
1               5                   10                  15
Gly Pro Ala Lys Pro Met Ser Met Ser Glu Arg Ala Arg Leu Ala Arg
            20                  25                  30
Ile Pro Leu Pro Glu Pro Gly Leu Lys Cys Pro Arg Cys Asp Ser Thr
```

```
         35                  40                  45
Asn Thr Lys Phe Cys Tyr Phe Asn Asn Tyr Ser Leu Ser Gln Pro Arg
    50                  55                  60
His Phe Cys Arg Ala Cys Arg Arg Tyr Trp Thr Arg Gly Gly Ala Leu
65                  70                  75                  80
Arg Asn Val Pro Val Gly Gly Gly Tyr Arg Arg His Ala Lys Arg Ala
                85                  90                  95
Lys Pro Lys Pro Ala Ser Ala Ala Gly Ser Ala Ser Ala Ala Thr Thr
            100                 105                 110
Thr Ala Gly Ser Thr Pro Ala Gly Ser Thr Thr Thr Thr Thr Thr Ser
        115                 120                 125
Ser Thr Cys Ala Thr Pro Asn Ala Pro Ala Leu Pro Ala Met Leu Gly
    130                 135                 140
Gly Asn Leu Ser Ile Leu Pro Pro Leu Leu Arg Leu Ala Asp Phe Asp
145                 150                 155                 160
Ala Met Ser Leu Gly Ser Thr Phe Ser Gly Met Ala Ala Ala Ala Gly
                165                 170                 175
Lys Pro Pro Pro Val Asp Ala Ala Gly Cys Tyr Ser Val Gly Ala Ala
            180                 185                 190
Thr Gly Leu Glu Gln Trp Arg Leu Gln Gln Met Gln Ser Phe Pro Phe
        195                 200                 205
Phe His Ala Met Asp His Gln Ala Ala Met Ala Ala Pro Pro Pro Ala
    210                 215                 220
Met Ala Met Pro Gly Met Phe Gln Leu Gly Leu Asp Gly Asp Gly His
225                 230                 235                 240
Gly Ser Gly Gly Gly Glu Asp Gly Gly Glu Leu His His Ala Met Pro
                245                 250                 255
Ser Ser Lys Arg Glu Gly Tyr Pro Arg Gly Met Tyr Gly Asp His His
            260                 265                 270
Leu Ala Gly Gly Tyr Thr Ser Tyr Ser Ser Ala Thr Thr Gly Asn His
        275                 280                 285
Leu Leu
    290
<210> 207
<211> 1068
<212> DNA
<213> Oryza sativa
<400> 207
atgccgccgc atcacggcgg cctcatggcg cctcggcctg acatggtagc agcggccgtc     60
gcggcgagcg gcggcggcgg tggcggcggc ggcccgaccg gcggcacggc ggtgcggccg    120
ggctcgatga cggaacgggc tcggctggcg aagatcccgc agccggagcc ggggctcaag    180
tgcccgcgct gcgagtccac caacaccaag ttctgctact tcaacaacta ctcgctctcg    240
cagccgcgcc acttctgcaa gacgtgccgc cgctactgga cgcgcggcgg agcgctccgc    300
aacgtccccg tcggcggcgg gtgccgccgc aacaagcgca ccaagtcgtc caagtcgtcc    360
tcgtcgacgt cggccgccgg ctcggcctcc gccaccggcg gcacgtcgtc gtccacatcg    420
tcgaccgcca cgggtggcag cagcagcgcc gcggcggccg cggcgatgat gccgccgcag    480
gcgcagctgc cgttcctggc ctcgttgcac cacccgctcg gcggcggcga tcactacagc    540
tccggtgcgt ccaggctagg gtttcccgga ttgagctcgc tggatcccgt cgactaccag    600
ctcggcggcg gcgccgccgc cgccgccgcc atcgggctag agcagtggcg cctcccgcag    660
atacagcaat tccccttctt gagccgcaac gacgccatgc cgccgccaat gtccggcatt    720
tacccgttcg acgcggaggc cgccgccgac gccgccggct tcgccggcca gttgctggcc    780
ggcaccaagg tgcccggctc gtcgggcctg atcacgcagc tcgcatccgt caagatggag    840
gacagcaacg ctcagtccgc ggcgatgaac agctcgccga gggagttctt gggcctcccc    900
ggcaacctcc aattctgggg cggtggcaac ggcgcgggac ccggcggcaa tggagacggc    960
gccaccggcg gcagcggcgc cggtgtcgct ccgggaggcg gcggcagcgg cggcggatgg   1020
gctgatctct ccggattcaa ctcgtcgtcg tcggggaaca tactgtga                1068
<210> 208
<211> 355
<212> PRT
<213> Oryza sativa
<400> 208
Met Pro Pro His His Gly Gly Leu Met Ala Pro Arg Pro Asp Met Val
1               5                   10                  15
```

```
Ala Ala Ala Val Ala Ala Ser Gly Gly Gly Gly Gly Gly Gly Gly Pro
            20                  25                  30
Thr Gly Gly Thr Ala Val Arg Pro Gly Ser Met Thr Glu Arg Ala Arg
        35                  40                  45
Leu Ala Lys Ile Pro Gln Pro Glu Pro Gly Leu Lys Cys Pro Arg Cys
    50                  55                  60
Glu Ser Thr Asn Thr Lys Phe Cys Tyr Phe Asn Asn Tyr Ser Leu Ser
65                  70                  75                  80
Gln Pro Arg His Phe Cys Lys Thr Cys Arg Arg Tyr Trp Thr Arg Gly
                85                  90                  95
Gly Ala Leu Arg Asn Val Pro Val Gly Gly Gly Cys Arg Arg Asn Lys
            100                 105                 110
Arg Thr Lys Ser Ser Lys Ser Ser Ser Ser Thr Ser Ala Ala Gly Ser
        115                 120                 125
Ala Ser Ala Thr Gly Gly Thr Ser Ser Ser Thr Ser Ser Thr Ala Thr
    130                 135                 140
Gly Gly Ser Ser Ser Ala Ala Ala Ala Ala Ala Met Met Pro Pro Gln
145                 150                 155                 160
Ala Gln Leu Pro Phe Leu Ala Ser Leu His His Pro Leu Gly Gly Gly
                165                 170                 175
Asp His Tyr Ser Ser Gly Ala Ser Arg Leu Gly Phe Pro Gly Leu Ser
            180                 185                 190
Ser Leu Asp Pro Val Asp Tyr Gln Leu Gly Gly Gly Ala Ala Ala Ala
        195                 200                 205
Ala Ala Ile Gly Leu Glu Gln Trp Arg Leu Pro Gln Ile Gln Gln Phe
    210                 215                 220
Pro Phe Leu Ser Arg Asn Asp Ala Met Pro Pro Pro Met Ser Gly Ile
225                 230                 235                 240
Tyr Pro Phe Asp Ala Glu Ala Ala Ala Asp Ala Ala Gly Phe Ala Gly
                245                 250                 255
Gln Leu Leu Ala Gly Thr Lys Val Pro Gly Ser Ser Gly Leu Ile Thr
            260                 265                 270
Gln Leu Ala Ser Val Lys Met Glu Asp Ser Asn Ala Gln Ser Ala Ala
        275                 280                 285
Met Asn Ser Ser Pro Arg Glu Phe Leu Gly Leu Pro Gly Asn Leu Gln
    290                 295                 300
Phe Trp Gly Gly Gly Asn Gly Ala Gly Pro Gly Gly Asn Gly Asp Gly
305                 310                 315                 320
Ala Thr Gly Gly Ser Gly Ala Gly Val Ala Pro Gly Gly Gly Gly Ser
                325                 330                 335
Gly Gly Gly Trp Ala Asp Leu Ser Gly Phe Asn Ser Ser Ser Ser Gly
            340                 345                 350
Asn Ile Leu
        355
<210> 209
<211> 1000
<212> DNA
<213> Triticum aestivum
<400> 209
ctcctagctc gtcgacaagc atgcccacaa agccactgat aacaagcgca tagcagcgca     60
cgctctctta tagtagaatg ttctgaactc cacaccagcc catgcaagac ttccagtcca    120
tcccgggcct cgccgggcgg ctgttcggcg gcgcggccgc ggcagacatc cggcgcgtgc    180
agggcccggc gtcccggtgc ggcgtgttct cgcaggcggc gtccgcgcag ccggaggcgg    240
ccgtcaagtg cccgcggtgc gagtccacca acaccaagtt ctgctactac aacaactaca    300
acctgtcgca gccgcgccac ttctgcaaga gctgccgccg gtactggacc aagggcggcg    360
tcctccgcaa cgtccccgtc ggcggcggct gccgcaaggc caagcgcagc tcctcgtcgg    420
cgtccgcacc gtcgacgccc gcggccacgg acgccaagag ccagcggcgc gcgtccgcgt    480
cgtcctcctc ccgctccaac agcggcagcg gcagcgccag ccccacggcc gctgcggaag    540
agacgacgac aacggagacc gagcccctc ctccgcccac gccgtcgtcc aactccaact    600
ccaacgcggt ctccttcgcc aaccgcatga cgaactaccc cttcgcggca gacgtgccac    660
ctctggcgcc gatattcgcc gaccaggccg ccgcgctcgc gtccctcttc gcgccgcctc    720
ctccaccgcc tctcccggtg ttcagcttct cggcggagcc caagatggag gaggcgatcg    780
ggtcactgct gctcccgggg caggaggcgt cgcaggagcc agaggagccc acctgcacct    840
```

ccaccgtcgc ggacatggcg ccgttcatgt cgctggacgc ggggatcttc gagctcggcg 900
acgcgtcgcc ggccgattac tggaacggcg ggagctgctg gacggacgtc caggacccgt 960
ccgtctacct accctagttt ggcttagttc ctcatcacga 1000

<210> 210
<211> 291
<212> PRT
<213> Triticum aestivum
<400> 210

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

<210> 211
<211> 1230
<212> DNA
<213> Zea mays
<400> 211

aggtggtggc gggggacagg ttgggggccc cgccaagccc atgtccatgg cggagcgcgc 60
gcgcctcgcg aggatcccac tgccggagcc gggactcaag tgcccgcgct gcgactcaac 120
caacaccaag ttctgctact tcaacaacta ctccctcacg cagccgcgcc acttctgccg 180
ggcctgccgc cgctactgga cgcgtggcgg cgcgctccgc aacgtgccgg tcggtggcgg 240
ataccgccgc cacgccaagc gcgccaagcc caagcagcag cagcagcacg cggccgccgg 300
gaccggagct gccaacggcg cgatgcagca gcctcccgct gggtctatgg cgtcgtcggc 360
cgccgcctgc accgcgacca cgacgatgac gaccaacgcg ctggacgccg ggcccggcgg 420
catgctgccc atgctgccgc cgctcgtccg cctagcagac ttcgacgcaa tgagcctcgg 480
ctccagcttc tccgggatat cgtccatggg gaagcccgga tccatcggcg cggcctgcta 540
cccgcactct gtcggcgggc tggagcagtg gagggtgcag cagatgcaga gcttcccgtt 600
cttgcatgcg atggaccagg gcccgctggg gccacctctg gccatggcga tggcggcgcc 660
aggagggatg ttccagctag gtctagacac caccagtgat aacagccgtg gcggtggcgg 720

cggcggctgc ggcgaagacg ggtcgtctgc gggagaggcg ctccatatga tgcaagcagc 780
caccaagagg gagagctacc cggcaccacc aagagccatg tacggcgacc aacaccacaa 840
tcacctcgct gctgctggtg gctacacttc ctattccacc aatgctgctg caggtaacca 900
tctcttgtaa tggccggccg atcgatcgat cgagagctca acaattcaag tgtcctgcta 960
tagctagcta ctacgacgtc gtgctacgat cgtatcggtt tcggttcgtt cctacaaata 1020
tagcctagag atagagtgtc tctgtctgtg tgatcgatgg tattgttatg atcatataga 1080
aaagaccagt gtagcatgca tgcacttgtt gcaatgtttg ctttcaagaa gaaactggag 1140
ggaggagagg ctcggtttgg atgctgatca tgcacaaata ttactagtgt ctacagctgc 1200
tacttcatta taaaaaaaaa aaaaaaaaaa 1230

<210> 212
<211> 302
<212> PRT
<213> Zea mays
<400> 212

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

<210> 213
<211> 1243
<212> DNA
<213> Solanum tuberosum
<400> 213

gaggacttat caacctttt atataaaaga gataaagatc aaagagatca aagaaagaaa 60
atcatggttt tctcatcttt tcctgtgtat ctagatcatc ccaatttgca tcagttacaa 120
cagccagacg gccatcaaca aggaaatact gggctggaga atccaactct gcagccccca 180
cctatgcagg tgggggctag tccgggctca atcagaccag gttctatggt ggatcgagcc 240
cggttagcta agattccact accggaagct ggattaaagt gtccaaggtg tgattcaaca 300
aacacaaagt tctgctactt caacaactac aacctttccc aaccaagaca cttctgcaag 360

acctgtcgcc ggtactggac tagaggggc gccttgagaa gcgtgccggt aggaggagga 420
tgccggagga acaagagaag caaaagtaat aacaacaaca acagctcaaa gacagctggc 480
agtaatgtca atactaatac tattgcttcc ggtacttcaa caagtgcaag tccttcaagc 540
tgcagcacgg aaataatgaa tggtcgccat cacttctctc atgagcaacc accgcagtta 600
actcctctca tggctgcttt ccaaaatctt aatcatcact atggcggatt tcaacctcct 660
cctttggttt caacacacca tgggaatgga accggcgcgc ttggtcatca tcatgaaatg 720
ggatttcaaa tagggagtag tactaatact aacaatttac ctgtaccccc aggaggagga 780
tctgatcatc agtggagatt accatctttg gcagcaaaca caaatttgta ccctttcaa 840
cacggtactg atcaaggaat tcatgaatca tcatctgtta acaataataa tattaatgct 900
catgatgacc aagggttaaa ttcgaccaaa cagtttttgg gaacaatgga aaataatact 960
aatcaatatt ggggtggaaa cgcatggaca gggttttctg gattaaactc atcttcttca 1020
gcaagccatc tactttgatt tcattcatgt aagtcaattt gtgtacttga ctctcaacga 1080
ttaaaatgtt gtatgtgttg ggaggggggt ctaaattaga tatatagtat actgggggca 1140
tgtttaagtc tttgttattc catggtcatt atcaccactt gtaattttac tggatgtttt 1200
ttttttataa cttaggtgtg tgaagttgta ttgtgagttt taa 1243

<210> 214
<211> 324
<212> PRT
<213> Solanum tuberosum
<400> 214

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

<210> 215

<211> 996
<212> PRT
<213> Arabidopsis thaliana
<400> 215

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

```
465                 470                 475                 480
Cys Thr Ala Ala Gly Gly Ala Cys Ala Ala Ala Thr Cys Ala Thr Cys
                485                 490                 495
Ala Gly Thr Thr Cys Cys Cys Thr Thr Thr Thr Thr Cys Ala Cys Cys
            500                 505                 510
Ala Ala Cys Thr Cys Thr Thr Thr Ala Cys Ala Ala Thr Cys Thr Cys
        515                 520                 525
Ala Cys Thr Cys Ala Ala Cys Thr Cys Gly Gly Ala Gly Gly Thr Ala
    530                 535                 540
Thr Thr Gly Gly Thr Thr Thr Gly Ala Ala Cys Thr Thr Ala Gly Cys
545                 550                 555                 560
Cys Gly Cys Thr Ala Cys Thr Ala Ala Thr Gly Gly Cys Ala Ala Cys
                565                 570                 575
Ala Ala Cys Cys Ala Ala Gly Cys Thr Cys Ala Cys Cys Ala Gly Ala
            580                 585                 590
Thr Cys Gly Gly Thr Thr Cys Cys Ala Gly Thr Thr Thr Gly Ala Thr
        595                 600                 605
Gly Ala Thr Gly Ala Gly Cys Gly Ala Thr Cys Thr Ala Gly Gly Gly
    610                 615                 620
Thr Thr Thr Cys Thr Cys Cys Ala Thr Gly Gly Ala Cys Gly Ala Ala
625                 630                 635                 640
Ala Thr Ala Cys Thr Thr Cys Ala Ala Cys Thr Cys Cys Gly Ala Thr
                645                 650                 655
Gly Ala Cys Gly Gly Gly Ala Ala Ala Cys Ala Thr Thr Cys Ala Thr
            660                 665                 670
Gly Ala Ala Ala Ala Cys Ala Ala Cys Ala Ala Cys Ala Ala Thr Ala
        675                 680                 685
Ala Thr Ala Ala Cys Ala Ala Thr Gly Ala Ala Ala Ala Cys Ala Ala
    690                 695                 700
Cys Cys Thr Ala Ala Thr Gly Gly Cys Ala Thr Cys Cys Gly Thr Thr
705                 710                 715                 720
Gly Gly Ala Thr Cys Thr Thr Thr Gly Ala Gly Cys Cys Cys Cys Thr
                725                 730                 735
Thr Thr Gly Cys Thr Cys Thr Cys Thr Thr Cys Gly Ala Thr Cys Cys
            740                 745                 750
Ala Ala Cys Gly Ala Cys Gly Gly Gly Gly Cys Thr Ala Thr Ala Cys
        755                 760                 765
Gly Cys Thr Thr Thr Cys Cys Ala Gly Ala Ala Cys Gly Ala Cys Gly
    770                 775                 780
Gly Thr Ala Ala Thr Ala Thr Cys Gly Gly Gly Ala Ala Cys Ala Ala
785                 790                 795                 800
Cys Gly Thr Thr Gly Gly Gly Ala Thr Ala Thr Cys Thr Gly Gly Thr
                805                 810                 815
Thr Cys Thr Thr Cys Thr Ala Cys Thr Thr Cys Cys Ala Thr Gly Gly
            820                 825                 830
Thr Thr Gly Ala Thr Thr Cys Thr Ala Gly Gly Gly Thr Thr Thr Ala
        835                 840                 845
Thr Cys Ala Gly Ala Cys Gly Cys Cys Thr Cys Cys Gly Gly Thr Gly
    850                 855                 860
Ala Ala Gly Ala Thr Gly Gly Ala Ala Gly Ala Ala Cys Ala Ala Cys
865                 870                 875                 880
Cys Thr Ala Ala Thr Thr Thr Gly Gly Cys Thr Ala Ala Cys Thr Thr
                885                 890                 895
Gly Thr Cys Thr Ala Gly Ala Cys Cys Gly Gly Thr Cys Thr Cys Cys
            900                 905                 910
Gly Gly Thr Thr Thr Gly Ala Cys Gly Thr Cys Thr Cys Cys Thr Gly
        915                 920                 925
Gly Gly Ala Ala Thr Cys Ala Ala Ala Cys Ala Ala Ala Thr Cys Ala
    930                 935                 940
Gly Thr Ala Cys Thr Thr Thr Thr Gly Gly Cys Cys Thr Gly Gly Thr
945                 950                 955                 960
Thr Cys Gly Gly Ala Thr Thr Thr Cys Thr Cys Gly Gly Gly Thr Cys
                965                 970                 975
```

Cys Thr Thr Cys Thr Ala Ala Thr Gly Ala Thr Cys Thr Cys Thr Thr
980 985 990
Gly Thr Gly Ala
995

<210> 216
<211> 275
<212> PRT
<213> Arabidopsis thaliana
<400> 216

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

<210> 217
<211> 993
<212> DNA
<213> Arabidopsis thaliana
<400> 217

atggttttct cctccatcca agcctatctt gattcatcca actggcaaca ggctcctccg 60
agcaattata atcatgacgg aacaggcgcc tcagcaaatg gaggtcatgt tcttcgtcct 120
cagctgcagc cacagcagca gccacagcag cagccgcatc ctaatgggag cgggggcgga 180
ggtggaggtg gaggcggctc gatccgagca ggatcaatgg tggacagagc aagacaagca 240
aacgtagcct tgccagaagc agcactgaaa tgtccgagat gcgaatccac caacaccaag 300
ttctgctact tcaacaacta cagcctcact caaccacgcc acttctgcaa gacctgccgg 360
agatactgga cacgtggcgg agctctccgc aacgtcccag tcggcggtgg ctgccggaga 420
aacaggcgta ccaaaagcaa cagcaacaac aacaataaca gcactgctac tagcaataac 480
accagtttct cctccgggaa tgcatccacc atcagcacga ttctctcctc ccactatgga 540
ggaaaccaag agagtatctt aagccagatt ttgtctccgg cgaggctaat gaatcctact 600
tacaatcatc tcggagatct cacaagtaat acaaaaacag acaacaacat gagcttgttg 660
aactatggag gattgagtca agacttgagg tcaatccaca tgggagcttc tggtggctcg 720
cttatgagct gtgttgatga atggagatcg gcgtcttatc atcagcagtc aagtatgggc 780

ggtgggaact tggaggattc ttctaatcct aatccatccg caaatgggtt ttactctttt 840
gagtcgccga ggataacttc agcgtcaatc tcttctgctt tagcatcgca gttttcttcg 900
gttaaagttg aagataatcc ttacaaatgg gttaatgtca atggtaattg ctcttcctgg 960
aatgatcttt ctgctttcgg ctcttctcgt tga 993

<210> 218
<211> 330
<212> PRT
<213> Arabidopsis thaliana
<400> 218

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

<210> 219
<211> 972
<212> DNA
<213> Arabidopsis thaliana
<400> 219

atggttttct catctcttcc agtgaatcag ttcgattccc aaaattggca gcagcaaggg 60
aaccaacatc agctagaatg tgtcacaact gaccagaacc ctaataatta cttacggcag 120
ctctcatcac caccgacttc tcaggttgca ggttcgagtc aagctagagt gaattcaatg 180
gtggaacgtg ctcggatcgc aaaagtccca ttgcctgaag cagctctaaa ttgccctaga 240
tgtgactcaa ccaatactaa gttctgttac ttcaataact atagccttac tcaacctcgc 300
catttctgca aaacatgtcg tcgctattgg acacgtggcg gttccttgag gaatgttcct 360
gttggaggag gctttaggag gaacaagaga agcaaatcca gatcgaaatc tacggtcgtg 420

gtctcgactg ataatactac tagtacttca tcacttactt ctcgcccaag ttactcaaac 480
cctagcaagt ttcatagcta cggtcaaatc ccggagttta attccaactt gcccatcttg 540
cctcctctcc aaagccttgg agattacaat tcaagcaaca ctggattaga ttttggtgga 600
actcaaataa gcaacatgat aagtggtatg agttctagtg gtgggatctt ggatgcatgg 660
agaatacctc catcacaaca agctcagcaa ttccctttct tgatcaacac taccggattg 720
gtgcaatctt caaacgcgtt atatccatta ctagaaggcg gggttagcgc cacgcaaaca 780
agaaatgtga aggcggaaga gaatgatcag gatcggggta gggatgggga tggagtgaat 840
aacttatcaa gaaacttttt gggtaatatc aacataaact caggcaggaa cgaggaatac 900
acatcatggg gaggtaacag ttcttggacc ggtttcacct ccaacaactc aacaggccat 960
ctctcattct aa 972

<210> 220
<211> 323
<212> PRT
<213> Arabidopsis thaliana
<400> 220

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

<210> 221
<211> 1200
<212> DNA
<213> Arabidopsis thaliana
<400> 221

atggtttttt cttcatttcc tacttatcct gatcattcat caaactggca acaacaacat 60

caaccaatca caaccaccgt tggattcacg ggaaataaca tcaaccaaca gtttcttcct 120
caccatcccc tcccaccgca acagcaacaa acgcctccgc agcttcacca caacaacggt 180
aacggcggag tcgctgttcc cggtggacct ggcgggttaa tccgaccagg ttcgatggcg 240
gaaagagcaa ggctagccaa cataccatta cctgaaacag ccttgaagtg tccaagatgt 300
gactcaacta acaccaaatt ctgttacttc aacaactaca gtctcactca acctcgccac 360
ttctgcaaag catgccgtcg ttactggaca cgtggcggtg ctctaaggag cgttcccgtc 420
ggtggcggtt gccgtagaaa caaaagaacc aaaaacagca gcggtggagg tggcggtagc 480
accagtagcg gtaacagcaa gtcacaagac agcgccacga gcaacgacca ataccaccac 540
cgagccatgg ctaacaatca gatgggacca ccttcttcgt catcgtctct aagctcgttg 600
ctgtcttctt acaacgcagg gttaatcccc ggacatgatc ataacagcaa taacaacaac 660
atacttggac ttggatcatc tttgcctcct cttaagctta tgcctccttt agacttcaca 720
gacaacttca ccttacaata cggtgccgtt tcagctcctt cttatcatat aggcggtgga 780
agcagtggag gagcggcggc tcttttaaac ggttttgacc agtggagatt cccggcaaca 840
aaccaacttc ctttaggcgg tttagacccg tttgatcaac aacatcaaat ggagcagcag 900
aatccaggtt acggattggt taccgggtcg ggtcagtatc gacctaagaa cattttccat 960
aaccttatct cctcttcttc gtctgcttca tcagctatgg ttacagccac cgcgtcgcaa 1020
ttagcttcag tgaaaatgga agatagtaac aatcagctca acttgtctag acaacttttt 1080
ggagacgaac aacagctctg gaatattcat ggcgctgctg cagcatccac cgcagctgca 1140
acaagttcgt ggagtgaagt ctctaataat ttcagttctt cttctactag caatatataa 1200

<210> 222
<211> 399
<212> PRT
<213> Arabidopsis thaliana
<400> 222

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

305 310 315 320
Asn Leu Ile Ser Ser Ser Ser Ser Ala Ser Ser Ala Met Val Thr Ala
325 330 335
Thr Ala Ser Gln Leu Ala Ser Val Lys Met Glu Asp Ser Asn Asn Gln
340 345 350
Leu Asn Leu Ser Arg Gln Leu Phe Gly Asp Glu Gln Gln Leu Trp Asn
355 360 365
Ile His Gly Ala Ala Ala Ala Ser Thr Ala Ala Ala Thr Ser Ser Trp
370 375 380
Ser Glu Val Ser Asn Asn Phe Ser Ser Ser Ser Thr Ser Asn Ile
385 390 395

<210> 223
<211> 52
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm07315
<400> 223
ggggacaagt ttgtacaaaa aagcaggctt aaacaatggg tggatcgatg gc 52

<210> 224
<211> 50
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm07316
<400> 224
ggggaccact ttgtacaaga aagctgggtc gttaatgatc cgacaaaaca 50

<210> 225
<211> 2193
<212> DNA
<213> Oryza sativa
<400> 225
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 taataccccct 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> 226
<211> 1344
<212> DNA
<213> Arabidopsis thaliana
<220>
<221> misc_feature
<222> (196)..(196)
<223> n is a, c, g, ou t
<220>
<221> misc_feature
<222> (461)..(461)
<223> n is a, c, g, ou t
<400> 226
atgatgatgg agactagaga tccagctatt aagcttttcg gtatgaaaat ccctttccg 60
tcggtttttg aatcggcagt tacggtggag gatgacgaag aagatgactg gagcggcgga 120
gatgacaaat caccagagaa ggtaactcca gagttatcag ataagaacaa caacaactgt 180
aacgacaaca gttttnacaa ttcgaaaccc gaaaccttgg acaaagagga agcgacatca 240
actgatcaga tagagagtag tgacacgcct gaggataatc agcagacgac acctgatggt 300
aaaaccctaa agaaaccgac taagattcta ccgtgtccga gatgcaaaag catggagacc 360
aagttctgtt attacaacaa ctacaacata aaccagcctc gtcatttctg caaggcttgt 420
cagagatatt ggactgctgg agggactatg aggaatgttc ntgtgggggc aggacgtcgt 480
aagaacaaaa gcttatcttc tcattaccgt cacatcacta tttccgaggc tcttgaggct 540
gcgaggcttg acccgggctt acaggcaaac acaagggtct tgagttttgg tctcgaagct 600
cagcagcagc acgttgctgc tcccatgaca cctgtgatga agctacaaga agatcaaaag 660
gtctcaaacg gtgctaggaa caggtttcac gggttagcgg atcaacggct tgtagctcgg 720
gtagagaatg gagatgattg ctcaagcgga tcctctgtga ccacctctaa caatcactca 780
gtggatgaat caagagcaca aagcggcagt gttgttgaag cacaaatgaa caacaacaac 840
aataacatga atggttatgc ttgcatccca ggtgttccat ggccttacac gtggaatcca 900
gcgatgcctc caccaggttt ttacccgcct ccagggtatc caatgccgtt ttacccttac 960
tggaccatcc caatgctacc accgcatcaa tcctcatcgc ctataagcca aaagtgttca 1020
aatacaaact ctccgactct cggaaagcat ccgagagatg aaggatcatc gaaaaaggac 1080
aacgagacag agcgaaaaca gaaggccggg tgcgttctgg tcccgaaaac gttgagaata 1140
gatgatccta acgaagcagc aaagagctcg atatggacaa cattgggaat caagaacgag 1200
gcgatgtgca aagccggtgg tatgttcaaa gggtttgatc ataagacaaa gatgtataac 1260
aacgacaaag ctgagaactc ccctgttctt tctgctaacc ctgctgctct atcaagatca 1320
cacaatttcc atgaacagat ttag 1344

<210> 227
<211> 447
<212> PRT
<213> Arabidopsis thaliana
<220>
<221> INSEGURO
<222> (66)..(66)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> INSEGURO
<222> (154)..(154)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 227
Met Met Met Glu Thr Arg Asp Pro Ala Ile Lys Leu Phe Gly Met Lys
1 5 10 15
Ile Pro Phe Pro Ser Val Phe Glu Ser Ala Val Thr Val Glu Asp Asp
20 25 30
Glu Glu Asp Asp Trp Ser Gly Gly Asp Asp Lys Ser Pro Glu Lys Val
35 40 45
Thr Pro Glu Leu Ser Asp Lys Asn Asn Asn Asn Cys Asn Asp Asn Ser
50 55 60

```
Phe Xaa Asn Ser Lys Pro Glu Thr Leu Asp Lys Glu Glu Ala Thr Ser
65                  70                  75                  80
Thr Asp Gln Ile Glu Ser Ser Asp Thr Pro Glu Asp Asn Gln Gln Thr
                85                  90                  95
Thr Pro Asp Gly Lys Thr Leu Lys Lys Pro Thr Lys Ile Leu Pro Cys
            100                 105                 110
Pro Arg Cys Lys Ser Met Glu Thr Lys Phe Cys Tyr Tyr Asn Asn Tyr
        115                 120                 125
Asn Ile Asn Gln Pro Arg His Phe Cys Lys Ala Cys Gln Arg Tyr Trp
    130                 135                 140
Thr Ala Gly Gly Thr Met Arg Asn Val Xaa Val Gly Ala Gly Arg Arg
145                 150                 155                 160
Lys Asn Lys Ser Leu Ser Ser His Tyr Arg His Ile Thr Ile Ser Glu
                165                 170                 175
Ala Leu Glu Ala Ala Arg Leu Asp Pro Gly Leu Gln Ala Asn Thr Arg
            180                 185                 190
Val Leu Ser Phe Gly Leu Glu Ala Gln Gln Gln His Val Ala Ala Pro
        195                 200                 205
Met Thr Pro Val Met Lys Leu Gln Glu Asp Gln Lys Val Ser Asn Gly
    210                 215                 220
Ala Arg Asn Arg Phe His Gly Leu Ala Asp Gln Arg Leu Val Ala Arg
225                 230                 235                 240
Val Glu Asn Gly Asp Asp Cys Ser Ser Gly Ser Ser Val Thr Thr Ser
                245                 250                 255
Asn Asn His Ser Val Asp Glu Ser Arg Ala Gln Ser Gly Ser Val Val
            260                 265                 270
Glu Ala Gln Met Asn Asn Asn Asn Asn Asn Met Asn Gly Tyr Ala Cys
        275                 280                 285
Ile Pro Gly Val Pro Trp Pro Tyr Thr Trp Asn Pro Ala Met Pro Pro
    290                 295                 300
Pro Gly Phe Tyr Pro Pro Pro Gly Tyr Pro Met Pro Phe Tyr Pro Tyr
305                 310                 315                 320
Trp Thr Ile Pro Met Leu Pro Pro His Gln Ser Ser Ser Pro Ile Ser
                325                 330                 335
Gln Lys Cys Ser Asn Thr Asn Ser Pro Thr Leu Gly Lys His Pro Arg
            340                 345                 350
Asp Glu Gly Ser Ser Lys Lys Asp Asn Glu Thr Glu Arg Lys Gln Lys
        355                 360                 365
Ala Gly Cys Val Leu Val Pro Lys Thr Leu Arg Ile Asp Asp Pro Asn
    370                 375                 380
Glu Ala Ala Lys Ser Ser Ile Trp Thr Thr Leu Gly Ile Lys Asn Glu
385                 390                 395                 400
Ala Met Cys Lys Ala Gly Gly Met Phe Lys Gly Phe Asp His Lys Thr
                405                 410                 415
Lys Met Tyr Asn Asn Asp Lys Ala Glu Asn Ser Pro Val Leu Ser Ala
            420                 425                 430
Asn Pro Ala Ala Leu Ser Arg Ser His Asn Phe His Glu Gln Ile
        435                 440                 445
<210> 228
<211> 63
<212> PRT
<213> Seqüência artificial
<220>
<223> Domínio DOF
<220>
<221> INSEGURO
<222> (50)..(50)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 228
Lys Pro Thr Lys Ile Leu Pro Cys Pro Arg Cys Lys Ser Met Glu Thr
1               5                   10                  15
Lys Phe Cys Tyr Tyr Asn Asn Tyr Asn Ile Asn Gln Pro Arg His Phe
            20                  25                  30
```

Cys Lys Ala Cys Gln Arg Tyr Trp Thr Ala Gly Gly Thr Met Arg Asn
35 40 45
Val Xaa Val Gly Ala Gly Arg Arg Lys Asn Lys Ser Leu Ser Ser
50 55 60

<210> 229
<211> 11
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 1
<400> 229
Lys Ala Leu Lys Lys Pro Asp Lys Ile Leu Pro
1 5 10

<210> 230
<211> 14
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 2
<400> 230
Asp Asp Pro Gly Ile Lys Leu Phe Gly Lys Thr Ile Pro Phe
1 5 10

<210> 231
<211> 11
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 3
<400> 231
Ser Pro Thr Leu Gly Lys His Ser Arg Asp Glu
1 5 10

<210> 232
<211> 16
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 4
<400> 232
Leu Gln Ala Asn Pro Ala Ala Leu Ser Arg Ser Gln Asn Phe Gln Glu
1 5 10 15

<210> 233
<211> 32
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 5
<400> 233
Lys Gly Glu Gly Cys Leu Trp Val Pro Lys Thr Leu Arg Ile Asp Asp
1 5 10 15
Pro Asp Glu Ala Ala Lys Ser Ser Ile Trp Thr Thr Leu Gly Ile Lys
20 25 30

<210> 234
<211> 1263
<212> DNA
<213> Oryza sativa
<400> 234
atgaacaacg ttgaagagaa agcagcatca gactcaaaag atgaaaatga aaagacagca 60
aatgatgaat caggccagga caaagtactt aagaagccag ataagattct cccttgccct 120
cggtgcaaca gtatggacac aaagttttgt tattacaaca actacaatgt taatcaaccc 180
aggcacttct gtaagaactg ccaaaggtat tggactgccg ggggaaccat gagaaatgta 240
cctgttggtg ctgggaggcg caaaagcaag agctcatcgt tgcactaccg tcacttactg 300
atggcccctg attgcatgat ggggtctaga gtggaaatat ccaagtcaat gaaccctgaa 360
gctttcgcat ctgcgcattc gacccctata caaccaattg gcagaaacga aacagttctc 420

```
aaatttgggc ctgaggtgcc actctgtgaa tcgatggcat cagtgctgaa cattcaggag    480
cagaatggaa ccaatgctgc agcagtacca acgggtgaaa atcaggaaga taactcttgc    540
atctcttcaa tcacatcaca caacgtgtta cctgaaaatg cagcccaagt tgacaagaac    600
agcacgccgg tgtattgcaa cggagtcggc ccagtgccgc agtactacct tggagctcct    660
tacatgtacc catggaacat aggatggaac aacgttccta tgatggtgcc aggtacaagc    720
atgccagagt ctgcttccca atctgagagc tgcagcacca gttcagctcc atggatgaac    780
atgaactccc ccatgatgcc ggttgcctcg aggctttctg caccaccatt tccataccct    840
ctagtgccac ctgcactatg gggttgctta tccagctggc cggccacggc atggaacata    900
ccgtggatca gaacgaatgg cggctgcatg tctccatcgt cgtcgagcaa cagcagctgt    960
tcaggcaatg gctcccctct ggggaagcat tccagggact cctctctccc actgaaggag   1020
gacaaggagg agaaatcact gtgggttccc aagacgctcc gcatcgacga tcccgacgag   1080
gcggcgaaga gctccatctg ggccaccctg gggatcaagc ctggagaccc tggcatcttc   1140
aagccgttcc agtccaaagg tgagagcaaa ggccaagcag catcagagac tcgtcctgct   1200
cgtgctctta aggcaaaccc agctgcattg tcgcggtcgc agtcgttcca ggagacttct   1260
tga                                                                 1263
<210> 235
<211> 420
<212> PRT
<213> Oryza sativa
<400> 235
Met Asn Asn Val Glu Glu Lys Ala Ala Ser Asp Ser Lys Asp Glu Asn
1               5                  10                  15
Glu Lys Thr Ala Asn Asp Glu Ser Gly Gln Asp Lys Val Leu Lys Lys
            20                  25                  30
Pro Asp Lys Ile Leu Pro Cys Pro Arg Cys Asn Ser Met Asp Thr Lys
        35                  40                  45
Phe Cys Tyr Tyr Asn Asn Tyr Asn Val Asn Gln Pro Arg His Phe Cys
    50                  55                  60
Lys Asn Cys Gln Arg Tyr Trp Thr Ala Gly Gly Thr Met Arg Asn Val
65                  70                  75                  80
Pro Val Gly Ala Gly Arg Arg Lys Ser Lys Ser Ser Ser Leu His Tyr
                85                  90                  95
Arg His Leu Leu Met Ala Pro Asp Cys Met Met Gly Ser Arg Val Glu
            100                 105                 110
Ile Ser Lys Ser Met Asn Pro Glu Ala Phe Ala Ser Ala His Ser Thr
        115                 120                 125
Pro Ile Gln Pro Ile Gly Arg Asn Glu Thr Val Leu Lys Phe Gly Pro
    130                 135                 140
Glu Val Pro Leu Cys Glu Ser Met Ala Ser Val Leu Asn Ile Gln Glu
145                 150                 155                 160
Gln Asn Gly Thr Asn Ala Ala Ala Val Pro Thr Gly Glu Asn Gln Glu
                165                 170                 175
Asp Asn Ser Cys Ile Ser Ser Ile Thr Ser His Asn Val Leu Pro Glu
            180                 185                 190
Asn Ala Ala Gln Val Asp Lys Asn Ser Thr Pro Val Tyr Cys Asn Gly
        195                 200                 205
Val Gly Pro Val Pro Gln Tyr Tyr Leu Gly Ala Pro Tyr Met Tyr Pro
    210                 215                 220
Trp Asn Ile Gly Trp Asn Asn Val Pro Met Met Val Pro Gly Thr Ser
225                 230                 235                 240
Met Pro Glu Ser Ala Ser Gln Ser Glu Ser Cys Ser Thr Ser Ser Ala
                245                 250                 255
Pro Trp Met Asn Met Asn Ser Pro Met Met Pro Val Ala Ser Arg Leu
            260                 265                 270
Ser Ala Pro Pro Phe Pro Tyr Pro Leu Val Pro Pro Ala Leu Trp Gly
        275                 280                 285
Cys Leu Ser Ser Trp Pro Ala Thr Ala Trp Asn Ile Pro Trp Ile Arg
    290                 295                 300
Thr Asn Gly Gly Cys Met Ser Pro Ser Ser Ser Ser Asn Ser Ser Cys
305                 310                 315                 320
Ser Gly Asn Gly Ser Pro Leu Gly Lys His Ser Arg Asp Ser Ser Leu
                325                 330                 335
Pro Leu Lys Glu Asp Lys Glu Glu Lys Ser Leu Trp Val Pro Lys Thr
```

```
            340                 345                 350
Leu Arg Ile Asp Asp Pro Asp Glu Ala Ala Lys Ser Ser Ile Trp Ala
        355                 360                 365
Thr Leu Gly Ile Lys Pro Gly Asp Pro Gly Ile Phe Lys Pro Phe Gln
    370                 375                 380
Ser Lys Gly Glu Ser Lys Gly Gln Ala Ala Ser Glu Thr Arg Pro Ala
385                 390                 395                 400
Arg Ala Leu Lys Ala Asn Pro Ala Ala Leu Ser Arg Ser Gln Ser Phe
                405                 410                 415
Gln Glu Thr Ser
            420
<210> 236
<211> 1365
<212> DNA
<213> Oryza sativa
<400> 236
atgcctaatt taggaaatgg tgtaaaaacc aataatgact tacctttagt ctcagacaag     60
cttttgattg tcaaaggtat tccttttttgt cccaacaata gtaagaaaaa tgatttacaa    120
ggcatcagca gaccggatgg aagaatagaa atcgattcca tgactgagga tgttaagact    180
gagcctgatg gatctgtccc tgagaagata ctcaagaagc cagataagat tctgccatgt    240
ccacgctgca atagcatgga aacaaagttc tgctacttca acaactacaa tgttcaccag    300
cccaggcact tctgcaggaa ctgccaaaga tattggaccg ctggtggagc tatgaggaat    360
gtcccagttg gtgctggaag acgcagaaat aagcatgtgt caaaatactg tcaggcgatg    420
atgacgtgca ataatactgt agctcctgga gatgtttctg atgtggttca ccatcaggtt    480
attacacatg gatcttctct ccttccagca acactgaagg aaaatgaaac acctacagaa    540
ttcatatcag aagtaccacc atgtaagtct tcagcttcaa tccttgatat tggagagccg    600
aatgatactg accttgttcc cttggcctct ggtgataaca aggaagaaaa atcatgtgca    660
tcttctgtgg tagtatccag ctgttcagag aatctgatgc cagataatgc aattatgaaa    720
gagccaaaca acaggtcagg atgttgtaat ggtgtggcat tgcctttccc tactggacct    780
gctttggtgc tcccctggag tcttggatgg aacagtgttg ctctcatgcc agctacccag    840
tgctcgatgc agcccgttct tgggttaaaa gatgggatac cctgcccgcc ttcatggcca    900
ccgcaactga tggtgccggc cccaggaatc tgtactcctg ttgttccaat ccctcttgtg    960
ccacctctgt ggagctgctt ccctggctgg cctaatggaa tgtggaatgc acaatgccct   1020
ggaggtaata ctactgtgct gccgtcaacc gctccaaaca aaatttcttg ttcaggaagc   1080
agttctctgg tgttgggaaa gcattcaaga gaagaaagct tgcaggaaga agagaagacg   1140
agaaattact tatgggtacc taaaactctt aggattgatg atccagctga ggctgcaaag   1200
agttcaatct gggcgaccct tggcatcaag cctgacgata aaggcatatt caagtctttc   1260
cagccaaatg ttgcgaaaaa tggcacggca ccagaatcgc ctcaggctct gcaggccaat   1320
ccagcagcat tttcgcgttc tcaatcgttt caagagacga cttga                   1365
<210> 237
<211> 454
<212> PRT
<213> Oryza sativa
<400> 237
Met Pro Asn Leu Gly Asn Gly Val Lys Thr Asn Asn Asp Leu Pro Leu
1               5                   10                  15
Val Ser Asp Lys Leu Leu Ile Val Lys Gly Ile Pro Phe Cys Pro Asn
            20                  25                  30
Asn Ser Lys Lys Asn Asp Leu Gln Gly Ile Ser Arg Pro Asp Gly Arg
        35                  40                  45
Ile Glu Ile Asp Ser Met Thr Glu Asp Val Lys Thr Glu Pro Asp Gly
    50                  55                  60
Ser Val Pro Glu Lys Ile Leu Lys Lys Pro Asp Lys Ile Leu Pro Cys
65                  70                  75                  80
Pro Arg Cys Asn Ser Met Glu Thr Lys Phe Cys Tyr Phe Asn Asn Tyr
                85                  90                  95
Asn Val His Gln Pro Arg His Phe Cys Arg Asn Cys Gln Arg Tyr Trp
            100                 105                 110
Thr Ala Gly Gly Ala Met Arg Asn Val Pro Val Gly Ala Gly Arg Arg
        115                 120                 125
Arg Asn Lys His Val Ser Lys Tyr Cys Gln Ala Met Met Thr Cys Asn
    130                 135                 140
Asn Thr Val Ala Pro Gly Asp Val Ser Asp Val Val His His Gln Val
```

```
145                 150                 155                 160
Ile Thr His Gly Ser Ser Leu Leu Pro Ala Thr Leu Lys Glu Asn Glu
                165                 170                 175
Thr Pro Thr Glu Phe Ile Ser Glu Val Pro Pro Cys Lys Ser Ser Ala
            180                 185                 190
Ser Ile Leu Asp Ile Gly Glu Pro Asn Asp Thr Asp Leu Val Pro Leu
        195                 200                 205
Ala Ser Gly Asp Asn Lys Glu Glu Lys Ser Cys Ala Ser Ser Val Val
    210                 215                 220
Val Ser Ser Cys Ser Glu Asn Leu Met Pro Asp Asn Ala Ile Met Lys
225                 230                 235                 240
Glu Pro Asn Asn Arg Ser Gly Cys Cys Asn Gly Val Ala Leu Pro Phe
                245                 250                 255
Pro Thr Gly Pro Ala Leu Val Leu Pro Trp Ser Leu Gly Trp Asn Ser
            260                 265                 270
Val Ala Leu Met Pro Ala Thr Gln Cys Ser Met Gln Pro Val Leu Gly
        275                 280                 285
Leu Lys Asp Gly Ile Pro Cys Pro Pro Ser Trp Pro Pro Gln Leu Met
    290                 295                 300
Val Pro Ala Pro Gly Ile Cys Thr Pro Val Val Pro Ile Pro Leu Val
305                 310                 315                 320
Pro Pro Leu Trp Ser Cys Phe Pro Gly Trp Pro Asn Gly Met Trp Asn
                325                 330                 335
Ala Gln Cys Pro Gly Gly Asn Thr Thr Val Leu Pro Ser Thr Ala Pro
            340                 345                 350
Asn Lys Ile Ser Cys Ser Gly Ser Ser Ser Leu Val Leu Gly Lys His
        355                 360                 365
Ser Arg Glu Glu Ser Leu Gln Glu Glu Glu Lys Thr Arg Asn Tyr Leu
    370                 375                 380
Trp Val Pro Lys Thr Leu Arg Ile Asp Asp Pro Ala Glu Ala Ala Lys
385                 390                 395                 400
Ser Ser Ile Trp Ala Thr Leu Gly Ile Lys Pro Asp Asp Lys Gly Ile
                405                 410                 415
Phe Lys Ser Phe Gln Pro Asn Val Ala Lys Asn Gly Thr Ala Pro Glu
            420                 425                 430
Ser Pro Gln Ala Leu Gln Ala Asn Pro Ala Ala Phe Ser Arg Ser Gln
        435                 440                 445
Ser Phe Gln Glu Thr Thr
    450
<210> 238
<211> 1461
<212> DNA
<213> Oryza sativa
<400> 238
atgggagggg gtggtaagtg caaaagaggg aaaagaggaa agattgcagc caagcgaaga     60
aggggccagt gttgcagcag cagcagcagc agcggaggcg caagaagacg agctagagag    120
agagaatcag aggagagctt gggtgagatg ggtgagtgca gaggaggagg aggaggagga    180
gatgggctga tcaagctgtt cgggaagacg atcccggtgc agccggatgc caaggatgtt    240
cagcagcata gtggcagtag cagcagctca accgaatccg acgtccaaga aaccgccgct    300
gtcgccgtcg ccgacccctc cccgcggtcg gaggtcgtcg acggcgagag cccgccgcag    360
ccgggcggcg aggcggcgag ccatcagcag cagcagaagg agatgaagct gaagaagccg    420
gacaagatcc tgccatgccc gcggtgcagc agcatggaca ccaagttctg ctacttcaac    480
aactacaacg tcaaccagcc tcgccacttc tgcaagcact gccagcgcta ctggaccgcc    540
ggcggcgcca tgcgcaacgt ccccgtcggc gccggccgcc gcaagaacaa gaacgccacc    600
gccgccgccc acttcctcca ccgcgtccgc gcctgcgccg ccgccgccgc catgcccgcg    660
gcgccccacg acgccaccaa cgccaccgtg ctcagcttcg gcggcggcgg aggcggacac    720
gacgcgccgc cggtcaccct ggacctcgcc gacaagatga cgcgcctcgg caaggagggg    780
ctcgtcgccc atgcccggaa cgccgacgcc gccgccgcgt gcagcgaggt gtcgagcaac    840
agggacgacg agcagatcgg caacactgta gcaaaacctg caaacgggtt gcagcagcat    900
cctcctcctc ctcatcatca tcatcattca gccatgaacg gtggcggcat ctggccctac    960
tacacctcgg ggatcgcgat cccgatatac ccggcggcgc cggcgtactg gggctgcatg   1020
attccacctc ctggagcttg gagcctccca tggccggcca cagtccagtc tcaggccatc   1080
tcatcatcat caccacctac aagtgctaca ccttcagtct cctccttcac actaggcaag   1140
```

catcccagag agggtggtga tcatgaggca agagatcacc atggcaatgg taaagtgtgg 1200
gtgccgaaga cgatccggat cgacaacgcc gacgaggttg cccggagctc aatccggtca 1260
ctcttcgcct tcagaggcgg cgacaaggcg gatgataaca acgacgacga tggcaccggc 1320
gtgcacaagc tcgccaccac ggtgttcgag ccaaagaggg acagcaagac ggcgaaacat 1380
ccggcgatca cgagcttgcc gctcttgcac accaaccccg tcgcgcttac ccgatccgcg 1440
accttccagg agggatcttg a 1461

<210> 239
<211> 486
<212> PRT
<213> Oryza sativa
<400> 239

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

```
Ser Ile Arg Ser Leu Phe Ala Phe Arg Gly Gly Asp Lys Ala Asp Asp
            420                 425                 430
Asn Asn Asp Asp Asp Gly Thr Gly Val His Lys Leu Ala Thr Thr Val
        435                 440                 445
Phe Glu Pro Lys Arg Asp Ser Lys Thr Ala Lys His Pro Ala Ile Thr
    450                 455                 460
Ser Leu Pro Leu Leu His Thr Asn Pro Val Ala Leu Thr Arg Ser Ala
465                 470                 475                 480
Thr Phe Gln Glu Gly Ser
                485
<210> 240
<211> 1461
<212> DNA
<213> Oryza sativa
<400> 240
atgggagggg gtggtaagtg caaaagaggg aaaagaggaa agattgcagc caagcgaaga     60
aggggccagt gttgcagcag cagcagcagc agcggaggcg caagaagacg agctagagag    120
agagaatcag aggagagctt gggtgagatg ggtgagtgca gaggaggagg aggaggagga    180
gatgggctga tcaagctgtt cgggaagacg atcccggtgc agccggatgc caaggatgtt    240
cagcagcata gtggcagtag cagcagctca accgaatccg acgtccaaga aaccgccgct    300
gtcgccgtcg ccgacccctc cccgcggtcg gaggtcgtcg acggcgagag cccgccgcag    360
ccgggcggcg aggcggcgag ccatcagcag cagcagaagg agatgaagct gaagaagccg    420
gacaagatcc tgccatgccc gcggtgcagc agcatggaca ccaagttctg ctacttcaac    480
aactacaacg tcaaccagcc tcgccacttc tgcaagcact gccagcgcta ctggaccgcc    540
ggcggcgcca tgcgcaacgt ccccgtcggc gccggccgcc gcaagaacaa gaacgccacc    600
gccgccgccc acttcctcca ccgcgtccgc gcctgcgccg ccgccgccgc catgcccgcg    660
gcgccccacg acgccaccaa cgccaccgtg ctcagcttcg gcggcggcgg aggcggacac    720
gacgcgccgc cggtcaccct ggacctcgcc gacaagatga cgcgcctcgg caaggagggg    780
ctcgtcgccc atgcccggaa cgccgacgcc gccgccgcgt gcagcgaggt gtcgagcaac    840
agggacgacg agcagatcgg caacactgta gcaaaacctg caaacgggtt gcagcagcat    900
cctcctcctc ctcatcatca tcatcattca gccatgaacg gtggcggcat ctggccctac    960
tacacctcgg ggatcgcgat cccgatatac ccggcggcgc cggcgtactg gggctgcatg   1020
attccacctc ctggagcttg gagcctccca tggccggcca cagtccagtc tcaggccatc   1080
tcatcatcat caccacctac aagtgctaca ccttcagtct cctccttcac actaggcaag   1140
catcccagag agggtggtga tcatgaggca agagatcacc atggcaatgg taaagtgtgg   1200
gtgccgaaga cgatccggat cgacaacgcc gacgaggttg cccggagctc aatccggtca   1260
ctcttcgcct tcagaggcgg cgacaaggcg gatgataaca acgacgacga tggcaccggc   1320
gtgcacaagc tcgccaccac ggtgttcgag ccaaagaggg acagcaagac ggcgaaacat   1380
ccggcgatca cgagcttgcc gctcttgcac accaaccccg tcgcgcttac ccgatccgcg   1440
accttccagg agggatcttg a                                             1461
<210> 241
<211> 440
<212> PRT
<213> Oryza sativa
<400> 241
Met Gly Glu Cys Lys Val Gly Gly Gly Gly Gly Gly Gly Asp Cys Leu
1               5                  10                  15
Ile Lys Leu Phe Gly Lys Thr Ile Pro Val Pro Glu Pro Gly Ala Cys
            20                  25                  30
Ala Ala Gly Asp Val Asp Lys Asp Leu Gln His Ser Gly Ser Ser Thr
        35                  40                  45
Thr Glu Pro Lys Thr Gln Glu Asn Thr Val Gln Asp Ser Thr Ser Pro
    50                  55                  60
Pro Pro Gln Pro Glu Val Val Asp Thr Glu Asp Ser Ser Ala Asp Lys
65                  70                  75                  80
Asn Ser Ser Glu Asn Gln Gln Gln Gln Gly Asp Thr Ala Asn Gln Lys
                85                  90                  95
Glu Lys Leu Lys Lys Pro Asp Lys Ile Leu Pro Cys Pro Arg Cys Ser
            100                 105                 110
Ser Met Asp Thr Lys Phe Cys Tyr Tyr Asn Asn Tyr Asn Ile Asn Gln
        115                 120                 125
Pro Arg His Phe Cys Lys Asn Cys Gln Arg Tyr Trp Thr Ala Gly Gly
    130                 135                 140
```

```
Ala Met Arg Asn Val Pro Val Gly Ala Gly Arg Arg Lys Ser Lys Ser
145                 150                 155                 160
Val Ser Ala Ala Ser His Phe Leu Gln Arg Val Arg Ala Ala Leu Pro
                165                 170                 175
Gly Asp Pro Pro Leu Tyr Ala Pro Val Lys Thr Asn Gly Thr Val Leu
            180                 185                 190
Ser Phe Gly Ser Asp Leu Ser Thr Leu Asp Leu Thr Glu Gln Met Lys
        195                 200                 205
His Leu Lys Asp Lys Phe Ile Pro Thr Thr Gly Ile Lys Asn Thr Asp
    210                 215                 220
Glu Met Pro Val Gly Leu Cys Ala Glu Gly Leu Ser Lys Thr Glu Glu
225                 230                 235                 240
Ser Asn Gln Thr Asn Leu Lys Glu Lys Val Ser Ala Asp Arg Ser Pro
                245                 250                 255
Asn Val Ala Gln His Pro Cys Met Asn Gly Gly Ala Met Trp Pro Phe
            260                 265                 270
Gly Val Ala Pro Pro Pro Ala Tyr Tyr Thr Ser Ser Ile Ala Ile Pro
        275                 280                 285
Phe Tyr Pro Ala Ala Ala Ala Ala Val Ala Ala Tyr Trp Gly Cys Met
    290                 295                 300
Val Pro Gly Ala Trp Asn Ala Pro Trp Pro Pro Gln Ser Gln Ser Gln
305                 310                 315                 320
Ser Val Ser Ser Ser Ser Ala Ala Ser Pro Val Ser Thr Met Thr Asn
                325                 330                 335
Cys Phe Arg Leu Gly Lys His Pro Arg Asp Gly Asp Glu Glu Leu Asp
            340                 345                 350
Ser Lys Gly Asn Gly Lys Val Trp Val Pro Lys Thr Val Arg Ile Asp
        355                 360                 365
Asp Val Asp Glu Val Ala Arg Ser Ser Ile Trp Ser Leu Ile Gly Ile
    370                 375                 380
Lys Gly Asp Lys Val Gly Ala Asp His Gly Arg Gly Cys Lys Leu Ala
385                 390                 395                 400
Lys Val Phe Glu Ser Lys Asp Glu Ala Lys Ala Ser Thr His Thr Ala
                405                 410                 415
Ile Ser Ser Leu Pro Phe Met Gln Gly Asn Pro Ala Ala Leu Thr Arg
            420                 425                 430
Ser Val Thr Phe Gln Glu Gly Ser
        435                 440
<210> 242
<211> 1803
<212> DNA
<213> Oryza sativa
<400> 242
acttgtcgga tctctctctc tctctctctc tccctctctc tctcgcccat gtaccatctc     60
catggacgac ctcgccgccg cctcccctcc tcacccgccg ccgccgccgc cggaatccca    120
cgtgcctccg cccccgcaga cgccggagaa ggattcatgt gaggatacag gagacatgag    180
gatcagtgaa gaaaagccat gcacagatca ggagttagat gctgatcaga tgaatagttc    240
tagctttaat agttccagcg agtgtgagaa tcaaacacct agcaatgatg aaatgactgg    300
atcagagtcc aaatctgagg cagctcaaac agagggtggt ggatcgagtg aagagaaggt    360
cctgaagaag ccagacaaga tcctgccttg tcctcgttgc aacagcatgg atacaaagtt    420
ctgctactac aacaactaca acattaacca gccaagacat ttttgcaaga gttgtcagag    480
atattggacg gcaggtggaa gcatgaggaa cctccctgtt ggtgctggta ggcgcaagag    540
taagagctcc actgcaaatt accgcagtat attaatcacg ggcagcaatc tagctgctcc    600
tgctggagat gctccccctct atcaactctc tataaaagga gatcaaacag caacggcagt    660
taaatttgca cctgattccc cactctgtaa ttccatggcc tctgtgctga aaattggaga    720
gcagagtaag aatgccaagc ctacctcaac agcacaaccc agaaatggag aaacccagac    780
ctgcccggct tcaggaacaa cttcagatag tccccggaat gaaccagtta atggagcagt    840
tagtgggcat caaaatggaa ttgttgggca tagtggagtc cctcccatgc atcccatacc    900
atgcttccct ggtcctcctt ttgtgtaccc atggagtcca gcatggaatg gcattcctgc    960
catggcacca ccggtatgca cagcaccagc tgaaccagca aattcttcag acaatggaag   1020
cacagctagt gttcagtgga gcatgccacc agtgatgccg gtaccaggat actttccggt   1080
aattccatct tcagtttggc ccttcatttc tccctggcca aatggtgcat ggagctcgcc   1140
atggattcaa cctaattgca gcgtgtcagc ttcatctcct acaagcacta gtacatgttc   1200
```

agacaacggc tctcctgtcc taggaaagca ctccagggac tccaaaccgc aaggagatga 1260
caaggcagag aaaaacttgt ggattccgaa aacgcttcgg atcgatgatc ctgacgaagc 1320
tgcaaagagc tcaatctgga caacccttgg cattgaacct ggtgaccgta gcatgttcag 1380
atcattccag tcgaaacctg aaagcaggga gcagatatcc ggtgctgcac gagtcctgca 1440
ggcgaatcca gcagctctat ctcgatctca atctttccag gagacaacgt gatgtatatt 1500
gaagaaatcg tgtgacaatt gtagaacatg ttactactat aatttaagca agtgctacag 1560
cctgaaggat gattggcgat gtaggcgctg ctgcaaacat ggaggcagcg tgatctgtac 1620
tattgaaacc tgaagagtac tattcttgac atttttctat aatttgcctg tggagcatgc 1680
accagtctac tgggttcatg atcaggctgt cagcatatgt tttgttgtac aataataata 1740
attgttaata gctttgccta aaggatactg atagtattct gctgatcctt cagctctgtg 1800
atc 1803

<210> 243
<211> 476
<212> PRT
<213> Oryza sativa
<400> 243

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

```
    370                 375                 380
Pro Val Leu Gly Lys His Ser Arg Asp Ser Lys Pro Gln Gly Asp Asp
385                 390                 395                 400
Lys Ala Glu Lys Asn Leu Trp Ile Pro Lys Thr Leu Arg Ile Asp Asp
                405                 410                 415
Pro Asp Glu Ala Ala Lys Ser Ser Ile Trp Thr Thr Leu Gly Ile Glu
            420                 425                 430
Pro Gly Asp Arg Ser Met Phe Arg Ser Phe Gln Ser Lys Pro Glu Ser
        435                 440                 445
Arg Glu Gln Ile Ser Gly Ala Ala Arg Val Leu Gln Ala Asn Pro Ala
    450                 455                 460
Ala Leu Ser Arg Ser Gln Ser Phe Gln Glu Thr Thr
465                 470                 475
<210> 244
<211> 1817
<212> DNA
<213> Hordeum vulgare
<400> 244
ctagacatga ttgccaagct cgaaaataac cctcacaaag ggaacaaaac tggtaccggg     60
ccccccctcg aggtcgacaa ggtcggtgtc catgcccccc ccccctcctc ctttgtgttc    120
ttgtgtatga aatcccggga attttgattg gtttctaata gagcctgtga aaaaaaaaaa    180
ggccatttcg aggcctgctt tctttggctt tggagccgat catggcgacg agtcgtgttt    240
ggtctagtca tcaggtctta agcatggagt tccctgcctg ttttgggagg tactagtaag    300
cgctaagcag gcaggtttag cttccattag caagggaatt tcgcctttgc tggtctctaa    360
taataatagc agtttcaaca gttttagtat tgcatgatct agtcgtctga acatctggac    420
ttctgaaaga caggcctttc taggttatta tcttgcctct tttaataggt cgttttcgcg    480
tgctcccaag gattttgctg aaagtggcta ctgatatgtg ttgccgtaac ctactggact    540
tcagaagttc agatgcgtcc gactctgagt acatcaccca ctgcaaatca aagcaccata    600
gcttcttttt ttcattaaca agatataaga actgggaata tgaacaaact ttgtcaaatt    660
ttacacttgg gtcagatggg ggaaagcatg ttgtaatcac cggtcaataa cccgaatatt    720
acaccgtgtt tgtggacctc taattcacaa caaaattatt tcaaagttag actatatata    780
tatacaagga actcatcgag attttctgaa tgtttacagg accttcagca cagaggaagc    840
accacggctg aaccgaaagt acaagaaagc gccccacagg actccacggg ctcgcctccg    900
cagccggagg ttgtggacgt cgaggatcca tctgctgtca agaactcagc agcagatcaa    960
caagaagaag aagaagaaca gggtgacacg gccaaccaga aggagaagct caagaagcct   1020
gacaagatcc tgccctgccc acgctgtagc agcatggaca ccaagttctg ctactacaac   1080
aactacaaca tcaaccagcc gcgccacttc tgcaagaact gccagaggta ctggacggcc   1140
ggcggggcca tgcgcaacgt gcccgtgggc gcgggccgcc gcaagagcaa gagcatatcg   1200
gcagcgtccc acttccttca gaggataagg gccgctctgc ccggtgatcc tctctgcacc   1260
ccggtcaaca ctaacggcac ggtgctcagc ttcggctccg acgcatccac cttggacgtc   1320
gtctcagaac agatgaagca catgaaggag ctcagctcag taacccggac cgagaacacc   1380
gatgccccgt cagtagggtc ttgtgctgaa ggatgggcaa agggagaaga gtcgagccag   1440
atgaactcga gggagagagt tgcagcagat agatcgccaa attttgcgca gcacccgtgc   1500
atgaacgggg cagccatgtg gccattcagt tgtgcaccat cgcctgccta tttcacccca   1560
aacgtagcaa ttccattcta tccagctgct gctgctgcct actggggctg catggttccg   1620
ggagcctgga acactccatg gcagccgcag ccgcagccgc agtctcagtg tcaatctagc   1680
tcaccaccta gtgccgcttc tccggtatca acaatgtcca gttgcttcca atcccggaag   1740
catcctagag atggtgatga ggaaagagat accaagggta atggcaaggt gtgggtgccc   1800
aagacgatcg agtcgac                                                  1817
<210> 245
<211> 270
<212> PRT
<213> Hordeum vulgare
<400> 245
Leu Lys Lys Pro Asp Lys Ile Leu Pro Cys Pro Arg Cys Ser Ser Met
1               5                   10                  15
Asp Thr Lys Phe Cys Tyr Tyr Asn Asn Tyr Asn Ile Asn Gln Pro Arg
            20                  25                  30
His Phe Cys Lys Asn Cys Gln Arg Tyr Trp Thr Ala Gly Gly Ala Met
        35                  40                  45
Arg Asn Val Pro Val Gly Ala Gly Arg Arg Lys Ser Lys Ser Ile Ser
    50                  55                  60
Ala Ala Ser His Phe Leu Gln Arg Ile Arg Ala Ala Leu Pro Gly Asp
```

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

<210> 246
<211> 1485
<212> DNA
<213> Cucurbita maxima
<400> 246
cagctgcaaa agaggaaact cttcatgatt cagaggatta tgcatgtgta aaaacagcaa 60
atgaggcgca catgaatcct gaagtgttat caacagatga aaatgataag ctcgcaacaa 120
ggaaaactga gaaagaacag aatgatgccc ccaactcgaa agagaaactg aagaaaccag 180
ataagatact tccatgtccc cgctgcaata gcatggaaac caagttttgt tattataata 240
attataatgt caatcaacca cgccattttt gcaaagcctg tcaaagatat tggactgaag 300
gcggtaccat caggaatgtc cctgttggag ctggccgccg aaaaaacaag aactcagcct 360
cacactaccg acacattaca atctcagagg ctctccgagc tgcacaaatt gatgttccta 420
ttgaggtcaa ccacctagca tcaaaaggca atggacgagt cctcaatttc agtgtaagcc 480
cacctgtatg tgaatctatg gtcaatgtat cacatcctgc agaaagaaag gtcttgaatg 540
gaacaaggaa tgaatttgag ggagccaaag gaccttgtga gggtggggaa actggtgatg 600
attgttctag tgcatcttca gtaacaatgt caagctcaat gaagaatgga gccaggaggt 660
tccctcaaga accacatatg cagaacatca atggttttcc ttctcaaatc ccatgtcttc 720
ctggtgttcc ttggccttgt tcatggactg caccaatacc tccaccagcc ttgtgccctc 780
ctggagttcc tttatcattc taccctgcaa catattggag ttgcagtgct tcaggttctt 840
ggaatattcc ttgggtcact ccacaaccct gtcctccaat ccctggtcca aattctccga 900
cgctaggcaa gcattcgaga gatggcgatg aactccaggc tgataattct gagatgaaag 960
atcctccaaa acagaaaaat ggatctgttt tggttcccaa aactttaagg attgatgatc 1020
caaacgaagc tgctaagagt tcaatatggg agacacttgg tattaagaat gattcaatca 1080
aagctgttga tctgtctaat gttttccaat caaagggcga cctaaagagt aacgtttctg 1140
aagtgttgtc tccagttttg caagccaacc ctgcagcctt gtcaagatct cttactttcc 1200
acgagcggtc gtgaatggta tttttgattt tccaggttca ttgtaaaaga aatctttaag 1260
ttaagaagct taagcatgaa gtggatacaa gcatcaatct caagcttctt ctgctcaaga 1320
agcaatccag tttttcaaga atcccggaat tgttctctct atagaaagag tagctgccat 1380
tatcaagcca tccttttgta cataacttta catatgtgaa taaccgatgt agagtggcag 1440
ctaaagattc tcagatgtat aaaaatgcag ccaaaagttt gcttc 1485

<210> 247
<211> 380
<212> PRT
<213> Cucurbita maxima
<400> 247
Met Asn Pro Glu Val Leu Ser Thr Asp Glu Asn Asp Lys Leu Ala Thr
1 5 10 15
Arg Lys Thr Glu Lys Glu Gln Asn Asp Ala Pro Asn Ser Lys Glu Lys

```
            20                  25                  30
Leu Lys Lys Pro Asp Lys Ile Leu Pro Cys Pro Arg Cys Asn Ser Met
        35                  40                  45
Glu Thr Lys Phe Cys Tyr Tyr Asn Asn Tyr Asn Val Asn Gln Pro Arg
    50                  55                  60
His Phe Cys Lys Ala Cys Gln Arg Tyr Trp Thr Glu Gly Gly Thr Ile
65                  70                  75                  80
Arg Asn Val Pro Val Gly Ala Gly Arg Arg Lys Asn Lys Asn Ser Ala
                85                  90                  95
Ser His Tyr Arg His Ile Thr Ile Ser Glu Ala Leu Arg Ala Ala Gln
            100                 105                 110
Ile Asp Val Pro Ile Glu Val Asn His Leu Ala Ser Lys Gly Asn Gly
        115                 120                 125
Arg Val Leu Asn Phe Ser Val Ser Pro Pro Val Cys Glu Ser Met Val
    130                 135                 140
Asn Val Ser His Pro Ala Glu Arg Lys Val Leu Asn Gly Thr Arg Asn
145                 150                 155                 160
Glu Phe Glu Gly Ala Lys Gly Pro Cys Glu Gly Gly Glu Thr Gly Asp
                165                 170                 175
Asp Cys Ser Ser Ala Ser Ser Val Thr Met Ser Ser Ser Met Lys Asn
            180                 185                 190
Gly Ala Arg Arg Phe Pro Gln Glu Pro His Met Gln Asn Ile Asn Gly
        195                 200                 205
Phe Pro Ser Gln Ile Pro Cys Leu Pro Gly Val Pro Trp Pro Cys Ser
    210                 215                 220
Trp Thr Ala Pro Ile Pro Pro Pro Ala Leu Cys Pro Pro Gly Val Pro
225                 230                 235                 240
Leu Ser Phe Tyr Pro Ala Thr Tyr Trp Ser Cys Ser Ala Ser Gly Ser
                245                 250                 255
Trp Asn Ile Pro Trp Val Thr Pro Gln Pro Cys Pro Pro Ile Pro Gly
            260                 265                 270
Pro Asn Ser Pro Thr Leu Gly Lys His Ser Arg Asp Gly Asp Glu Leu
        275                 280                 285
Gln Ala Asp Asn Ser Glu Met Lys Asp Pro Pro Lys Gln Lys Asn Gly
    290                 295                 300
Ser Val Leu Val Pro Lys Thr Leu Arg Ile Asp Asp Pro Asn Glu Ala
305                 310                 315                 320
Ala Lys Ser Ser Ile Trp Glu Thr Leu Gly Ile Lys Asn Asp Ser Ile
                325                 330                 335
Lys Ala Val Asp Leu Ser Asn Val Phe Gln Ser Lys Gly Asp Leu Lys
            340                 345                 350
Ser Asn Val Ser Glu Val Leu Ser Pro Val Leu Gln Ala Asn Pro Ala
        355                 360                 365
Ala Leu Ser Arg Ser Leu Thr Phe His Glu Arg Ser
    370                 375                 380
<210> 248
<211> 1191
<212> DNA
<213> Arabidopsis thaliana
<400> 248
atgtggctct ctcatctctt catgtccctc tctaagctga cgtgtaattt ctccattttc     60
tctgtcttta tggcttgtgg ctctattggt atgtctcaag ttagagatac tccggttaaa    120
ttgtttggct ggacaattac accggtttct catgatccat actcttcttc gtcccatgtt    180
cttcctgatt cttcctcgtc ttcctcgtct tcttctctat cacttcgacc acacatgatg    240
aataaccaat ctgttactga caatacaagt cttaagctgt catctaatct taacaacgag    300
tcaaaagaaa catctgagaa cagtgatgac caacacagcg agatcacaac aattacatcg    360
gaagaagaga aaacaactga actgaagaaa ccagacaaga ttcttccatg tccgagatgc    420
aacagcgcag acaccaaatt ctgttactac aacaactaca acgttaacca gccacgtcac    480
ttctgtagaa aatgccagag gtattggacc gctggtggat ccatgaggat cgtcccggtt    540
ggctcaggcc gtcgcaagaa caagggatgg gtttcttcag accagtacct gcacatcact    600
tccgaggata ctgacaatta caatagctcc tcaacaaaga ttctaagctt cgagtcttcg    660
gactctttgg taactgagag gcctaagcat caatcaaacg aagtgaagat aaacgctgaa    720
cctgtttcac aagaacccaa caacttccaa gggttacttc ctccccaagc atcccctgtt    780
```

tcgcctcctt ggccttacca ataccctcca aaccctagtt tctaccacat gcccgtctac 840
tggggctgcg cgataccggt ttggtctacc ctcgacactt ctacatgtct tgggaaaagg 900
acaagagacg aaacttctca tgaaactgtt aaagagagta aaaatgcttt tgagagaaca 960
agcttgcttt tggaatctca gagcatcaaa aatgaaacaa gtatggctac aaataaccat 1020
gtgtggtatc cagtaccgat gacccgcgag aagacacaag aattcagctt tttcagtaat 1080
ggagctgaaa caaagagcag caacaacaga ttcgtccctg aaacgtatct taacctgcaa 1140
gcaaaccctg cagccatggc aagatctatg aacttcagag agagcatata a 1191

<210> 249
<211> 396
<212> PRT
<213> Arabidopsis thaliana
<400> 249

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

<210> 250

<211> 1200
<212> DNA
<213> Arabidopsis thaliana
<400> 250

```
atgtctaaat ctagagatac ggagataaag ttgtttggga ggacaatcac atctctttta     60
gatgtgaatt gttatgatcc gtcgtcgttg tcccctgttc acgatgtttc ttctgatcca    120
agcaaggagg attcgtcttc ttcttcatct tcttgttctc caactattgg accaatcagg    180
gttccggtta aaaaaagtga gcaagagagt aacaaattca aagatccata tatattatcc    240
gatctaaacg aaccaccaaa agcagtatct gagatttcat caccaagaag ttccaagaac    300
aactgtgatc aacagagcga gatcacaaca acaactacca caagtactac atcaggagag    360
aaatcaacgg ctctcaagaa accggacaag cttattccat gtcctagatg tgaaagcgca    420
aacaccaaat tctgttatta caacaactac aacgtgaacc agccacgtta cttctgcagg    480
aactgtcaga ggtattggac agctggtgga tctatgagga acgttcctgt tggctcaggt    540
cgtcgcaaga acaaaggatg gccttcttca aaccattact tgcaagtcac ttctgaggat    600
tgtgataata ataactcggg gacgatcctt agtttcggtt cttcggagtc ttcggttaca    660
gagactggta agcatcagtc aggtgataca gcaaagataa gtgctgattc agtttctcaa    720
gaaaataaaa gctaccaagg gtttcttcct ccgcaagtaa tgttacctaa taattcttct    780
ccttggcctt accaatggag tccaacgggt cctaacgcta gtttctaccc tgtccccttc    840
tactggggat gcacggttcc gatataccct acctcagaga cttcatcatg tttaggaaaa    900
cggtcaagag atcaaactga aggaagaatc aatgatacta atacaacaat aactactaca    960
agagcaagat tggtctcaga atctcttaga atgaatatcg aagctagtaa gagcgctgtg   1020
tggtctaagt taccgacaaa acccgagaaa aaaacgcaag gattcagttt gttcaatgga   1080
tttgacacaa agggaaacag caacagaagt agcttggtct ccgaaacttc tcacagtcta   1140
caagcaaacc ctgcagcgat gtctagagct atgaacttca gggagagcat gcaacaataa   1200
```

<210> 251
<211> 399
<212> PRT
<213> Arabidopsis thaliana
<400> 251

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

Ala Ser Phe Tyr Pro Val Pro Phe Tyr Trp Gly Cys Thr Val Pro Ile
275 280 285
Tyr Pro Thr Ser Glu Thr Ser Ser Cys Leu Gly Lys Arg Ser Arg Asp
290 295 300
Gln Thr Glu Gly Arg Ile Asn Asp Thr Asn Thr Thr Ile Thr Thr Thr
305 310 315 320
Arg Ala Arg Leu Val Ser Glu Ser Leu Arg Met Asn Ile Glu Ala Ser
325 330 335
Lys Ser Ala Val Trp Ser Lys Leu Pro Thr Lys Pro Glu Lys Lys Thr
340 345 350
Gln Gly Phe Ser Leu Phe Asn Gly Phe Asp Thr Lys Gly Asn Ser Asn
355 360 365
Arg Ser Ser Leu Val Ser Glu Thr Ser His Ser Leu Gln Ala Asn Pro
370 375 380
Ala Ala Met Ser Arg Ala Met Asn Phe Arg Glu Ser Met Gln Gln
385 390 395

<210> 252
<211> 1347
<212> DNA
<213> Arabidopsis thaliana
<400> 252

atgatgatgg agactagaga tccagctatt aagcttttcg gtatgaaaat ccctttccg 60
tcggtttttg aatcggcagt tacggtggag gatgacgaag aagatgactg gagcggcgga 120
gatgacaaat caccagagaa ggtaactcca gagttatcag ataagaacaa caacaactgt 180
aacgacaaca gttttaacaa ttcgaaaccc gaaaccttgg acaaagagga agcgacatca 240
actgatcaga tagagagtag tgacacgcct gaggataatc agcagacgac acctgatggt 300
aaaaccctaa agaaaccgac taagattcta ccgtgtccga gatgcaaaag catggagacc 360
aagttctgtt attacaacaa ctacaacata aaccagcctc gtcatttctg caaggcttgt 420
cagagatatt ggactgctgg agggactatg aggaatgttc ctgtgggggc aggacgtcgt 480
aagaacaaaa gctcatcttc tcattaccgt cacatcacta tttccgaggc tcttgaggct 540
gcgaggcttg acccgggctt acaggcaaac acaagggtct tgagttttgg tctcgaagct 600
cagcagcagc acgttgctgc tcccatgaca cctgtgatga agctacaaga agatcaaaag 660
gtctcaaacg gtgctaggaa caggtttcac gggttagcgg atcaacggct tgtagctcgg 720
gtagagaatg gagatgattg ctcaagcgga tcctctgtga ccacctctaa caatcactca 780
gtggatgaat caagagcaca aagcggcagt gttgttgaag cacaaatgaa caacaacaac 840
aacaataaca tgaatggtta tgcttgcatc ccaggtgttc catggcctta cacgtggaat 900
ccagcgatgc ctccaccagg tttttacccg cctccagggt atccaatgcc gttttaccct 960
tactggacca tcccaatgct accaccgcat caatcctcat cgcctataag ccaaaagtgt 1020
tcaaatacaa actctccgac tctcggaaag catccgagag atgaaggatc atcgaaaaag 1080
gacaacgaga cagagcgaaa acagaaggcc gggtgcgttc tggtcccgaa aacgttgaga 1140
atagatgatc ctaacgaagc agcaaagagc tcgatatgga caacattggg aatcaagaac 1200
gaggcgatgt gcaaagccgg tggtatgttc aaagggtttg atcataagac aaagatgtat 1260
aacaacgaca aagctgagaa ctcccctgtt ctttctgcta accctgctgc tctatcaaga 1320
tcacacaatt tccatgaaca gatttag 1347

<210> 253
<211> 448
<212> PRT
<213> Arabidopsis thaliana
<400> 253

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

```
Pro Arg Cys Lys Ser Met Glu Thr Lys Phe Cys Tyr Tyr Asn Asn Tyr
        115                 120                 125
Asn Ile Asn Gln Pro Arg His Phe Cys Lys Ala Cys Gln Arg Tyr Trp
    130                 135                 140
Thr Ala Gly Gly Thr Met Arg Asn Val Pro Val Gly Ala Gly Arg Arg
145                 150                 155                 160
Lys Asn Lys Ser Ser Ser Ser His Tyr Arg His Ile Thr Ile Ser Glu
                165                 170                 175
Ala Leu Glu Ala Ala Arg Leu Asp Pro Gly Leu Gln Ala Asn Thr Arg
            180                 185                 190
Val Leu Ser Phe Gly Leu Glu Ala Gln Gln Gln His Val Ala Ala Pro
        195                 200                 205
Met Thr Pro Val Met Lys Leu Gln Glu Asp Gln Lys Val Ser Asn Gly
    210                 215                 220
Ala Arg Asn Arg Phe His Gly Leu Ala Asp Gln Arg Leu Val Ala Arg
225                 230                 235                 240
Val Glu Asn Gly Asp Asp Cys Ser Ser Gly Ser Ser Val Thr Thr Ser
                245                 250                 255
Asn Asn His Ser Val Asp Glu Ser Arg Ala Gln Ser Gly Ser Val Val
            260                 265                 270
Glu Ala Gln Met Asn Asn Asn Asn Asn Asn Asn Met Asn Gly Tyr Ala
        275                 280                 285
Cys Ile Pro Gly Val Pro Trp Pro Tyr Thr Trp Asn Pro Ala Met Pro
    290                 295                 300
Pro Pro Gly Phe Tyr Pro Pro Pro Gly Tyr Pro Met Pro Phe Tyr Pro
305                 310                 315                 320
Tyr Trp Thr Ile Pro Met Leu Pro Pro His Gln Ser Ser Ser Pro Ile
                325                 330                 335
Ser Gln Lys Cys Ser Asn Thr Asn Ser Pro Thr Leu Gly Lys His Pro
            340                 345                 350
Arg Asp Glu Gly Ser Ser Lys Lys Asp Asn Glu Thr Glu Arg Lys Gln
        355                 360                 365
Lys Ala Gly Cys Val Leu Val Pro Lys Thr Leu Arg Ile Asp Asp Pro
    370                 375                 380
Asn Glu Ala Ala Lys Ser Ser Ile Trp Thr Thr Leu Gly Ile Lys Asn
385                 390                 395                 400
Glu Ala Met Cys Lys Ala Gly Gly Met Phe Lys Gly Phe Asp His Lys
                405                 410                 415
Thr Lys Met Tyr Asn Asn Asp Lys Ala Glu Asn Ser Pro Val Leu Ser
            420                 425                 430
Ala Asn Pro Ala Ala Leu Ser Arg Ser His Asn Phe His Glu Gln Ile
        435                 440                 445
<210> 254
<211> 1374
<212> DNA
<213> Arabidopsis thaliana
<400> 254
atggctgatc cggcgattaa gctctttgga aagacgattc ctttacctga gcttggtgtt     60
gttgattctt cttctagcta taccggattt ttaaccgaaa ctcagattcc tgttcggtta    120
tcagattcgt gtaccggcga tgatgatgat gaagagatgg gtgattccgg tttaggacga    180
gaagaaggtg atgatgttgg tgatggtgga ggagagagcg agactgataa aaaggaagaa    240
aaagatagtg agtgtcagga agagtcattg aggaatgaat ctaatgatgt tactactact    300
acatcgggta taactgaaaa aacggaaaca acaaaagctg caaagacgaa tgaagagtca    360
ggtggtactg cttgctctca agaggggaag ttaaagaaac ctgataagat tctaccgtgt    420
ccgcgatgta acagcatgga aaccaagttc tgttactaca acaactataa tgttaaccaa    480
cctcgccatt tctgcaagaa atgtcagaga tattggacag ctggtggaac gatgaggaat    540
gttccggttg gtgctgggag acgtaagaat aagagtccag cttctcatta taaccgtcat    600
gtaagtataa catctgcgga agctatgcag aaggtggcga gaactgatct tcaacatcct    660
aatggtgcaa atcttctcac ttttggctct gattctgtgc tttgtgaatc tatggcttct    720
ggattgaatc ttgttgagaa gtcattgttg aagacacaaa ctgtattgca agaacccaat    780
gaaggcttga agattacggt tccgttaaac cagacaaacg aagaagctgg aacagtcagc    840
ccgttaccaa aagttccatg ctttccagga ccaccaccaa cttggcctta cgcttggaac    900
ggagtttcgt ggacgatttt accgttttac cctccaccgg cttactggag ctgcccgggg    960
```

gtttcaccgg gggcatggaa cagcttcaca tggatgccac aacccaattc accatctggt 1020
tccaatccaa attctcctac actaggtaaa cattcacgtg acgagaacgc tgctgaacca 1080
ggaaccgctt ttgatgaaac cgagtcactt ggtagggaga aaagcaaacc cgagagatgc 1140
ttgtgggttc ccaagacgct gaggattgat gatccagagg aagctgctaa aagttccatc 1200
tgggaaacat tagggatcaa aaaagacgaa aatgcggata ctttcggagc tttcagatca 1260
tcaaccaaag aaaaaagcag tctttctgaa ggaagacttc cgggaagaag accggagttg 1320
caagcgaatc ctgctgctct ttctaggtca gcaaacttcc atgagagctc atag 1374

<210> 255
<211> 457
<212> PRT
<213> Arabidopsis thaliana
<400> 255

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

405 410 415
Ala Phe Arg Ser Ser Thr Lys Glu Lys Ser Ser Leu Ser Glu Gly Arg
420 425 430
Leu Pro Gly Arg Arg Pro Glu Leu Gln Ala Asn Pro Ala Ala Leu Ser
435 440 445
Arg Ser Ala Asn Phe His Glu Ser Ser
450 455

<210> 256
<211> 57
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm07303
<400> 256
ggggacaagt ttgtacaaaa aagcaggctt aaacaatgat gatggagact agagatc 57

<210> 257
<211> 53
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm07304
<400> 257
ggggaccact ttgtacaaga aagctgggtc atatgtaact ctaaatctgt tca 53

<210> 258
<211> 654
<212> DNA
<213> Oryza sativa
<400> 258
cttctacatc ggcttaggtg tagcaacacg actttattat tattattatt attattatta 60
ttatttttaca 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> 259
<211> 1828
<212> DNA
<213> Oryza sativa
<400> 259
gcttgagtca tagggagaaa acaaatcgat catatttgac tcttttccct ccatctctct 60
taccggcaaa aaaagtagta ctggtttata tgtaaagtaa gattctttaa ttatgtgaga 120
tccggcttaa tgcttttctt ttgtcacata tactgcattg caacaattgc catatattca 180
cttctgccat cccattatat agcaactcaa gaatggattg atatatcccc tattactaat 240
ctagacatgt taaggctgag ttgggcagtc catcttccca acccaccacc ttcgtttttc 300
gcgcacatac ttttcaaact actaaatggt gtgtttttta aaaatatttt caatacaaaa 360
gttgctttaa aaaattatat tgatccattt ttttaaaaaa aatagctaat acttaattaa 420
tcacgtgtta aaagaccgct ccgttttgcg tgcaggaggg ataggttcac atcctgcatt 480
accgaacaca gcctaaatct tgttgtctag attcgtagta ctggatatat taaatcatgt 540
tctaagttac tatatactga gatgaataga ataagtaaaa ttagacccac cttaagtctt 600
gatgaagtta ctactagctg cgtttgggag gacttcccaa aaaaaaaagt attagccatt 660
agcacgtgat taattaagta ctagtttaaa aaacttaaaa aataaattaa tatgattctc 720
ttaagtaact ctcctataga aaacttttac aaaattacac cgtttaatag tttggaaaat 780
atgtcagtaa aaaataagag agtagaagtt atgaaagtta gaaaaagaat tgttttagta 840
gtatacagtt ataaactatt ccctctgttc taaaacataa gggattatgg atggattcga 900
catgtaccag taccatgaat cgaatccaga caagtttttt atgcatattt attctactat 960
aatatatcac atctgctcta aatatcttat atttcgaggt ggagactgtc gctatgtttt 1020
tctgcccgtt gctaagcaca cgccaccccc gatgcgggga cgcctctggc cttcttgcca 1080
cgataattga atggaacttc cacattcaga ttcgataggt gaccgtcgac tccaagtgct 1140

ttgcacaaaa caactccggc ctcccggcca ccagtcacac gactcacggc actaccaccc 1200
ctgactccct gaggcggacc tgccactgtt ctgcatgcga agctatctaa aattctgaag 1260
caaagaaagc acagcacatg ctccgggaca cgcgccaccc ggcggaaaag ggctcggtgt 1320
ggcgatctca cagccgcata tcgcatttca caagccgccc atctccaccg gcttcacgag 1380
gctcatcgcg gcacgaccgc gcacggaacg cacgcggccg acccgcgcgc ctcgatgcgc 1440
gagcccatcc gccgcgtcct ccctttgcct ttgccgctat cctctcggtc gtatcccgtt 1500
tctctgtctt ttgctccccg gcgcgcgcca gttcggagta ccagcgaaac ccggacacct 1560
ggtacacctc cgccggccac aacgcgtgtc ccccctacgt ggccgcgcag cacatgccca 1620
tgcgcgacac gtgcacctcc tcatccaaac tctcaagtct caacggtcct ataaatgcac 1680
ggatagcctc aagctgctcg tcacaaggca agaggcaaga ggcaagagca tccgtattaa 1740
ccagcctttt gagacttgag agtgtgtgtg actcgatcca gcgtagtttc agttcgtgtg 1800
ttggtgagtg attccagcca agtttgcg 1828

<210> 260
<211> 1243
<212> DNA
<213> Oryza sativa
<400> 260
aaaaccaccg agggacctga tctgcaccgg ttttgatagt tgagggaccc gttgtgtctg 60
gttttccgat cgagggacga aaatcggatt cggtgtaaag ttaagggacc tcagatgaac 120
ttattccgga gcatgattgg gaagggagga cataaggccc atgtcgcatg tgtttggacg 180
gtccagatct ccagatcact cagcaggatc ggccgcgttc gcgtagcacc cgcggtttga 240
ttcggcttcc cgcaaggcgg cggccggtgg ccgtgccgcc gtagcttccg ccggaagcga 300
gcacgccgcc gccgccgacc cggctctgcg tttgcaccgc cttgcacgcg atacatcggg 360
atagatagct actactctct ccgtttcaca atgtaaatca ttctactatt ttccacattc 420
atattgatgt taatgaatat agacatatat atctatttag attcattaac atcaatatga 480
atgtaggaaa tgctagaatg acttacattg tgaattgtga aatggacgaa gtacctacga 540
tggatggatg caggatcatg aaagaattaa tgcaagatcg tatctgccgc atgcaaaatc 600
ttactaattg cgctgcatat atgcatgaca gcctgcatgc gggcgtgtaa gcgtgttcat 660
ccattaggaa gtaaccttgt cattacttat accagtacta catactatat agtattgatt 720
tcatgagcaa atctacaaaa ctggaaagca ataagaaata cgggactgga aaagactcaa 780
cattaatcac caaatatttc gccttctcca gcagaatata tatctctcca tcttgatcac 840
tgtacacact gacagtgtac gcataaacgc agcagccagc ttaactgtcg tctcaccgtc 900
gcacactggc cttccatctc aggctagctt tctcagccac ccatcgtaca tgtcaactcg 960
gcgcgcgcac aggcacaaat tacgtacaaa acgcatgacc aaatcaaaac caccggagaa 1020
gaatcgctcc cgcgcgcggc ggcgacgcgc acgtacgaac gcacgcacgc acgcccaacc 1080
ccacgacacg atcgcgcgcg acgccggcga caccggccgt ccacccgcgc cctcacctcg 1140
ccgactataa atacgtaggc atctgcttga tcttgtcatc catctcacca ccaaaaaaaa 1200
aaggaaaaaa aaacaaaaca caccaagcca aataaaagcg aca 1243

<210> 261
<211> 8
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 1
<220>
<221> INSEGURO
<222> (2)..(3)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 261
Phe Xaa Xaa Lys Tyr Asn Phe Asp
1 5

<210> 262
<211> 8
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 2
<220>
<221> VARIANTE
<222> (1)..(1)
<223> /substituir="Leu"
<220>
<221> INSEGURO

<222> (3)..(3)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 262
Pro Leu Xaa Gly Arg Tyr Glu Trp
1 5
<210> 263
<211> 10
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 3
<220>
<221> INSEGURO
<222> (2)..(2)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> VARIANTE
<222> (4)..(4)
<223> /substituir="Glu"
<220>
<221> INSEGURO
<222> (7)..(9)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 263
Glu Xaa Glu Asp Phe Phe Xaa Xaa Xaa Glu
1 5 10
<210> 264
<211> 8
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 4
<220>
<221> INSEGURO
<222> (2)..(2)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 264
Tyr Xaa Gln Leu Arg Ser Arg Arg
1 5
<210> 265
<211> 9
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 5
<220>
<221> VARIANTE
<222> (5)..(5)
<223> /substituir="Ile"
<220>
<221> VARIANTE
<222> (6)..(6)
<223> /substituir="Arg"
<220>
<221> INSEGURO
<222> (8)..(8)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<220>
<221> VARIANTE
<222> (9)..(9)
<223> /substituir="Arg"
<400> 265
Met Gly Lys Tyr Met Lys Lys Xaa Lys

1 5
<210> 266
<211> 8
<212> PRT
<213> Seqüência artificial
<220>
<223> motivo 6
<220>
<221> INSEGURO
<222> (2)..(2)
<223> Xaa pode ser qualquer aminoácido ocorrendo naturalmente
<400> 266
Ser Xaa Gly Val Arg Thr Arg Ala
1 5
<210> 267
<211> 585
<212> DNA
<213> Oryza sativa
<400> 267
atgggcaagt acatgcgcaa ggccaaggtg gtggtctccg gcgaggtggt ggccgccgcc 60
gtcatggagc tcgccgcggc gccgctcggg gtgcgcaccc gcgcccgctc cctcgcgctg 120
cagaagaggc agggcgggga gtacctcgag ctcaggagcc gcaggctcga gaagctccct 180
cctcccccgc cgccgccgcc gaggaggagg gcgacggctg cggctgcgac tgctgatgcg 240
acggcgacgg agagcgcgga ggcggaggtg tcgttcgggg gggagaacgt cctcgagctg 300
gaggccatgg aaaggaatac cagggagacg acaccttgca gcttgatcag ggaccccgat 360
acgattagca cccctggatc taccacaagg cgcagccact cgagttctca ttgcaaggtg 420
caaacacccg tgcgccacaa cattattcca gcatcagcag agctggaagc gttcttcgcc 480
gccgaagagc aacggcaacg acaggctttc atcgacaagt ataactttga tcctgtgaat 540
gactgccctc ttcccggccg atttgaatgg gtcaagctag actga 585
<210> 268
<211> 194
<212> PRT
<213> Oryza sativa
<400> 268
Met Gly Lys Tyr Met Arg Lys Ala Lys Val Val Val Ser Gly Glu Val
1 5 10 15
Val Ala Ala Ala Val Met Glu Leu Ala Ala Ala Pro Leu Gly Val Arg
20 25 30
Thr Arg Ala Arg Ser Leu Ala Leu Gln Lys Arg Gln Gly Gly Glu Tyr
35 40 45
Leu Glu Leu Arg Ser Arg Arg Leu Glu Lys Leu Pro Pro Pro Pro Pro
50 55 60
Pro Pro Pro Arg Arg Arg Ala Thr Ala Ala Ala Ala Thr Ala Asp Ala
65 70 75 80
Thr Ala Thr Glu Ser Ala Glu Ala Glu Val Ser Phe Gly Gly Glu Asn
85 90 95
Val Leu Glu Leu Glu Ala Met Glu Arg Asn Thr Arg Glu Thr Thr Pro
100 105 110
Cys Ser Leu Ile Arg Asp Pro Asp Thr Ile Ser Thr Pro Gly Ser Thr
115 120 125
Thr Arg Arg Ser His Ser Ser Ser His Cys Lys Val Gln Thr Pro Val
130 135 140
Arg His Asn Ile Ile Pro Ala Ser Ala Glu Leu Glu Ala Phe Phe Ala
145 150 155 160
Ala Glu Glu Gln Arg Gln Arg Gln Ala Phe Ile Asp Lys Tyr Asn Phe
165 170 175
Asp Pro Val Asn Asp Cys Pro Leu Pro Gly Arg Phe Glu Trp Val Lys
180 185 190
Leu Asp

<210> 269
<211> 573
<212> DNA

<213> Zea mays
<400> 269
atgggcaagt acatgcgcaa ggccaaggct tccagcgagg ttgtcatcat ggatgtcgcc 60
gccgctccgc tcggagtccg cacccgagcg cgcgccctcg cgctgcagcg tctgcaggag 120
caacagacgc agtgggaaga aggtgctggc ggcgagtacc tggagctaag gaaccggagg 180
ctcgagaagc tgccgccgcc gccggcgacc actaggaggt cgggcgggag gaaagcggca 240
gccgaggccg ccgcaactaa ggaggctgag gcgtcgtacg gggagaacat gctcgagttg 300
gaggccatgg agaggattac cagggagacg acgccttgca gcttgattaa cacccagatg 360
actagcactc ctgggtccac gagatccagc cactcttgcc accgcagggt gaacgctcct 420
ccggtgcacg ccgtcccaag ttctagggag atgaatgagt acttcgctgc cgaacagcga 480
cggcaacagc aggatttcat tgacaagtac aacttcgatc ctgcaaacga ctgccctctc 540
ccaggcaggt ttgagtgggt gaagctagac tga 573
<210> 270
<211> 190
<212> PRT
<213> Zea mays
<400> 270
Met Gly Lys Tyr Met Arg Lys Ala Lys Ala Ser Ser Glu Val Val Ile
1 5 10 15
Met Asp Val Ala Ala Ala Pro Leu Gly Val Arg Thr Arg Ala Arg Ala
20 25 30
Leu Ala Leu Gln Arg Leu Gln Glu Gln Gln Thr Gln Trp Glu Glu Gly
35 40 45
Ala Gly Gly Glu Tyr Leu Glu Leu Arg Asn Arg Arg Leu Glu Lys Leu
50 55 60
Pro Pro Pro Pro Ala Thr Thr Arg Arg Ser Gly Gly Arg Lys Ala Ala
65 70 75 80
Ala Glu Ala Ala Ala Thr Lys Glu Ala Glu Ala Ser Tyr Gly Glu Asn
85 90 95
Met Leu Glu Leu Glu Ala Met Glu Arg Ile Thr Arg Glu Thr Thr Pro
100 105 110
Cys Ser Leu Ile Asn Thr Gln Met Thr Ser Thr Pro Gly Ser Thr Arg
115 120 125
Ser Gly His Ser Cys His Arg Arg Val Asn Ala Pro Pro Val His Ala
130 135 140
Val Pro Ser Ser Arg Glu Met Asn Glu Tyr Phe Ala Ala Glu Gln Arg
145 150 155 160
Arg Gln Gln Gln Asp Phe Ile Asp Lys Tyr Asn Phe Asp Pro Ala Asn
165 170 175
Asp Cys Pro Leu Pro Gly Arg Phe Glu Trp Val Lys Leu Asp
180 185 190
<210> 271
<211> 755
<212> DNA
<213> Triticum aestivum
<400> 271
atgggcaagt acatgcgcaa gcccaaggtc tccggcgagg tggccgtcat ggaggtcgcc 60
gccgcgccgc taggggtccg cacccgcgca cgagcgctcg cgatgcagag gcagccgcag 120
ggggcggcgg tggccaagga ccaggggggag tacctggagc tcaggagtcg gaagctcgag 180
aagctgcccc cgccgccgcc ggcggcgagg aggagggcgg ccgcggcgga gcgtgtcgag 240
gccgaggccg aggccgacga ggtgtccttc ggtgagaacg tgctcgagtc ggaggccatg 300
gggaggggta ccagggagac gacgccctgc agcttgatta gggactcggg aacgataagc 360
actcctggat ccacaacaag accgagccac tcgaattccc atcgcagggt gcaagctcca 420
gcgcgccata ttattccatg ttcagcagag atgaatgagt tcttctctgc tgcggagcaa 480
ccgcaacagc aagccttcat tgacaagtac aactttgatc ctgtgaacga ctgtcctctc 540
ccaggccgat acgagtgggt gaagctagac tgataattct ccaggaagga gagcaccatg 600
tacttctccg ctccctccac cttagcgtcg tggtaaaggc cgccccgtcg tgtagctttg 660
tttccgttgt aaaaagaata gttagctgta gtagcctcaa tggcgttaca tacagagtaa 720
tgctgattac acctaatcct caaaccatgt acgtt 755
<210> 272
<211> 190
<212> PRT
<213> Triticum aestivum

<400> 272
Met Gly Lys Tyr Met Arg Lys Pro Lys Val Ser Gly Glu Val Ala Val
1 5 10 15
Met Glu Val Ala Ala Ala Pro Leu Gly Val Arg Thr Arg Ala Arg Ala
20 25 30
Leu Ala Met Gln Arg Gln Pro Gln Gly Ala Ala Val Ala Lys Asp Gln
35 40 45
Gly Glu Tyr Leu Glu Leu Arg Ser Arg Lys Leu Glu Lys Leu Pro Pro
50 55 60
Pro Pro Pro Ala Ala Arg Arg Arg Ala Ala Ala Ala Glu Arg Val Glu
65 70 75 80
Ala Glu Ala Glu Ala Asp Glu Val Ser Phe Gly Glu Asn Val Leu Glu
85 90 95
Ser Glu Ala Met Gly Arg Gly Thr Arg Glu Thr Thr Pro Cys Ser Leu
100 105 110
Ile Arg Asp Ser Gly Thr Ile Ser Thr Pro Gly Ser Thr Thr Arg Pro
115 120 125
Ser His Ser Asn Ser His Arg Arg Val Gln Ala Pro Ala Arg His Ile
130 135 140
Ile Pro Cys Ser Ala Glu Met Asn Glu Phe Phe Ser Ala Ala Glu Gln
145 150 155 160
Pro Gln Gln Gln Ala Phe Ile Asp Lys Tyr Asn Phe Asp Pro Val Asn
165 170 175
Asp Cys Pro Leu Pro Gly Arg Tyr Glu Trp Val Lys Leu Asp
180 185 190

<210> 273
<211> 666
<212> DNA
<213> Oryza sativa
<400> 273
atggggaagt acatgcggaa ggggaaggtg tcgggggagg tggcggtgat ggaggtgggc 60
ggggcgctgc tcggcgtccg cacccgctcc cgcacgctcg cgctgcagcg gacgacctcg 120
tcgcagaagc cgccggagaa gggggagggg gaccccggtg cgggcgcggg cgcgggggcg 180
gagtacctcg agctcaggag ccggaggctc gagaagccgc ctccgcacac gccgccggcc 240
aaggagaagg agaccgccag gagggcttcc gccgccgccg ccgccgccgt gaggatgccg 300
gcggcgccgc aagcggccga ggagttcgag gcggaggtcg aggtgtcctt cggcgacaac 360
gttcttgacc tcgacggcga cgccatggag aggagtacca gggagacaac gccttgcagt 420
ttaattagga gctcagaaat gataagcacc cctggctcca caactaaaac caacacctcg 480
atcagttccc ggcgcagaat ggagacctct gtttgtcgtt acgttccgag ttctcttgag 540
atggaagagt tctttgcagc tgctgaacaa cagcaacatc aggctttcag agagaggtat 600
aacttctgtc ctgtgaacga ctgcccactt cctggacggt acgaatggac aaggctagac 660
tgctag 666

<210> 274
<211> 221
<212> PRT
<213> Oryza sativa
<400> 274
Met Gly Lys Tyr Met Arg Lys Gly Lys Val Ser Gly Glu Val Ala Val
1 5 10 15
Met Glu Val Gly Gly Ala Leu Leu Gly Val Arg Thr Arg Ser Arg Thr
20 25 30
Leu Ala Leu Gln Arg Thr Thr Ser Ser Gln Lys Pro Pro Glu Lys Gly
35 40 45
Glu Gly Asp Pro Gly Ala Gly Ala Gly Ala Gly Ala Glu Tyr Leu Glu
50 55 60
Leu Arg Ser Arg Arg Leu Glu Lys Pro Pro Pro His Thr Pro Pro Ala
65 70 75 80
Lys Glu Lys Glu Thr Ala Arg Arg Ala Ser Ala Ala Ala Ala Ala Ala
85 90 95
Val Arg Met Pro Ala Ala Pro Gln Ala Ala Glu Glu Phe Glu Ala Glu
100 105 110
Val Glu Val Ser Phe Gly Asp Asn Val Leu Asp Leu Asp Gly Asp Ala
115 120 125

```
Met Glu Arg Ser Thr Arg Glu Thr Thr Pro Cys Ser Leu Ile Arg Ser
    130                 135                 140
Ser Glu Met Ile Ser Thr Pro Gly Ser Thr Thr Lys Thr Asn Thr Ser
145                 150                 155                 160
Ile Ser Ser Arg Arg Arg Met Glu Thr Ser Val Cys Arg Tyr Val Pro
                165                 170                 175
Ser Ser Leu Glu Met Glu Glu Phe Phe Ala Ala Ala Glu Gln Gln Gln
            180                 185                 190
His Gln Ala Phe Arg Glu Arg Tyr Asn Phe Cys Pro Val Asn Asp Cys
        195                 200                 205
Pro Leu Pro Gly Arg Tyr Glu Trp Thr Arg Leu Asp Cys
    210                 215                 220
<210> 275
<211> 642
<212> DNA
<213> Zea mays
<400> 275
atggggaagt acatgcgcaa gggcaaggtg tccggggagg tcgccgtcat ggaggtaccc     60
ggcggcgcgc tgctcggcgt ccgcacccgc tcccgcacgc tcgcgctgca gcgcgcgcag    120
aggccgctcg acaagggcga cgcggaggac gccgccgcgg agtacctcga gctcaggagc    180
cggaggctcg agaagccgca caaggagcat ccgtcgccgc ccgcgaccgc gaccaagagg    240
ggcgccggga ggaaggccgc cgccgccgcc gcggtgcagc acgtgctgat gcaggacgag    300
gtcgaggtcg aggtctcgtt cggggacaac gtgcttgact tggacaccat ggaaaggagt    360
accagagaga caacaccgtg cagcctgatt aggaacccag agatgataag caccccagga    420
tccacaacta aaagcaaaac cagcagcaac tcgacgactt cccgccgcag aacggaggaa    480
accccgagct gccggttcat accgagctcg ctcgagatgg aggagttctt ctcggcggcc    540
gagcaacagg agcagcatag cttcagggag aagtacaact tctgtcccgt gaacgactgt    600
cctctccctg gccggtacga atgggcgagg ctagactgct ag                       642
<210> 276
<211> 213
<212> PRT
<213> Zea mays
<400> 276
Met Gly Lys Tyr Met Arg Lys Gly Lys Val Ser Gly Glu Val Ala Val
1               5                   10                  15
Met Glu Val Pro Gly Gly Ala Leu Leu Gly Val Arg Thr Arg Ser Arg
            20                  25                  30
Thr Leu Ala Leu Gln Arg Ala Gln Arg Pro Leu Asp Lys Gly Asp Ala
        35                  40                  45
Glu Asp Ala Ala Ala Glu Tyr Leu Glu Leu Arg Ser Arg Arg Leu Glu
    50                  55                  60
Lys Pro His Lys Glu His Pro Ser Pro Pro Ala Thr Ala Thr Lys Arg
65                  70                  75                  80
Gly Ala Gly Arg Lys Ala Ala Ala Ala Ala Ala Val Gln His Val Leu
                85                  90                  95
Met Gln Asp Glu Val Glu Val Glu Val Ser Phe Gly Asp Asn Val Leu
            100                 105                 110
Asp Leu Asp Thr Met Glu Arg Ser Thr Arg Glu Thr Thr Pro Cys Ser
        115                 120                 125
Leu Ile Arg Asn Pro Glu Met Ile Ser Thr Pro Gly Ser Thr Thr Lys
    130                 135                 140
Ser Lys Thr Ser Ser Asn Ser Thr Thr Ser Arg Arg Arg Thr Glu Glu
145                 150                 155                 160
Thr Pro Ser Cys Arg Phe Ile Pro Ser Ser Leu Glu Met Glu Glu Phe
                165                 170                 175
Phe Ser Ala Ala Glu Gln Gln Glu Gln His Ser Phe Arg Glu Lys Tyr
            180                 185                 190
Asn Phe Cys Pro Val Asn Asp Cys Pro Leu Pro Gly Arg Tyr Glu Trp
        195                 200                 205
Ala Arg Leu Asp Cys
    210
<210> 277
<211> 642
```

<212> DNA
<213> Sorghum bicolor
<400> 277
atggggaagt acatgcgcaa gggcaaggtg tccggggagg tcgccgtcat ggaggtcccc 60
ggcggcgcgc tgctcggcgt ccgcacccgc tcccgcacgc tcgcgctgca gcgcgcgcag 120
aggccgctcg acaagggcga cgccgaggac gccgctgcgg agtacctcga gctcaggagc 180
cggaggctcg agaagccgca caaggacccg ttaccgccgc cgtctgcgcc tgcgaccaag 240
aggggcgccg ggaggaaggt cgccaccgcc gccgccgccg ccgcggcgcc gcacgggctg 300
gcggaggacg acgtcgaggt ctccttcggc gagaacgtgc tcgacttcga cgccatggaa 360
aggagtacca gagagacaac accgtgcagt ttgattagga acccagagat gataagcacc 420
ccaggatcca caactaagag taaaaccagc aactcgatga cctcccgtcg cagaatggaa 480
acctcaatct gccgtttcat accaagttcg catgagatgg aagagttctt ctcagcagct 540
gaaaaacagg agcagcaaag cttcagggag aagtataact tctgtcctgt gaacgactgt 600
cctcttccgg gtcggtatga atgggcgagg ctagactgct ag 642
<210> 278
<211> 213
<212> PRT
<213> Sorghum bicolor
<400> 278
Met Gly Lys Tyr Met Arg Lys Gly Lys Val Ser Gly Glu Val Ala Val
1 5 10 15
Met Glu Val Pro Gly Gly Ala Leu Leu Gly Val Arg Thr Arg Ser Arg
20 25 30
Thr Leu Ala Leu Gln Arg Ala Gln Arg Pro Leu Asp Lys Gly Asp Ala
35 40 45
Glu Asp Ala Ala Ala Glu Tyr Leu Glu Leu Arg Ser Arg Arg Leu Glu
50 55 60
Lys Pro His Lys Asp Pro Leu Pro Pro Pro Ser Ala Pro Ala Thr Lys
65 70 75 80
Arg Gly Ala Gly Arg Lys Val Ala Thr Ala Ala Ala Ala Ala Ala Ala
85 90 95
Pro His Gly Leu Ala Glu Asp Asp Val Glu Val Ser Phe Gly Glu Asn
100 105 110
Val Leu Asp Phe Asp Ala Met Glu Arg Ser Thr Arg Glu Thr Thr Pro
115 120 125
Cys Ser Leu Ile Arg Asn Pro Glu Met Ile Ser Thr Pro Gly Ser Thr
130 135 140
Thr Lys Ser Lys Thr Ser Asn Ser Met Thr Ser Arg Arg Arg Met Glu
145 150 155 160
Thr Ser Ile Cys Arg Phe Ile Pro Ser Ser His Glu Met Glu Glu Phe
165 170 175
Phe Ser Ala Ala Glu Lys Gln Glu Gln Gln Ser Phe Arg Glu Lys Tyr
180 185 190
Asn Phe Cys Pro Val Asn Asp Cys Pro Leu Pro Gly Arg Tyr Glu Trp
195 200 205
Ala Arg Leu Asp Cys
210
<210> 279
<211> 397
<212> DNA
<213> Saccharum officinarum
<220>
<221> misc_feature
<222> (342)..(343)
<223> n is a, c, g, ou t
<400> 279
gtcgacccac gcgtccggta cttcgctgct gaacagcgac gccaacaaca ggctttcatt 60
gacaagtaca actttgatcc tgtaaatgac tgccctctcc caggcaggtt tgaatgggtg 120
aagctagact gattgattca gaggacaaga gagcagcagc atggaactca cctccgctcc 180
ctccaccacc gcagtgttgt ggcagaggcg cataccgtcg tgttagcttt gtttctgttg 240
taaaaactta gtgttagcct gtagccttaa ttgtcgtgtg ttacagtaca aaactgatgc 300
tgagttacaa caccctgatc tgatctgatc cctcaactca gnngtaaccc ttaacagctt 360
attctgtaag gaaccttacc cacccttgtt accagtt 397

<210> 280
<211> 37
<212> PRT
<213> Saccharum officinarum
<400> 280
Tyr Phe Ala Ala Glu Gln Arg Arg Gln Gln Gln Ala Phe Ile Asp Lys
1 5 10 15
Tyr Asn Phe Asp Pro Val Asn Asp Cys Pro Leu Pro Gly Arg Phe Glu
20 25 30
Trp Val Lys Leu Asp
35
<210> 281
<211> 654
<212> DNA
<213> Oryza sativa
<400> 281
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> 282
<211> 1243
<212> DNA
<213> Oryza sativa
<400> 282
aaaaccaccg agggacctga tctgcaccgg ttttgatagt tgagggaccc gttgtgtctg 60
gttttccgat cgagggacga aaatcggatt cggtgtaaag ttaagggacc tcagatgaac 120
ttattccgga gcatgattgg gaagggagga cataaggccc atgtcgcatg tgtttggacg 180
gtccagatct ccagatcact cagcaggatc ggccgcgttc gcgtagcacc cgcggtttga 240
ttcggcttcc cgcaaggcgg cggccggtgg ccgtgccgcc gtagcttccg ccggaagcga 300
gcacgccgcc gccgccgacc cggctctgcg tttgcaccgc cttgcacgcg atacatcggg 360
atagatagct actactctct ccgtttcaca atgtaaatca ttctactatt ttccacattc 420
atattgatgt taatgaatat agacatatat atctatttag attcattaac atcaatatga 480
atgtaggaaa tgctagaatg acttacattg tgaattgtga aatggacgaa gtacctacga 540
tggatggatg caggatcatg aaagaattaa tgcaagatcg tatctgccgc atgcaaaatc 600
ttactaattg cgctgcatat atgcatgaca gcctgcatgc gggcgtgtaa gcgtgttcat 660
ccattaggaa gtaaccttgt cattacttat accagtacta catactatat agtattgatt 720
tcatgagcaa atctacaaaa ctggaaagca ataagaaata cgggactgga aaagactcaa 780
cattaatcac caaatatttc gccttctcca gcagaatata tatctctcca tcttgatcac 840
tgtacacact gacagtgtac gcataaacgc agcagccagc ttaactgtcg tctcaccgtc 900
gcacactggc cttccatctc aggctagctt tctcagccac ccatcgtaca tgtcaactcg 960
gcgcgcgcac aggcacaaat tacgtacaaa acgcatgacc aaatcaaaac caccggagaa 1020
gaatcgctcc cgcgcgcggc ggcgacgcgc acgtacgaac gcacgcacgc acgcccaacc 1080
ccacgacacg atcgcgcgcg acgccggcga caccggccgt ccacccgcgc cctcacctcg 1140
ccgactataa atacgtaggc atctgcttga tcttgtcatc catctcacca ccaaaaaaaa 1200
aaggaaaaaa aaacaaaaca caccaagcca aataaaagcg aca 1243
<210> 283
<211> 315
<212> DNA
<213> Nicotiana tabacum
<400> 283
ctgccattct ttagagggga tgcttgttta agaacaaaaa atatatcact ttcttttgtt 60
ccaagtcatt gcgtattttt ttaaaaatat ttgttccttc gtatatttcg agcttcaatc 120
actttatggt tctttgtatt ctggctttgc tgtaaatcgt agctaacctt cttcctagca 180
gaaattatta atacttggga tatttttta gaatcaagta aattacatat taccaccaca 240
tcgagctgct tttaaattca tattacagcc atataggctt gattcatttt gcaaaatttc 300

caggatattg acaac 315

<210> 284
<211> 53
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm00472
<400> 284
ggggacaagt ttgtacaaaa aagcaggctt cacaatgggc aagtacatgc gca 53

<210> 285
<211> 53
<212> DNA
<213> Seqüência artificial
<220>
<223> iniciador: prm00473
<400> 285
ggggaccact ttgtacaaga aagctgggtg gagcagagag gtccatggtg ccc 53

REIVINDICAÇÕES

1. Proteína SYR isolada selecionada a partir do grupo, caracterizada pelo fato de consistir de:

(a) um polipeptídeo como dado em SEQ ID NO 44;

(b) um polipeptídeo com uma seqüência de aminoácidos que tem pelo menos, em ordem crescente de preferência, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ou 99% ou mais de identidade de seqüência com uma seqüência de aminoácidos como dada em SEQ ID NO 44;

(c) um derivado de um polipeptídeo como definido em (a) ou (b),

ou uma sequência de ácido nucleico isolada compreendendo:

(a) uma seqüência de ácidos nucleicos representada por SEQ ID NO: 43, ou a fita complementar desta;

(b) uma seqüência de ácidos nucleicos codificando a seqüência de aminoácidos representada por SEQ ID NO: 44;

(c) uma seqüência de ácidos nucleicos capaz de hibridizar (preferivelmente em condições estringentes) com uma seqüência de ácidos nucleicos de (a) ou (b), cuja seqüência hibridizante codifica preferivelmente uma proteína SYR;

(d) um ácido nucleico que é uma variante alélica para as seqüências de ácidos nucleicos de acordo com (a) ou (b);

(e) um ácido nucleico que é uma variante alternativa de união para as seqüências de ácidos nucleicos de acordo com (a) ou (b);

(f) uma seqüência de ácidos nucleicos que tem 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de seqüência com a seqüência definida em (a) ou (b).

2. Método para aumentar produção de semente e/ou aumentar taxa de crescimento de plantas em relação a correspondentes plantas do tipo

selvagem, caracterizado pelo fato de compreender modular expressão em uma planta de um ácido nucleico codificando um polipeptídeo SYR ou um homólogo do mesmo, e opcionalmente selecionar plantas tendo características de crescimento melhoradas; contanto que dita proteína SYR ou homólogo desta não seja a proteína de SEQ ID NO: 26.

3. Método de acordo com a reivindicação 2, caracterizado pelo fato de que dita expressão modulada é efetuada introduzindo uma modificação genética preferivelmente no local de um gene codificando um polipeptídeo SYR ou um homólogo do mesmo, preferivelmente em que dita modificação genética é efetuada por um de: ativação de T-DNA, TILLING, mutagênese direcionada por sítio, recombinação homóloga ou evolução direta.

4. Método para aumentar produção de semente e/ou aumentar taxa de crescimento em relação a correspondentes plantas do tipo selvagem, caracterizado pelo fato de compreender introduzir e expressar em uma planta um ácido nucleico SYR ou uma variante do mesmo; contanto que dita proteína SYR ou homólogo desta não seja a proteína de SEQ ID NO: 26.

5. Método de acordo com a reivindicação 4, caracterizado pelo fato de que dito ácido nucleico codifica um homólogo da proteína SYR de SEQ ID NO: 2, preferivelmente dito ácido nucleico codifica um ortólogo ou parálogo da proteína SYR de SEQ ID NO: 2.

6. Método de acordo com a reivindicação 4, caracterizado pelo fato de que dita variante é uma porção de um ácido nucleico SYR ou uma seqüência capaz de hibridizar com um ácido nucleico SYR, cuja porção ou seqüência hibridizante codifica um polipeptídeo de cerca de 65 a cerca de 200 aminoácidos, compreendendo um domínio rico em leucina, precedido pelo motivo conservado de tripeptídeo 1 (um de SEQ ID NO: 6, 7, 8 ou 9) e seguido pelo motivo conservado 2 (SEQ ID NO: 10) e preferivelmente também pelo motivo conservado 3 (SEQ ID NO: 11).

7. Método de acordo com qualquer uma das reivindicações 4 a 6, caracterizado pelo fato de que dito ácido nucleico compreende os motivos conservados de SEQ ID NO: 6, SEQ ID NO: 10 e SEQ ID NO: 11, em que o motivo de SEQ ID NO: 10 é VLAFMPT e em que o motivo de SEQ ID NO: 11 é PYL, preferivelmente em que dito ácido nucleico compreende a seqüência de SEQ ID NO: 1.

8. Método de acordo com qualquer uma das reivindicações 4 a 7, caracterizado pelo fato de que dito ácido nucleico SYR ou variante do mesmo é superexpressado em uma planta.

9. Método de acordo com qualquer uma das reivindicações 4 a 8, caracterizado pelo fato de que dito ácido nucleico SYR ou variante do mesmo é de origem vegetal, preferivelmente de uma planta monocotiledônea, mais preferivelmente da família Poaceae, o mais preferivelmente o ácido nucleico é de Oryza sativa.

10. Método de acordo com qualquer uma das reivindicações 4 a 9, caracterizado pelo fato de que dito ácido nucleico SYR ou variante do mesmo é operavelmente ligado a um promotor constitutivo, preferivelmente em que dito promotor constitutivo é um promotor GOS2 ou um promotor de proteína de alta mobilidade.

11. Método de acordo com qualquer uma das reivindicações 1 a 10, caracterizado pelo fato de que dita produção aumentada de sementes é selecionada a partir de: peso total aumentado de sementes, número aumentado de sementes cheias, taxa de enchimento de sementes ou índice de colheita aumentado, e em que dita taxa de crescimento aumentada compreende pelo menos produção aumentada de sementes obtida sem atraso em época de florescimento.

12. Método de acordo com qualquer uma das reivindicações 1 a 11, caracterizado pelo fato de que ditas plantas são crescidas em condições sem estresse, ou em que ditas plantas são crescidas em condições de estresse

abiótico, preferivelmente em que ditas condições de estresse abiótico são condições de estresse osmótico.

13. Célula vegetal, caracterizada pelo fato de ser obtenível por um método como definido em qualquer uma das reivindicações 1 a 12.

14. Construção, caracterizada pelo fato de compreender:

(i) um ácido nucleico SYR ou uma variante do mesmo;

(ii) uma ou mais seqüências de controle capazes de dirigir expressão da seqüência de ácidos nucleicos de (i); preferivelmente em que dita seqüência de controle é um promotor constitutivo ou um promotor de Proteína de Grupo de Alta Mobilidade (HMGP), e opcionalmente

(iii) uma seqüência de término de transcrição,

contanto que dito ácido nucleico SYR não codifique a proteína de SEQ ID NO: 26.

15. Célula vegetal, caracterizada pelo fato de ser transformada com uma construção como definida na reivindicação 14.

16. Método para produzir uma planta transgênica tendo produção aumentada comparada com correspondentes plantas do tipo selvagem, caracterizado pelo fato de compreender:

(i). introduzir e expressar em uma planta ou célula vegetal um ácido nucleico SYR ou variante do mesmo; e

(ii). cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal,

com a condição de que dito ácido nucleico SYR ou variante do mesmo não codifique a proteína de SEQ ID NO: 26.

17. Método para melhorar características de crescimento de plantas em relação a plantas de controle, caracterizado pelo fato de compreender aumentar expressão em uma planta de um ácido nucleico codificando um polipeptídeo FG-GAP ou um homólogo do mesmo, e opcionalmente selecionar plantas tendo características de crescimento

melhoradas.

18. Método de acordo com a reivindicação 17, caracterizado pelo fato de que dita expressão aumentada é efetuada introduzindo uma modificação genética preferivelmente no local de um gene codificando um polipeptídeo FG-GAP ou um homólogo do mesmo, preferivelmente em que dita modificação genética é efetuada por um de: ativação de T-DNA, TILLING, mutagênese direcionada por sítio, recombinação homóloga ou evolução direta.

19. Método para melhorar características de crescimento em relação a correspondentes plantas do tipo selvagem, caracterizado pelo fato de compreender introduzir e expressar em uma planta um ácido nucleico FG-GAP ou uma variante do mesmo.

20. Método de acordo com a reivindicação 19, caracterizado pelo fato de que dito ácido nucleico codifica um homólogo da proteína FG-GAP de SEQ ID NO: 46, preferivelmente dito ácido nucleico codifica um ortólogo ou parálogo da proteína FG-GAP de SEQ ID NO: 46, preferivelmente em que dito ácido nucleico compreende um ou mais dos motivos conservados de SEQ ID NO: 50 a 52.

21. Método de acordo com a reivindicação 19, caracterizado pelo fato de que dita variante é uma porção de um ácido nucleico FG-GAP ou uma seqüência capaz de hibridizar com um ácido nucleico FG-GAP, cuja porção ou seqüência hibridizante codifica um polipeptídeo compreendendo um peptídeo sinal, um ou mais domínios FG-GAP e um domínio transmembrana localizado na metade C-terminal da proteína.

22. Método de acordo com qualquer uma das reivindicações 21 a 25, caracterizado pelo fato de que dita característica de crescimento melhorada é produção aumentada, preferivelmente produção aumentada de sementes, preferivelmente em que dita produção aumentada de sementes é selecionada a partir de: peso total aumentado de sementes, número aumentado

de sementes cheias ou índice de colheita aumentado.

23. Célula vegetal, caracterizada pelo fato de ser obtenível por um método como definido em qualquer uma das reivindicações 17 a 22.

24. Construção, caracterizada pelo fato de compreender:

(i) um ácido nucleico FG-GAP ou uma variante do mesmo;

(ii) uma ou mais seqüências de controle capazes de dirigir expressão da seqüência de ácidos nucleicos de (i), preferivelmente em que dita seqüência de controle é um promotor constitutivo, preferivelmente um promotor constitutivo é um promotor GOS2; e opcionalmente

(iii) uma seqüência de término de transcrição,

contanto que dita construção não seja uma construção de gene tipo pPZP.

25. Célula vegetal, caracterizada pelo fato de ser transformada com uma construção como definida na reivindicação 24.

26. Método para produzir uma planta transgênica tendo produção aumentada comparada com correspondentes plantas do tipo selvagem, caracterizado pelo fato de compreender:

(a) introduzir e expressar em a planta ou célula vegetal um ácido nucleico FG-GAP ou variante do mesmo;

(b) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

27. Proteína FG-GAP isolada, caracterizada pelo fato de ser selecionada a partir do grupo consistindo de:

(a) uma proteína codificada pelo ácido nucleico de SEQ ID NO: 72;

(b) uma proteína compreendendo uma seqüência sinal, um ou mais domínios FG-GAP e um domínio transmembrana localizado na metade C-terminal da proteína, em que dita proteína compreende pelo menos um de SEQ ID NO: 73 a SEQ ID NO: 76;

(c) um fragmento ativo de uma seqüência de aminoácidos como definido em (a) ou (b), cujo fragmento ativo compreende uma seqüência sinal, um ou mais domínios FG-GAP e um domínio transmembrana localizado na metade C-terminal da proteína,

ou um ácido nucleico isolado codificando uma proteína FG-GAP, selecionado a partir do grupo consistindo de:

(i) o ácido nucleico como representado em SEQ ID NO: 72;

(ii) um ácido nucleico codificando uma proteína como definido em (a) a (c) acima;

(iii) uma seqüência de ácidos nucleicos capaz de hibridizar (preferivelmente em condições estringentes) com uma seqüência de ácidos nucleicos de (i) ou (ii) acima, cuja seqüência hibridizante preferivelmente codifica uma proteína compreendendo um peptídeo sinal, um ou mais domínios FG-GAP e um domínio de transmembrana localizado na metade C-terminal da proteína;

(iv) um ácido nucleico que é uma variante alélica para as seqüências de ácidos nucleicos de acordo com (i) a (iii);

(v) um ácido nucleico que é uma variante alternativa de união para as seqüências de ácidos nucleicos de acordo com (i) a (iii);

(vi) uma porção de uma seqüência de ácidos nucleicos de acordo com qualquer um de (i) a (v) acima, cuja porção preferivelmente codifica uma proteína compreendendo um peptídeo sinal, um ou mais domínios FG-GAP e um domínio transmembrana localizado na metade C-terminal da proteína.

28. Proteína CYP90B isolada, caracterizada pelo fato de ser selecionada a partir do grupo consistindo de:

(i) uma proteína codificada pelo ácido nucleico de SEQ ID NO: 117;

(ii) uma proteína compreendendo o seguinte: (i) domínios

CYP A a D; (ii) um domínio âncora hidrofóbico N-terminal; (iii) um domínio de transição; e (iv) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição, e tendo em ordem crescente de preferência pelo menos 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% de identidade com a seqüência de aminoácidos de SEQ ID NO: 118,

ou um ácido nucleico isolado codificando uma proteína CYP90B, selecionado a partir do grupo consistindo de:

(i) um ácido nucleico como representado por SEQ ID NO: 117;

(ii) um ácido nucleico codificando uma proteína como definido em (i) e (ii) acima;

(iii) um ácido nucleico tendo em ordem crescente de preferência pelo menos 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais de identidade com o ácido nucleico representado por SEQ ID NO: 117;

(iv) uma seqüência de ácidos nucleicos capaz de hibridizar em condições estringentes com uma seqüência de ácidos nucleicos de (i) a (iii) acima, cuja seqüência hibridizante codifica uma proteína compreendendo (a) domínios CYP A a D; (b) um domínio âncora hidrofóbico N-terminal; (c) um domínio de transição; e (d) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição, e tendo em ordem crescente de preferência pelo menos 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais de identidade com a seqüência de aminoácidos de SEQ ID NO: 118;

(v) um ácido nucleico que é uma variante alélica ou uma variante de união das seqüências de ácidos nucleicos de acordo com (i) a (iv);

(vi) uma porção de uma seqüência de ácidos nucleicos de

acordo com qualquer uma de (i) a (v) acima, cuja porção codifica uma proteína compreendendo: (i) domínios CYP A a D; (ii) um domínio âncora hidrofóbico N-terminal; (iii) um domínio de transição; e (iv) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição, e tendo em ordem crescente de preferência pelo menos 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% de identidade com a seqüência de aminoácidos de SEQ ID NO: 118.

29. Método para aumentar produção vegetal em relação a plantas de controle adequadas, caracterizado pelo fato de compreender aumentar expressão não constitutiva em uma planta de um ácido nucleico codificando ácido nucleico codificando uma monooxigenase CYP90B de citocromo P450 (CYP) ou um homólogo do mesmo, e opcionalmente selecionar plantas tendo produção aumentada, em que dito polipeptídeo CYP90B ou homólogo compreende o seguinte: (a) domínios CYP A a D; (b) um domínio âncora hidrofóbico N-terminal; (c) um domínio de transição; e (d) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição.

30. Método de acordo com a reivindicação 29, caracterizado pelo fato de que dito polipeptídeo CYP90B ou homólogo do mesmo compreende adicionalmente: (i) uma seqüência com mais do que 50% de identidade com SEQ ID NO: 78; e (ii) atividade enzimática de 22-alfa hidroxilase esteróide.

31. Método de acordo com a reivindicação 29 ou 30, caracterizado pelo fato de que dita expressão não constitutiva aumentada é efetuada introduzindo uma modificação genética preferivelmente no local de um gene codificando um polipeptídeo CYP90B ou um homólogo do mesmo, preferivelmente em que dita modificação genética é efetuada por um de: ativação de T-DNA, TILLING, mutagênese direcionada por sítio ou evolução

direta.

32. Método para aumentar produção vegetal em relação a plantas de controle adequadas, caracterizado pelo fato de compreender introduzir e expressar não constitutivamente em uma planta um ácido nucleico CYP90B ou uma variante do mesmo, preferivelmente em que dita variante é uma porção de um ácido nucleico CYP90B, cuja porção codifica um polipeptídeo compreendendo: (a) domínios CYP A a D; (b) um domínio âncora hidrofóbico N-terminal; (c) um domínio de transição; e (d) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição, ou preferivelmente em que dita variante é uma seqüência capaz de hibridizar com um ácido nucleico CYP90B, cuja seqüência hibridizante codifica um polipeptídeo compreendendo: (a) domínios CYP A a D; (b) um domínio âncora hidrofóbico N-terminal; (c) um domínio de transição; e (d) dentro do domínio A, a seqüência consenso Phe-Ala-Gly-His-Glu-Thr-Ser-Ser, admitindo uma mudança de aminoácido em qualquer posição, ou preferivelmente em que dita variante codifica um ortólogo ou parálogo da proteína CYP90B de SEQ ID NO: 78.

33. Método de acordo com a reivindicação 32, caracterizado pelo fato de que dito ácido nucleico CYP90B ou variante do mesmo é de origem vegetal, preferivelmente de uma planta monocotiledônea, ainda preferivelmente da família Poaceae, mais preferivelmente do gênero *Oryza*, o mais preferivelmente de *Oryza sativa*.

34. Método de acordo com qualquer uma das reivindicações 31 a 33, caracterizado pelo fato de que dito ácido nucleico CYP90B ou variante do mesmo é operavelmente ligado a um promotor não constitutivo, preferivelmente em que dito promotor não constitutivo é um promotor específico de semente, preferivelmente dito promotor específico de semente é um promotor específico de endosperma, ainda preferivelmente o promotor

específico de endosperma é um promotor de prolamina, preferivelmente em que dito promotor específico de endosperma é um promotor de prolamina RP6 de arroz, mais preferivelmente o promotor específico de endosperma é representado por uma seqüência de ácidos nucleicos substancialmente similar a SEQ ID NO: 109, o mais preferivelmente o promotor específico de endosperma é como representado por SEQ ID NO: 109, ou em que dito promotor não constitutivo é um promotor específico de semente, preferivelmente dito promotor específico de semente é um promotor específico de embrião/aleurona, ainda preferivelmente o promotor específico de embrião/aleurona é um promotor de oleosina, preferivelmente em que dito promotor específico de embrião/aleurona é um promotor de 18 kDa de oleosina de arroz, mais preferivelmente o promotor específico de embrião/aleurona é representado por uma seqüência de ácidos nucleicos substancialmente similar a SEQ ID NO: 110, o mais preferivelmente o promotor específico de embrião/aleurona é como representado por SEQ ID NO:110.

35. Método de acordo com qualquer uma das reivindicações 29 a 34, caracterizado pelo fato de que dita produção aumentada é selecionada a partir de um ou ambos de: produção total de sementes aumentada ou índice de colheita aumentado (HI), cada em relação a plantas de controle adequadas.

36. Método de acordo com qualquer uma das reivindicações 29 a 34, caracterizado pelo fato de que dita produção aumentada é selecionada a partir de um ou mais de: peso de mil sementes aumentado (TKW), área de semente aumentada ou comprimento de semente aumentado, cada em relação a plantas de controle adequadas.

37. Método de acordo com qualquer uma das reivindicações 29 a 36, caracterizado pelo fato de que dita planta é uma planta monocotiledônea.

38. Célula vegetal, caracterizada pelo fato de ser obtenível por

um método como definido em qualquer uma das reivindicações 29 a 37.

39. Célula vegetal, caracterizada pelo fato de ser transformada com uma construção compreendendo um ácido nucleico CYP90B ou variante do mesmo sob o controle de um promotor de prolamina RP6.

40. Método para produzir uma planta transgênica tendo produção aumentada em relação à planta de controle adequada, caracterizado pelo fato de compreender:

(a) introduzir e expressar não constitutivamente em uma planta ou célula vegetal um ácido nucleico CYP90B ou variante do mesmo;

(b) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

97. Planta transgênica, caracterizada pelo fato de ter produção aumentada em relação à planta de controle adequada, dito produção aumentada resultando de um ácido nucleico CYP90B ou uma variante do mesmo introduzido e expresso não constitutivamente em dita planta.

41. Método para aumentar número de sementes em plantas em relação a plantas de controle adequadas, caracterizado pelo fato de compreender aumentar preferivelmente expressão no tecido de meristema apical de broto de uma planta de um ácido nucleico codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo, e opcionalmente selecionar plantas tendo número de sementes aumentado.

42. Método de acordo com a reivindicação 41, caracterizado pelo fato de que dita expressão aumentada é efetuada introduzindo uma modificação genética preferivelmente no local de um gene codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo, preferivelmente em que dita modificação genética efetuada é mutagênese direcionada por sítio ou evolução direta.

43. Método para aumentar número de sementes em plantas em

relação a plantas de controle adequadas, caracterizado pelo fato de compreender introduzir e expressar preferencialmente no tecido de meristema apical de broto de uma planta um ácido nucleico codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo, e/ou em que dito ácido nucleico introduzido é uma variante de união ou uma variante alélica de um ácido nucleico representado por SEQ ID NO: 129 ou SEQ ID NO: 131, cuja variante de união ou variante alélica codifica um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo, e/ou em que dito ácido nucleico introduzido é capaz de hibridizar com um ácido nucleico como representado por SEQ ID NO: 129 ou SEQ ID NO: 131 ou com uma variante de união ou variante alélica como definida na reivindicação 110, em que dita seqüência hibridizante codifica um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo, e/ou em que dito ácido nucleico CDC27 codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo codifica um ortólogo ou parálogo de um polipeptídeo CDC27 representado por SEQ ID NO: 130 ou SEQ ID NO: 132.

44. Método de acordo com qualquer uma das reivindicações 41 a 43, caracterizado pelo fato de que dito ácido nucleico CDC27 codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo, ou o próprio polipeptídeo, é de origem vegetal, preferivelmente de uma planta dicotiledônea, preferivelmente da família Brassicaceae, ainda preferivelmente de Arabidopsis thaliana.

45. Método de acordo com qualquer uma das reivindicações 41 a 44, caracterizado pelo fato de que dito ácido nucleico CDC27 codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo é operavelmente ligado a um

promotor de meristema apical de broto, preferivelmente a um promotor de meristema apical de broto precoce, preferivelmente em que dito promotor de meristema apical de broto é um promotor OSH1.

46. Célula vegetal, caracterizada pelo fato de ser obtenível por um método como definido em qualquer uma das reivindicações 41 a 45.

47. Célula vegetal transformada com uma construção, caracterizada pelo fato de compreender:

(i) um ácido nucleico CDC27 codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo; e

(ii) uma ou mais seqüências de controle capazes de aumentar preferencialmente expressão da seqüência de ácidos nucleicos de (i) em tecido de meristema apical de broto de uma planta, preferivelmente em que dita seqüência de controle é um promotor OSH1; e opcionalmente

(iii) uma seqüência de término de transcrição

48. Método para produzir uma planta transgênica tendo número de sementes aumentado em relação a plantas de controle adequadas, caracterizado pelo fato de compreender:

(a) introduzir e expressar em uma planta um ácido nucleico CDC27 codificando um polipeptídeo CDC27 tendo pelo menos um domínio TPR inativo na região NH_2 terminal do polipeptídeo operavelmente ligado a um promotor meristema apical de broto específico; e

(b) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

49. Método para aumentar produção de semente em uma planta monocotiledônea em relação à produção de sementes de plantas de controle adequadas, caracterizado pelo fato de compreender aumentar preferencialmente expressão em tecido de endosperma de uma planta monocotiledônea de um ácido nucleico codificando um polipeptídeo

compreendendo um domínio AT-hook e um domínio DUF296, preferivelmente em que dito polipeptídeo compreende adicionalmente um dos seguintes motivos: Motivo 1: QGQ V/I GG; ou Motivo 2: ILSLSGSFLPPPAPP; ou Motivo 3: NATYERLP; ou Motivo 4: SFTNVAYERLPL com nenhuma ou uma mudança de aminoácido em qualquer posição; ou Motivo 5: GRFEILSLTGSFLPGPAPPGSTGLTI YLAGGQGQVVGGSVVG com nenhuma, uma ou duas mudança de aminoácido em qualquer posição.

50. Método de acordo com a reivindicação 49, caracterizado pelo fato de que dita expressão aumentada é efetuada introduzindo uma modificação genética preferivelmente no local de um gene codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296.

51. Método para aumentar produção de semente em uma planta monocotiledônea em relação à produção de sementes de plantas de controle adequadas, caracterizado pelo fato de compreender introduzir e expressar preferencialmente em tecido de endosperma de uma planta monocotiledônea de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296, e/ou em que dito polipeptídeo compreende adicionalmente um dos seguintes motivos: Motivo 1: QGQ V/I GG; ou Motivo 2: ILSLSGSFLPPPAPP; ou Motivo 3: NATYERLP; ou Motivo 4: SFTNVAYERLPL com nenhuma ou uma mudança de aminoácido em qualquer posição; ou Motivo 5: GRFEILSLTGSFLPGPAPPGSTGLTIY LAGGQGQVVGGSVVG com nenhuma, uma ou duas mudança de aminoácido em qualquer posição.

52. Método de acordo com a reivindicação 51, caracterizado pelo fato de que dito ácido nucleico é operavelmente ligado a um promotor específico de endosperma, preferivelmente a um promotor de prolamina.

53. Método de acordo com qualquer uma das reivindicações 127 a 133, caracterizado pelo fato de que dito ácido nucleico é uma porção ou

uma variante alélica ou uma variante de união ou uma seqüência capaz de hibridizar com uma seqüência de acordo com qualquer uma de SEQ ID NO: 152, SEQ ID NO: 153, SEQ ID NO: 154, SEQ ID NO: 156, SEQ ID NO: 158, SEQ ID NO: 160, SEQ ID NO: 162, SEQ ID NO: 164, SEQ ID NO: 166, SEQ ID NO: 168 e SEQ ID NO: 170, em que dita porção, variante alélica, variante de união ou seqüência hibridizante codifica um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296, e/ou em que dita porção, variante alélica, variante de união ou seqüência hibridizante codifica um ortólogo ou parálogo da proteína AT-hook de SEQ ID NO: 153.

54. Método de acordo com qualquer uma das reivindicações 49 a 53, caracterizado pelo fato de que dito ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 é de origem vegetal, preferivelmente de uma planta monocotiledônea, ainda preferivelmente da família Poaceae, mais preferivelmente do gênero *Oryza*, o mais preferivelmente de *Oryza sativa*.

55. Método de acordo com qualquer uma das reivindicações 49 a 54, caracterizado pelo fato de que dita produção aumentada é selecionada a partir de um ou mais de: peso total de sementes aumentado, número aumentado de sementes cheias, número total aumentado de sementes, número aumentado de flores por panícula, índice de colheita aumentado (HI).

56. Célula vegetal, caracterizada pelo fato de ser obtenível por um método como definido em qualquer uma das reivindicações 49 a 55, em que dita planta ou parte da mesma compreende um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 cujo ácido nucleico é operavelmente ligado a um promotor específico de endosperma.

57. Construção de gene, caracterizada pelo fato de compreender:

(a) Um ácido nucleico codificando um polipeptídeo

compreendendo um domínio AT-hook e um domínio DUF296;

(b) Uma ou mais seqüências de controle capazes de dirigir expressão da seqüência de ácidos nucleicos de (i) em tecido de endosperma de uma planta monocotiledônea, preferivelmente em que dita seqüência de controle é um promotor de prolamina; e opcionalmente

(c) uma seqüência de término de transcrição.

58. Célula vegetal, caracterizada pelo fato de ser transformada com uma construção como definida na reivindicação 57.

59. Método para produzir uma planta transgênica monocotiledônea tendo produção aumentada de sementes em relação a plantas de controle adequadas, caracterizado pelo fato de compreender:

(a) introduzir e aumentar preferencialmente expressão em tecido de endosperma de uma planta monocotiledônea de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296; e

(b) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

60. Método para aumentar produção vegetal em relação a plantas de controle adequadas, caracterizado pelo fato de compreender aumentar expressão em uma planta de um ácido nucleico codificando um polipeptídeo de fator de transcrição de domínio DOF (ligação de DNA com um dedo) compreendendo característica (i) como segue, e adicionalmente ou característica (ii) ou (iii) como seguem:

(i) pelo menos 60% identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e

(ii) pelo menos 70% identidade de seqüência com o domínio DOF representado por SEQ ID NO: 200; ou

(iii) Motivo I: KALKKPDKILP (SEQ ID NO: 229) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em

qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

(iv) Motivo II: DDPGIKLFGKTIPF (SEQ ID NO: 230) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição,

e/ou em que dito polipeptídeo compreendendo característica (i) e característica (iii) compreende adicionalmente qualquer um, quaisquer dois ou todos os três dos seguintes motivos:

(i) Motivo III: SPTLGKHSRDE (SEQ ID NO: 231) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

(ii) Motivo IV: LQANPAALSRSQNFQE (SEQ ID NO: 232) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

(iii) Motivo V: KGEGCLWVPKTLRIDDPDEAAKSSIW TTLGIK (SEQ ID NO: 233) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas, três, quatro ou cinco mudança(s) não-conservativa(s) em qualquer posição,

e/ou em que dito polipeptídeo compreendendo característica (i) e característica (iii) compreende tanto Motivo I como II.

61. Método de acordo com a reivindicação 60, caracterizado pelo fato de que dita expressão aumentada é efetuada introduzindo uma modificação genética preferivelmente no local de um gene codificando um polipeptídeo de fator de transcrição DOF, preferivelmente em que dita modificação genética é efetuada por um de: ativação de T-DNA, TILLING e recombinação homóloga.

62. Método para aumentar produção vegetal em relação a plantas de controle adequadas, caracterizado pelo fato de compreender introduzir e expressar em uma planta um ácido nucleico ou variante do mesmo codificando um polipeptídeo de fator de transcrição DOF compreendendo característica (i) como segue, e adicionalmente ou característica (ii) ou (iii) como seguem:

(i) pelo menos 60% identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e

(ii) pelo menos 70% identidade de seqüência com o domínio DOF representado por SEQ ID NO: 200; ou

(iii) Motivo I: KALKKPDKILP (SEQ ID NO: 229) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

(iv) Motivo II: DDPGIKLFGKTIPF (SEQ ID NO: 230) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição,

e/ou em que dito polipeptídeo compreendendo característica (i) e característica (iii) compreende adicionalmente qualquer um, quaisquer dois ou todos os três dos seguintes motivos:

(i) Motivo III: SPTLGKHSRDE (SEQ ID NO: 231) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

(ii) Motivo IV: LQANPAALSRSQNFQE (SEQ ID NO: 232) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

(iii) Motivo V: KGEGCLWVPKTLRIDDPDEAAKSSIWTTLIK (SEQ ID NO: 233) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas, três, quatro ou cinco mudança(s) não-conservativa(s) em qualquer posição,

e/ou em que dito polipeptídeo compreende característica (i) e característica (iii) compreende tanto Motivo I como II.

63. Método de acordo com a reivindicação 62, caracterizado pelo fato de que dito ácido nucleico ou variante do mesmo fator de transcrição DOF é superexpressado em uma planta.

64. Método de acordo com qualquer uma das reivindicações 62 a 63, caracterizado pelo fato de que dito ácido nucleico ou variante do mesmo é de origem vegetal, preferivelmente de uma planta dicotiledônea, ainda preferivelmente da família Brassicaceae, mais preferivelmente o ácido nucleico é de Arabidopsis thaliana.

65. Método de acordo com qualquer uma das reivindicações 62 a 64, caracterizado pelo fato de que dita variante codifica um homólogo de uma proteína de fator de transcrição DOF de SEQ ID NO: 199 ou SEQ ID NO: 227, preferivelmente em que dito homólogo de uma proteína de fator de transcrição DOF de SEQ ID NO: 199 é representado por qualquer uma de SEQ ID NO: 202, SEQ ID NO: 204, SEQ ID NO: 206, SEQ ID NO: 208, SEQ ID NO: 210, SEQ ID NO: 212, SEQ ID NO: 214, SEQ ID NO: 216, SEQ ID NO: 218, SEQ ID NO: 220 e SEQ ID NO: 222, e/ou em que dito homólogo de uma proteína de fator de transcrição DOF de SEQ ID NO: 227 é representado por qualquer uma de SEQ ID NO: 235, SEQ ID NO: 237, SEQ ID NO: 239, SEQ ID NO: 241, SEQ ID NO: 243, SEQ ID NO: 245, SEQ ID NO: 247, SEQ ID NO: 249, SEQ ID NO: 251, SEQ ID NO: 253 e SEQ ID NO: 255.

66. Método de acordo com qualquer uma das reivindicações 62 a 65, caracterizado pelo fato de que dito ácido nucleico ou variante do

mesmo codificando um polipeptídeo de fator de transcrição DOF é operavelmente ligado a um promotor constitutivo, preferivelmente em que dito promotor constitutivo é um promotor GOS2, preferivelmente um promotor GOS2 de arroz, e/ou em que dito ácido nucleico ou variante do mesmo codificando um polipeptídeo de fator de transcrição DOF é operavelmente ligado a um promotor específico de semente, e/ou em que dito promotor específico de semente é um promotor específico de endosperma, preferivelmente um promotor de prolamina, e/ou em que dito promotor constitutivo dirige expressão de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF compreendendo ditas características (i) e (ii), e/ou em que dito promotor específico de semente dirige expressão de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF compreendendo ditas características (i) e (iii).

67. Método de acordo com qualquer uma das reivindicações 59 a 66, caracterizado pelo fato de que dita produção aumentada é selecionada a partir de um ou mais de: número aumentado de sementes cheias, peso de semente aumentado, número aumentado de flores por panícula, taxa de enchimento de sementes aumentada, índice de colheita aumentado (HI), peso de mil sementes aumentado (TKW), biomassa de raiz aumentada, comprimento de raiz aumentado e diâmetro de raiz aumentado.

68. Método de acordo com qualquer uma das reivindicações 59 a 66, caracterizado pelo fato de que dita produção aumentada ocorre em condições de estresse suave à seco

69. Célula vegetal, caracterizada pelo fato de ser obtenível por um método como definido em qualquer uma das reivindicações 59 a 68.

70. Construção, caracterizada pelo fato de compreender:

(i) um ácido nucleico ou variante do mesmo codificando um polipeptídeo de fator de transcrição DOF como definido em reivindicação 62;

(ii) uma ou mais seqüências de controle capazes de dirigir

expressão da seqüência de ácidos nucleicos de (a) preferivelmente em que dita seqüência de controle é um promotor constitutivo, mais preferivelmente em que dita seqüência de controle é um promotor específico de semente; e opcionalmente

(iii) uma seqüência de término de transcrição.

71. Célula vegetal, caracterizada pelo fato de ser transformada com uma construção como definida na reivindicação 70.

72. Método para produzir uma planta transgênica tendo produção aumentada em relação à planta do tipo selvagem correspondente, caracterizado pelo fato de compreender:

(a) introduzir e expressar em a planta, parte vegetal ou célula vegetal um ácido nucleico ou variante do mesmo codificando um polipeptídeo de fator de transcrição DOF como definido na reivindicação 62;

(b) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

73. Método de acordo com a reivindicação 72, caracterizado pelo fato de que dita produção aumentada ocorre em condições de estresse suave à seco.

74. Método para aumentar produção de semente em plantas em relação a plantas de controle adequadas, caracterizado pelo fato de compreender reduzir preferencialmente expressão de um gene CKI endógeno em tecido de endosperma de uma planta, e/ou reduzir preferencialmente expressão é efetuado por regulação negativa mediada por RNA de expressão gênica, e/ou em que dita regulação negativa mediada por RNA é efetuada por co-supressão, e/ou em que dita regulação negativa mediada por RNA é efetuada por uso de seqüências de ácidos nucleicos de CKI anti-sentido, e/ou em que dito reduzir preferencialmente expressão é efetuado usando uma repetição invertida de um gene CKI ou fragmento deste, preferivelmente capaz de formar uma estrutura em grampo, e/ou em que dito reduzir

preferencialmente expressão é efetuado usando ribozimas com especificidade para um ácido nucleico de CKI, e/ou em que dito reduzir preferencialmente expressão é efetuado por mutagênese por inserção.

75. Método de acordo com a reivindicação 74, caracterizado pelo fato de que dito reduzir preferencialmente expressão de um gene CKI endógeno em tecido de endosperma de uma planta é efetuado por um promotor específico de endosperma, preferivelmente um promotor de prolamina.

76. Método de acordo com qualquer uma das reivindicações 74 a 75, caracterizado pelo fato de que dito gene CKI endógeno é um gene CKI como encontrado em uma planta em sua forma natural ou é um gene CKI isolado subseqüentemente introduzido em uma planta.

77. Método de acordo com qualquer uma das reivindicações 74 a 76, caracterizado pelo fato de que dito gene/nucleico de CKI é de uma fonte vegetal ou fonte artificial, preferivelmente que seqüências de ácidos nucleicos de CKI de plantas monocotiledôneas são usadas para transformação de plantas monocotiledôneas, ainda preferivelmente que seqüências de CKI da família Poaceae são usadas para transformação em plantas da família Poaceae, mais preferivelmente que seqüências de ácidos nucleicos de CKI de arroz são usadas para transformar plantas de arroz.

78. Método de acordo com a reivindicação 77, caracterizado pelo fato de que dita seqüência de ácidos nucleicos de CKI de arroz compreende um comprimento suficiente de nucleotídeos substancialmente contíguos de SEQ ID NO: 267 (OsCKI4) ou compreende um comprimento suficiente de nucleotídeos substancialmente contíguos de uma seqüência de ácidos nucleicos codificando um ortólogo ou parálogo de OsCKI4 (SEQ ID NO: 267), e/ou em que ditos ortólogos ou parálogos de OsCKI4 são representados por SEQ ID NO: 270, SEQ ID NO: 272, SEQ ID NO: 274, SEQ ID NO: 276, SEQ ID NO: 278 e SEQ ID NO: 280, e/ou em que ditos

nucleotídeos substancialmente contíguos de uma seqüência de ácidos nucleicos codificando um ortólogo ou parálogo de OsCKI4 (SEQ ID NO: 267) são nucleotídeos substancialmente contíguos de seqüências de ácidos nucleicos representadas por SEQ ID NO: 269, SEQ ID NO: 271, SEQ ID NO: 273, SEQ ID NO: 275, SEQ ID NO: 277 e SEQ ID NO: 279.

79. Método de acordo com qualquer uma das reivindicações 74 a 78, caracterizado pelo fato de que dita produção aumentada de sementes é selecionada a partir de um ou mais dos seguintes: a) biomassa de semente aumentada; b) número aumentado de flores por planta; c) número aumentado de sementes (cheios); e d) índice de colheita aumentado.

80. Método de acordo com qualquer uma das reivindicações 74 a 80, caracterizado pelo fato de que dita produção aumentada ocorre em condições suaves de estresse.

81. Célula vegetal, caracterizada pelo fato de ser obtenível por um método de acordo com qualquer uma das reivindicações 74 a 80.

82. Método para produzir uma planta transgênica tendo produção aumentada de sementes em relação a plantas de controle adequadas, caracterizado pelo fato de compreender:

(a) introduzir e expressar em a planta, parte vegetal ou célula vegetal uma construção de gene, compreendendo uma ou mais seqüências de controle capazes de dirigir preferencialmente expressão de uma seqüência de ácidos nucleicos de CKI sentido e/ou anti-sentido CKI em tecido de endosperma de planta de forma a silenciar um gene CKI endógeno em tecido de endosperma de uma planta; e

(b) cultivar a planta, parte vegetal ou célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

136. Método de acordo com qualquer uma das reivindicações 127 a 135, caracterizado pelo fato de que dito ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 é de origem vegetal, preferivelmente de uma planta monocotiledônea, ainda preferivelmente da família Poaceae, mais preferivelmente do gênero Oryza, o mais preferivelmente de Oryza sativa.

137. Método de acordo com qualquer uma das reivindicações 127 a 136, caracterizado pelo fato de que dita produção aumentada é selecionada a partir de um ou mais de: peso total de sementes aumentado, número aumentado de sementes cheias, número total aumentado de sementes, número aumentado de flores por panícula, índice de colheita aumentado (HI).

138. Planta ou parte da mesma incluindo célula vegetal obtenível por um método como definido em qualquer uma das reivindicações 127 a 137, caracterizada pelo fato de que dita planta ou parte da mesma compreende um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296 cujo ácido nucleico é operavelmente ligado a um promotor específico de endosperma.

139. Construção de gene, caracterizada pelo fato de compreender:

(a) Um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296;

(b) Uma ou mais seqüências de controle capazes de dirigir expressão da seqüência de ácidos nucleicos de (i) em tecido de endosperma de uma planta monocotiledônea; e opcionalmente

(c) uma seqüência de término de transcrição.

140. Uso de uma construção como definida na reivindicação 139, caracterizado pelo fato de ser para aumentar produção de sementes em plantas monocotiledôneas.

141. Construção de acordo com a reivindicação 139,

caracterizada pelo fato de que dita seqüência de controle é um promotor de prolamina.

142. Planta, parte vegetal ou célula vegetal, caracterizada pelo fato de ser transformada com uma construção como definida na reivindicação 139.

143. Método para produzir uma planta transgênica monocotiledônea tendo produção aumentada de sementes em relação a plantas de controle adequadas, caracterizado pelo fato de compreender:

(a) introduzir e aumentar preferencialmente expressão em tecido de endosperma de uma planta monocotiledônea de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296; e

(b) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

144. Planta transgênica monocotiledônea, caracterizada pelo fato de ter produção aumentada de sementes em relação a plantas de controle adequadas, dita produção aumentada de sementes resultando de expressão preferencial em tecido de endosperma de uma planta monocotiledônea de um ácido nucleico codificando um polipeptídeo compreendendo um domínio AT-hook e um domínio DUF296.

145. Planta transgênica monocotiledônea de acordo com a reivindicação 138, 142 ou 144, caracterizada pelo fato de que dita planta é um cereal, tal como arroz, milho, cana-de-açúcar, trigo, cevada, milheto, centeio, sorgo, gramíneas ou aveia.

146. Partes que podem ser colhidas de uma planta como definida em qualquer uma das reivindicações 138, 142, 144 ou 145, caracterizadas pelo fato de que ditas partes que podem ser colhidas são sementes.

147. Produtos, caracterizados pelo fato de serem derivados de

uma planta como definida na reivindicação 145 e/ou de partes que podem ser colhidas de uma planta como definidas na reivindicação 146.

148. Uso de um ácido nucleico codificando um polipeptídeo, compreendendo um domínio AT-hook e um domínio DUF296, cujo ácido nucleico é operavelmente ligado a um promotor específico de endosperma, caracterizado pelo fato de ser para aumentar produção de sementes em uma planta monocotiledônea comparada com produção de sementes em uma planta de controle adequada.

149. Uso de acordo com a reivindicação 148, caracterizado pelo fato de que dita produção aumentada de sementes é selecionada a partir de um ou mais de: peso total de sementes aumentado, número aumentado de sementes cheias, número total aumentado de sementes número aumentado de flores por panícula, índice de colheita aumentado (HI).

150. Método para aumentar produção vegetal em relação a plantas de controle adequadas, caracterizado pelo fato de compreender aumentar expressão em uma planta de um ácido nucleico codificando um polipeptídeo de fator de transcrição de domínio DOF (ligação de DNA com um dedo) compreendendo característica (i) como segue, e adicionalmente ou característica (ii) ou (iii) como seguem:

(i) pelo menos 60% identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e

(ii) pelo menos 70% identidade de seqüência com o domínio DOF representado por SEQ ID NO: 200; ou

(iii) Motivo I: KALKKPDKILP (SEQ ID NO: 229) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

Motivo II: DDPGIKLFGKTIPF (SEQ ID NO: 230) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em

qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição.

151. Método de acordo com a reivindicação 150, caracterizado pelo fato de que dito polipeptídeo compreendendo característica (i) e característica (iii) compreende adicionalmente qualquer um, quaisquer dois ou todos os três dos seguintes motivos:

- Motivo III: SPTLGKHSRDE (SEQ ID NO: 231) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou
- Motivo IV: LQANPAALSRSQNFQE (SEQ ID NO: 232) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou
- Motivo V: KGEGCLWVPKTLRIDDPDEAAKSSIW TTLGIK (SEQ ID NO: 233) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas, três, quatro ou cinco mudança(s) não-conservativa(s) em qualquer posição.

152. Método de acordo com a reivindicação 150 ou 151, caracterizado pelo fato de que dito polipeptídeo compreendendo característica (i) e característica (iii) compreende tanto Motivo I como II.

153. Método de acordo com qualquer uma das reivindicações 150 a 152, caracterizado pelo fato de que dita expressão aumentada é efetuada introduzindo uma modificação genética preferivelmente no local de um gene codificando um polipeptídeo de fator de transcrição DOF.

154. Método de acordo com a reivindicação 153, caracterizado pelo fato de que dita modificação genética é efetuada por um de: ativação de T-DNA, TILLING e recombinação homóloga.

155. Método para aumentar produção vegetal em relação a

plantas de controle adequadas, caracterizado pelo fato de compreender introduzir e expressar em uma planta um ácido nucleico ou variante do mesmo codificando um polipeptídeo de fator de transcrição DOF compreendendo característica (i) como segue, e adicionalmente ou característica (ii) ou (iii) como seguem:

(i) pelo menos 60% identidade de seqüência ou com o domínio DOF representado por SEQ ID NO: 200 ou SEQ ID NO: 228; e

(ii) pelo menos 70% identidade de seqüência com o domínio DOF representado por SEQ ID NO: 200; ou

(iii) Motivo I: KALKKPDKILP (SEQ ID NO: 229) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

Motivo II: DDPGIKLFGKTIPF (SEQ ID NO: 230) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição.

156. Método de acordo com a reivindicação 155, caracterizado pelo fato de que dito polipeptídeo compreendendo característica (i) e característica (iii) compreende adicionalmente qualquer um, quaisquer dois ou todos os três dos seguintes motivos:

- Motivo III: SPTLGKHSRDE (SEQ ID NO: 231) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

- Motivo IV: LQANPAALSRSQNFQE (SEQ ID NO: 232) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas ou três mudança(s) não-conservativa(s) em qualquer posição; e/ou

- Motivo V: KGEGCLWVPKTLRIDDPDEAAKSSIWTTLIK (SEQ ID NO: 233) sem nenhuma mudança; ou com uma ou mais mudanças conservativas em qualquer posição; ou com uma, duas, três, quatro ou cinco mudança(s) não-conservativa(s) em qualquer posição.

157. Método de acordo com a reivindicação 155 ou 156, caracterizado pelo fato de que dito polipeptídeo compreende característica (i) e característica (iii) compreende tanto Motivo I como II.

158. Método de acordo com reivindicações 155 a 157, caracterizado pelo fato de que dito ácido nucleico ou variante do mesmo fator de transcrição DOF é superexpressado em uma planta.

159. Método de acordo com qualquer uma das reivindicações 155 a 158, caracterizado pelo fato de que dito ácido nucleico ou variante do mesmo é de origem vegetal, preferivelmente de uma planta dicotiledônea, ainda preferivelmente da família Brassicaceae, mais preferivelmente o ácido nucleico é de Arabidopsis thaliana.

160. Método de acordo com qualquer uma das reivindicações 155 a 159, caracterizado pelo fato de que dita variante codifica um homólogo de uma proteína de fator de transcrição DOF de SEQ ID NO: 199 ou SEQ ID NO: 227.

161. Método de acordo com a reivindicação 160, caracterizado pelo fato de que dito homólogo de uma proteína de fator de transcrição DOF de SEQ ID NO: 199 é representado por qualquer uma de SEQ ID NO: 202, SEQ ID NO: 204, SEQ ID NO: 206, SEQ ID NO: 208, SEQ ID NO: 210, SEQ ID NO: 212, SEQ ID NO: 214, SEQ ID NO: 216, SEQ ID NO: 218, SEQ ID NO: 220 e SEQ ID NO: 222.

162. Método de acordo com a reivindicação 160, caracterizado pelo fato de que dito homólogo de uma proteína de fator de transcrição DOF de SEQ ID NO: 227 é representado por qualquer uma de SEQ ID NO: 235, SEQ ID NO: 237, SEQ ID NO: 239, SEQ ID NO: 241, SEQ ID NO: 243,

SEQ ID NO: 245, SEQ ID NO: 247, SEQ ID NO: 249, SEQ ID NO: 251, SEQ ID NO: 253 e SEQ ID NO: 255.

163. Método de acordo com qualquer uma das reivindicações 155 a 162, caracterizado pelo fato de que dito ácido nucleico ou variante do mesmo codificando um polipeptídeo de fator de transcrição DOF é operavelmente ligado a um promotor constitutivo.

164. Método de acordo com a reivindicação 163, caracterizado pelo fato de que dito promotor constitutivo é um promotor GOS2, preferivelmente um promotor GOS2 de arroz.

165. Método de acordo com qualquer uma das reivindicações 155 a 162, caracterizado pelo fato de que dito ácido nucleico ou variante do mesmo codificando um polipeptídeo de fator de transcrição DOF é operavelmente ligado a um promotor específico de semente.

166. Método de acordo com a reivindicação 165, caracterizado pelo fato de que dito promotor específico de semente é um promotor específico de endosperma, preferivelmente um promotor de prolamina.

167. Método de acordo com a reivindicação 163 ou 164, caracterizado pelo fato de que dito promotor constitutivo dirige expressão de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF compreendendo ditas características (i) e (ii).

168. Método de acordo com a reivindicação 165 ou 166, caracterizado pelo fato de que dito promotor específico de semente dirige expressão de um ácido nucleico codificando um polipeptídeo de fator de transcrição DOF compreendendo ditas características (i) e (iii).

169. Método de acordo com qualquer uma das reivindicações 150 a 168, caracterizado pelo fato de que dita produção aumentada é selecionada a partir de um ou mais de: número aumentado de sementes cheias, peso de semente aumentado, número aumentado de flores por panícula, taxa de enchimento de sementes aumentada, índice de colheita

aumentado (HI), peso de mil sementes aumentado (TKW), biomassa de raiz aumentada, comprimento de raiz aumentado e diâmetro de raiz aumentado.

170. Método de acordo com qualquer uma das reivindicações 150 a 169, caracterizado pelo fato de que dita produção aumentada ocorre em condições de estresse suave à seco

171. Planta ou parte da mesma, caracterizada pelo fato de ser obtenível por um método como definido em qualquer uma das reivindicações 150 a 170.

172. Construção, caracterizada pelo fato de compreender:

(i) um ácido nucleico ou variante do mesmo codificando um polipeptídeo de fator de transcrição DOF como definido em reivindicação 155;

(ii) uma ou mais seqüências de controle capazes de dirigir expressão da seqüência de ácidos nucleicos de (a), e opcionalmente

(iii) uma seqüência de término de transcrição.

173. Construção de acordo com a reivindicação 172, caracterizada pelo fato de que dita seqüência de controle é um promotor constitutivo.

174. Construção de acordo com a reivindicação 172, caracterizada pelo fato de que dita seqüência de controle é um promotor específico de semente.

175. Planta, parte vegetal ou célula vegetal, caracterizada pelo fato de ser transformada com uma construção como definida em qualquer uma das reivindicações 172 a 174.

176. Método para produzir uma planta transgênica tendo produção aumentada em relação à planta do tipo selvagem correspondente, caracterizado pelo fato de compreender:

(a) introduzir e expressar em a planta, parte vegetal ou célula vegetal um ácido nucleico ou variante do mesmo codificando um polipeptídeo

de fator de transcrição DOF como definido na reivindicação 155;

(b) cultivar a célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

177. Método de acordo com a reivindicação 176, caracterizado pelo fato de que dita produção aumentada ocorre em condições de estresse suave à seco.

178. Planta transgênica, caracterizada pelo fato de ter produção aumentada em relação à planta do tipo selvagem correspondente, dita produção aumentada resultando de um ácido nucleico ou variante do mesmo codificando um polipeptídeo de fator de transcrição DOF como definido na reivindicação 155 introduzido em dita planta.

179. Planta transgênica de acordo com a reivindicação 171, 175 ou 178, caracterizada pelo fato de que dita planta é uma planta monocotiledônea, tal como cana-de-açúcar ou que a planta é um cereal, tal como arroz, milho, trigo, cevada, milheto, centeio, aveia ou sorgo.

180. Partes que podem ser colhidas de uma planta, caracterizadas por serem de acordo com qualquer uma das reivindicações 171, 175, 178 ou 179.

181. Partes que podem ser colhidas de uma planta como definida em reivindicação 180, caracterizadas pelo fato de que ditas partes que podem ser colhidas são sementes.

182. Produtos, caracterizados pelo fato de serem derivados de uma planta como definida na reivindicação 179 e/ou de partes que podem ser colhidas de uma planta como definidas nas reivindicações 180 ou 181.

183. Uso de um ácido nucleico ou variante do mesmo codificando um polipeptídeo, ou uso de tal polipeptídeo, o polipeptídeo sendo de fator de transcrição DOF, caracterizado pelo fato de ser para aumentar produção vegetal em relação a plantas de controle adequadas.

184. Uso de acordo com a reivindicação 183, caracterizado

pelo fato de que dita produção aumentada é selecionada a partir de um ou mais de: número aumentado de sementes cheias, peso de semente aumentado, número aumentado de flores por panícula, taxa de enchimento de sementes aumentada, índice de colheita aumentado (HI), peso de mil sementes aumentado (TKW), biomassa de raiz aumentada, comprimento de raiz aumentado e diâmetro de raiz aumentado.

185. Uso de acordo com reivindicações 183 ou 184, caracterizado pelo fato de que dita produção aumentada ocorre em condições de estresse suave à seco.

186. Uso de um ácido nucleico ou variante do mesmo codificando um polipeptídeo, ou uso de tal polipeptídeo, o polipeptídeo sendo polipeptídeo de fator de transcrição DOF, caracterizado pelo fato de ser como um marcador molecular.

187. Método para aumentar produção de semente em plantas em relação a plantas de controle adequadas, caracterizado pelo fato de compreender reduzir preferencialmente expressão de um gene CKI endógeno em tecido de endosperma de uma planta.

188. Método de acordo com a reivindicação 187, caracterizado pelo fato de que dito reduzir preferencialmente expressão é efetuado por regulação negativa mediada por RNA de expressão gênica.

189. Método de acordo com a reivindicação 188, caracterizado pelo fato de que dita regulação negativa mediada por RNA é efetuada por co-supressão.

190. Método de acordo com a reivindicação 188, caracterizado pelo fato de que dita regulação negativa mediada por RNA é efetuada por uso de seqüências de ácidos nucleicos de CKI anti-sentido.

191. Método de acordo com a reivindicação 187, caracterizado pelo fato de que dito reduzir preferencialmente expressão é efetuado usando uma repetição invertida de um gene CKI ou fragmento deste, preferivelmente

capaz de formar uma estrutura em grampo.

192. Método de acordo com a reivindicação 187, caracterizado pelo fato de que dito reduzir preferencialmente expressão é efetuado usando ribozimas com especificidade para um ácido nucleico de CKI.

193. Método de acordo com a reivindicação 187, caracterizado pelo fato de que dito reduzir preferencialmente expressão é efetuado por mutagênese por inserção.

194. Método de acordo com qualquer uma das reivindicações 187 a 193, caracterizado pelo fato de que dito reduzir preferencialmente expressão de um gene CKI endógeno em tecido de endosperma de uma planta é efetuado por um promotor específico de endosperma, preferivelmente um promotor de prolamina.

195. Método de acordo com qualquer uma das reivindicações 187 a 194, caracterizado pelo fato de que dito gene CKI endógeno é um gene CKI como encontrado em uma planta em sua forma natural ou é um gene CKI isolado subseqüentemente introduzido em uma planta.

196. Método de acordo com qualquer uma das reivindicações 187 a 197, caracterizado pelo fato de que dito gene/nucleico de CKI é de uma fonte vegetal ou fonte artificial, preferivelmente que seqüências de ácidos nucleicos de CKI de plantas monocotiledôneas são usadas para transformação de plantas monocotiledôneas, ainda preferivelmente que seqüências de CKI da família Poaceae são usadas para transformação em plantas da família Poaceae, mais preferivelmente que seqüências de ácidos nucleicos de CKI de arroz são usadas para transformar plantas de arroz.

197. Método de acordo com a reivindicação 196, caracterizado pelo fato de que dita seqüência de ácidos nucleicos de CKI de arroz compreende um comprimento suficiente de nucleotídeos substancialmente contíguos de SEQ ID NO: 267 (OsCKI4) ou compreende um comprimento suficiente de nucleotídeos substancialmente contíguos de uma seqüência de

ácidos nucleicos codificando um ortólogo ou parálogo de OsCKI4 (SEQ ID NO: 267).

198. Método de acordo com a reivindicação 197, caracterizado pelo fato de que ditos ortólogos ou parálogos de OsCKI4 são representados por SEQ ID NO: 270, SEQ ID NO: 272, SEQ ID NO: 274, SEQ ID NO: 276, SEQ ID NO: 278 e SEQ ID NO: 280.

199. Método de acordo com a reivindicação 197 ou 198, caracterizado pelo fato de que ditos nucleotídeos substancialmente contíguos de uma seqüência de ácidos nucleicos codificando um ortólogo ou parálogo de OsCKI4 (SEQ ID NO: 267) são nucleotídeos substancialmente contíguos de seqüências de ácidos nucleicos representadas por SEQ ID NO: 269, SEQ ID NO: 271, SEQ ID NO: 273, SEQ ID NO: 275, SEQ ID NO: 277 e SEQ ID NO: 279.

200. Método de acordo com qualquer uma das reivindicações 187 a 199, caracterizado pelo fato de que dita produção aumentada de sementes é selecionada a partir de um ou mais dos seguintes: a) biomassa de semente aumentada; b) número aumentado de flores por planta; c) número aumentado de sementes (cheios); e d) índice de colheita aumentado.

201. Método de acordo com qualquer uma das reivindicações 187 a 200, caracterizado pelo fato de que dita produção aumentada ocorre em condições suaves de estresse.

202. Planta ou parte da mesma, caracterizada pelo fato de ser obtenível por um método de acordo com qualquer uma das reivindicações 187 a 201.

203. Método para produzir uma planta transgênica tendo produção aumentada de sementes em relação a plantas de controle adequadas, caracterizado pelo fato de compreender:

(a) introduzir e expressar em a planta, parte vegetal ou célula vegetal uma construção de gene, compreendendo uma ou mais seqüências de

controle capazes de dirigir preferencialmente expressão de uma seqüência de ácidos nucleicos de CKI sentido e/ou anti-sentido CKI em tecido de endosperma de planta de forma a silenciar um gene CKI endógeno em tecido de endosperma de uma planta; e

(b) cultivar a planta, parte vegetal ou célula vegetal em condições promovendo crescimento e desenvolvimento vegetal.

204. Uso de ácido nucleico CKI, caracterizado pelo fato de ser para a redução ou eliminação substancial de expressão de gene CKI endógeno em tecido de endosperma de planta para aumentar produção de sementes em plantas em relação a plantas de controle adequadas.

205. Uso de acordo com a reivindicação 204, caracterizado pelo fato de que dita produção aumentada é selecionada a partir de um ou mais de: selecionado a partir de um ou mais dos seguintes: a) biomassa de semente aumentada; b) número aumentado de flores por planta; c) número aumentado de sementes (cheios); e d) índice de colheita aumentado.

206. Uso de acordo com a reivindicação 204 ou 205, caracterizado pelo fato de que dita produção de sementes ocorre em condições suaves de estresse.

MEGVGARQRRNPLIPRP*NGSKR*HLQHQHQPNAAEKKTAATSN**YFS**IEAFLVLVFLTMSLLIL
PLVLPPLPPPPSLLLLLPVCLLILLV**VLAFMPT**DVRSMAS**SYL**

FIGURA 1

clustal w (1.83) Alinhamento múltiplo de seqüências

```
SEQID34     ------------------------------------------------------------
SEQID35     ------------------------------------------------------------
SEQID20     ----------MIREISNLQKD-----IINIQDS-YSNNRVMDV-GRNNR---KNMSFR
SEQID21     ----------MIREFSSLQND-----IINIQEH-YSLNNNMDVRGDHNR---KNTSFR
SEQID22     ----------MSIEQPEADSRLSEGPLINLQDR--YLSGIMEARGRRNSAPLQVERKN
SEQID23     ----------MNSDNSESRQRLSKG-IINLQDR--YPTSIMD-RG---------VRKI
SEQID24     ----------MNMDMESSEAKLRSSKG-FINLEEHQQYFNNIME--G----------
SEQID25     ------------------------------------ME--G----------
SEQID15     ----------------------------MDSQFGALERG----GSRQRR------SP
SEQID16     ----------------------------MDSQFGAMDRG----GSRQRS------SP
SEQID2      ------------------------------------MEGV----GARQRR------NP
SEQID14     ----------------------------MASRSSAMEGG----AAIQR---------
SEQID19     ----------------------------MASRSSALEGGG---AAIQR---------
SEQID17     ------------------------------------MEGG----GQIQR---------
SEQID18     ------------------------------------MEGG----GQIQR---------
SEQID13     ----------------------------MLLEHLMITMEEQMFREQQMQRG------GR
SEQID12     MYLLSPRNGDEEDEQEEIQELISDDEPPNLKLASCATAASSSSSSGSDMEKGRGKACGGG


SEQID34     ----------MVR----------CFSLGSVLILIALAASMVVLPL
SEQID35     ----------MIMVAS------KEKTNSG--------GCMFRYSVLILSLLALSILVLPL
SEQID20     -SSPE--KSKQELRRSFSAQK--RMMIPA--------NYFSLESLFLLVGLTASLLILPL
SEQID21     GSAPAPIMGKQELFRTLSSQNSPRRLISA--------SYFSLESMVVLVGLTASLLILPL
SEQID22     PTPPMAEGKKMEYNRTPLSRENSRRLIPA--------SYFSLESLLLLICLTASLLILPL
SEQID23     ATPPVEK-RKVEYHRS-YSQGASRKLFSA--------SYFTLESLLLLVCLTASLLILPL
SEQID24     --------NKMEHKRS-FTQGHGKKMLSM--------NYFSLESIILLLGLTASLLLLPL
SEQID25     --------NKMEHKRS-FTQGHGKKMLSM--------NYFSLESIILLLGLTASLLLLPL
SEQID15     VLARPNTTKRHIQQQ--RANAADKKVVMP--------NYFSIEAFFVLACLTVSLLILPL
SEQID16     VLARPNTAKRQMQQQ--RANAADKKVVIP--------NYFGVEAFFVLACLTVSLLILPL
SEQID2      LIPRPNGSKRHLQHQH-QPNAAEKKTAATS-------NYFSIEAFLVLVFLTMSLLILPL
SEQID14     ---RN-AVKRHLQQRQQEADFLDKKVIAS--------TYFSIGAFLVLACLTVSLLILPL
SEQID19     ---RNNAVKRHLQQRQQEADFHDKKVIAS--------TYFSIGAFLVLACLTFSLLILPL
SEQID17     ---RNNAVKRHLQQRQQEADFLDKKVIAS--------TYFSIEAFLVLACLTVSLLILPL
SEQID18     ---RNNAVKRHLQQRQQEADFLDKKVIAS--------TYFSIEAFLVLACLTVSLLILPL
SEQID13     HHQHHTTREQEQQQKQQQRRRLMNNATNGGGGDGGSRCYFSTEAILVLACVTVSLLVLPL
SEQID12     STAPPPPPPSSSGKSGGGGGSNIREAAASGGGGGVWGKYFSVESLLLLVCVTASLVILPL
                                                    :   :..:*  :: *:::***
```

FIGURA 2

```
SEQID34      MLPPLPPPPLALLFFPVGIMAALVVLAFSPSENV----KNVVV----------------
SEQID35      VMPPLPPPPLLLLLVPVFIMLLLFFIAFSPSKKV----PNKAS----------------
SEQID20      VLPPLPPPPFMLLLVPIGIMVLLVVLAFMPSSHSNANTDVTCN----------------
SEQID21      ILPPLPPPPFMLLLIPIGIMVLLMVLAFMPSSNS-KHVSSSST----------------
SEQID22      ILPPLPPPPFMLLLLPIGILAVLMILAFMPSNVR----DLTYT----------------
SEQID23      VLPPLPPPPFLLLLVPIXILAVLLVLAFMPSNVR----DITST----------------
SEQID24      MLPPLPPPPFMLLLVPIFILVVLMILAFMPSNVR----NVTCS----------------
SEQID25      MLPPLPPPPFMLLLVPIFILVVLMILAFMPSNVR----NVTCS----------------
SEQID15      VLPPLPPPPSLLLFVPVCLLILLMVLAFMPTDMR----SMATS----------------
SEQID16      VLPPLPPPPSLLLLLPVCLLILLMVLAFMPTDVR----SMATS----------------
SEQID2       VLPPLPPPPSLLLLLPVCLLILLVVLAFMPTDVR----SMASS----------------
SEQID14      XXXXXXXXXXXXLLWLPVCLLVLLVVLAFMPTDVR----SMASS----------------
SEQID19      VLPPLPPPPSLLLWLPVCLLVLLVVLAFMPTDVR----SVAAS----------------
SEQID17      VLPXLPAPASLLLWLPVWLLELLIVLAFMPTDVR----SMASS----------------
SEQID18      VLPPLPPPPSLLLWLPVCLLILLIVLAFMPTDVR----SMASS----------------
SEQID13      ILPPLPPPPTLLLLLPVCLLALLVVLAFMPTDMR----TMASS----------------
SEQID12      VLPPLPPPPSMLMLVPVAMLVLLLALAFMPTTTS----SSSSAGGGGGGGRNGATTGHAP
                        *: .*: ::  *. :** *:

SEQID34      YSSSSSGIANSKR
SEQID35      FVS----------
SEQID20      FM-----------
SEQID21      FM-----------
SEQID22      YV-----------
SEQID23      YV-----------
SEQID24      YL-----------
SEQID25      YL-----------
SEQID15      YL-----------
SEQID16      YL-----------
SEQID2       YL-----------
SEQID14      YL-----------
SEQID19      YL-----------
SEQID17      YL-----------
SEQID18      YL-----------
SEQID13      YL-----------
SEQID12      YL-----------
             :
```

FIGURA 2 (Cont.)

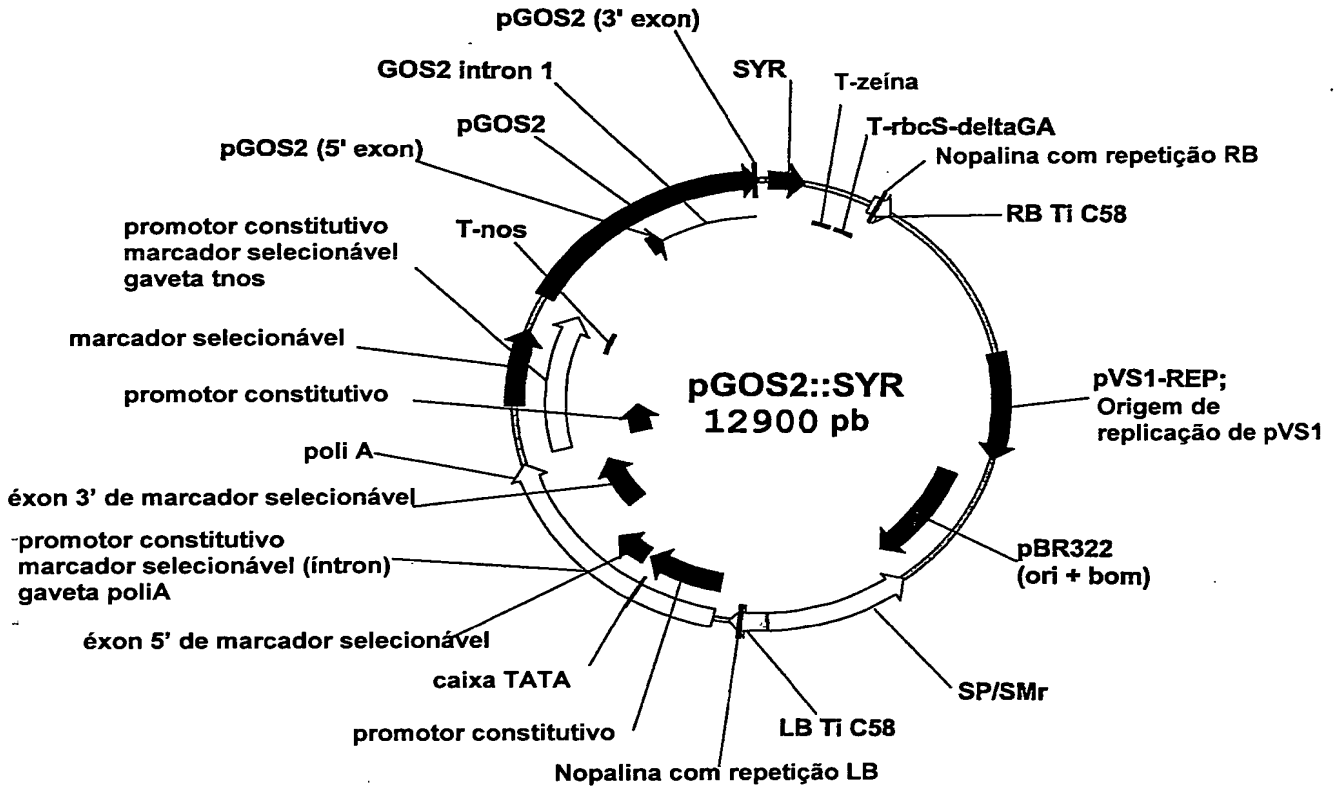


FIGURA 3

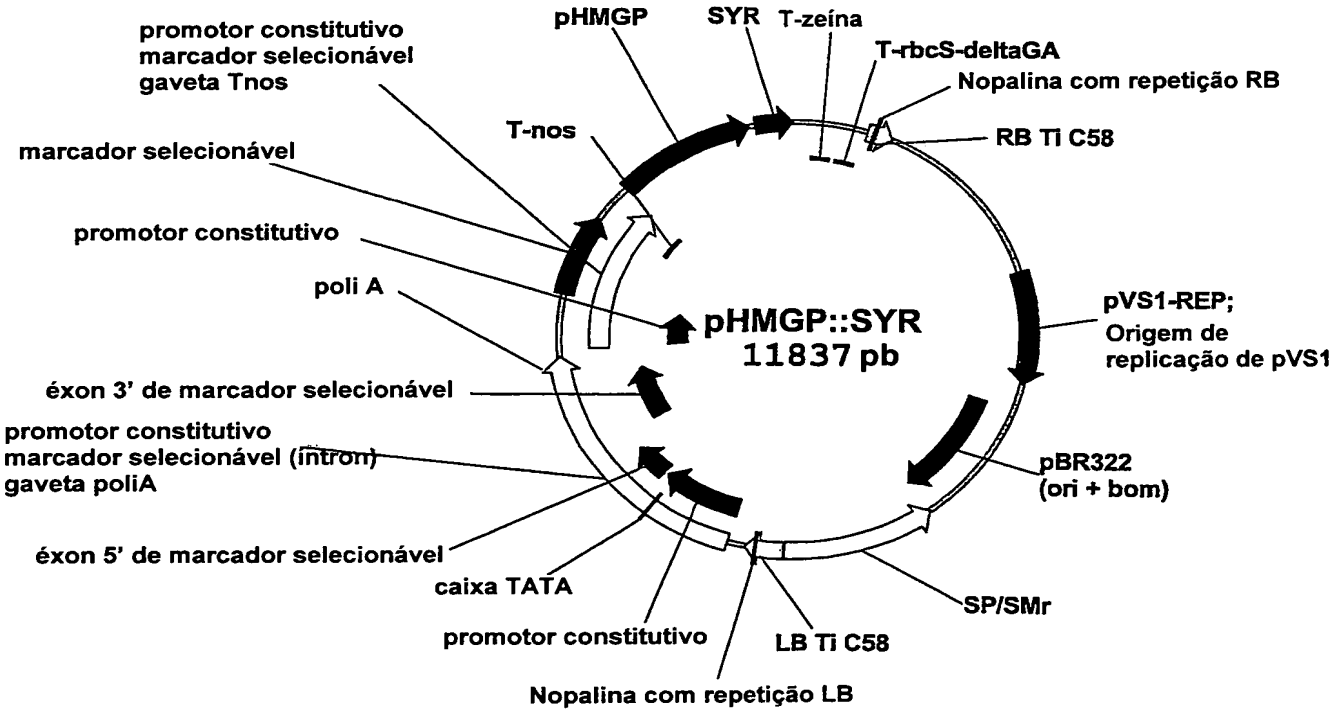


FIGURA 4

SEQ ID NO: 1, seqüência de codificação de OsSYR

ATGGAAGGTGTAGGTGCTAGGCAGAGGAGGAACCCTCTGATACCCAGACCAAACGGTTCAAA
GAGGCATCTGCAGCATCAGCATCAGCCAAATGCTGCCGAGAAGAAGACCGCCGCGACATCGA
ATTACTTCAGTATCGAGGCGTTCCTCGTGCTCGTCTTCCTCACCATGTCATTGCTCATACTT
CCATTGGTGCTTCCCCCATTGCCTCCGCCGCCATCGCTGCTGCTGCTGCTGCCAGTCTGCCT
GCTCATCCTGCTGGTTGTGCTGGCCTTCATGCCAACGGATGTGCGGAGCATGGCTTCCTCTT
ACTTGTAAATACATCTCCTAGGGGAATTTATTTTGTTTT**TGA**

SEQ ID NO: 2, seqüência deduzida de proteínas de OsSYR

MEGVGARQRRNPLIPRPNGSKRHLQHQHQPNAAEKKTAATSNYFSIEAFLVLVFLTMSLLIL
PLVLPPLPPPPSLLLLLPVCLLILLVVLAFMPTDVRSMASSYL

SEQ ID NO: 3, prm08170, códon de início em negrito, sítio AttB1 em itálico

*GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACA***ATG**GAAGGTGTAGGTGCTAGG

SEQ ID NO: 4, prm08171,reverso, complementar, sítio AttB2 em itálico

*GGGGACCACTTTGTACAAGAAAGCTGGGT*CAAAAACAAAAATAAATTCCCC

SEQ ID NO: 5, seqüência de promotor GOS2 de arroz

AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAA
ATATAAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATC
CACCTACTTTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCT
TAGTAATTAAGTGGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCAT
GAAGTTAAATTATTCGAGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTT
TCTAGCTGAACTCAATGGGTAAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTC
TGAACGTATTGGCAAAGATTTAAACATATAATTATATAATTTTATAGTTTGTGCATTCGTCA
TATCGCACATCATTAAGGACATGTCTTACTCCATCCCAATTTTTATTTAGTAATTAAAGACA
ATTGACTTATTTTTATTATTTATCTTTTTTCGATTAGATGCAAGGTACTTACGCACACACTT
TGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAATACACGTTCAACTAGCAACACATCTCT
AATATCACTCGCCTATTTAATACATTTAGGTAGCAATATCTGAATTCAAGCACTCCACCATC
ACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTTACAGAATAGCATGAAAAG
TATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAAAAAAGAATTTTGCTCGTGCGCGA
GCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCCACAGAACAACCCA
CAAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAGGCTTTG
CGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGCATCCTCCTCCTCCCATCTATAAAT
TCCTCCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCA
AGGACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTTCGATCCATATCTTCCGGTCGAGT
TCTTGGTCGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCCTTCGGTTGTTC
TTGGATTTATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATC
TGTGATGATTCCTGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTT
CGGTTTGATTAGTAGTATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTA
GGGTACGGAATCTTGCGATTTTGTGAGTACCTTTTGTTTGAGGTAAAATCAGAGCACCGGTG
ATTTTGCTTGGTGTAATAAAAGTACGGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTC
GATTTGACGAAGCTATCCTTTGTTTATTCCCTATTGAACAAAAATAATCCAACTTTGAAGAC
GGTCCCGTTGATGAGATTGAATGATTGATTCTTAAGCCTGTCCAAAATTTCGCAGCTGGCTT
GTTTAGATACAGTAGTCCCCATCACGAAATTCATGGAAACAGTTATAATCCTCAGGAACAGG

FIGURA 5

GGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTTTTTTCCCAAATATCTTAAAAA
GTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATGCTTTTATAGCGTTATCC
TAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGAAGAACTTATCCGA
TTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATTCATTTGGAT
TATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAACTGTC
CTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGTA
GAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAA
TCGGGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCT
TGCCACTTTCACCAGCAAAGTTC

SEQ ID NO: 6, motivo conservado 1a
YFS

SEQ ID NO: 7, motivo conservado 1b
YFT

SEQ ID NO: 8, motivo conservado 1c
YFG

SEQ ID NO: 9, motivo conservado 1d
YLG

SEQ ID NO: 10, motivo conservado 2
(V/A/I)LAFMP(T/S)

SEQ ID NO: 11, motivo conservado 4
(S/P)YL

SEQ ID NO: 12, homólogo 1 de SYR de arroz (XP_472637), codificado por SEQ ID NO: 27
MYLLSPRNGDEEDEQEEIQELISDDEPPNLKLASCATAASSSSSSGSDMEKGRGKACGGGST
APPPPPPSSSGKSGGGGGSNIREAAASGGGGGVWGKYFSVESLLLLVCVTASLVILPLVLPP
LPPPPSMLMLVPVAMLVLLLALAFMPTTTSSSSSAGGGGGGGRNGATTGHAPYL

SEQ ID NO: 13, homólogo 2 de SYR de arroz, seqüência deduzida de proteínas (AP008218)
MLLEHLMITMEEQMFREQQMQRGGRHHQHHTTREQEQQQKQQQRRRLMNNATNGGGGDGGSR
CYFSTEAILVLACVTVSLLVLPLILPPLPPPPTLLLLLPVCLLALLVVLAFMPTDMRTMASS
YL

SEQ ID NO: 14, homólogo de SYR de milho (AY110705), codificado por SEQ ID NO: 28
MASRSSAMEGGAAIQRRNAVKRHLQQRQQEADFLDKKVIASTYFSIGAFLVLACLTVSLLIL
PLXXXXXXXXXXXLLWLPVCLLVLLVVLAFMPTDVRSMASSYL

SEQ ID NO: 15, homólogo de SYR de trigo, seqüência deduzida de proteínas (CK211328)
MDSQFGALERGGSRQRRSPVLARPNTTKRHIQQQRANAADKKVVMPNYFSIEAFFVLACLTV
SLLILPLVLPPLPPPPSLLLFVPVCLLILLMVLAFMPTDMRSMATSYL

FIGURA 5 (Cont.)

SEQ ID NO: 16, homólogo de SYR de cevada (CB871444), codificado por SEQ ID NO: 36

MDSQFGAMDRGGSRQRSSPVLARPNTAKRQMQQQRANAADKKVVIPNYFGVEAFFVLACLTV
SLLILPLVLPPLPPPPSLLLLLPVCLLILLMVLAFMPTDVRSMATSYL

SEQ ID NO: 17, homólogo 1 de SYR de cana-de-açúcar codificado por SEQ ID NO: 37 (CA165713)

MEGGGQIQRRNNAVKRHLQQRQQEADFLDKKVIASTYFSIEAFLVLACLTVSLLILPLVLPX
LPAPASLLLWLPVWLLELLIVLAFMPTDVRSMASSYL

SEQ ID NO: 18, homólogo 2 de SYR de cana-de-açúcar codificado por SEQ ID NO: 38 (CA242805)

MEGGGQIQRRNNAVKRHLQQRQQEADFLDKKVIASTYFSIEAFLVLACLTVSLLILPLVLPP
LPPPPSLLLWLPVCLLILLIVLAFMPTDVRSMASSYL

SEQ ID NO: 19, homólogo de SYR de sorgo, codificado por SEQ ID NO: 39 (CX611532)

MASRSSALEGGGAAIQRRNNAVKRHLQQRQQEADFHDKKVIASTYFSIGAFLVLACLTFSLL
ILPLVLPPLPPPPSLLLWLPVCLLVLLVVLAFMPTDVRSVAASYL

SEQ ID NO: 20, homólogo 1 de SYR de Arabidopsis thaliana, codificado por SEQ ID NO: 40 (NM_115853)

MIREISNLQKDIINIQDSYSNNRVMDVGRNNRKNMSFRSSPEKSKQELRRSFSAQKRMMIPA
NYFSLESLFLLVGLTASLLILPLVLPPLPPPPFMLLLVPIGIMVLLVVLAFMPSSHSNANTD
VTCNFM

SEQ ID NO: 21, homólogo 2 de SYR de Arabidopsis thaliana, codificado por SEQ ID NO: 41 (NM_180078)

MIREFSSLQNDIINIQEHYSLNNNMDVRGDHNRKNTSFRGSAPAPIMGKQELFRTLSSQNSP
RRLISASYFSLESMVVLVGLTASLLILPLILPPLPPPPFMLLLIPIGIMVLLMVLAFMPSSN
SKHVSSSSTFM

SEQ ID NO: 22, homólogo de SYR de uva (CF404276), codificado por SEQ ID NO: 29

MSIEQPEADSRLSEGPLINLQDRYLSGIMEARGRRNSAPLQVERKNPTPPMAEGKKMEYNRT
PLSRENSRRLIPASYFSLESLLLLICLTASLLILPLILPPLPPPPFMLLLLPIGILAVLMIL
AFMPSNVRDLTYTYV

SEQ ID NO: 23, homólogo de SYR de citros (CF830612), codificado por SEQ ID NO: 30

MNSDNSESRQRLSKGIINLQDRYPTSIMDRGVRKIATPPVEKRKVEYHRSYSQGASRKLFSA
SYFTLESLLLLVCLTASLLILPLVLPPLPPPPFLLLLVPIXILAVLLVLAFMPSNVRDITST
YV

SEQ ID NO: 24, homólogo 1 de SYR de tomate (ai774560), codificado por SEQ ID NO: 32

MNMDMESSEAKLRSSKGFINLEEHQQYFNNIMEGNKMEHKRSFTQGHGKKMLSMNYFSLESI
ILLLGLTASLLLLPLMLPPLPPPPFMLLLVPIFILVVLMILAFMPSNVRNVTCSYL

FIGURA 5 (Cont.)

SEQ ID NO: 25, homólogo 2 de SYR de tomate, seqüência parcial (bg125370), codificada por SEQ ID NO: 31

MEGNKMEHKRSFTQGHGKKMLSMNYFSLESIILLLGLTASLLLLPLMLPPLPPPPFMLLLVP
IFILVVLMILAFMPSNVRNVTCSYL

SEQ ID NO: 26, proteína ARGOS de Arabidopsis thaliana (AY305869), codificada por SEQ ID NO: 42

MDVGRNNRKNMSFRSSPEKSKQELRRSFSAQKRMMIPANYFSLESLFLLVGLTASLLILPLV
LPPLPPPPFMLLLVPIGIMVLLVVLAFMPSSHSNANTDVTCNFM

SEQ ID NO: 27, homólogo de SYR de Oryza sativa mRNA, (XM_472637), proteína de codificação de SEQ ID NO: 12

ATGTACTTGTTGAGCCCAAGAAATGGCGACGAGGAGGACGAACAGGAGGAAATCCAGGAGCT
GATCAGCGACGACGAGCCGCCCAATCTCAAGTTGGCATCCTGCGCCACTGCAGCCAGCAGCA
GCAGCAGCAGCGGCAGCGACATGGAGAAGGGAAGAGGTAAAGCCTGCGGCGGCGGGAGTACG
GCGCCGCCGCCGCCGCCGCCGTCGTCGTCAGGTAAATCCGGCGGCGGCGGCGGCAGCAATAT
CAGGGAGGCGGCGGCTAGCGGCGGCGGCGGCGGCGTGTGGGGCAAGTACTTCTCGGTGGAGT
CGCTGCTCCTGCTGGTGTGCGTGACGGCGTCGCTGGTGATCCTCCCGCTCGTGCTGCCGCCG
CTGCCCCCGCCGCCGTCGATGCTGATGCTGGTGCCGGTGGCGATGCTGGTGCTGCTGCTGGC
GCTGGCGTTCATGCCGACGACGACGTCGTCGTCGTCCGCCGGCGGCGGCGGCGGCGGCG
GCCGCAATGGGGCGACGACGGGACATGCTCCCTACTTGTAG

SEQ ID NO: 28 homólogo de SYR de milho mRNA (AY110705), proteína de codificação de SEQ ID NO: 14

TTCACATTACACTCATGACTCGTCTTAGGACGAATCCTGCAGCTGCAAAACACAGAAAATCT
GGCCAAGACCTATTTATCTATTTACAGGTAAGAGGAGGCCATGCTGCGCACATCTGTCGGCA
TGAAGGCCAGTACAACCAGCAAGACGAGCAGGCAGACCGGCAGCCACAGCAGNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNCCAGCGGCAGTATCAGCAGCGAGACGGTGAGGCAGGCGAG
CACGAGGAATGCCCCGATGCTGAAGTAGGTGGACGCGATGACCTTCTTGTCGAGGAAATCCG
CCTCCTGCTGACGCTGCTGCAGATGCCGCTTCACGGCATTCCTCCTTTGTATTGCCGCCCCT
CCTTCCATCGCGCTAGATCGGCTTGCCATGTGTTCTTTGGTGTAGGCGGAGGTGGAGACCAG
TCGCAGTGAGTGGTTGCCAAAGTAAGCAAGAAGTGAAAGGCGGTGTAGAACCGCCTTGNGTT
TTCTGAACGTTTTGTAATCAGATCTGAGCTCGGGTGGTCGAA

SEQ ID NO: 29 seqüência de mRNA 3' de clone de cDNA de Vitis vinifera CSECS024D03 (complemento reverso CF404276), codifica SEQ ID NO: 22

TACTTTGAACCGGTTGGAACTTGGTGACCGGTAAAGGAGAAATTTAAATGAGTATTGAGCAG
CCTGAAGCAGATTCAAGACTGTCTGAAGGACCTTTGATCAATCTGCAAGATCGATATTTGAG
TGGCATCATGGAAGCGAGAGGAAGGAGGAACTCTGCTCCTCTGCAAGTTGAGAGGAAAAACC
CTACTCCTCCTATGGCCGAAGGAAAGAAGATGGAATATAATAGAACTCCCCTCTCACGAGAG
AACAGCAGGAGACTGATCCCAGCAAGCTATTTCAGCTTGGAGTCATTGCTTTTGCTCATCTG
TCTCACGGCTTCATTGCTGATCCTTCCCCTGATACTACCACCATTGCCACCCCCTCCTTTCA
TGCTGCTTCTGCTCCCCATTGGCATTCTAGCAGTGCTTATGATCTTGGCTTTCATGCCTTCT
AATGTCAGAGATTTGACTTATACATATGTGTAAATGGTGGTGTTCAAAAGTGCACCTCTTCT
CATCATACAATTTTTGTTATTTGTCTTGATATTCAGATGAAAATCAGTTATTTTATTTCTTG
ATTAAAAAAAAAAAAAAAAA

FIGURA 5 (Cont.)

SEQ ID NO: 30, clone de cDNA de Citrus reticulata CR_CEa05B13, seqüência de mRNA codificando SEQ ID NO: 23 (nt 313-693)

AGGGTTTCTTCAAAGATAGGTAGCCATTTGCACATTTGAATCTGCTTGTTGGATATTGTCAA
GGAGGCTGTTGGAATTAGGCCACATTTCAGAATCTGGTTTCATCCTGGATCGCTGGGCATTT
GAAGGCATTTTGTGATCATCGCTGTTTAAAATTTGGCCGCATATTAGAATCTGGGTTCTCAT
CGGTTTTCCGTACATTTACCTTGACCACATTTTGATATCTGGGTTGAGCTGCATTTTAGCCT
TCGTATTTAAAAGGACTTGATCTAATCTGGGGTCTTGGTGAGCCGGGGCAACTGATCATAGT
AAATGAATTCTGATAATTCTGAGTCGAGACAGAGACTATCAAAGGGCATTATAAACTTGCAA
GATCGATATCCGACCAGCATTATGGATCGTGGTGTAAGAAAAATTGCAACTCCTCCGGTCGA
GAAGAGGAAAGTTGAGTATCACCGAAGTTACTCGCAAGGGGCATCCAGAAAACTGTTTTCGG
CAAGCTATTTCACCCTGGAATCATTGCTTTTGCTCGTATGTCTGACGGCCTCATTGCTGATC
CTGCCATTGGTGCTTCCGCCCTTGCCGCCCCCGCCATTCCTGCTGCTTCTGGTTCCTATANG
TATTCTAGCCGTGCTTTTGGTCTTGGCATTCATGCCTTCTAATGTAAGAGATATAACTTCCA
CGTACGTGTAAATGGTGTTGCT

SEQ ID NO: 31 seqüência de clone de cDNA de Lycopersicon esculentum cTOF8M10 5' seqüência parcial de mRNA (BG125370) codificando SEQ ID NO: 25

GAAAAAAATGTATTTAATCATTATGTAAAAAACAAGTGAATCTACTTTGATATTTTCTTCTA
AATTAAACCACACAATTAAAGATATGAGCAAGTCACATTCCTAACATTAGAAGGCATAAAAG
CTAAGATCATAAGAACAACAAGAATGAAAATTGGGACTAACAACAACATAAAAGGTGGTGGT
GGCAATGGTGGAAGCATCAATGGCAAAAGTAACAAAGATGCTGTAAGACCAAGTAACAAAAT
AATTGACTCTAAGCTAAAATAATTCATTGACAACATTTTCTTGCCATGTCCTTGTGTAAATG
ATCTCTTATGCTCCATCTTATTGCCTTCCATAATGTTGTTGAAATATTGTTGATGTTCCTCC
AAATTAATAA

SEQ ID NO: 32 seqüência 5' de clone de cDNA de Lycopersicon esculentum cTOF8M10, 5' seqüência de mRNA (AI774560), codificando SEQ ID NO: 24

TTTGTTAAAGATTGGCACATTTTCAAGTTCAGTATTCATTCGATTTTTGATATCTACATAAA
AAAAAAGTGTCCTGGTACTACTCAATATTCCTCAGAACGACTTCATATTCAGGTCTCGAATT
CAAAACCTCACATCAAGATTCTTAGGAAATTTCAAGATTGGTTGAAAAACTCATATCCTTCT
CTAAGTTTCAAGATTGGTTCCAAATTAAAACTCGAGACTTCTGAGTAAGAGCGTACGACTAG
TAATGAACATGGACATGGAATCATCAGAGGCAAAATTGAGATCATCAAAAGGGTTTATTAAT
TTGGAGGAACATCAACAATATTTCAACAACATTATGGAAGGCAATAAGATGGAGCATAAGAG
ATCATTTACACAAGGACATGGCAAGAAAATGTTGTCAATGAATTATTTTAGCTTAGAGTCAA
TTATTTTGTTACTTGGTCTTACAGCATCTTTGTTACTTTTGCCATTGATGCTTCCACCATTG
CCACCACCACCTTTTATGTTGTTGTTAGTCCCAATTTTCATTCTTGTTGTTCTTATGATCTT
AGCTTTTATGCCTTCTAATGTTAGGAATGTGACTTGCTCATATCTTTAATTGTGTGGTTTAA
TTTA

SEQ ID NO: 33, PRO0170 – promotor de proteína de grupo de alta mobilidade

CATGCGGCTAATGTAGATGCTCACTGCGCTAGTAGTAAGGTACTCCAGTACATTATGGAATA
TACAAAGCTGTAATACTCGTATCAGCAAGAGAGAGGCACACAAGTTGTAGCAGTAGCACAGG
ATTAGAAAAACGGGACGACAAATAGTAATGGAAAAACAAAAAAAAACAAGGAAACACATGGC
AATATAAATGGAGAAATCACAAGAGGAACAGAATCCGGGCAATACGCTGCGAAAGTACTCGT
ACGTAAAAAAAAGAGGCGCATTCATGTGTGGACAGCGTGCAGCAGAAGCAGGGATTTGAAAC
CACTCAAATCCACCACTGCAAACCTTCAAACGAGGCCATGGTTTGAAGCATAGAAAGCACAG-

FIGURA 5 (Cont.)

GTAAGAAGCACAACGCCCTCGCTCTCCACCCTCCCACCCAATCGCGACGCACCTCGCGGATC
GGTGACGTGGCCTCGCCCCCCAAAAATATCCCGCGGCGTGAAGCTGACACCCCGGGCCCACC
CACCTGTCACGTTGGCACATGTTGGTTATGGTTCCCGGCCGCACCAAAATATCAACGCGGCG
CGGCCCAAAATTTCCAAAATCCCGCCCAAGCCCCTGGCGCGTGCCGCTCTTCCACCCAGGTC
CCTCTCGTAATCCATAATGGCGTGTGTACCCTCGGCTGGTTGTACGTGGGCGGGTTACCCTG
GGGGTGTGGGTGGATGACGGGTGGGCCCGGAGGAGGTCCGGCCCCGCGCGTCATCGCGGGGC
GGGGTGTAGCGGGTGCGAAAAGGAGGCGATCGGTACGAAAATTCAAATTAGGAGGTGGGGGG
CGGGGCCCTTGGAGAATAAGCGGAATCGCAGATATGCCCCTGACTTGGCTTGGCTCCTCTTC
TTCTTATCCCTTGTCCTCGCAACCCCGCTTCCTTCTCTCCTCTCCTCTTCTCTTCTCTTCTC
TGGTGGTGTGGGTGTGTCCCTGTCTCCCCTCTCCTTCCTCCTCTCCTTTCCCCTCCTCTCTT
CCCCCCTCTCACAAGAGAGAGAGCGCCAGACTCTCCCCAGGTGAGGTGAGACCAGTCTTTTT
GCTCGATTCGACGCGCCTTTCACGCCGCCTCGCGCGGATCTGACCGCTTCCCTCGCCCTTCT
CGCAGGATTCAGCC

SEQ ID NO: 34
MVRCFSLGSVLILIALAASMVVLPLMLPPLPPPPLALLFFPVGIMAALVVLAFSPSENVKNV
VVYSSSSSGIANSKR

SEQ ID NO: 35
MIMVASKEKTNSGGCMFRYSVLILSLLALSILVLPLVMPPLPPPPLLLLLVPVFIMLLLFFI
AFSPSKKVPNKASFVS

SEQ ID NO: 36, CB871444 HC03O01y CH clone de cDNA de Hordeum vulgare HC03O01 3-INICIADOR, seqüência de mRNA de complemento reverso codificando SEQ ID NO: 16

CTTCCGAGAAAGATGCTTCATATTGCACTCATCTTATGCCAACACACAATCAGAAACAAAAA
CCACGCAGCAAGCCTATTTACAAGTAAGAGGTGGCCATGCTCCGCACATCAGTCGGCATGAA
GGCCAGCACCATCAACAGGATGAGCAAGCAGACCGGCAAGAGCAGCAGTAGCGATGGCGGTG
GGGGCAACGGAGGCAGCACCAATGGCAGTATGAGCAATGAAACGGTGAGGCAGGCGAGCACG
AAGAACGCTTCGACGCCGAAGTAGTTCGGTATGACAACCTTCTTGTCAGCAGCATTTGCCCT
CTGCTGCTGCATTTGCCTCTTTGCAGTATTTGGTCTGGCTAACACAGGGCTGCTCCTCTGCC
TACTACCGCCTCTGTCCATTGCACCGAACTGGCTATCCATCTATTCTTTGGTGAGTGCAGAT
CAGCGGCTATCCAGGAACTGAGATGAATAAGAACTTGTGGAATCTGAAGGTTTTTAACAGAT
CTGAAGTCTTTGTGAGTCCGTGACTGAACTGCAGATCATCTGCTGTGGAAGAACTGCACCGC
AAGTGGCAATACCTTCGATCCTGAAACTTGCATTTTGGTGATGCCCGTACCACGC

SEQ ID NO: 37, CA165713 SCUTRZ3071B09.g RZ3, clone de cDNA de Saccharum officinarum seqüência de mRNA SCUTRZB09 5', codificando SEQ ID NO: 17

GGAGAACGTGCTTGTTCAGGTGGTTCAACAGTGCGGACCAGAGAACAATGCGAGGTTCAGGA
TTAAAGGTATTCTGGCTTCAGAGGCAGTTCTTCCAAAGCAAGTGACAGGCGATTCCATCACC
GGAGCTAAGATCTTATTACAACATTCAGAAACACAGGAGTCCTCCCACCGCCTTTCACTTCT
TGCTTACTTCTGCAACCACTCGCCGTGACTGATCTCCACGCACACCAAAGAACACAACACGT
GGCAAGCCGATCTAGCGCGATGGAAGGAGGGGGGCAAATACAGAGGAGGAATAATGCCGTGA
AGCGGCACCTGCAGCAGCGGCAGCAGGAGGCGGATTTCCTCGACAAGAAGGTCATCGCGTCC
ACCTATTTCAGCATCGAGGCGTTCCTCGTGCTCGCCTGCCTCACCGTCTCGCTGCTGATACT
GCCGCTTGTGCTGCCGNCGCTGCCGGCGCCGGCGTCGCTGCTGCTGTGGCTGCCGGTCTGGC
TGCTCGAACTGCTGATTGTGCTTGCCTTCATGCCGACAGATGTGCGCAGCATGGCCTCCTCT
TACTTGTAAAAAAATAGATAAATAGGCCACCTTTGGCAATTTTCTGGGGTTTGGAGGTG

FIGURA 5 (Cont.)

SEQ ID NO: 38, complemento reverso de Saccharum officinarum A242805 SCSFFL3090G07 . b, clone de cDNA de Saccharum officinarum FL3SCSFFL3090G07 3' seqüência de mRNA, codificando SEQ. ID NO: 18 SCSFFL3090G07 3'

GATTTTTTCCAAGCCAGTGACCGGCGATTCATCAACCGGAGCTGAGATCTTATACAACATTC
AGAAACACAGGAGTCCTCGCACCGCTTTCCACTCTTGGCTAATTTTGCAACCACTCGCCGTG
ACTGATCTCCACGCACACCAAAGAACACAACACGTGGCAACCCGATCTAGCGCGATGGAAGG
AGGGGGGCAAATACAGAGGAGGAATAATGCCGTGAAGCGGCACCTGCAGCAGCGGCAGCAGG
AGGCGGATTTCCTCGACAAGAAGGTCATCGCGTCCACCTATTTCAGCATCGAGGCGTTCCTC
GTGCTCGCCTGCCTCACCGTCTCGCTGCTGATACTGCCGCTGGTGCTGCCGCCGCTGCCGCC
GCCGCCGTCGCTGCTGCTGTGGCTGCCGGTCTGCCTGCTCATCCTGCTGATTGTGCTGGCCT
TCATGCCGACAGATGTGCGCAGCATGGCCTCCTCTTACTTGTAATAAATAGATAAATAGGCC
ACCTTGGTCAATATTCTGTGATTTGGAGGTGATTCGTCCTGAGATGAGTCTCGATTGATGTC
AGCTACTCCCAAGGGGAAATGCATGTAACACTTGGTGGCCGACGGTGGCAAGATAATCATGC
TACCATGTCAGTTAAACC

SEQ ID NO: 39, CX611532 ANR1_25 E08>b1_A002 Raízes anaeróbicas clone de cDNA de Sorghum bicolor, ANR1_25_EOB _A002 3' seqüência de mRNA codificando SEQ ID NO: 19

GCTGTTGGTGTTGTTCTTGAGATCTCTTTCTTGATCTTGTGTGGGATTAAAGAGGGATTCTT
GCCTTCCTACGGGAGAAAGAGAAAAGGGGAAGAACGTGCTTGTTCCGGTGGTTCAACAGTGC
GGAGACCCGGAGAACAATGCGAGGTTCAGGATTAAAGGTGTTCTGGCTTCAGGTGCAGTTCT
TCCAAAGCAGGTGACAGGCGATTCGATCACCGGAGCTCAGATCTGACGAAAACAAAACACAG
TCCTCCTCCCACCGCCTTTCAGTTCTTGCTTACTTCTGCAACCACTCACTGCGACCGTACAC
CAAAGAACACTGCACATGGCAAGCCGATCTAGCGCGCTGGAAGGAGGGGGGGCAGCAATACA
GCGGAGGAATAATGCCGTGAAGCGGCACCTGCAGCAGCGGCAGCAGGAGGCGGATTTCCACG
ACAAGAAGGTCATCGCGTCCACCTACTTCAGCATCGGCGCGTTCCTGGTGCTCGCCTGCCTC
ACCTTCTCGCTGCTCATCCTGCCGCTGGTGCTGCCGCCGCTGCCGCCGCCGCCGTCGCTGCT
GCTGTGGCTGCCGGTCTGCCTGCTCGTCCTGCTGGTTGTGCTGGCCTTCATGCCGACAGATG
TGCGCAGCGTGGCGGCCTCTTACTTGTAAATAGCCAGATAAATAGGCCACCACCTTTGGCCA
GTTTTCTGTGTTTTCGGGGGTGATTCGTCCTAAGATGAGTCATGATCGAGTGTAATGTGAAG
CATCTTTTCCAGGGGTAGTAGATTTCAATGAAGT

SEQ ID NO: 40, NM_115853.3 | mRNA de proteína desconhecida de Arabidopsis thaliana (AT3G59900), cds completo, codificando SEQ ID NO: 20

TTGTCTTCCTCATTTCCCTACTAGTACTTGTTTCACACAGTTTCTTGATCCAACCAAAACCA
ATACACAAAGCTTCTCAAACTCCTTCACCTCAAAGCTTCTTCCTTTACATCTGAATCGTTGA
GTTAACTCGGATTTGTTCTGCATCCTCTGTTTCTGAATCGTGGGCCATCCTTATTTTGTCTC
GAATTCTTCACCAATTGCTTCGATCAAGCTGCATTGGTTAACCAGTTGCCCTAAAGATCAGA
TCTTTGAGCAAAATTTTGTCACTGATCTTCTAAATCCAAACCAGACACAGCAAAACAACCTC
TGTAGATGATTCGAGAAATCTCAAACTTACAAAAAGATATTATAAACATTCAAGACAGTTAT
TCGAACAACCGAGTCATGGACGTCGGAAGAAACAACCGGAAAAACATGAGCTTTCGAAGTTC
GCCGGAGAAAAGCAAGCAAGAGTTACGGCGGAGTTTCTCGGCGCAGAAAAGGATGATGATCC
CGGCGAATTATTTCAGTTTAGAGTCTCTGTTCCTATTGGTTGGTCTAACGGCATCTCTGTTA
ATACTTCCGTTAGTTTTGCCGCCGTTACCTCCGCCTCCGTTTATGCTGCTATTGGTTCCCAT
TGGGATTATGGTTTTACTCGTCGTTCTTGCCTTCATGCCTTCTTCTCATTCTAATGCTAATA
CAGATGTAACTTGCAATTTCATGTAAATCTGAAATTTATTATATGATGAT

FIGURA 5 (Cont.)

SEQ ID NO: 41, NM_180078.2 | mRNA de proteína desconhecida de Arabidopsis thaliana (AT2G44080), cds completo, codificando SEQ ID NO: 21

CCCACTTTATTTCTTCTTCTCTGGTTTTCTTACCAAAAGAAACTTTCTTCGTCTTCCTCTGT
ATTTAAGCTTTAACACCCTGTTTTTGGTTTCCAACGTTCAATCTTCATCTTCTTCTCGCTGA
AGGTGTGTTTGGCTCTAACGGTTTAAAGCTTTCTTGAAACACCAATTGAATCTTTTCTCTCT
ACCGGCAAAAAAAAAAGATTAGTCCTTTTAGGTCTGGAAACGCCAAGATCACTCGTTCTAAA
CCTTAGATTTTGTCTGCATTTCGGGATAATCATTTCATCGTCAGGGTTCTTCAACCAAACTA
CATTTACAGAAGAAGAAGAAGAAGAAAAGTTCGTTACTTTTTATGCGTTTGGATAAACAAAC
TCAAGTTTCTTCTTCATACATCGATCTGATTTTCCAGATCAAACTTCGAAAAGAGAAAAAGC
CTTCTTTAAATGATTCGTGAGTTCTCCAGTCTACAAAACGACATCATAAACATTCAAGAACA
TTATTCTCTCAACAACAACATGGACGTGAGAGGAGATCATAACCGGAAAAACACGAGTTTTC
GTGGTTCAGCTCCAGCTCCGATTATGGGGAAGCAAGAATTGTTTCGGACATTGTCGTCGCAG
AACAGTCCAAGGAGGCTAATATCAGCGAGTTACTTCAGTTTAGAATCAATGGTTGTGCTTGT
TGGTCTCACAGCATCTCTCTTGATCTTACCGTTGATTCTTCCACCATTGCCTCCTCCTCCTT
TTATGCTGCTTTTGATTCCTATTGGGATTATGGTTTTGCTTATGGTTCTTGCTTTCATGCCT
TCTTCTAATTCCAAACATGTTTCTTCTTCTTCCACTTTTATGTAATAAACGTTTCTTTAATT
GAAGAAAGAAATCCTTAAACAAA

SEQ ID NO: 42, AY305869;1 | mRNA de proteína auxina-induzível de Arabidopsis thaliana
(ARGOS), cds completo, codificando SEQ ID NO: 26

TTGTCTTCCTCATTTCCCTACTAGTACTTGTTTCACACAGTTTCTTGATCCAACCAAAACCA
ATACACAAAGCTTCTCAAACTCCTTCACCTCAAAGCTTCTTCCTTTACATCTGAATCGTTGA
GTTAACTCGGATTTGTTCTGCATCCTCTGTTTCTGAATCGTGGGCCATCCTTATTTTGTCTC
GAATTCTTCACCAATTGCTTCGATCAAGCTGCATTGGTTAACCAGTTGCCCTAAAGATCAGA
TCTTTGAGCAAAATTTTGTCACTGATCTTCTAAATCCAAACCAGACACAGCAAAACAACCTC
TGTAGATGATTCGAGAAATCTCAAACTTACAAAAAGATATTATAAACATTCAAGACAGTTAT
TCGAACAACCGAGTCATGGACGTCGGAAGAAACAACCGGAAAAACATGAGCTTTCGAAGTTC
GCCGGAGAAAAGCAAGCAAGAGTTACGGCGGAGTTTCTCGGCGCAGAAAAGGATGATGATCC
CGGCGAATTATTTCAGTTTAGAGTCTCTGTTCCTATTGGTTGGTCTAACGGCATCTCTGTTA
ATACTTCCGTTAGTTTTGCCGCCGTTACCTCCGCCTCCGTTTATGCTGCTATTGGTTCCCAT
TGGGATTATGGTTTTACTCGTCGTTCTTGCCTTCATGCCTTCTTCTCATTCTAATGCTAATA
CAGATGTAACTTGCAATTTCATGTAAATCTGAAATTTATTATATGATGAT

SEQ ID NO: 43, ácido nucléico codificando proteína SYR de Zea mays
ZM57451407

ATGCGAGGTTCGGGATTTAAGATGTTCGGCTTTAGGGGCAGTTCTTCTGAAGCAGGGGACGG
GCGATTCGACCACCGGAGCTCAGATCTGATTACAAAACGTTCAGAAAACACAAGGCGTTCTC
ACACCGCCTTTCACTTCTTGCTTACTTTGGCAACCACTCACTGCGACTGGTCTCCACCTCCA
CCTACACCAAGAACACATGGCAAGCCGATCTACGCGATGGAAGGAGGGGCGGCAATACAAAG
GAGGAATGCCGTGAAGCGGCATCTGCAGCAGCGTCAGCAGGAGGCGGATTTCCTCGACAAGA
AGGTCATCGCGTCCACCTACTTCAGCATCGGGGCGTTCCTCGTGCTCGCCTGCCTCACCGTC
TCGCTGCTGATACTGCCGCTGGTGCTGCCTCCCCTGCCGCCGCCGCCGTCGCTGCTGTTGTG
GCTGCCGGTCTGCCTGCTCGTCTTGCTGGTTGTACTGGCCTTCATGCCGACAGATGTGCGCA
GCATGGCCTCCTCTTACCTGTAA

SEQ ID NO: 44, proteína SYR de Zea mays, ZM57451407

MEGGAAIQRRNAVKRHLQQRQQEADFLDKKVIASTYFSIGAFLVLACLTVSLLILPLVLPPL
PPPPSLLLWLPVCLLVLLVVLAFMPTDVRSMASSYL

FIGURA 5 (Cont.)

MKSRARQCLLVCLLCLSLTNLSYGENKFRERKATDDELGYPDIDEDALLN
TQCPKKLELRWQTEVTSSVYATPLIADINSDGKLDIVVPSFVHYLEVLEG
ADGDKMPGWPAFHQSNVHSSPLLFDIDKDGVREIALATYNAEVLFFRVSG
FLMSDKLEVPRRKVHKNWHVGLNPDPVDRSHPDVHDDVLEEEAMAMKSST
TQTNATTTTPNVTVSMTKEVHGANSYVSTQEDQKRPENNQTEAIVKPTPE
LHNSSMDAGANNLAANATTAGSRENLNRNVTTNEVDQSKISGDKNETVIK
LNTSTGNSSETLGTSGNSSTAETVTKSGRRLLEEDGSKESVDSHSDSKDN
SEGVRMATVENDGGLEADADSSFELLRENDELADEYSYDYDDYVDEKMWG
DEEWVEGQHENSEDYVNIDAHILCTPVIADIDKDGVQEMIVAVSYFFDPE
YYDNPEHLKELGGIDIKNYIASSIVVFNLDTKQVKWIKELDLSTDKANFR
AYIYSSPTVVDLDGDGYLDILVGTSFGLFYAMDHRGNIREKFPLEMAEIQ
GAVVAADINDDGKIELVTTDSHGNIAAWTTQGVEIWEAHLKSLVPQGPSI
GDVDGDGHTEVVVPTSSGNIYVLSGKDGSIVRPYPYRTHGRVMNQLLLVD
LNKRGEKKKGLTIVTTSFDGYLYLIDGPTSCTDVVDIGETSYSMVLADNV
DGGDDLDLIVSTMNGNVFCFSTPSPHHPLKAWRSSDQGRNNKANRYDREG
VFVTHSTRGFRDEEGKNFWAEIEIVDKYRYPSGSQAPYNVTTTLLVPGNY
QGERRITQSQIYDRPGKYRIKLPTVGVRTTGTVMVEMADKNGLHFSDEFS
LTFHMYYYKLLKWLLVLPMLGMFGLLVILRPQEAVPLPSFSRNTDL

FIGURA 6

CLUSTAL W (1.83) **alinhamento múltiplo de seqüências**

```
SEQID55         MRKRDLAILMLSGFAIFFT----------------------LQHEGDFAFKEAWFHLY
SEQID59         MRKRDLGILLLAAFAVFFS----------------------LQHDGDLSFREAWYHLS
SEQID46         MKSRARQCLLVCLLCLSLTNLSYGE---NKFRERKATDDELGYPDIDEDALLNTQCPKKL
SEQID57         MRPLLAFAAVCALLVAAAAPAAAEEEKANKFRQREATDDMLGYPHLDEDALLKTKCPKHV
                *:       : . :    :                              . : *  :.    :

SEQID55         D-EYPVKYEADRLPPPIVADLNGDGKKEVLVATNDAKIQVLEPHSR--RVDEGFSEARVL
SEQID59         DADYPIKHDADRLPSPLVADLNGDGKPEVLIPTHDAKIQVLQPHPRPSPDDASFHDARLM
SEQID46         ELRWQTEVTSSVYATPLIADINSDGKLDIVVPSFVHYLEVLEGADGDKMPGWPAFHQSNV
SEQID57         ELRWQTEVSSSIYATPLIADINSDGKLEVVVPSFVHYLEVLEGSDGDKLPGWPAFHQSNV
                :  :  :  :.  ..*::**:*.*** :::::.:    ::**:        .    .   :

SEQID55         AEITLLPDKIRVASGRRAVAMATGVIDRYYKNG-TPQKQVVV---VVTSGWSVLCFDHNL
SEQID59         ADVSLLPSNVRLSSGRRPVAMAVGTVDRHYAHAPSPSKQLLV---VVTSGWSVMCFDHNL
SEQID46         HSSPLLFDIDKDGVREIALATYNAEVLFFRVSGFLMSDKLEVPRRKVHKNWHVGLNPDPV
SEQID57         HSSPLLYDIDKDGTREIVLATYNGVVNFFRVSGYMMMDKLEVPRRKVHKDWYVGLNTDPV
                 . .** .  : .  .  :*   . : .   .    .:: *    * ..* *    . :

SEQID55         KKLWETNLQEDFPHNAHHREIAISISN---------------------------------
SEQID59         KKLWEANLQDDFPHAAHHREVAISITN---------------------------------
SEQID46         DRSHPDVHDDVLEEEAMAMKSSTTQTNATTTTPNVTVSMTKEVHGANSYVSTQEDQKRPE
SEQID57         DRSHPDVHDSSIAKKAASEESHPNIQDKPVVNE----------------SSKESQSRST
                .:       :. : . *   :   .  :
```

FIGURA 7

```
SEQID55         ----------------YTLKHGDTGLVIVG------------------------------
SEQID59         ----------------YTLKHGDAGLVIVG------------------------------
SEQID46         NNQTEAIVKPTPELHNSSMDAGANNLAANATTAGSRENLNRNVTTNEVDQSKISGDKNET
SEQID57         NDSTTRGVD--------SMKHASKEEPVESKPNSTRGQENMDVLNNLN------------
                              ::. .       .

SEQID55         -------------------------------GRMEMQPYNHMDPFEELGMTAQNADQHRRS
SEQID59         -------------------------------GRMEMQ-HHSAELFDEFMVSEHNREEHRRS
SEQID46         VIKLNTSTGNSSETLGTSGNSSTAETVTKSGRRLLEEDGSKESVDSHSDSKDNSEGVRMA
SEQID57         ----STDAGNNS------SLSTTTENASHVQRRLLQTDEKSNQAGSSETDASDTGTAKAA
                                          *  ::     :   .     :    : :

SEQID55         ATEN----QASEDSGAINLRHFSVYAFAGKTGLLRWSKKTDDVEAHTSDASQLIPQHNYK
SEQID59         ASEK----QAS-ETGNTDLRHFALYAFAGRTGELRWSRKNENIPSQPSDASVLIPQHNYK
SEQID46         TVENDGGLEADADSSFELLRENDELADEYSYDYDDYVDEKMWGDEEWVEGQHENSEDYVN
SEQID57         TVENSEPLEADADASFNLFRDVEDLPDEYNYDYDDYVDETMWGDEDWKEQQHEKAEDYVS
                : *:    :*. ::.   :*.    .     .   :  :.     .  : .   .:.  .

SEQID55         LDVHALNSRHPGEFECREFRESILSVMPHRWDRREDTLLKLAHFR---------------
SEQID59         LDAHALNSRHPGQFECREFRESVLGVMPHHWDRREDTFLQLAHFR---------------
SEQID46         IDAHILCTPVIADIDKDGVQEMIVAVSYFFDPEYYDNPEHLKELGGIDIKNYIASSIVVF
SEQID57         IDAHILSTPVIADIDRDGIQEMVISVSYFFDHEYYDKPEHLKELGGIDIGKYIASSIVVF
                :*.* * :   .:::   .:* ::.*  .   .  *.  :* .:

SEQID55         ----RHKRKTLKKQAGSKSTAYPFHKPEEHTPAGKDLSRKIPKLIGKAARYAGSAKPKKG
SEQID59         ----RHKRKALKKTPG-KAVVNNVHKPSEHNPPGKDVSNRLANVIGKAADMANSNKIKKS
SEQID46         NLDTKQVKWIKELDLSTDKANFRAYIYSSPTVVDLDGDGYLDILVGTSFGLFYAMDHRGN
SEQID57         NLDTRQVKWTAELDLSTDSGNFTAHAYSSPTVVDLDGDGNLDILVGTSFGLFYVIDHRGK
                    :: :   :   . .      :  .. . . * .  :  ::*.:       . :

SEQID55         --MQYIPTITNYTKLWWVPNV-------VVAHQKEGIEAIHLPTGRTLCKLSLL----EG
SEQID59         QRTLYVPTITNYTQVWWVPNV-------VVAHEKEGIEAVHLASGRTICKLHLT----EG
SEQID46         IREKFPLEMAEIQGAVVAADINDDGKIELVTTDSHGNIAAWTTQGVEIWEAHLKSLVPQG
SEQID57         VRNKFPLEMAEIHAPVIAADINDDGKIEMVTADVHGNVAAWTAEGEEIWEVHLKSLIPQR
                    :   :::      ..::       :*: : .*  *   . *  : :  *     :

SEQID55         GLHADINGDG-VLDHVQTVGGNVGERTVVSGS----------MEVLKPCWAVATSGVPIR
SEQID59         GLHADINGDG-VLDHVQVVGANGIEQTVVSGS----------MEVLKPCWAVATSGVPVR
SEQID46         PSIGDVDGDGHTEVVVPTSSGNIYVLSGKDGSIVRPYPYRTHGRVMNQLLLVDLNKRGEK
SEQID57         PTVGDVNGDGRTEVVVPTVSGNIYVLSGKDGSKIQPFPYRTHGRIMSPVLLLDMSKHDEK
                   .*::*** .   * . ..*    :  .**           .::.    :  .    :

SEQID55         EQLFNVSICHHSPFNFLHYGGDYSRHFAQARDTSTLEIATPILIPRDDGHKHRKGSHGDV
SEQID59         EQLFNVSICHYNNFNLFHHG-DFSRSFGRTFDTTGLEVATPILLQRDDGHKHRRGSHGDI
SEQID46         KKGLTIVTTSFDGYLYLIDGPTSCTDVVDIGETSYSMVLADNVDGGDDLDLIVSTMNGNV
SEQID57         SKGLTLATTSFDGYLYLIEGSSGCADVVDIGETSYSMVLADNVDGGDDLDLIVTTMNGNV
                .: :.:    .. :  :  *   .  .    :*:   : :  :   ** .      :*::

SEQID55         IFLTN----------RGEVTSYTPDVHGHDAVWQWQLQTEATWSNLPSPSGLTESGTVVP
SEQID59         IFLTS----------RGEVTSYSPGLLGHDAIWRWQLSTGATWSNLPSPSGMMEN-IVVP
SEQID46         FCFSTPSSHHPLKAWRSSDQGRNNKANRYDREGVFVTHSTRGFRDEEGKNFWAEIEIVDK
SEQID57         FCFSTPSPHHPLKEWRSSNQGRNNAAYRYNREGIYVKHGSRTFRDEEGKHFWVEFEIVDK
                : ::.          *..  . .    ::    :       : :  .    *   *
```

FIGURA 7 (Cont.)

```
SEQID55         TLKPFSLRIHD--NQPMILAGGDQAAVIISPGGSILASIELPSQPTHALITDDFSNDGLT
SEQID59         TLKAFSLRAYD--PKQVIIAGGDLEAVVISPSGGLLASIELPAPPTHALVLEDFNGDGLT
SEQID46         YRYPSGSQAPYNVTTTLLVPGNYQGERRITQS----QIYDRPGKYRIKLPTVGVRTTGTV
SEQID57         YRVPYGNQAPYNVTVTLLVPGNYQGERRIVVN----AAYNEPGKQRMKLPTVPVRTTGTV
                   . . :        :::.*.      *  .       : *.     *    .   * .

SEQID55         DVIVMTSNGVYGFVQTRQPGALFFSSLVGCLLVVMAV-IFVTQHLNSIQ-GKPRPSSSF-
SEQID59         DIILVTSGGVYGFVQTRHPGALFFSTLVGCLIVVIGV-IFVSLHLNSSNSGKPRASTDYR
SEQID46         MVEMADKNGLHFSDEFSLTFHMYYYKLLKWLLVLPMLGMFGLLVILRPQEAVPLPSFSRN
SEQID57         LVEMVDKNGFYFSDEFSLTFHMHYYKLLKWLVLLPMLGMFSVLVILRPQEGAPLPSFSRN
                 : :  ..*.:   :    .  :.: .*:  *::::  : :*    :   : . * .* .

SEQID55         ---
SEQID59         ---
SEQID46         TDL
SEQID57         ID-
```

FIGURA 7 (Cont.)

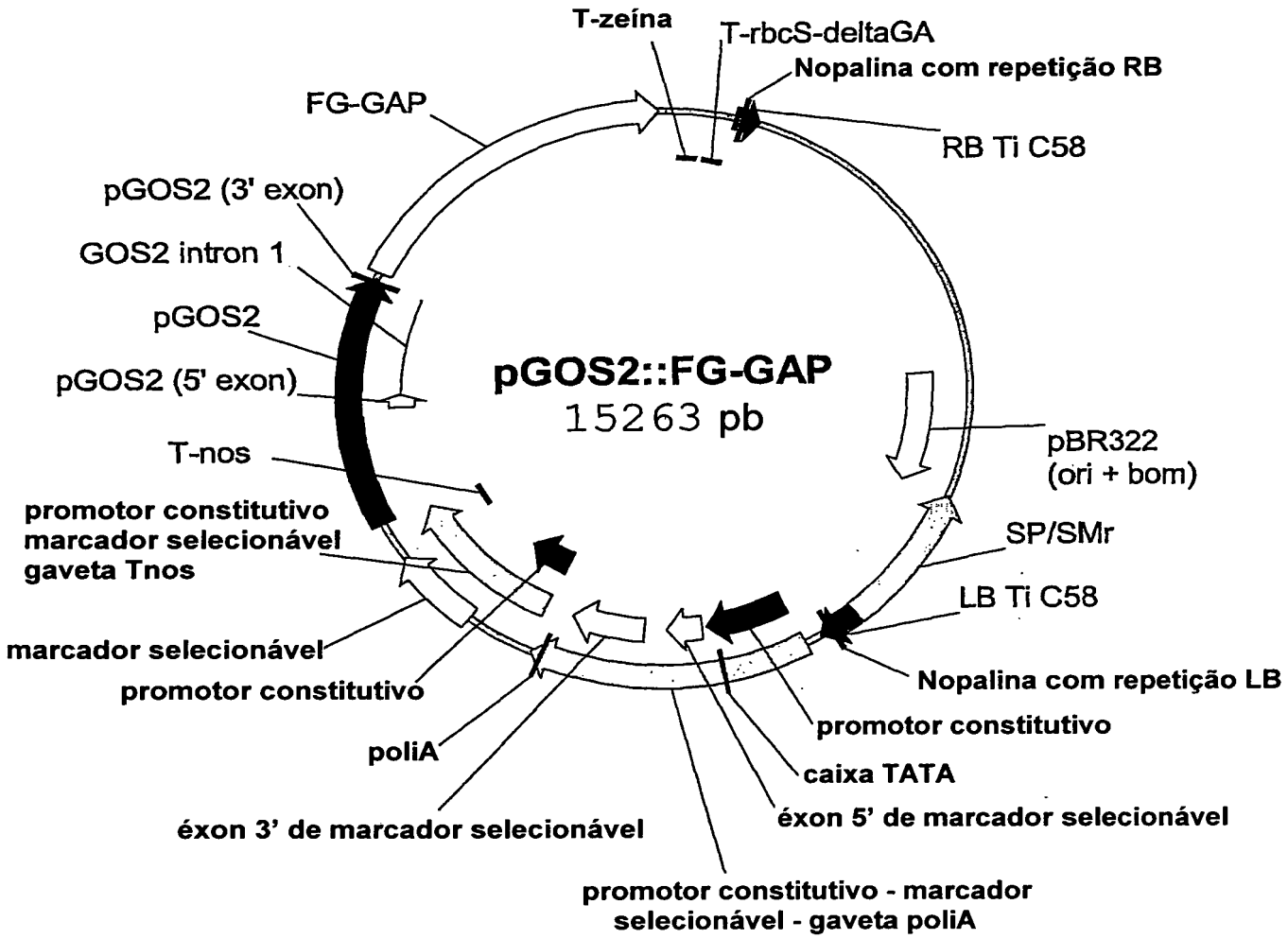


FIGURA 8

SEQ ID NO: 45, seqüência de codificação de AtFG-GAP

ATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTTAAACA**ATG**AAATC
TCGAGCGAGGCAGTGTCTGCTGGTTTGTTTGCTCTGTCTTTCCTTAACGAATCTCTCCTATG
GAGAAAATAAGTTCAGAGAGCGTAAAGCCACCGATGACGAGCTGGGCTACCCCGATATTGAT
GAAGATGCTTTATTGAATACTCAGTGCCCGAAAAAATTGGAGCTGCGATGGCAAACTGAAGT
CACTTCTAGCGTTTATGCTACACCCTTGATTGCTGATATTAACAGTGATGGAAAGCTTGACA
TTGTTGTTCCATCTTTTGTTCATTACCTCGAAGTTCTTGAAGGAGCTGATGGAGACAAGATG
CCAGGTTGGCCTGCTTTTCACCAGTCAAATGTGCACTCGAGTCCTCTTCTATTTGATATCGA
CAAAGATGGTGTTAGAGAAATTGCTCTGGCTACCTACAATGCCGAAGTGCTCTTTTTCAGGG
TATCAGGCTTTTTGATGTCAGATAAGCTAGAAGTGCCACGTAGAAAAGTGCACAAGAACTGG
CATGTGGGACTTAATCCTGATCCTGTTGACCGTTCACATCCTGATGTTCATGATGATGTGCT
TGAGGAGGAAGCTATGGCAATGAAGTCATCGACCACTCAAACGAATGCAACTACCACAACAC
CAAATGTTACAGTCTCGATGACCAAAGAAGTTCATGGCGCTAATTCATATGTGTCAACTCAA
GAGGATCAAAAGAGACCAGAGAATAATCAAACAGAAGCTATTGTAAAGCCTACTCCAGAGCT
ACATAATTCCTCCATGGATGCTGGAGCAAATAATTTGGCAGCAAATGCTACTACAGCTGGCT
CAAGAGAAAACCTCAATAGAAATGTAACCACCAATGAGGTGGATCAAAGCAAAATTAGTGGA
GATAAGAATGAAACTGTTATTAAATTAAATACTAGTACGGGTAATTCCTCAGAAACTCTGGG
GACATCTGGAAACAGTAGTACGGCAGAGACAGTAACCAAAAGTGGGAGGCGACTTCTGGAAG
AGGATGGTTCGAAAGAATCTGTGGACAGCCATTCGGACAGTAAAGACAACAGTGAGGGTGTC
CGCATGGCGACAGTAGAAAATGATGGAGGCTTAGAAGCTGACGCAGATTCATCGTTTGAGTT
GTTGCGTGAGAATGATGAGTTAGCTGATGAATACAGTTATGATTATGACGATTATGTTGATG
AGAAAATGTGGGGTGATGAGGAATGGGTTGAGGGGCAGCACGAGAACTCAGAAGATTATGTG
AATATTGACGCCCATATACTATGCACTCCTGTAATTGCTGACATAGACAAAGATGGAGTACA
GGAGATGATTGTTGCTGTTTCCTATTTCTTCGACCCCGAGTACTATGATAATCCAGAACATC
TGAAAGAGCTTGGTGGTATCGACATTAAAAATTATATTGCTAGTTCAATTGTGGTTTTCAAT
CTTGATACTAAACAAGTCAAGTGGATCAAAGAGCTAGATTTGAGTACGGATAAAGCAAACTT
CCGTGCTTATATTTATTCTTCACCAACGGTTGTTGATTTGGATGGCGATGGTTATTTGGATA
TCCTTGTCGGAACTTCCTTTGGCTTATTCTACGCCATGGATCATCGTGGAAACATCAGAGAA
AAATTCCCACTGGAAATGGCTGAAATTCAAGGAGCAGTGGTTGCGGCCGACATAAATGATGA
TGGAAAGATTGAACTTGTAACTACTGATTCACACGGAAATATAGCAGCATGGACCACCCAAG
GAGTGGAAATTTGGGAAGCACATCTTAAGAGCCTTGTTCCCCAGGGTCCTTCTATAGGCGAT
GTTGATGGTGACGGACACACGGAGGTTGTGGTTCCTACATCATCAGGAAACATATACGTTCT
TAGTGGCAAGGATGGTTCTATTGTCCGTCCTTACCCATACAGGACTCATGGAAGAGTGATGA
ACCAACTTCTTCTTGTTGATCTGAACAAGCGAGGTGAGAAAAAGAAGGGTCTCACCATCGTT
ACTACATCCTTTGACGGTTACCTGTATCTCATAGATGGACCCACCTCGTGCACTGACGTTGT
TGACATTGGCGAAACTTCATACAGCATGGTCTTGGCTGATAATGTTGACGGTGGAGATGATC
TCGATCTTATTGTCTCAACTATGAATGGAAACGTCTTTTGCTTCTCAACGCCTTCTTCTCAC
CATCCCCTTAAGGCTTGGAGATCTAGTGATCAAGGCAGGAACAATAAGGCCAATCGTTATGA
TCGTGAAGGCGTTTTTGTCACGCATTCGACCAGAGGTTTCCGTGATGAGGAAGGCAAAAACT
TCTGGGCTGAGATCGAGATCGTTGATAAATATAGATATCCATCTGGTTCACAAGCACCCTAC
AACGTTACTACGACGCTGTTGGTTCCAGGCAATTACCAGGGAGAGAGGAGGATAACGCAGAG
CCAGATCTATGACCGCCCTGGAAAATACCGGATAAAACTACCAACTGTGGGAGTGAGAACAA
CAGGAACTGTAATGGTGGAGATGGCAGATAAGAATGGACTCCATTTCTCAGACGAATTCTCA
CTAACTTTCCATATGTATTACTACAAGCTTCTGAAATGGCTTCTTGTTCTCCCGATGCTCGG
GATGTTCGGTCTGCTCGTGATACTACGGCCTCAAGAAGCCGTGCCTCTCCCGTCTTTTTCCC
GCAACACAGATTTA**TGA**

FIGURA 9

SEQ ID NO: 46, seqüência deduzida de proteínas de AtFG-GAP

MKSRARQCLLVCLLCLSLTNLSYGENKFRERKATDDELGYPDIDEDALLNTQCPKKLELR
WQTEVTSSVYATPLIADINSDGKLDIVVPSFVHYLEVLEGADGDKMPGWPAFHQSNVHSS
PLLFDIDKDGVREIALATYNAEVLFFRVSGFLMSDKLEVPRRKVHKNWHVGLNPDPVDRS
HPDVHDDVLEEEAMAMKSSTTQTNATTTTPNVTVSMTKEVHGANSYVSTQEDQKRPENNQ
TEAIVKPTPELHNSSMDAGANNLAANATTAGSRENLNRNVTTNEVDQSKISGDKNETVIK
LNTSTGNSSETLGTSGNSSTAETVTKSGRRLLEEDGSKESVDSHSDSKDNSEGVRMATVE
NDGGLEADADSSFELLRENDELADEYSYDYDDYVDEKMWGDEEWVEGQHENSEDYVNIDA
HILCTPVIADIDKDGVQEMIVAVSYFFDPEYYDNPEHLKELGGIDIKNYIASSIVVFNLD
TKQVKWIKELDLSTDKANFRAYIYSSPTVVDLDGDGYLDILVGTSFGLFYAMDHRGNIRE
KFPLEMAEIQGAVVAADINDDGKIELVTTDSHGNIAAWTTQGVEIWEAHLKSLVPQGPSI
GDVDGDGHTEVVVPTSSGNIYVLSGKDGSIVRPYPYRTHGRVMNQLLLVDLNKRGEKKKG
LTIVTTSFDGYLYLIDGPTSCTDVVDIGETSYSMVLADNVDGGDDLDLIVSTMNGNVFCF
STPSSHHPLKAWRSSDQGRNNKANRYDREGVFVTHSTRGFRDEEGKNFWAEIEIVDKYRY
PSGSQAPYNVTTTLLVPGNYQGERRITQSQIYDRPGKYRIKLPTVGVRTTGTVMVEMADK
NGLHFSDEFSLTFHMYYYKLLKWLLVLPMLGMFGLLVILRPQEAVPLPSFSRNTDL

SEQ ID NO: 47, códon de início em negrito, sítio AttB1 em itálico

*GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACA***ATG**AAATCTCGAGCGAGG

SEQ ID NO: 48, prm06644, reverso, complementar, sítio AttB2 em itálico

*GGGGACCACTTTGTACAAGAAAGCTGGGT*CCTGTTTACAGATGGTACCTAGT

SEQ ID NO: 49, gaveta de expressão de FG-GAP com a combinação de promotor-gene, seqüência de promotor em itálico, códon de início e de parada da ORF de FG-GAP em negrito e sublinhada

AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAA
ATATAAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATC
CACCTACTTTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCT
TAGTAATTAAGTGGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCAT
GAAGTTAAATTATTCGAGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTT
TCTAGCTGAACTCAATGGGTAAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTC
TGAACGTATTGGCAAAGATTTAAACATATAATTATATAATTTTATAGTTTGTGCATTCGTCA
TATCGCACATCATTAAGGACATGTCTTACTCCATCCCAATTTTTATTTAGTAATTAAAGACA
ATTGACTTATTTTTATTATTTATCTTTTTTCGATTAGATGCAAGGTACTTACGCACACACTT
TGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAATACACGTTCAACTAGCAACACATCTCT
AATATCACTCGCCTATTTAATACATTTAGGTAGCAATATCTGAATTCAAGCACTCCACCATC
ACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTTACAGAATAGCATGAAAAG
TATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAAAAAAGAATTTTGCTCGTGCGCGA
GCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCCACAGAACAACCCA
CAAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAGGCTTTG
CGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGCATCCTCCTCCTCCCATCTATAAAT
TCCTCCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCA
AGGACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTTCGATCCATATCTTCCGGTCGAGT
TCTTGGTCGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCCTTCGGTTGTTC
TTGGATTTATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATC

FIGURA 9 (Cont.)

TGTGATGATTCCTGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTT
CGGTTTGATTAGTAGTATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTA
GGGTACGGAATCTTGCGATTTTGTGAGTACCTTTTGTTTGAGGTAAAATCAGAGCACCGGTG
ATTTTGCTTGGTGTAATAAAAGTACGGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTC
GATTTGACGAAGCTATCCTTTGTTTATTCCCTATTGAACAAAAATAATCCAACTTTGAAGAC
GGTCCCGTTGATGAGATTGAATGATTGATTCTTAAGCCTGTCCAAAATTTCGCAGCTGGCTT
GTTTAGATACAGTAGTCCCCATCACGAAATTCATGGAAACAGTTATAATCCTCAGGAACAGG
GGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTTTTTTCCCAAATATCTTAAAAA
GTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATGCTTTTATAGCGTTATCC
TAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGAAGAACTTATCCGA
TTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATTCATTTGGAT
TATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAACTGTC
CTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGTA
GAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAA
TCGGGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCT
*TGCCACTTTCACCAGCAAAGTTC*ATTTAAATCAACTAGGGATATCACAAGTTTGTACAAAAA
AGCAGGCTTAAACA**ATG**AAATCTCGAGCGAGGCAGTGTCTGCTGGTTTGTTTGCTCTGTCTT
TCCTTAACGAATCTCTCCTATGGAGAAAATAAGTTCAGAGAGCGTAAAGCCACCGATGACGA
GCTGGGCTACCCCGATATTGATGAAGATGCTTTATTGAATACTCAGTGCCCGAAAAAATTGG
AGCTGCGATGGCAAACTGAAGTCACTTCTAGCGTTTATGCTACACCCTTGATTGCTGATATT
AACAGTGATGGAAAGCTTGACATTGTTGTTCCATCTTTTGTTCATTACCTCGAAGTTCTTGA
AGGAGCTGATGGAGACAAGATGCCAGGTTGGCCTGCTTTTCACCAGTCAAATGTGCACTCGA
GTCCTCTTCTATTTGATATCGACAAAGATGGTGTTAGAGAAATTGCTCTGGCTACCTACAAT
GCCGAAGTGCTCTTTTTCAGGGTATCAGGCTTTTTGATGTCAGATAAGCTAGAAGTGCCACG
TAGAAAAGTGCACAAGAACTGGCATGTGGGACTTAATCCTGATCCTGTTGACCGTTCACATC
CTGATGTTCATGATGATGTGCTTGAGGAGGAAGCTATGGCAATGAAGTCATCGACCACTCAA
ACGAATGCAACTACCACAACACCAAATGTTACAGTCTCGATGACCAAAGAAGTTCATGGCGC
TAATTCATATGTGTCAACTCAAGAGGATCAAAAGAGACCAGAGAATAATCAAACAGAAGCTA
TTGTAAAGCCTACTCCAGAGCTACATAATTCCTCCATGGATGCTGGAGCAAATAATTTGGCA
GCAAATGCTACTACAGCTGGCTCAAGAGAAAACCTCAATAGAAATGTAACCACCAATGAGGT
GGATCAAAGCAAAATTAGTGGAGATAAGAATGAAACTGTTATTAAATTAAATACTAGTACGG
GTAATTCCTCAGAAACTCTGGGGACATCTGGAAACAGTAGTACGGCAGAGACAGTAACCAAA
AGTGGGAGGCGACTTCTGGAAGAGGATGGTTCGAAAGAATCTGTGGACAGCCATTCGGACAG
TAAAGACAACAGTGAGGGTGTCCGCATGGCGACAGTAGAAAATGATGGAGGCTTAGAAGCTG
ACGCAGATTCATCGTTTGAGTTGTTGCGTGAGAATGATGAGTTAGCTGATGAATACAGTTAT
GATTATGACGATTATGTTGATGAGAAAATGTGGGGTGATGAGGAATGGGTTGAGGGGCAGCA
CGAGAACTCAGAAGATTATGTGAATATTGACGCCCATATACTATGCACTCCTGTAATTGCTG
ACATAGACAAAGATGGAGTACAGGAGATGATTGTTGCTGTTTCCTATTTCTTCGACCCCGAG
TACTATGATAATCCAGAACATCTGAAAGAGCTTGGTGGTATCGACATTAAAAATTATATTGC
TAGTTCAATTGTGGTTTTCAATCTTGATACTAAACAAGTCAAGTGGATCAAAGAGCTAGATT
TGAGTACGGATAAAGCAAACTTCCGTGCTTATATTTATTCTTCACCAACGGTTGTTGATTTG
GATGGCGATGGTTATTTGGATATCCTTGTCGGAACTTCCTTTGGCTTATTCTACGCCATGGA
TCATCGTGGAAACATCAGAGAAAAATTCCCACTGGAAATGGCTGAAATTCAAGGAGCAGTGG
TTGCGGCCGACATAAATGATGATGGAAAGATTGAACTTGTAACTACTGATTCACACGGAAAT
ATAGCAGCATGGACCACCCAAGGAGTGGAAATTTGGGAAGCACATCTTAAGAGCCTTGTTCC
CCAGGGTCCTTCTATAGGCGATGTTGATGGTGACGGACACACGGAGGTTGTGGTTCCTACAT
CATCAGGAAACATATACGTTCTTAGTGGCAAGGATGGTTCTATTGTCCGTCCTTACCCATAC

FIGURA 9 (Cont.)

AGGACTCATGGAAGAGTGATGAACCAACTTCTTCTTGTTGATCTGAACAAGCGAGGTGAGAA
AAAGAAGGGTCTCACCATCGTTACTACATCCTTTGACGGTTACCTGTATCTCATAGATGGAC
CCACCTCGTGCACTGACGTTGTTGACATTGGCGAAACTTCATACAGCATGGTCTTGGCTGAT
AATGTTGACGGTGGAGATGATCTCGATCTTATTGTCTCAACTATGAATGGAAACGTCTTTTG
CTTCTCAACGCCTTCTCCTCACCATCCCCTTAAGGCTTGGAGATCTAGTGATCAAGGCAGGA
ACAATAAGGCCAATCGTTATGATCGTGAAGGCGTTTTTGTCACGCATTCGACCAGAGGTTTC
CGTGATGAGGAAGGCAAAAACTTCTGGGCTGAGATCGAGATCGTTGATAAATATAGATATCC
ATCTGGTTCACAAGCACCCTACAACGTTACTACGACGCTGTTGGTTCCAGGCAATTACCAGG
GAGAGAGGAGGATAACGCAGAGCCAGATCTATGACCGCCCTGGAAAATACCGGATAAAACTA
CCAACTGTGGGAGTGAGAACAACAGGAACTGTAATGGTGGAGATGGCAGATAAGAATGGACT
CCATTTCTCAGACGAATTCTCACTAACTTTCCATATGTATTACTACAAGCTTCTGAAATGGC
TTCTTGTTCTCCCGATGCTCGGGATGTTCGGTCTGCTCGTGATACTACGGCCTCAAGAAGCC
GTGCCTCTCCCGTCTTTTTCCCGCAACACAGATTTA**TGA**T

SEQ ID NO: 50, motivo conservado 1
FDGYLYLI(D/E)G

SEQ ID NO: 51, motivo conservado 2
DGXX(D/E)

SEQ ID NO: 52, motivo conservado 3
DXDXDGXX(D/E)

SEQ ID NO: 53: seqüência conservada de domínio FG-GAP
ggssvaagDlnGDGrpDlvvgaPgadggtdgsvyll

SEQ ID NO: 54, homólogo 1 de AtFG-GAP, seqüência de codificação (NM_114965)
ATTCTTCATCGTTCTCTCTGCTCTGGTAAGGAGATTGCTGCTACCTACTGATGCAAATCTGC
CATTGATACATCTGCTTGCGGTTAGGGCAAAGATGAGGAAACGCGATTTGGCTATTTTGATG
CTCTCTGGATTTGCTATATTCTTCACTCTTCAGCACGAGGGTGATTTTGCGTTCAAAGAAGC
ATGGTTTCATTTGTATGATGAATACCCAGTCAAATACGAAGCTGATCGTCTCCCACCACCTA
TTGTAGCTGATCTTAATGGTGATGGAAAGAAGGAGGTTCTCGTTGCTACTAATGATGCCAAA
ATTCAGGTTCTGGAGCCTCATTCCAGGCGTGTGGATGAAGGTTTTAGTGAAGCACGTGTTCT
TGCGGAAATCACTCTTTTGCCTGACAAGATCCGTGTTGCGTCAGGGAGACGTGCTGTGGCCA
TGGCCACAGGTGTTATTGATAGGTACTATAAAAATGGAACTCCCCAGAAGCAGGTTGTGGTC
GTTGTTACCTCAGGTTGGTCTGTGCTCTGCTTTGATCACAACCTGAAAAAGCTGTGGGAAAC
GAATCTGCAGGAGGATTTCCCACATAATGCACACCATAGAGAGATAGCAATTTCGATAAGCA
ATTACACATTGAAGCATGGTGATACGGGTTTGGTTATTGTTGGTGGACGGATGGAGATGCAG
CCATATAATCACATGGACCCATTTGAGGAACTTGGCATGACAGCACAAAATGCTGATCAACA
CAGAAGAAGTGCTACTGAGAATCAGGCCTCTGAAGATTCTGGAGCCATAAACTTGCGTCACT
TTTCTGTCTATGCATTTGCTGGCAAGACTGGCCTTCTTCGATGGAGTAAAAAGACTGATGAT
GTTGAAGCTCACACCTCAGATGCATCACAATTAATTCCACAACACAATTACAAGCTTGATGT
GCATGCTCTAAATAGCCGTCACCCAGGAGAGTTTGAGTGCAGGGAATTTAGAGAATCAATTC
TTAGTGTCATGCCCCATCGCTGGGACCGACGTGAAGATACATTATTGAAGCTTGCTCATTTC
AGGCGACACAAGAGGAAAACATTGAAGAAGCAGGCTGGTAGTAAGTCTACAGCTTATCCGTT
TCACAAACCTGAGGAACACACACCCGCTGGAAAGGACTTGTCAAGAAAGATTCCAAAATTGA

FIGURA 9 (Cont.)

TTGGAAAGGCTGCACGCTATGCTGGCTCGGCAAAACCCAAGAAGGGTATGCAATACATTCCG
ACGATAACTAATTACACGAAGCTTTGGTGGGTTCCTAATGTTGTTGTGGCTCATCAAAAGGA
AGGAATTGAAGCTATTCATTTGCCTACTGGTCGAACACTTTGCAAGCTTTCTCTACTTGAAG
GTGGACTTCACGCCGACATAAACGGAGATGGTGTTCTCGATCATGTCCAGACTGTCGGAGGC
AATGTTGGAGAGAGAACTGTAGTGAGCGGGTCAATGGAAGTGTTGAAGCCTTGCTGGGCAGT
GGCAACCTCAGGCGTTCCCATCCGGGAACAGCTCTTTAACGTCTCGATCTGCCATCACTCCC
CTTTTAACTTCTTGCACTATGGAGGAGATTACTCACGACACTTCGCCCAGGCAAGAGACACC
TCTACTCTGGAGATCGCAACTCCCATTCTCATCCCCAGAGATGATGGACACAAACACCGCAA
AGGCAGTCACGGAGATGTAATCTTCTTGACAAACCGTGGAGAGGTGACATCATACACGCCTG
ATGTGCACGGCCACGACGCAGTCTGGCAATGGCAGCTTCAGACAGAAGCCACATGGTCGAAT
CTCCCGTCTCCATCGGGTTTAACTGAATCAGGGACGGTGGTCCCAACCCTGAAACCATTCTC
GTTGCGCATTCATGATAACCAGCCTATGATCCTTGCCGGGGGAGATCAAGCAGCGGTCATCA
TCTCTCCGGGAGGAAGCATATTGGCTTCCATCGAATTACCGTCTCAACCGACTCATGCACTT
ATCACTGATGACTTCTCGAACGATGGTCTAACGGATGTGATTGTGATGACCTCAAATGGGGT
TTACGGGTTTGTTCAAACCAGACAGCCAGGGGCTCTGTTCTTCAGTTCGTTGGTGGGTTGTC
TCTTAGTAGTTATGGCAGTTATTTTCGTTACTCAGCACTTAAACTCCATTCAAGGTAAGCCT
CGACCATCATCTAGCTTTTAATAGAAATCTCGTAGAGTTTTCTTCCTCTCTACGGAAGTTAA
ACCGGCAGGTCCGCTTTAGCAGCTTCACCCACATCGGTTTTGGTAAACCGGAAATATTACTG
ATATACCAGTGTAGATTCAGTGCCTTATTTCCTGGTTTCAGCTGTATAACTTTATACTTTGT
AGAATTCAAGTAAAAAATGGTAGAGGCAAATAGAGGACTTTATGTTTTTGTCAATTTTGATG
TTTTAATGGAATTATGTGACATTTAATTTAT

SEQ ID NO: 55, homólogo 1 de AtFG-GAP, seqüência deduzida de proteínas (NP_190674)

MRKRDLAILMLSGFAIFFTLQHEGDFAFKEAWFHLYDEYPVKYEADRLPPPIVADLNGDGKK
EVLVATNDAKIQVLEPHSRRVDEGFSEARVLAEITLLPDKIRVASGRRAVAMATGVIDRYYK
NGTPQKQVVVVVTSGWSVLCFDHNLKKLWETNLQEDFPHNAHHREIAISISNYTLKHGDTGL
VIVGGRMEMQPYNHMDPFEELGMTAQNADQHRRSATENQASEDSGAINLRHFSVYAFAGKTG
LLRWSKKTDDVEAHTSDASQLIPQHNYKLDVHALNSRHPGEFECREFRESILSVMPHRWDRR
EDTLLKLAHFRRHKRKTLKKQAGSKSTAYPFHKPEEHTPAGKDLSRKIPKLIGKAARYAGSA
KPKKGMQYIPTITNYTKLWWVPNVVVAHQKEGIEAIHLPTGRTLCKLSLLEGGLHADINGDG
VLDHVQTVGGNVGERTVVSGSMEVLKPCWAVATSGVPIREQLFNVSICHHSPFNFLHYGGDY
SRHFAQARDTSTLEIATPILIPRDDGHKHRKGSHGDVIFLTNRGEVTSYTPDVHGHDAVWQW
QLQTEATWSNLPSPSGLTESGTVVPTLKPFSLRIHDNQPMILAGGDQAAVIISPGGSILASI
ELPSQPTHALITDDFSNDGLTDVIVMTSNGVYGFVQTRQPGALFFSSLVGCLLVVMAVIFVT
QHLNSIQGKPRPSSSF

SEQ ID NO: 56, homólogo 1 de OsFG-GAP, seqüência de codificação (NM_185137)

ATGCGTCCCCTCCTCGCCTTCGCGGCGGTATGCGCCCTTCTCGTGGCTGCCGCAGCGCCGGC
GGCCGCGGAGGAGGAGAAGGCGAACAAGTTCCGGCAGCGCGAGGCCACCGACGACATGCTTG
GATACCCCCACCTTGATGAAGATGCTTTATTGAAGACCAAGTGTCCAAAACATGTAGAGCTG
AGATGGCAGACTGAAGTTAGTTCCAGCATTTATGCAACTCCGTTGATCGCTGATATCAACAG
CGATGGAAAGTTGGAAGTAGTGGTGCCATCATTTGTTCATTACCTGGAAGTTCTTGAAGGCT
CTGATGGGGACAAATTGCCAGGATGGCCTGCATTTCACCAGTCAAATGTTCATTCAAGTCCA
CTTCTATATGATATTGACAAGGACGGGACACGGGAAATAGTTTTGGCAACTTACAATGGTGT
AGTGAATTTCTTCAGGGTATCAGGTTATATGATGATGGACAAGCTAGAAGTACCTCGTAGGA
AGGTACACAAAGACTGGTATGTTGGGCTGAATACAGATCCCGTTGACCGCTCCCATCCAGAT

FIGURA 9 (Cont.)

GTTCATGACAGCTCAATTGCAAAGAAAGCTGCTTCTGAAGAATCTCACCCAAATATTCAGGA
CAAGCCAGTTGTGAATGAATCTTCTAAGGAATCCCAGTCACGGAGCACAAATGACTCAACAA
CACGAGGAGTTGATTCCATGAAACATGCGTCTAAGGAAGAGCCAGTTGAAAGTAAACCTAAT
TCTACCCGAGGACAGGAGAATATGGATGTGTTAAACAATCTAAACAGCACAGATGCTGGGAA
CAACTCTAGTTTAAGTACTACAACAGAGAATGCATCACATGTTCAGAGAAGGTTGCTTCAAA
CAGATGAGAAAAGTAATCAAGCAGGAAGCTCAGAAACTGATGCAAGTGATACCGGAACAGCA
AAAGCAGCTACTGTTGAAAATAGCGAGCCTCTAGAGGCTGATGCTGATGCATCATTTAATTT
GTTTCGGGATGTAGAAGATCTGCCTGATGAGTACAATTACGATTATGATGACTACGTTGATG
AAACCATGTGGGGAGACGAGGACTGGAAAGAACAACAACATGAAAAGGCAGAAGATTATGTG
AGCATAGATGCTCACATCTTGTCCACCCCAGTGATTGCAGATATTGACAGAGATGGCATACA
AGAAATGGTGATTTCTGTATCTTACTTCTTTGACCACGAGTATTATGATAAACCAGAACATC
TAAAGGAGTTAGGAGGGATTGACATTGGCAAATATATTGCAAGCAGTATAGTTGTTTTTAAC
CTTGACACAAGACAAGTCAAATGGACTGCAGAACTTGATTTGAGTACAGATAGTGGAAATTT
TACTGCCCATGCATATTCTTCGCCGACCGTGGTTGATTTGGATGGTGATGGAAATTTGGATA
TTCTTGTTGGAACTTCCTTTGGCTTGTTTTATGTTATCGATCATCGTGGTAAGGTTAGGAAC
AAGTTCCCACTTGAGATGGCGGAGATTCATGCACCAGTCATTGCAGCAGATATCAATGATGA
TGGGAAAATCGAGATGGTCACTGCTGATGTCCATGGCAATGTAGCAGCATGGACTGCAGAGG
GAGAAGAAATCTGGGAGGTTCATCTTAAGAGTCTTATCCCACAGCGCCCTACTGTTGGTGAT
GTTAATGGAGATGGTCGCACTGAAGTTGTGGTCCCAACCGTGTCAGGAAACATATATGTTCT
TAGTGGAAAGGATGGCTCAAAAATTCAGCCTTTCCCATATAGAACGCATGGAAGGATCATGA
GTCCTGTTCTATTGCTTGACATGAGCAAACATGATGAAAAGTCTAAAGGCCTCACCCTTGCT
ACTACATCCTTTGATGGTTACTTGTATTTGATTGAGGGCTCAAGTGGCTGTGCAGATGTTGT
TGACATTGGAGAGACCTCGTACAGTATGGTTTTGGCTGATAATGTTGATGGTGGAGATGACC
TCGATCTTATTGTTACTACCATGAATGGCAATGTCTTCTGCTTCTCCACTCCCTCTCCGCAC
CATCCACTTAAGGAATGGAGATCATCAAACCAGGGAAGAAACAATGCTGCATATCGTTACAA
CCGTGAAGGTATTTATGTTAAACACGGTTCCAGAACATTCCGTGACGAAGAGGGCAAGCATT
TCTGGGTAGAGTTTGAGATTGTGGACAAGTACAGGGTTCCTTATGGGAACCAAGCTCCTTAT
AACGTGACGGTTACTCTACTTGTCCCTGGGAATTATCAAGGAGAAAGGCGCATTGTGGTTAA
TGCAGCTTATAATGAACCAGGCAAGCAGCGGATGAAGCTTCCCACAGTTCCTGTGAGAACCA
CAGGAACTGTGCTTGTGGAAATGGTTGACAAAAACGGGTTCTACTTCTCTGACGAGTTCTCG
CTCACCTTCCACATGCATTATTACAAGCTTCTGAAATGGCTCGTGCTTCTTCCAATGCTTGG
GATGTTTAGCGTTCTTGTCATCCTGCGGCCACAAGAAGGCGCTCCGTTGCCATCATTTTCAA
GGAATATTGATTAGTGATTTGCACAGTGAAACGTCAGGTCCTCATAGAAGCAGATAAGCAGA
ACAAGATATTAGCAGGATATCCAATCACCCATGAAGGAGCTTGTCCACATCATACTGCATCG
GGGAGCTCAATCTGCACAGAAAAAACCCCTCACGATGCTGCATCCCATCAGTTAGCTAGTTC
GGGAAGTACAGCTTTTAGTCATGCAAAAAGAATTTTGTCACACTTGAGATTGTTTTCTAACT
TATGATATTTACACAAGGAACTGGCAGCTGTCTGATCCTGGTTT

SEQ ID NO: 57, homólogo 1 de OsFG-GAP, seqüência deduzida de proteínas (NP_910026)

MRPLLAFAAVCALLVAAAAPAAAEEEKANKFRQREATDDMLGYPHLDEDALLKTKCPKHVEL
RWQTEVSSSIYATPLIADINSDGKLEVVVPSFVHYLEVLEGSDGDKLPGWPAFHQSNVHSSP
LLYDIDKDGTREIVLATYNGVVNFFRVSGYMMMDKLEVPRRKVHKDWYVGLNTDPVDRSHPD
VHDSSIAKKAASEESHPNIQDKPVVNESSKESQSRSTNDSTTRGVDSMKHASKEEPVESKPN
STRGQENMDVLNNLNSTDAGNNSSLSTTTENASHVQRRLLQTDEKSNQAGSSETDASDTGTA
KAATVENSEPLEADADASFNLFRDVEDLPDEYNYDYDDYVDETMWGDEDWKEQQHEKAEDYV
SIDAHILSTPVIADIDRDGIQEMVISVSYFFDHEYYDKPEHLKELGGIDIGKYIASSIVVFN

FIGURA 9 (Cont.)

LDTRQVKWTAELDLSTDSGNFTAHAYSSPTVVDLDGDGNLDILVGTSFGLFYVIDHRGKVRN
KFPLEMAEIHAPVIAADINDDGKIEMVTADVHGNVAAWTAEGEEIWEVHLKSLIPQRPTVGD
VNGDGRTEVVVPTVSGNIYVLSGKDGSKIQPFPYRTHGRIMSPVLLLDMSKHDEKSKGLTLA
TTSFDGYLYLIEGSSGCADVVDIGETSYSMVLADNVDGGDDLDLIVTTMNGNVFCFSTPSPH
HPLKEWRSSNQGRNNAAYRYNREGIYVKHGSRTFRDEEGKHFWVEFEIVDKYRVPYGNQAPY
NVTVTLLVPGNYQGERRIVVNAAYNEPGKQRMKLPTVPVRTTGTVLVEMVDKNGFYFSDEFS
LTFHMHYYKLLKWLVLLPMLGMFSVLVILRPQEGAPLPSFSRNID

SEQ ID NO: 58, homólogo 2 de OsFG-GAP, seqüência de codificação (AK068943)

GAGTGTCTAGGCCAGTCCAGTCCAGTCCAGTCCAGGACGGTCCAGTCCGGTCCCGTAGAGCG
CCGGCGCGTGTAGCTTCCTCCTCCGCTCCAGGCTCCAGCGAGGGAGGGAGGGAGACCACCGT
CTCCGCCGGCTTTGAGAGAGAAAGAGAAAGAGAGAGAGAGAGCGGCGAGATGCGGAAGCG
GGATCTGGGCATCCTCCTCCTCGCCGCCTTCGCCGTCTTCTTCTCGCTCCAGCACGACGGCG
ACCTCTCCTTCCGCGAGGCCTGGTACCACCTCTCCGATGCCGACTACCCCATCAAGCACGAC
GCCGACCGCCTCCCCTCTCCCCTAGTCGCCGACCTCAACGGCGACGGCAAACCCGAGGTCCT
CATCCCCACCCACGACGCCAAGATCCAGGTCCTCCAGCCCCACCCCAGGCCCTCCCCCGACG
ATGCCTCCTTCCACGACGCCCGCCTCATGGCTGATGTCTCCCTCCTCCCCTCCAACGTTCGC
CTCTCCTCCGGGAGGCGCCCCGTCGCCATGGCCGTTGGCACCGTCGACCGCCACTACGCCCA
CGCCCCCTCCCCCTCCAAGCAGCTCCTCGTCGTCGTCACCTCCGGATGGTCCGTCATGTGCT
TCGACCACAACCTCAAAAAGCTCTGGGAGGCTAATCTCCAGGACGATTTCCCCCACGCCGCC
CACCATCGCGAGGTTGCCATTTCCATTACCAACTACACTCTCAAGCATGGGGATGCAGGTTT
GGTCATCGTCGGAGGAAGGATGGAAATGCAACATCATTCAGCAGAGCTTTTCGATGAATTTA
TGGTGTCAGAACACAACAGGGAAGAGCACCGTAGAAGCGCCAGTGAGAAGCAGGCTTCTGAG
ACAGGCAACACAGACCTGCGTCATTTTGCCCTTTATGCTTTTGCTGGCCGCACTGGTGAATT
AAGATGGAGCCGAAAGAATGAGAATATCCCATCGCAACCATCCGATGCTTCAGTGCTGATAC
CACAACACAATTACAAGCTTGATGCCCATGCCCTTAATAGTCGTCATCCTGGTCAGTTCGAA
TGTCGGGAATTTAGAGAATCAGTTCTTGGGGTCATGCCTCATCATTGGGATAGGAGAGAGGA
TACTTTTCTGCAACTTGCCCATTTTAGGAGGCATAAAAGGAAAGCACTGAAGAAAACACCTG
GAAAAGCTGTTGTAAATAACGTGCACAAGCCCAGTGAACATAATCCACCTGGAAAGGATGTT
TCCAATAGGTTAGCAAATGTGATTGGGAAAGCTGCAGATATGGCTAATTCAAATAAAATCAA
GAAGTCACAAAGGACGCTTTATGTTCCGACAATCACCAACTATACTCAAGTTTGGTGGGTTC
CTAATGTTGTTGTTGCCCACGAAAAAGAAGGGATAGAGGCTGTTCATCTAGCTTCTGGACGT
ACAATCTGCAAGCTTCATTTAACAGAAGGAGGCCTTCATGCAGATATTAATGGAGATGGGGT
TCTAGACCATGTTCAGGTTGTTGGTGCAAATGGCATCGAGCAAACAGTTGTTAGTGGTTCAA
TGGAAGTGCTGAAACCTTGTTGGGCAGTAGCTACATCTGGTGTGCCAGTGCGGGAGCAACTT
TTTAATGTTTCTATCTGCCATTACAACAATTTCAATTTGTTTCATCATGGTGACTTTTCAAG
AAGTTTTGGGAGGACATTTGATACAACTGGCTTAGAGGTGGCGACTCCTATTCTGCTCCAGA
GAGATGATGGTCATAAACACAGGAGAGGAAGCCACGGGGATATCATCTTTCTTACGAGTCGT
GGGGAGGTGACCTCGTACTCTCCAGGTCTACTTGGTCATGATGCAATATGGAGATGGCAACT
GTCAACAGGTGCAACATGGTCCAACCTTCCGTCTCCATCAGGGATGATGGAAAACATTGTAG
TTCCAACTTTGAAGGCTTTCTCTCTGCGAGCCTATGACCCAAAACAGGTAATCATCGCGGGT
GGTGATCTGGAGGCTGTGGTGATTTCTCCTTCTGGTGGTTTATTGGCATCCATTGAACTCCC
TGCACCTCCAACCCATGCGCTGGTACTTGAAGATTTCAATGGTGATGGTTTAACTGATATTA
TTCTGGTAACATCGGGGGGAGTCTATGGGTTTGTGCAGACAAGACATCCTGGGGCTCTTTTC
TTCAGCACGCTTGTGGGTTGCTTGATAGTTGTTATCGGAGTGATATTTGTTTCACTGCACCT
CAACTCCTCGAACAGCGGTAAACCTAGGGCTTCAACCGACTATAGGTGACGACATTATGTCA
GTCGACAGAATTCCTTGTCAATATTGAGGTGTTGTTATAGCATTTGATTTAGATTTTTCATA
TAATCAACTGCCAACTGGGTTATGTAAACTTGTGCGGCGAATGCATTTTTGGCGT

FIGURA 9 (Cont.)

SEQ ID NO: 59, homólogo 2 de OsFG-GAP, seqüência deduzida de proteínas (XP_463771)

MRKRDLGILLLAAFAVFFSLQHDGDLSFREAWYHLSDADYPIKHDADRLPSPLVADLNGDGK
PEVLIPTHDAKIQVLQPHPRPSPDDASFHDARLMADVSLLPSNVRLSSGRRPVAMAVGTVDR
HYAHAPSPSKQLLVVVTSGWSVMCFDHNLKKLWEANLQDDFPHAAHHREVAISITNYTLKHG
DAGLVIVGGRMEMQHHSAELFDEFMVSEHNREEHRRSASEKQASETGNTDLRHFALYAFAGR
TGELRWSRKNENIPSQPSDASVLIPQHNYKLDAHALNSRHPGQFECREFRESVLGVMPHHWD
RREDTFLQLAHFRRHKRKALKKTPGKAVVNNVHKPSEHNPPGKDVSNRLANVIGKAADMANS
NKIKKSQRTLYVPTITNYTQVWWVPNVVVAHEKEGIEAVHLASGRTICKLHLTEGGLHADIN
GDGVLDHVQVVGANGIEQTVVSGSMEVLKPCWAVATSGVPVREQLFNVSICHYNNFNLFHHG
DFSRSFGRTFDTTGLEVATPILLQRDDGHKHRRGSHGDIIFLTSRGEVTSYSPGLLGHDAIW
RWQLSTGATWSNLPSPSGMMENIVVPTLKAFSLRAYDPKQVIIAGGDLEAVVISPSGGLLAS
IELPAPPTHALVLEDFNGDGLTDIILVTSGGVYGFVQTRHPGALFFSTLVGCLIVVIGVIFV
SLHLNSSNSGKPRASTDYR

SDEQ ID NO: 60, homólogo de FG-GAP de Triticum aestivum, seqüência parcial (CK207217)

GCGTTCCGGAGCTTTNAAAAGAGTCAGGAGGGATTGACATAGGACCATATATTGCATGCAGT
ATAGTTGTGGGTAACCTTGACACAAAACAAGTTAAATGGACAGCAGAACTCGATTTGAGTAC
CGAAAGCGGGAAATTCCTTGCCCATGCATATTCGTCTCCAACTGTGGTTGATTTGGATGGTG
ATGGAAATTTGGATATCCTTGTCGGAACTTCCTATGGCTTGTTTTATGTTCTTGATCATCAC
GGTAAGACTAGGAAAAATTTCCCCCTTGAGATGGCTGAGATCCATGCACCAGTCATTGCAGC
AGACATCAATGATGATGGTAAGATCGAGATGGTCACTGCTGATGTGCATGGTAATGTAGCAG
CTTGGACTGCAGAGGGAGACGAAATCTGGGAGGTGCATCTGAAGAGCCTTGTTCCACAGCGA
CCTACTGTCGGGGACGTCAATGGAGATGGCCACACTGATGTTGTGGTCCCAACTGTATCAGG
AAACATCTACGTTCTTAGTGGAAAGGATGGCTCAAAAGTTCAGCCTTTCCCATATAGAACAC
ATGGAAGGATCATGAGTCCGGTCTTGTTAGTTGACATGAGCAAACGTGGAGAAAAGACGCAA
GGACTAACCCTTGCTACTACGTCCTTTGATGGTTATTTGTATTTGATCGAGGGCTCTAGTGG
CTGTGCAGATGTTGTCGACATTGGAGAGACCTCGTACACTATGGGTTTGGCTGACAATGTGG
ATGGCGGAGATGACCTTGATCTGATTGTTACTACCATGATTGGCAATGTCTTTTGCTTTTTC
ACTCCCTCACCGGAACATCCTCTCAAGGAATGGAGATCATCAAACCAGGAAAGGAATAATGC
TGCATATAGGCACCACCCGTCAAGGTATTTATGTAAGGCATGGTTCAGGACAACCCGGGAAT
GAAAAGGGTAAACATTTCGGGGTCGATTTGAAAATTGGAGACAAGTACAGGGTTCCCTATGG
GAATTAACTCCTTAAAATGTCAGGGTACCTTACTCGTCCTGGGAATTACCAGGGAAACAGGG
GTATTGGGGTTTGCCAATTTTAAAATGAACCGGGCACAAGGAATTGGGT

SEQ ID NO: 61, homólogo de FG-GAP de Zea mays, seqüência parcial (AY111316)

ACAGGCTTATTATAGGTTTGTTTTCTACTAATAGTTATGAAGAGCGTTTTGACACAGTATCC
GAGAGTTTTGCTTCTCCTCTCCGAGGGTCGCTTCTCTTAGAATCTGACAATTACCGTTTCAA
AGGCCTACTGTAAACTCTCAATGGTGACAATATGCTTGTTCTTTGGGTGAGCCAATATTGAA
AATGGCTGTGTACAAAATTGTTCTGCAGAACTAAATGGGTACATTNNNNNNNNNNNNNNNNN
NNNNNNNNAAATGGTTAACTGTTGGGATACTACACTGAACAAGACATTTAGGTCATGGAGAT
GTTGGGGAGTCACTTCGTCAGCTTCAATGGGTCACCGGCATTTCACCACTAGTCGATATTCC
GTGAAAACGATGGCAGCGGCGCACCTTCTTGGGGCCTCAAGATNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAGTAATGCATGTGGAAGGTGAG
TGAGAACTCGTCGGAGAAGTACAGCCCANNNNNATCGATCATCTCCACGACCACAGTTCCAG
TAGTTCTAACAGGAACAGTGGGAAGCATCATCCGTTGCTTCCCTGGTTGATGATACACACCA

FIGURA 9 (Cont.)

CTGACAACAATGCGCCTATCTCCTTGATAATTCCCCGGAACAAGTAAAGTTACCGTCACGTT
ATAAGGAGCTTGCTTCCCATAGGGCACCCTGTACTTGTCAACGATTTCAAACTCTATCCAGA
AATTCTTGCCCTCTTCGTCACGGAAGGTCCTGGAACCATGCTTAACATAGATACCTTCACGG
TTGTACTGATATGCAGCATTATTCCTTCCTTGGTTTGATGACCTCCATTCCTTAAGCGGATG
GTGTGGTGATTGAGTGGAGAAGCAGAAGACGTTGCCATTCATAGTGGTAACAATAAGATCAA
GGTCATCTCCACCATCCACGTTATCAGCCAAAACCATAGTGTACGAGGTCTCTCCAATGTCA
ACAACATCTGCACAGCCACTCGATCCCTCAATNNNNNNNNNNNNNNNNNNNNNNNACGTTGT
TGCAAGAGTTAGCCCTTTTGCATTGTCTCCATGTTTGCTCATGTCGAGAAGTAGAACAGGAC
TCATGATTCTTCCATGTGTTCTATATGGGAAAGGTT

SEQ ID NO: 62, homólogo de FG-GAP de Solanum tuberosum, seqüência parcial (BG598275)

GTTTCAATCTAGATACCAAGCAAGTTAAATGGACTGCACAGTTGGACTTAAGTACTGACGAC
GGGAAATTCCGTGCCTATATATACTCTTCTCCTACGGTAGTTGATTTGGATGGTGATGGAAA
CATGGACATTCTAGTGGGGACCTCCTATGGCTTCTTTTATGTGTTGGATCACAACGGCAAAG
TGAGGGAAAAGTTCCCTCTCGAAATGGCTGAAATCCAAGGAGCAGTAGTTGCAGCTGATATC
AATGATGATGGAAAGATTGAACTAGTTACAACAGATTCACATGGAAATGTTGCTGCTTGGAC
CGCACAAGGCACAGAAATTTGGGAAACGCATCTCAAGAGCCTTGTTCCTCAGGGACCGGTTA
TTGGCGATGTGGATGGAGATGGCCATACAGATGTCGTTGTCCCAACACTTTCTGGGAATATA
TATGTTCTGAATGGCAAGGACGGCTCATTTGTACGTCCATATCCTTATAGGACTCATGGTAG
GGTGATGAATCGAGCACTTCTTGTCGACTTGAGCAAACGTGGGGAGAAGAAAAAAGGGCTTA
CAATTGTCACAATGTCATTTGATGGTTATTTGTATCTCATAGATGGACCAACATCATGTGCT
GATGTTGTAGATATTGGTGAAACTTCATACAGCATGGTCTTGGCTGATAATGTTGATGGTGG
CGATGATCTTGATCTTATTGTAACAACCATGAATGGTA

EQ ID NO: 63, homólogo de FG-GAP de Aquilegia sp., seqüência parcial (DT735817)

GTGAATATTGATTCACACATCCTTTGCACTCCTGTGATTGCAGACATCGACAATGACGGAGT
ATCAGAAATGGTTGTCGCTGTATCCTATTTCTTTGATCATGAGTATTATGACAACCCAGAGC
ATCTCTCTGAGTTGGGTGGCATAGATATTGGGAAATATGTAGCGGGAGGAATTGTAGTATTT
GATCTTGATACAAGACAAGTTAAGTGGACAACAGAGCTAGATCTTAGTACAGACACAGGGGA
CTTCCGTGCTTATATATATTCTTCTCCAACGGTGGTCGATTTGGATGGAGATGGAAATTTGG
ACATTCTTGTTGGGACATCTTTTGGCTTGTTTTATGTTTTGGACCATAATGGCAAGATAAGA
AACAAGTTCCCTCTCGAAATGGCTGAGATTCAGGGCTCTGTCATTGCGGCGGATATAAATGA
TGATGGAAAAATTGAATTGGTTACAACTGATACTCATGGAAATGTTGCTGCATGGACCCCAG
AAGGAGAAGAAATTTGGGAAGTACATCTTAAGAGTCTTGTTCCACAACGTCCAACAGTTGGT
GATGTTGATGGCGATGGTCATACTGATGTGGTGGTTCCTACACTATCAGGGAACATATACGT
TCTCAGTGGCAAGGATGGCTCTTTTGTTCACCCATACCCATATCGTACTCATGGGAGAGTCA
TGAATCAAGTTCTTTTAGTAGATCTAAGTAAACGTGAAGAGAAACAGAAGGGACTCACTCTT
GTCACAACATCATTTGATGGCTATTTGTACCTTATTGACGGACCATCATCTTGTGCTGATGT
TGTTGATATTGGCGAGACTTCATATAGTATGGTCTTGGCAGACAATGTTGATGGTGGGGATG
ATCTTGATCTAATTGTTACAACTATGAATGGGAATG

FIGURA 9 (Cont.)

SEQ ID NO: 64, homólogo de FG-GAP de Brassica napus, seqüência parcial (CX192752)

ATTTGGGAAGTGCATCTAAAGAGTCTTGTTCCCCAGGGTCCTTCAATAGGCGATGTTGATGG
TGATGGACATACTGATGTCGTGGTTCCTACAACGTCAGGAAACATTTATGTTCTTAGTGGCA
AGGATGGTTCGATTATACGTCCGTACCCGTACAGAACTCATGGAAGAGTGATGAACCAACTT
CTTCTTGTTGATCTGAACAAGCGAGGTGAGAAAAAGAAGGGGCTCACCATTGTTACCACATC
CTTTGACGGTTACATGTACCTCATAGATGGACCCACCTCATGCACGGACGCTGTTGATATTG
GTGAAACTTCATACAGCATGGTCTTGGCTGATAATGTTGACGGTGGAGATGATCTTGATCTA
ATCGTCTCAACTATGAATGGAAACGTCTTTTGCTTCTCTACACCTTCTCCTCACCATCCCCT
CAAGGCGTGGAGATCGACTGATCAAGGGAGGAACAATAAGGCAATCCGTTACGGTCGTGAAG
GTGTCTTTGTCACTCATTCAACCAGAGGCTTCCGTGACGAGGAAGGCAAAAACTTCTGGGCT
GAGATCGAGATTGTTGATAACTACAGATACCCATCTGGTTCACAAGCACCCTACAACGTTAC
TACGACGTTGTTGGTACCAGGCAACTACCAAGGAGATAGGAGGATAACACATAGCCAGATCT
ATGACCGACCAGGAAAATACAGAATT

SEQ ID NO: 65, homólogo de FG-GAP de Citrus sinensis, seqüência parcial (CX674859)

GTTGATGGGGATGGCCATACTGATGTTGTAGTTCCAACACTATCGGGGAACATTTACGTTCT
TAGTGGCAAGGATGGCAGTAAAGTCCGTCCTTATCCTTATAGAACTCATGGAAGGGTGATGA
ATCAAGTCCTTCTTGTTGATTTAACTAAACGCGGGGAGAAAAGCAAGGGACTCACAATTGTT
ACAACATCTTTTGATGGCTATTTGTACCTTATAGATGGCCCAACATCATGTGCTGATGTAGT
CGACATTGGCGAAACTTCATATAGCATGGTCTTAGCTGACAATGTTGACGGTGGAGATGATC
TTGATCTTATTGTTACCACAATGAACGGCAATGTTTTTTGCTTCTCAACTCCTGCTCCACAC
CATCCCCTCAAGGCATGGAGATCAATTAACCAAGGAAGAAACAATGTTGCAATCCGTTACAA
CCGTGAAGGAATCTATGTTACACATCCATCAAGAGCTTTCCGTGATGAGGAAGGCAGAAACT
TCTGGGTGGAGATTGAGATTGTAGATGAATATAGATTCCCATCTGGGTCCCAAGCTCCATAT
AATGTCACTACAACATTGTTGGTCCCCGGCAATTACCAAGGTGAGAGGAGGATTAAGCAAAG
CCAAATCTTTGCACGTCGTGGAAAATATAGAATCAAATTGCCGACAGTCGGGGTAAGGACGA
CAGGGACTGTTTTGGTGGAGATGGTTGACAAGAACGGACTTTATTTCTCAGACGAATTCTCG
CTTACATTCCATATGTATTACTATAAACTACTAAAGTGGCTCCTAGTCCTCCCGATGCTCGG
GATGTTTGGTGTGCTTGTCATCCTTCGT

SEQ ID NO: 66, homólogo de FG-GAP de Asparagus officinalis, seqüência parcial (CV288972)

GGGTGGAAAGGATGGATCGTTTGTTCGACCCTTCCCCTATAGGACACGAGGGAGATTGATGA
GTCCAGTTCTTCTGGCTGATTTAAGCAAACGTGATGAAAAGTTAAAGGGCCTAACTCTTGTT
ACAACTGCATTTGATGGCTATTTGTACCTAATTGATGGGTCCACTGGTTGTACGGATGTTGT
TGACATTGGCGAGACATCATACACCATGGTCTTGGCAGATAATGTTGATGGTGGGGACGATC
TTGACCTTATTGTTACTACTATGAATGGCAATGTCTTTTGCTTTTCCACACCCTCACCGCAT
CATCCTCTGAAGGAATGGAGATCATCCAACCAGGGAAGAAACAATGCAGCAAGTAGATATAA
CCGTGAAGGAATATACATATCACATGGATCCAGAGCTTTCCGCGATGAGGAAGGTAAACATT
TCTGGGTGGAGCTGGAGATAATAGACAAGTACAGGTTTCCCTCAGGGCATCAAGGGCCTTAT
AATGTCACAACAACATTATTAGTTCCTGGAAATTACCAAGGAGAAAGGCGAATTGTCGTCAA
CAATGTATATAATCAACCTGGAAAGCAACGGATAAAGCTGCCAACAGTTCCTGTTCGAACAA
CAGGGACTGTGCTGGTGGAGATGGTTGACAAGAATGGATTGTATTTCTCTGATGAATTTTCT
CTTACATTTCACATGCATTACTACAAGCTACTGAAGTGGCTTCTGTTTCTACCGATGCTCGG
AATGTTTGGAGTTCTTGTCATTCTACGTCCCCAGGAAGGTGCACCGTTACCATCGTTTTCAC
GAAATTTGGATTGATATTAATGATATCATCCAATGAAGATGAGAATACTCCAAAGTGCGATG
CTATCTGACATGTATGCCTTTGTACATCTGAATGTGTG

FIGURA 9 (Cont.)

SEQ ID NO: 67, homólogo 1 de FG-GAP de Populus sp., seqüência parcial (CN520999)

TGAATCAAGTTCTTCTTCTTGACTTAAGTAAACNNNNNNGAGAAAAACAAGGGACTCACACT
TGTTACAACATCATTCGATGGTTATCTTTACCTTATAGATGGACCAACTTCTTGTGCTGATG
TTGTTGATATTGGTGAAACTTCATATAGCATGGTCTTGGCAGACAATGTTGATGGTGGAGAT
GATCTTGATCTCATAGTTTCAACAATGAATGGAAATGTCTTTTGCTTTTCAACTCCTGTTCC
ACATCATCCCCTGAAGGCTTGGAGATCTAATAATCAAGGAAGAAACAATGTAGCAAACCGCT
ACAACCGTGAAGGGGTGTATATTAAACCTTCATCAAGAAGTTTCCGTGATGAGGAGGGGAAG
AGCTTCTGGGTGGAATTTGAGATTGTGGACAAGTATAGAATCCCGTCTGGGTCTCAAGCACC
TTATAATGTCACTACAACCCTGTTAGTTCCTGGCAATTATCAAGGTGAACGACGGATAAAGC
AAAATCAAATCTTTGACCGTCCAGGAAAATATCGGATAAAACTTCCAACAGTTGGAGTGAGA
ACTACTGGAACTGTTTTGGTGGAGATGGTTGATAAGAATGGACTCTATTTCTCAGATGACTT
CTCGCTTACATTCCACATGCATTACTATAAACTGCTGAAGTGNCTCCTTGTCCTCCCAATGC
TTGNNNTGTTTGGTGTGCTTGTCATCCTTCGTCCACAAGAGGCCGTGCCCTTGCCANNNNNN
TNNNGGAATACTGACNNGTGANNNNCATNNNNNAC

SEQ ID NO: 68, homólogo 2 de FG-GAP de Populus sp., seqüência parcial (CX176799)

GGCACGAGCACACTTGTTACAACATCATTTGATGGTTATTTGTACCTTATAGACGGACCAAC
TTCTTGTGCTGATGTTGTTGATATTGGTGAAACTTCATACAGCATGGTCTTGGCAGATAATG
TTGATGGTGGAGATGATCTAGATCTCATAGTCTCAACAATGAATGGGAATGTCTTTTGCTTT
TCAACTCCTGTTCCACATCACCCTCTCAAGGCTTGGAGATCTTCTAATCAAGGAAGAAACAA
CGTGGTGAACCGCTACAACCGTGAAGGGGTTTATGTTACACCTTCATCAAGAAGTTTTCGTG
ATGAGGAGGGAAAGAGCTTCTGGGTGGAATTTGAGATTGTAGACAAGTATAGATTCCCATCT
GGGTCTCAAGCACCTTATAATGTCACTACAACCCTTTTAGTTCCTGGCAATTATCAAGGTGA
GAGACGAATAAAGCAAAGCCAAATCTTTGACCGTCCAGGAAATTATCGGGTAAAACTTCCAA
CAGTTGGAGTGAGGACTACTGGAACTGTTTTGGTGGAGATGGTTGATAAGAACGGGCTCTAT
TTCTCAGATGACTTCTCCCTTACGTTTCACATGCATTACTATAAACTGCTGAAGTGGCTCCT
AGTCCTCCCAATGCTTGGAATGTTTTGTGTGCTTGTCATCCTTCGTCCACAGAGGCCATGCC
CCTTACATCATTTTCAGGAATACTGACTTGTGATATCATCT

SEQ ID NO: 69, homólogo de FG-GAP de Euphorbia esula, seqüência parcial (DV130386)

CTGATACCCNNTGAAATGTTGCTGCTTGGACTTCACAAGGAAAGGAAATTTGGGAGAGGCAT
CTCAAAAGTCTTATTTCCCAGGGCCCATCAGTTGGTGATGTGGATGGGGATGGCCATACCGA
TGTGGTAGTCCCTACTATATCTGGGAATATATATGTTTTGAGCGGCAAGGATGGGTCAAATG
TTCGCCCATACCCATATAGAACCCATGGGAGAGTAATGAATCAAGTTCTCCTTGTTGATTTA
AGTAAACGTGGGGAGAAAAGCACCGGACTCAGCCTTGTAACAACTTCATTTGATGGATACTT
GTACCTTATAGATGGGCCCACATCCTGCGCTGATGTCGTTGATATTGGTGAAACCTCATATA
GCATGGTCTTAGCGGACAATGTAGACGGAGGAGATGACCTTGACCTAGTAGTCACAACAATG
AATGGGAACGTCTTTTGCTTCGCGACTCCTGTTCCACATCATCCCCTTAAGGCATGGAGATC
GGCTAATCAAGGTAGAAATAATGTGGCAAACAGATTTAACCGTGAAGGGGTCTATGTTACAC
CTTCATCAAGAGCTTTTCGAGACGAGGAAGGAAAGAATTTCTGGGTGGAAATCGAAATCGTG
GATAAATATAGATACCCTTCTGGTTCTCAAGCACCTTATAAAGTCACTACAACCCTGTTAGT
TCCGGGTAATTATCAAGGTGAG

FIGURA 9 (Cont.)

SEQ ID NO: 70, homólogo de FG-GAP de Ceratopteris richardii, seqüência parcial, (CV736049)

CCACGCGTCCGGGTTACAACAGATGCCCGGGGAAATGTTGCTGCATGGACAAATAAAGGAAA
AGAAATTTGGGAAGTTCACTTAAAAAGCTTGATTGCTCAGGGTCCGACTGTCGGAGATATTG
ATGGAGACGGAACATTGGATATTGTTGTGCCAACTGTGTCAGGAAATATATATGTTTTGCAT
GGAAAGACTGGTGTTACTATGAAGCCATATCCATTTCGTACCCATGGGCGTGTGATGGCTCC
TGTTCTTCTGGTGGACTTGGCCAAGCTAGGCACCGAGAGAAGAGGTCTGGTCCTGGTGGTTT
CATCATTTGATGGGTACCTCTACCTCATAGATGGTCCCACAGCCTGTGCTGATGTGGTTGAT
GTAGGAGAGACATCCTACACAATGGTGCTTGCAGACAACGTTGATGGAGGAGATGATTTGGA
TTTAATAGTGACGACCATGAATGGGAATGTTTTCTGCTTCTCTACACCTGCTCCTTATCATC
CTTTACGGTCATGGACCTCACAGAATCAGGGTAGGAGCAATCTTGCATCTCGTATAAGACAT
GAAGGGATTTATGCTACATCTGGCTCGAGAAGTTTCAGGGATGAAGGTGGTGAAAGTTTTTG
GGTGCATTTCAATATTATTGATGAGAACCGCTTACCTTCGGGGCCTTACAATATTACGGCTA
CCCTTCTTGTCCCCCACAACTACATGGGTCCAAGGCGCATTACAGAGAGTCAGTTAATTCAA
CAACCAGGAACGCACAAGTTCAAACTGCCATGTGTATCTGTACGATCATCTGGGACTGTAGT
TCTAGAGATGGTGGATAAACATGGGTTCTATTTCACTGATGAGTATTCATTGACATTCCACA
TGCACTATTATCGGCTACTGAAATGGATGGTAGTGCTTCCAATGCTTGGAATGCTTTTTGTC
TTGTTGTGGCTTCATCCTGAGGATGCCATATCAGTTCTTCCTTCCTTTTCAAGGGATCATTA
AAATGACATATGAGCTCAATTTTTGGTACTCCTAAATCCAAACCTTATTGTAAATGTATACA
TTTTGACATTTATCCAGTCTACTAAAGAAATGCA

SEQ ID NO: 71, homólogo de FG-GAP de Welwitschia mirabilis, seqüência parcial, (DT601669)

ATGACAATCATGAACACTTGAAGGAGCTGGGTGATATTGATATTAGCAAATATGTTTCAGGT
GCTATAGTGGTATTTAATCTTGATACAAAGCAAGTGAAGTGGAGTACTCAGTTGGATCTTAG
CACAACCTCTGGGACTTTTAATGCATACATCTATTCTTCTCCTACTGTGGTTGACCTGGATG
GTGACGGCAATTTGGATATTATTGTTGGCACGTCATTTGGTTTCTTTTATGTATTGGATCAC
CATGGAAAAAATAGAGAAGGCTTTCCATTACAGATGGGTGAGATTCAAGGACAAGTAATTGC
TGCTGATATAAATGATGATGGGAAGATTGAAATGGTTACAACAGATACCCGTGGAAATGTGG
CTGCTTGGACTTCTCAGGGGAAAGAAATATGGGAGATACATTTAAAAAGTCTTATTGCTCAG
GGGCCTACAGTAGGTGATATTGATGGAGATGGTCATACAGATTTGGTGGTTCCTACAGTATC
AGGAAATATTTATGTATTAAATGGGAAAGATGGATCATTGGTGAAGCCATTTCCTTATCGTA
CTCATGGAAGAGTCATGAGTCCGGTACTTTTGGTTGACCTTAGCAAACGTGGGGAAAAGCAA
AAGGGCTTAACTCTTGCAGCATTATCATTTGATGGTTACTTCTACTTAATTGATGGTCAAAC
AGCTTGTGCAGATGTTGTTGATATTGGAGAGACTTCCTACTCTATGGCTTTGGCGGACAATG
TAGATGGCGGCGATGATCTTGATTTTATTATCACAACTATGAATGGCAATGTGTTCTGCTTT
TCAACCCCTGCTCCACATCATCCA

SEQ ID NO: 72, seqüência genômica codificando homólogo de FG-GAP

ATGAGGAAGCGTGATTTGGCGATTCTTATGCTCTGCGCTTTCGCTATTTTCTTCTCTCTTCA
GGTATTGTACCCTTTAACCCCCTAATTCACTTTCTCACTCTTAAACCCTAACCCTCATTTTC
AATTCAATTGATTCAGCAAGATGGTGGTGTTTCATTCAAAGACGCGTGGATGCATCTAACAG
ATGAATACCCAATCAAATACGAAGCCGAACGCCTTCCTCCTCCTGTTGTCGCCGATCTCAAC
GGCGATGGTAAGAAAGAAGTTCTCGTAGCTACCCACGATGCCAAAATTCAGGTTCTTTCATT
TCCCAATTTTCCTCAATTTTTTATTCACTTATAAAATTCCAGGTTTTATTTGGTTTCCTTCA
CAACTGGACATTGTCAAAAAAACATTTTTTTTCTTTCAGTAAACCATCAATGTAGTTGATGA

FIGURA 9 (Cont.)

ACTATAACTCTCTTCAATTTAGTTTGTAAGTAGTCTCTAAAATTCATCAATTAATCCCTGCT
ATTAATGTTTGGTTCTACACCGAAGAAAGTTGACTTTGAATGAGTTGATTGTGTAAACTCTA
TTTGGGCTAAAATTGATTTGAAGGTAAAGTGATTTATGTTTGGACTTTGGATACATTCATGT
AAAAGTGAATTGAACAGGTGAAAAATCAACTCTAGAATCAGAAGCTACAATTTCTAACTTTA
AGTAGAATGTAGAATCAATTCTGGAGGTAAAATCAATTTTACTCTAGAAGAACCAAATATGT
CAATCAATTTTGGGTCCTCCAAAATTAGAACCAAACATACAGAAAGTAGAGACTAAAATTGA
TGGATTTTCATATACTTTAGGGACTTAATTGACCGGTTTTATATACTTTAAAGACTACTTAT
CATATTTATATAGTTTAGGGACTAGATTGATGGTACAAGTAATATAGTAGTAACTGGATTTT
GTATTTATGGTCGGTGTCTTAGATTTTGGAGCCACATAGTAGGCGTGTTGATGAAGGATTCA
GTGAGGCACGTGTGTTAGCGGAGGTGTCTCTGTTGCCTGACAAAGTGCGTGTCATGTCTGGG
AGACGACCTGTTGCTATGGCCACTGGCTTTATTGACCGCCATAGAATTGGACAACCACATAA
GCAGGTTTTGGTTGTAGTAACATCGGGTTGGTTTGTAATGTGTTTTGATTCCAACCTCCAAA
AGTTGTGGGAAAATAATTTGCAGGTATGATGTTTATGTTGTGCCATGTTTCAGTTGATAGAA
TTTTCAGTTAAACATTTTGATTATCTTGTTATGTTTTAGTTTACCAACTGTTTTAGTGGTTG
TTAACAGGAGGATTTCCCCCATAATGCTCACCACAGGGAAGTTTCAATATCTATAAGCAATT
ATACTCTCAAGCATGGAGATACAGGATTGATCATTGTTGGTGGGAGAATGGAAATGCAGCCT
CATGTATGTGTCACTTTCCTTCTTGTTTTCTGGTTTGTTATTCTTTTTAGCAGAGTATGTGT
GTCTGTGTTCTATCTGCTGATGGAAAAACTTAGAATTTAATGTGGATGGTGTTTACCATCTT
TTATGAAAGCGAAAACTCTCAACTTCATGAGTAGCATAAGATTCTGGCTATCTTGTTTTCTT
GGAAGTGATGAAAATGAAAGTTCCTGAGTAATTTTATTTGATAAAGTATGATGGTAATGTGA
GGGTTTGCTGCTGCTTCTCGAGTTTCAGAATTTTTTATTTGTTATTTGATTCTTTCTATGGT
AATGATTTTGGAATTGTTCATCTTTGGAATTATTGGCATTATCTTCTAATCTGATGGTTCTG
TTGGCTTATGTTGATTGCTTTCCATTTTCAGATTTTTATGGACCCTTTTGAAGAAATGGGAA
TGGGAGCTAGATTTGCTGAGCAGCATCGAAGAAGTGCTACTGAAAAGGAGGTATAGATACGT
TTCTAGTTGATGCTAAATTTTATGCCCATTTTTTTACTATCTGTAGCATGCTAATTTTGGTT
TAAACTTTGAGACACTATTTGCTTTATTTATTTCACTTTAATTTTGGGTCTTCAAGGTTGAC
CTTATTTTATGCTCCTAAACAGTCAATTCAAATTCCTATTTTGATATAAATAGATAATATAA
AAGTTCACCATGTCTTGGCATGATAAAGCCAATAAATAAAATCTGTGCACCTTGCAGTAAGT
AGTCTATAAATCTGAACCTTCATGATCAGGTAGCTAAATAAACAGGAGACAGTTTCACTCGT
TTTGTTTAGATTTTTGTTTTTACAAAAAAATGTATCAGGTGATTTTAGTTTTCTTCCCATTT
TAAGGTTGAAGACTTTAACTTGTTTATTGTGTTCCTTTTTCTGGATAGGCTTCTGAAAACAC
CGGAACTGTGGATTTACGCCATTTTGCATTTTATGCATTTGCTGGTCGATCAGGTGTAGAAC
GATGGAGCAGAAAAACAGAGGTGGTGTTCTTTTTTTTTTGTTTTCAGTGCGCTCTATTATTT
GGGTGTTGTCAACATGCAAATATGAAACAATTATTTATTTTTAATTTAAAAATGCATTTTCT
TGTTTTTTAAAAGAGCTGTCTTTTCTATCTGGAAATTCATTAGGTGATTAAGTGAACTTTCT
CTGTCCTCCCTCTCTGTACATAACTTTATTCAGGGGGTAGCAACTTAGTTCTAGACAGTATG
TTTTGTCCTGCAACAGACATGAAAACTGAAGAAAAGTTATATAGTTAAGCTCACAAAGTTCA
GAACATAATGAGTTAGAGCCTAAATGTCAATAATAGATTATAAATTATACCTTTGAAGATAT
TGGAGTCTTGTTGCAAGCATCTGCACTTGGATTTTAACGCACTGGTGCACAAATATGTATTT
TTTGAAAAGAAAATCTATCAAACAAATGGAACCACCTCCCAACCAGCATAAACTTTAGTCTT
TTCTGTCTTTGGGTGGTAGTACCAGATAGATGTTTTACAATCAACATAGCGGTCCAGTTTGA
TTGAATATGTGATTTACGCTAGTCCAACATTTATTTTGTTAAATGATCTTGTGTATCCTTGT
CCTTTTTTTTTTCTTTTCTTTTCTAATTGCTAGTATTGTGATTTAATTGTTTAGAACATTGA
AGCAGCAGCTTCTTCAGACGCATCACAGTTAATTCCACAGCATAACTACAAACTTGATGTTC
ATGCTCTGAATAGACGTCAACCTGGAGAGGTACTTAGCTATAGCTGTTCAATTTAATCCTAA
ATTCCTAATGGTATATTGAATTGATGGAATATGGTAAATTGGAGTGGCATAAAATACCATTC
CATCCTTTTTAAGCAGGGGAATCATTGCTTTTCTTTTTCATACATCCATTTTATTCTAATCT

FIGURA 9 (Cont.)

AGAAATATGCATGTTGATTTTTATGTTCTCCCCTTTAGCTCAACGTACTGATTTTGTTATAT
GTAAGTAATACTGTAATTTCAGATTTATTACTTTTTCCATTCCTTTCCTCATCAATCTTAAG
TAATACCGTAGAATAGATTCCTAGTCATATCTTTATGTTCTTCACCTTTTCTCTTCAAATGA
TCAAACACTAAAGCAAATATTTGTTTGTTAATATGGAGTTTACTATGAGAACTCTAGTACAA
TAGTACTCATCCAAATCTACTTGAGTAACTTGGTCATCTCCTTCCTTTGTGTGTGGAAATAG
CAGCATGTGCACATAAAGGAGTGTTTCAATCAAATTTAATATTATTGTGAAACCTTTGTATA
TGTTTATCCATATCTACCCAATAGATAGCTTCCTGTGACATGTGAAATTATAGTTTGAATGC
AGGGAATTCAGAGAATCAATCCTGGGAGTTATGCCACATCAATGGGTACACTCTATTTTCAA
CCTTGATTAACGTTTGGTACTAAAGCATAATCTCAAATGTCATAGTTAACTGATATTAGATA
CATGCACTCATCAGGATAGGAGAGAAGATACTTTATTGAAGTTGGTCCACTTCAATCGGCAT
AAGAGGAAAACATTGAAGAAAACACCTGGAAAGACTATCAATTACCCTTTCGACAAGCCTGA
GGAAAACCATCCACCAGGGAAGGACTCAACCAAAAAAATTTCAAACATAATTGGGAAAGCAG
CAAATTTTGCTGGTTCAGCAAAATCCAAGAAGGTAATTAGAAACCATGTCATGCAATGCAAC
TGGTAACAATTCATTCAATTGTTGACATAATGATAATTAGTTTTAAGATGGCTACAACTGGA
CAACTCAAAAATTCTTTTCCAAACATGATTTCAGCTACAAGCTTTTGTTGTCTTTATCCAGT
AACAGAGCTGTCTGTTTACTTTCTAACTACTGCATGTCCAATTTTTTGTTATGTGATCTAAA
TCTTTGTGTTGGTAATATTCTCAGCATCTGCTGATTTTAATTCATTTTGGCTAAACCTATTT
TGAATTTGTTATGATAACCATTGCTTAGCATTTTTAATTTTTTTTTTAAACTAAGCACTGGT
CAGATAAACCATCTCACTTGTTATATTTGTATTTGTTCTTAAAGATTTAAGTCTAGCTACAT
TTTTTATTTCTCATTTAAATTTTTTGAATCCTGAATTTAGGGAGCCAAGATATTTAGTTATA
AAAATAATGATTAATAATAAATATGCAGTTCTCCCCATTGTCGTGTTAATTTGGTAGATAGC
ATAGTAGTCAATCTTGGATAGCGGTGCGTAGTGGCCAACCCCAAAAATGGGATAGCGTATAG
CGGTATGACAAATAGCGGTCACCGATGTTTGATTATTTTTTGGACAATTATATGAATAATAG
CAAAGTATAAACAAACTTATATATTTCTCAACAAAAAAAACTTATATTAAATATAACACTCA
AATCTCACGAGACATTCATAAATCAAAACACCGAAAACAAAAGACAGGAACAAATAAAAAAA
ATCATAAAACAGGAGAAACCATAAGATAGAAGATAGAAGAGTGTGTTTCGCGGCACAGGTGA
GAAGAGAGAAAAAGTTGAAAAATGTTAGGGTTTACGAAATCCTCAAAATGCCCTAAAAAAGG
TTTTAAAAATAGGGTCAATTTGGAAATTTCCTACTAAAATAAATAGCGGCTGCTATCCAACC
CGCACCGCTATAGCAGCTGAAGCGGCCGCTATTTGATTCCGCGCTATTTGATACCGCTATGC
TACACGCGGCCGCTATAGCGCCACTATCGGCTGCTATTGACTACTATGGTAGATAGAAACTG
TTAGGTTACCTAGTAGGGCTGATGATGTGTTAGTGGGCTCAAGTAATGAGTGAGAGGCCCAT
TATTTGGAGGGTGTATGAACGTGGGAAGAAGAGAGGAGAGAAGGGACTAAGGGTGCTTCCTA
GCTGGCAGTTAGTTAGTATGAGGCAGAGGGAGTGAATATATTTCTTCTGAGGAGCTGTTGCT
CTAGTTTATCCGTCATAGTGATTTCATTGTTATTCTAGTTTAGCCAGGCTTTGGGAGTATTT
GGTATATCTACCTCGAAGATGCTAGGTCCTAAAATGATGCAGTTGGAAGGTTCGATTGCGGG
AGTCTCTCTGTTGGTGATGATTTAAACACAATTTCATATACTCCTGGTGTGGTTGAAGCGTT
GGGGGTTGCCTTTGGATTCATCAGATACTGCTGGTGTGCATTTGGGTGGGGGTGGGGGGTCA
TATGGTGCAGATATTGGGCAAATGTGTGGTTGGCTTCTGCATATCGGCCCTTATGCAGTCTT
GCTGGATACATATGTGTTGGGGTTAGAAGAGGTCATGCCAAAGTACCAATTACTCTCTTAAT
ACCTGTCTAATAAGACTCTAATGCTTTAAATTGACGAGGATTTGCATTTTTCACCTTTCACT
GGCAAACATATGGGTTAAACTGGTAGATGTTTAGCGTGTTCAAGTACAAGCTGTGTTCCCAT
AGATGAAGTACTTCATGTTTGGGGAGTTTGATCGTTAAACTACCTAATTGTTTGTTTAAATG
AATTTTGAATTCCCTCTTCTAATAATTTTAGTGCTTCCAAAGATTATAATGCATATAAAACA
ATTTGAGGATTATCTTGTTCTTAGGTGAAAAACAGTTTTTTTCCCATTTGAAATTAGATAAA
CAGAGTAGTTGTAGATTTCTACTGTGGAAAGCATAACATGTCTAAGCTGGTTAAATATACTT
AATTGATGGTTTTTGTCATTAAAGTAATTTCTTTGCATACATTTTTATGTTCTCATGTTATT
GCTTAATCGGAAAGGGGTTCTATAAAATGGATGGCATTTTATGGGATTCCGAGAAGCCTTTT

FIGURA 9 (Cont.)

TCTTTTTTTTGCCAGGGTGGGGGGTTTGCATTATTTGTCATAATGTTTTTTCTTTTGGAATG
CAGTATCCACCATATGTTCCCACCATAACCAACTATACTAAGGTTTGGTGGGTTCCTAATGT
TGTTGTGGCTCATCTAAAGGAGGGGATAGAAGTCCTTCATCTGGCATCTGGTCGAACACTAT
GTAAGGTAAATATAAAGCCACTGATCAGCTTGAAAATCAAAATACAAAAAAAAAAAAAAAAA
GTTTACAAGTAGTTGTACAACATTTGCTCCTGTGGTTGTAGATTTGTAGTATTAAACTGCCA
TGGTTTCAGCCAAAATCTTATTTAGCATTCACGTTACTTTTGAAATTTTTCGTATTCTGTGA
AATGACACTTCTATTTATCTCAGCTTCACCTTCAGGAAGGTGGTCTACATGCTGATATTAAT
GGTGATGGAGTTCTGGATCATGTGCAGGTTCACTTTTTTTCCTTCACATTTTTTGATAAACA
AACAAGCCAAATCTTTATTAATTTAGCATGATAATTGTTAACTTAACGGGGGATTTTCCTCT
GCCTGATGTCTGAAAGATTTATCTAATCAGGCTGTTGGAGGAAATGGTGCTGAGCAAACTGT
AGTTAGTGGGTCCATGGATGTTCTACGTCCTTGTTGGGCTGTTGCAACGTCTGGCGTACCAG
TACGGGAACAACTCTTCAATGTATCTATTTGTCATTATACCCATTTTAACTTATTCCAACAT
GGAGAACTTTATAGAGGCTTCAACCGAGGTTCAGATATGTCTTCTTTGGAGGTAGCAACACC
CATTCTCATTCCTAGAAGCGATGGTCACAAGCACCGGAAGGGAAGCCACGGGGATGTTATCT
TCTTGACAAACCGGGGAGAGGTACGCTCTTTTTTGTTCGACTAAATCTGGTTATGTAACATT
TGATAGTTTGTTTGAGCACCAATCAAACAAACCTAGTTCTTAACTTCTGTTAACAAATTATT
TCAAGTCACGCACGACTAATATATTTTGGTGGTATTTGTGAGATAATCCGTACCTCCGTACT
GACTTTATGAAAATTGCTCCTTGCAGATAACTTCGCACACTCCTGGTTTGCATGGTCATGAT
GCTGTTTGGCAGTGGCAACAATCAACTGGTGTCACATGGTCAAACCTACCTTCCCCAGCAGG
GATGATGGAAGGTGGTTTGGTGATTCCCACACTAAAGCCTTTTCCTTTGCGGTTGCATGACA
ATCATGAGATGATCCTTGCAGCTGGAGAACAAGAGGCTGTGGTAATATCACCCGGAGGTAGC
ATATTGGCTACAATTGAACTCCCTGGTTCACCTACACATGTATTGATCCGTGAAGACTTCTC
AAATGATGGGCTCACTGACCTTATTCTCGTCACCTCATCCGGAGTGTACGGCTTTGTCCAGA
CTCGACAACCTGGTGCTCTCTTCTTCAGTGTGCTGATCGGCTGTCTCATAGTCGTGATGGGA
ATTATATTTGTTACCCAGCACATAAATTCCATGAAGGGGAAGCCTCGTCCATCATCTGGTCC
TAGGTGAAATGCAGGACCTGAAATTTATACATTAACTTTGCAGGCGCTTACAGATGAAGCAG
AAGACAAAGCGGTGTGCGCATCGTATGGAGCTTCAAGATTGTTGGAAATGCATAGAATGTAA
GGAGATTTTGAGGAGGTGTTGTACTGTCTAGGATGGATGATTGTAAAATTTAGACAGAAGTG
ACAGAATGTCCTTCATGAGAGAAACATGATCCATTAATATAAGGAAACTATAGTCTTTGCCT
ACGTATCTGGCTGTTGTTTAGAAGTTATACTGCCCACACTTGACAATTGTCACCCGTTCCGA
ATTAGTTGATACTTGATAGTTGATGCATAACAGAAACTCACCTTGGTCTTTACACTTTTCGC
TAAAAGGATTTTCTCTTGTCATTACACAATTCTTGTAAATACGACCATCGTTTATTTGCTTC
AGATTTAGCAGTGTAAAACATTAAATCTAGCAACCAACTTGCTCCAATCTGTCTTGGTCTTC
TTCAGAATACACTTCTTAAAAGTCAATTAAGAAATGACTTCCTGATTCCATCAAATTACCAA
TGTGGTTGATCCCATCATACATCAAGTGCTATGTACGTAATCCCTGTTAGTCTTTGGCATCA
CCTTAAGTGTTTTAGCTATTGCTTTTCCTGTGCGTCGAGAAATGTTTCAAATAACCATCTGA
ATGATTCAATATTGCATCATATAATATTGCAAACCCAAAATTAGTGTGTTCCTATAATGGAT
AAAGCTTGAAAATAATGCGAGTGGTCGATTAGTGCGATTGGTGCAATGTTGGATTCATTAAC
AACTGGTATTGAAAATATTGCATAAAGCTATCAACTTCTCCATATACCATGCTACATTGTCT
TTTGCCTGATGATGATTTTTGACATCCAACCGTGTAAATACAATATTAGT

SEQ ID NO: 73, homólogo de FG-GAP de Medicago sativa, peptídeo compreendendo um domínio FG-GAP

IQQDGGVSFKDAWMHLTDEYPIKYEAERLPPPVVADLNGDGKKEVLVATHDAKIQVL

SEQ ID NO: 74, homólogo de FG-GAP de Medicago sativa, peptídeo compreendendo o motivo conservado 2

LSQLHLQEGGLHADINGDGVLDHVQ

FIGURA 9 (Cont.)

SEQ ID NO: 75, peptídeo de homólogo de FG-GAP de Medicago sativa

QAVGGNGAEQTVVSGSMDVLRPCWAVATSGVPVREQLFNVSICHYTHFNLFQHGELYRGFNR
GSDMSSLEVATPILIPRSDGHKHRKGSHGDVIFLTNRGEVRSFLF

SEQ ID NO: 76, homólogo de FG-GAP de Medicago sativa, peptídeo compreendendo um domínio transmembrana

QITSHTPGLHGHDAVWQWQQSTGVTWSNLPSPAGMMEGGLVIPTLKPFPLRLHDNHEMILAA
GEQEAVVISPGGSILATIELPGSPTHVLIREDFSNDGLTDLILVTSSGVYGFVQTRQPGALF
FSVLIGCLIVVMGIIFVTQHINSMKGKPRPSS

FIGURA 9 (Cont.)

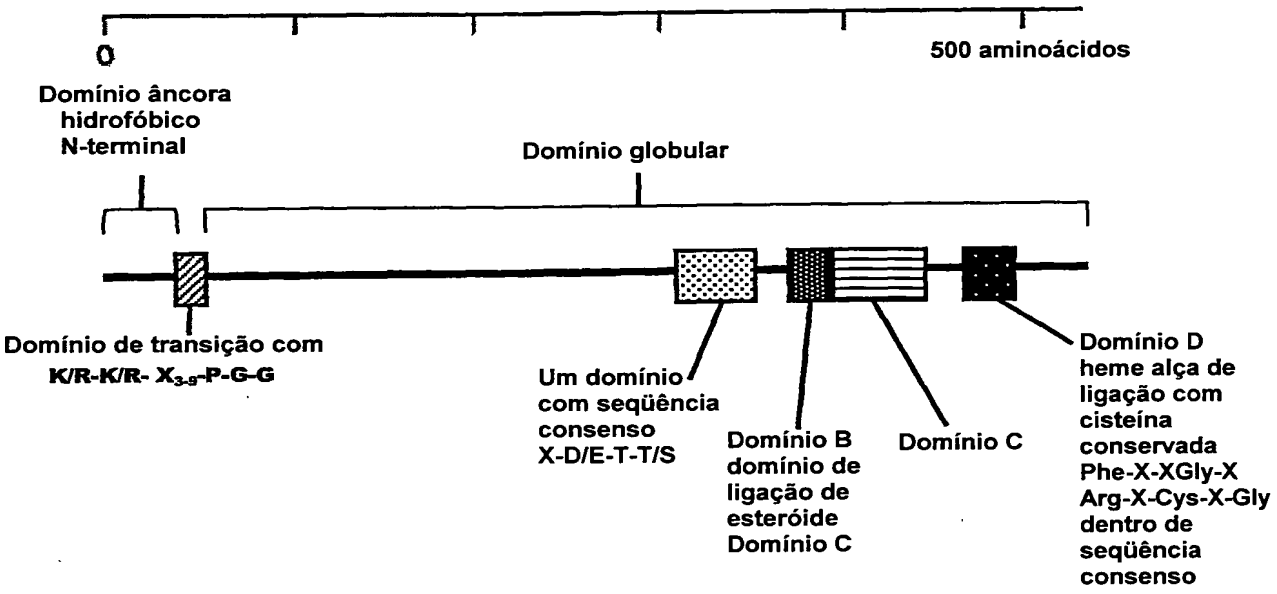

0
500 aminoácidos
Domínio âncora hidrofóbico N-terminal
Domínio globular
Domínio de transição com K/R-K/R- X3-9-P-G-G
Um domínio com seqüência consenso X-D/E-T-T/S
Domínio B domínio de ligação de esteróide Domínio C
Domínio C
Domínio D heme alça de ligação com cisteína conservada Phe-X-XGly-X Arg-X-Cys-X-Gly dentro de seqüência consenso

FIGURA 10

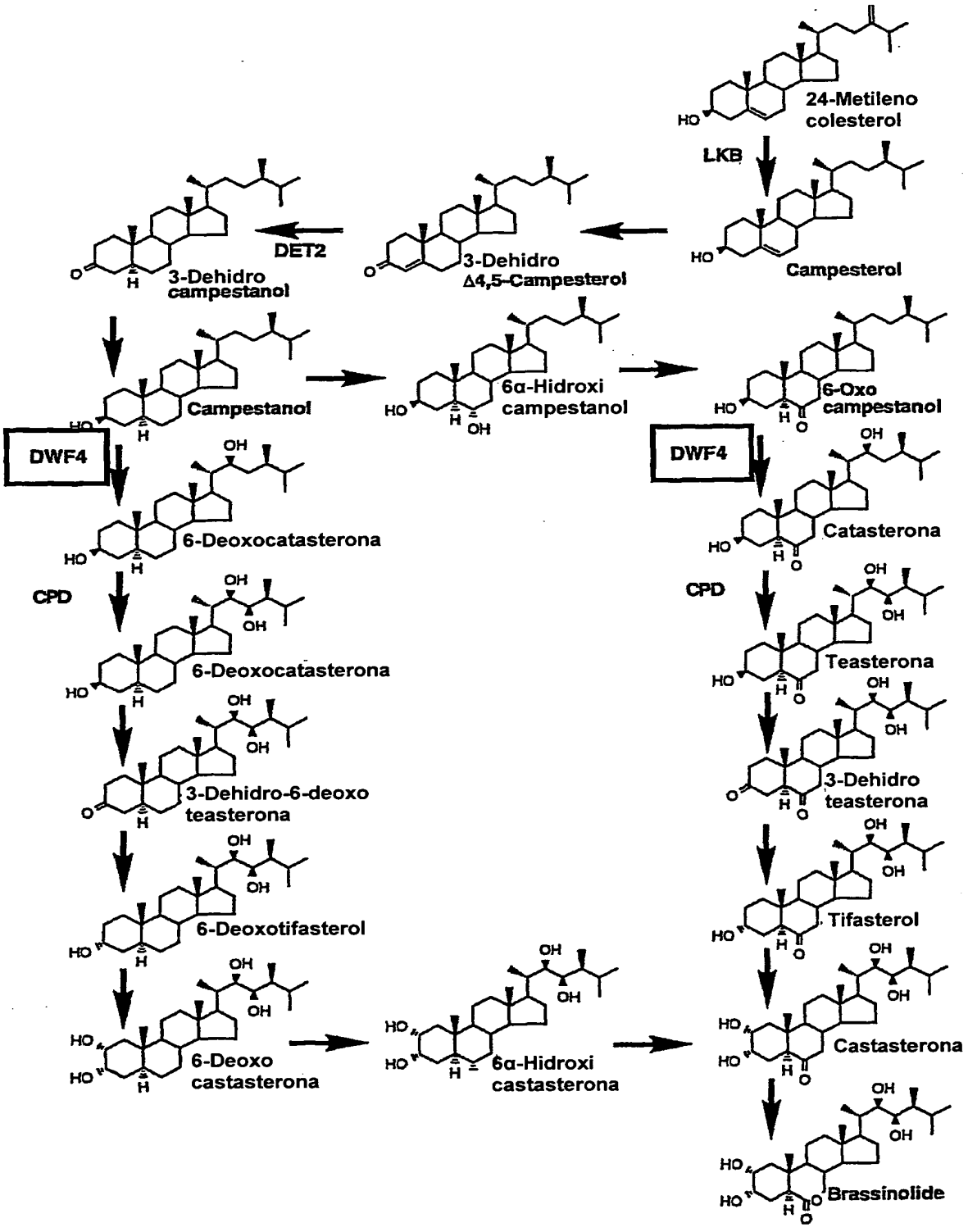

24-Metileno colesterol
LKB
Campesterol
DET2
3-Dehidro Δ4,5-Campesterol
3-Dehidro campestanol
Campestanol
6α-Hidroxi campestanol
6-Oxo campestanol
DWF4
DWF4
6-Deoxocatasterona
Catasterona
CPD
CPD
6-Deoxocatasterona
Teasterona
3-Dehidro-6-deoxo teasterona
3-Dehidro teasterona
6-Deoxotifasterol
Tifasterol
6-Deoxo castasterona
6α-Hidroxi castasterona
Castasterona
Brassinolide

FIGURA 11

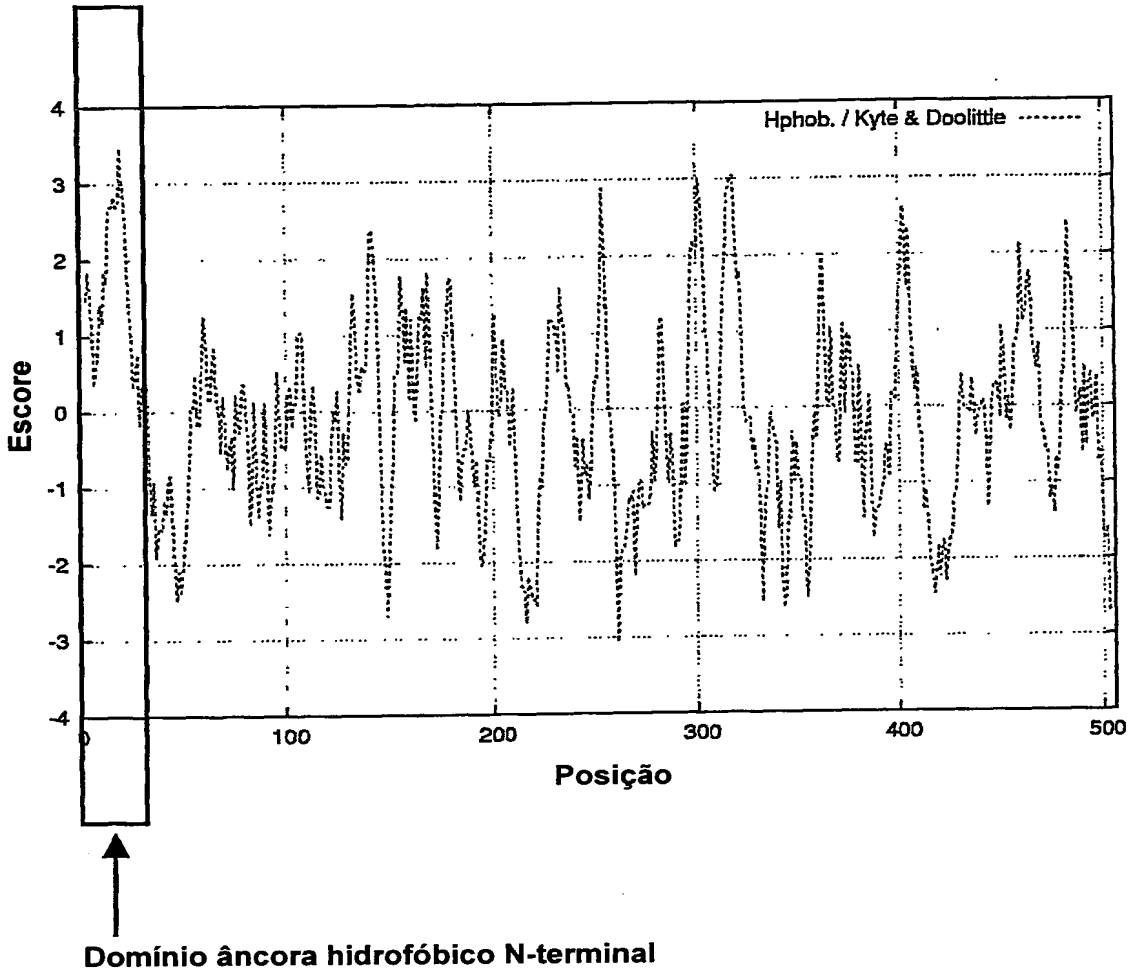

Hphob. / Kyte & Doolittle
Escore
4
3
2
1
0
-1
-2
-3
-4
100
200
300
400
500
Posição
Domínio âncora hidrofóbico N-terminal

FIGURA 12

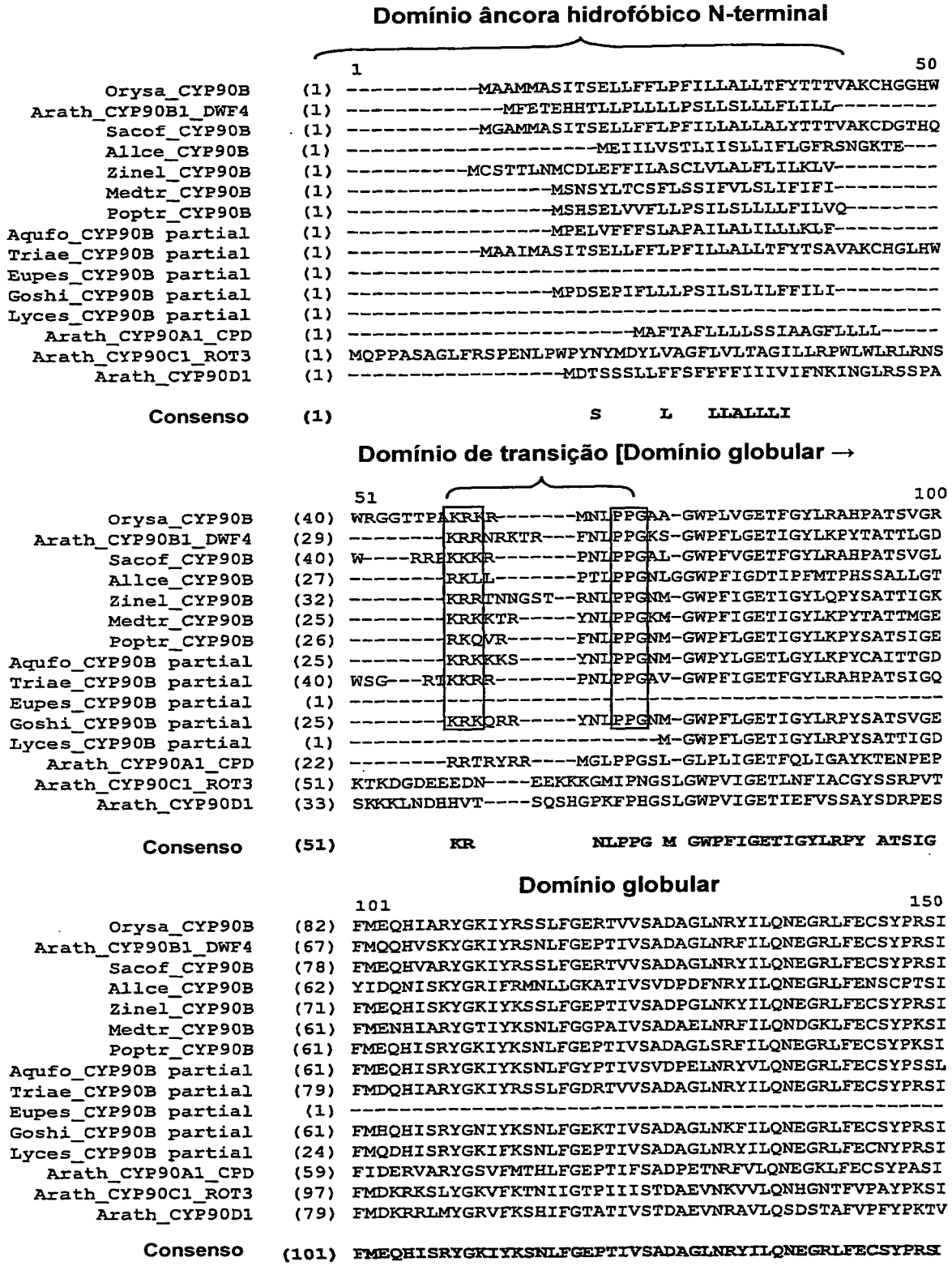

Domínio âncora hidrofóbico N-terminal
1
50
Orysa_CYP90B (1) ----------MAAMMASITSELLFFLPFILLALLTFYTTTVAKCHGGHW
Arath_CYP90B1_DWF4 (1) ----------MFETEHHTLLPLLLLPSLLSLLLFLILL---------
Sacof_CYP90B (1) ----------MGAMMASITSELLFFLPFILLALLALYTTTVAKCDGTHQ
Allce_CYP90B (1) --------------------MEIILVSTLIISLLIFLGFRSNGKTE---
Zinel_CYP90B (1) ----------MCSTTLNMCDLEFFILASCLVLALFLILKLV---------
Medtr_CYP90B (1) ----------MSNSYLTCSFLSSIFVLSLIFIFI---------
Poptr_CYP90B (1) ----------MSHSELVVFLLPSILSLLLLFILVQ--------
Aqufo_CYP90B partial (1) ----------MPELVFFFSLAPAILALILLLKLF---------
Triae_CYP90B partial (1) ----------MAAIMASITSELLFFLPFILLALLTFYTSAVAKCHGLHW
Eupes_CYP90B partial (1) ----------
Goshi_CYP90B partial (1) ----------MPDSEPIFLLLPSILSLILFFILI---------
Lyces_CYP90B partial (1) ----------
Arath_CYP90A1_CPD (1) ----------MAFTAFLLLLSSIAAGFLLLL-----
Arath_CYP90C1_ROT3 (1) MQPPASAGLFRSPENLPWPYNYMDYLVAGFLVLTAGILLRPWLWLRLRNS
Arath_CYP90D1 (1) ----------MDTSSSLLFFSFFFFIIIVIFNKINGLRSSPA
Consenso (1) S L LLALLLI
Domínio de transição [Domínio globular →
51
100
Orysa_CYP90B (40) WRGGTTPAKRKR-------MNLPPGAA-GWPLVGETFGYLRAHPATSVGR
Arath_CYP90B1_DWF4 (29) --------KRRNRKTR---FNLPPGKS-GWPFLGETIGYLKPYTATTLGD
Sacof_CYP90B (40) W----RREKKKR-------PNLPPGAL-GWPFVGETFGYLRAHPATSVGL
Allce_CYP90B (27) --------RKLL-------PTLPPGNLGGWPFIGDTIPFMTPHSSALLGT
Zinel_CYP90B (32) --------KRRTNNGST--RNLPPGNM-GWPFIGETIGYLQPYSATTIGK
Medtr_CYP90B (25) --------KRKKTR-----YNLPPGKM-GWPFIGETIGYLKPYTATTMGE
Poptr_CYP90B (26) --------RKQVR------FNLPPGNM-GWPFLGETIGYLKPYSATSIGE
Aqufo_CYP90B partial (25) --------KRKKKS-----YNLPPGNM-GWPYLGETLGYLKPYCAITTGD
Triae_CYP90B partial (40) WSG---RTKKRR-------PNLPPGAV-GWPFIGETFGYLRAHPATSIGQ
Eupes_CYP90B partial (1) ----------
Goshi_CYP90B partial (25) --------KRKQRR-----YNLPPGNM-GWPFLGETIGYLRPYSATSVGE
Lyces_CYP90B partial (1) ----------M-GWPFLGETIGYLRPYSATTIGD
Arath_CYP90A1_CPD (22) --------RRTRYRR----MGLPPGSL-GLPLIGETFQLIGAYKTENPEP
Arath_CYP90C1_ROT3 (51) KTKDGDEEEDN----EEKKKGMIPNGSLGWPVIGETLNFIACGYSSRPVT
Arath_CYP90D1 (33) SKKKLNDHHVT----SQSHGPKFPHGSLGWPVIGETIEFVSSAYSDRPES
Consenso (51) KR NLPPG M GWPFIGETIGYLRPY ATSIG
Domínio globular
101
150
Orysa_CYP90B (82) FMEQHIARYGKIYRSSLFGERTVVSADAGLNRYILQNEGRLFECSYPRSI
Arath_CYP90B1_DWF4 (67) FMQQHVSKYGKIYRSNLFGEPTIVSADAGLNRFILQNEGRLFECSYPRSI
Sacof_CYP90B (78) FMEQHVARYGKIYRSSLFGERTVVSADAGLNRYILQNEGRLFECSYPRSI
Allce_CYP90B (62) YIDQNISKYGRIFRMNLLGKATIVSVDPDFNRYILQNEGRLFENSCPTSI
Zinel_CYP90B (71) FMEQHISKYGKIYKSSLFGEPTIVSADPGLNKYILQNEGRLFECSYPRSI
Medtr_CYP90B (61) FMENHIARYGTIYKSNLFGGPAIVSADAELNRFILQNDGKLFECSYPKSI
Poptr_CYP90B (61) FMEQHISRYGKIYKSNLFGEPTIVSADAGLSRFILQNEGRLFECSYPKSI
Aqufo_CYP90B partial (61) FMEQHISRYGKIYKSNLFGYPTIVSVDPELNRYVLQNEGRLFECSYPSSL
Triae_CYP90B partial (79) FMDQHIARYGKIYRSSLFGDRTVVSADAGLNRYILQNEGRLFECSYPRSI
Eupes_CYP90B partial (1) ----------
Goshi_CYP90B partial (61) FMHQHISRYGNIYKSNLFGEKTIVSADAGLNKFILQNEGRLFECSYPRSI
Lyces_CYP90B partial (24) FMQDHISRYGKIFKSNLFGEPTIVSADAGLNRYILQNEGRLFECNYPRSI
Arath_CYP90A1_CPD (59) FIDERVARYGSVFMTHLFGEPTIFSADPETNRFVLQNEGKLFECSYPASI
Arath_CYP90C1_ROT3 (97) FMDKRKSLYGKVFKTNIIGTPIIISTDAEVNKVVLQNHGNTFVPAYPKSI
Arath_CYP90D1 (79) FMDKRRLMYGRVFKSHIFGTATIVSTDAEVNRAVLQSDSTAFVPFYPKTV
Consenso (101) FMEQHISRYGKIYKSNLFGEPTIVSADAGLNRYILQNEGRLFECSYPRSI

FIGURA 13

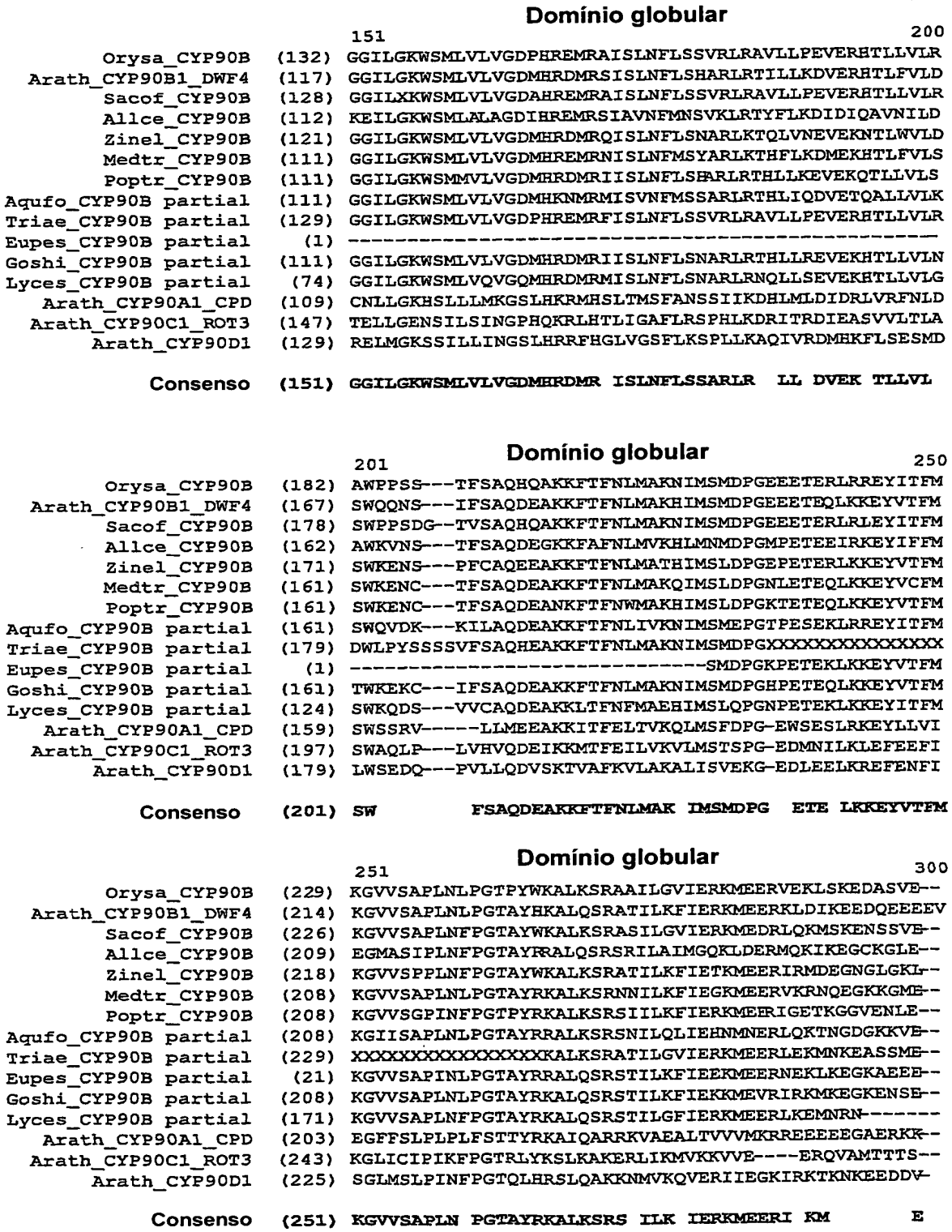


FIGURA 13 (Cont.)

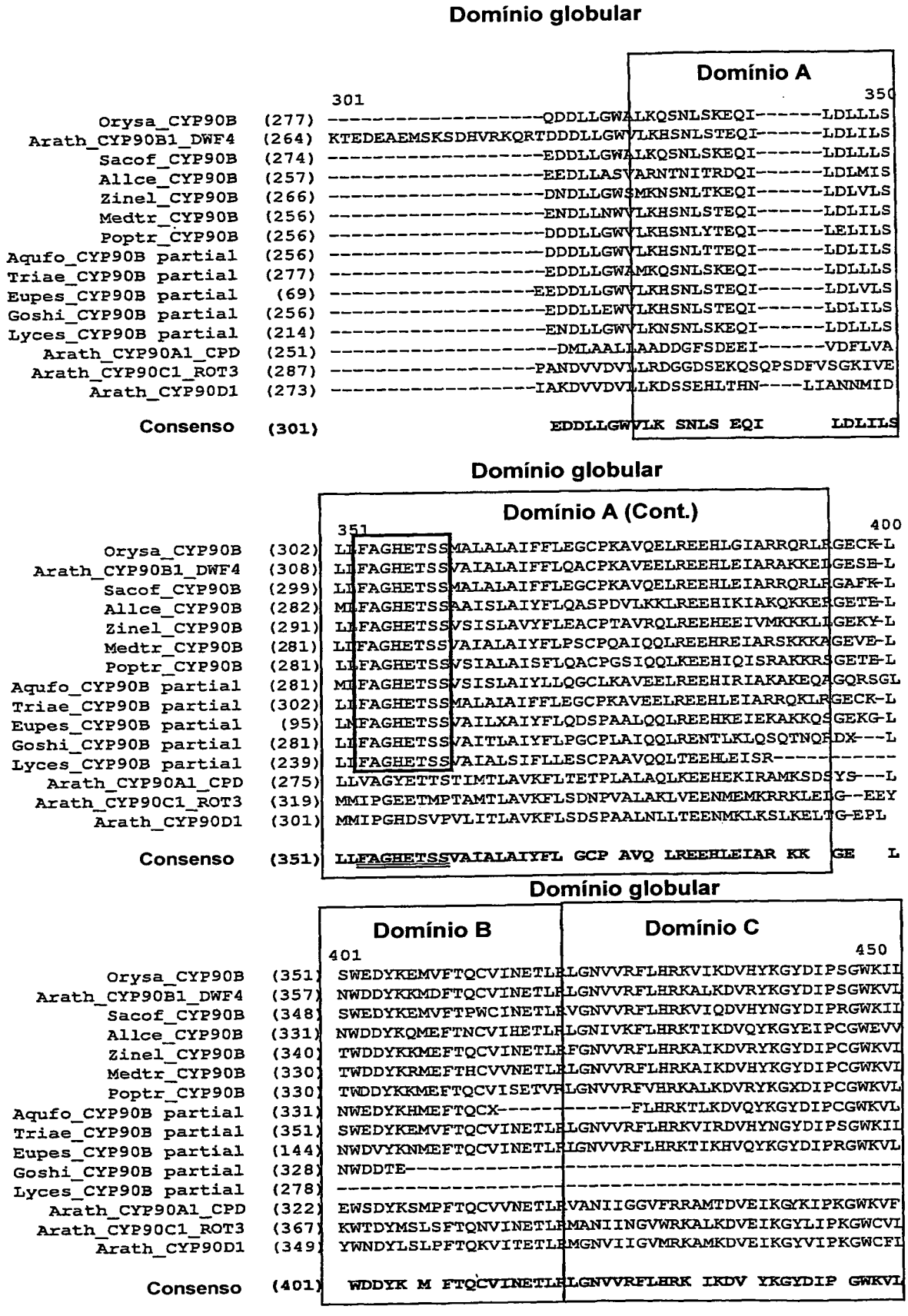


FIGURA 13 (Cont.)

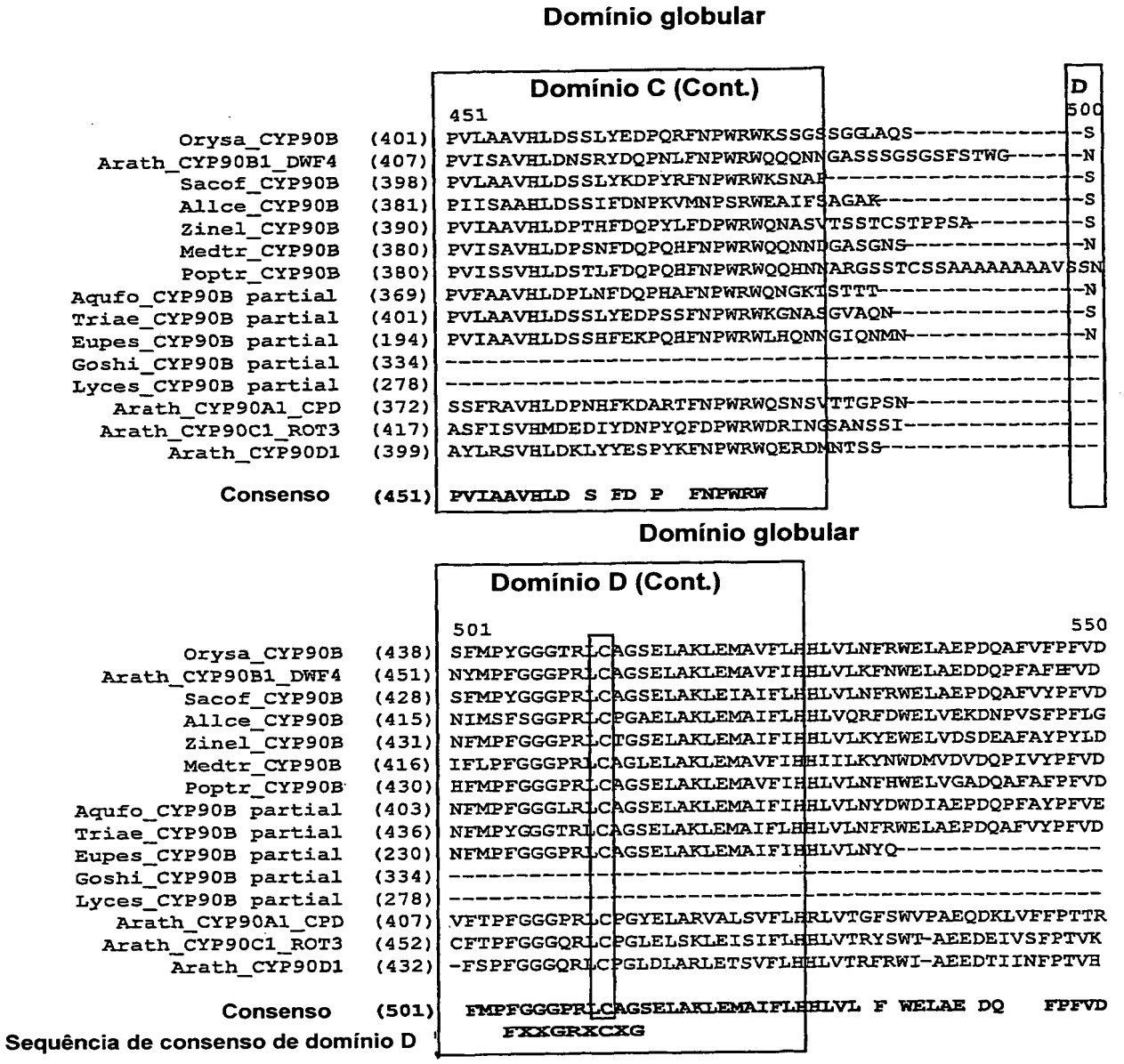


FIGURA 13 (Cont.)

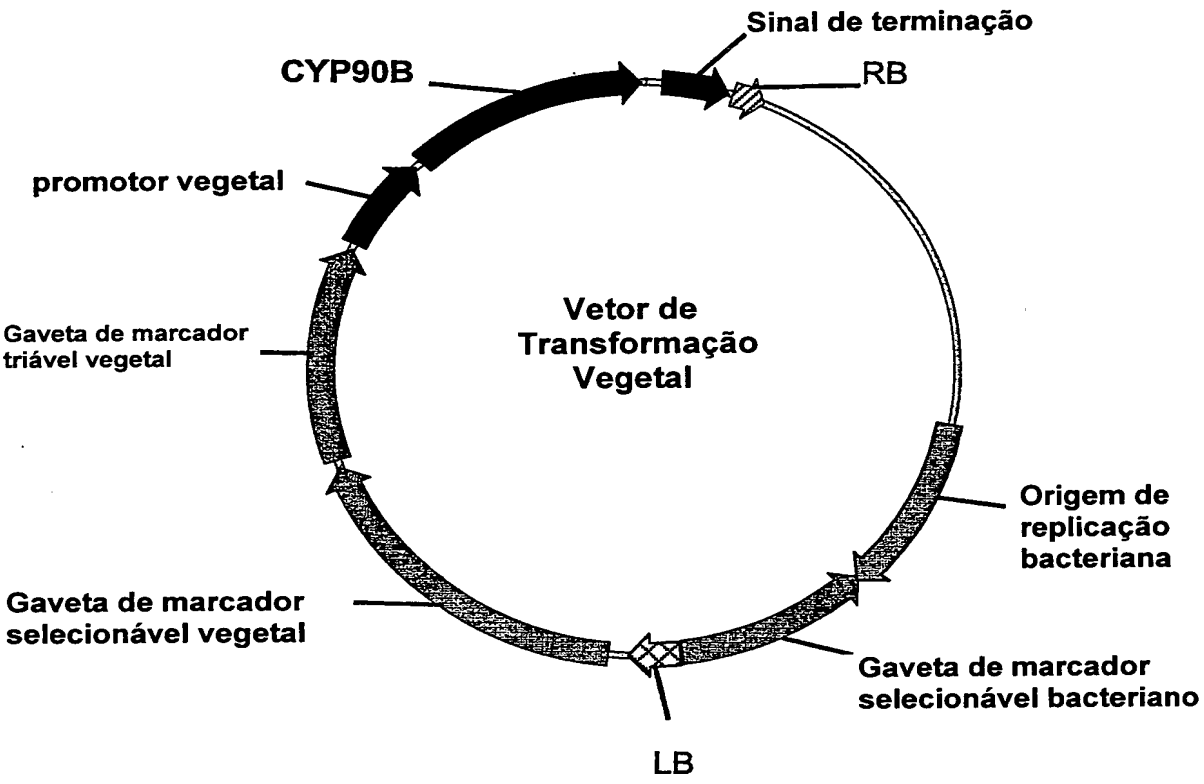

Sinal de terminação
CYP90B
RB
promotor vegetal
Vetor de
Transformação
Vegetal
Gaveta de marcador
triável vegetal
Origem de
replicação
bacteriana
Gaveta de marcador
selecionável vegetal
Gaveta de marcador
selecionável bacteriano
LB

FIGURA 14

SEQ ID NO: 77, seqüência de ácidos nucléicos CYP90B de Oryza sativa AB206579.1

ATGGCCGCCATGATGGCGTCCATAACCAGCGAGCTGCTCTTCTTTCTCCCCTTCATCCTCCT
TGCCCTGCTCACGTTCTACACCACCACCGTGGCCAAATGCCACGGCGGGCACTGGTGGCGAG
GTGGGACGACGCCGGCGAAGAGGAAGCGGATGAACCTGCCGCCCGGCGCCGCCGGGTGGCCG
CTCGTCGGCGAGACGTTCGGCTACCTCCGCGCCCACCCCGCCACCTCCGTCGGCCGCTTCAT
GGAGCAGCACATCGCACGGTACGGGAAGATATACCGGTCGAGCCTGTTCGGGGAGCGGACGG
TGGTGTCGGCGGACGCGGGGCTCAACCGGTACATCCTGCAGAACGAGGGGAGGCTGTTCGAG
TGCAGCTACCCGCGCAGCATCGGCGGCATCCTGGGCAAGTGGTCCATGCTGGTCCTCGTCGG
GGACCCGCACCGCGAGATGCGCGCCATCTCCCTCAACTTCCTCTCCTCCGTCCGCCTCCGCG
CCGTCCTCCTCCCCGAGGTCGAGCGCCACACCCTCCTCGTCCTCCGCGCCTGGCCCCCTTCC
TCCACCTTCTCCGCTCAGCACCAAGCCAAGAAGTTCACGTTCAACCTGATGGCGAAGAACAT
AATGAGCATGGACCCGGGGGAGGAAGAGACGGAGCGGCTGCGGCGGGAGTACATCACCTTCA
TGAAGGGCGTGGTCTCCGCGCCGCTCAACCTGCCCGGGACGCCCTACTGGAAGGCTCTCAAG
TCGCGTGCTGCCATTCTCGGAGTAATAGAGAGGAAAATGGAAGAGCGGGTTGAGAAGCTGAG
CAAGGAGGATGCAAGCGTAGAGCAAGACGATCTTCTCGGATGGGCTCTGAAACAATCTAACC
TTTCAAAAGAGCAAATCCTGGACCTCTTGCTGAGCTTGCTCTTCGCCGGGCACGAGACGTCG
TCCATGGCGCTCGCCCTCGCCATCTTCTTCCTTGAAGGCTGCCCCAAGGCTGTCCAAGAACT
GAGGGAGGAGCATCTTGGGATTGCAAGGAGACAAAGGCTAAGAGGGGAGTGCAAATTGAGCT
GGGAAGACTACAAAGAGATGGTTTTCACGCAATGTGTCATAAACGAGACGTTGCGGCTAGGA
AACGTGGTCAGGTTCCTGCACCGGAAGGTCATCAAGGACGTGCACTACAAGGGTTATGACAT
TCCAAGCGGATGGAAGATCCTGCCGGTGTTAGCCGCGGTGCATCTGGACTCGTCCCTGTACG
AGGACCCCCAGCGCTTCAATCCCTGGAGATGGAAGAGTAGCGGATCATCCGGCGGCTTGGCT
CAGAGCAGCAGCTTCATGCCGTACGGCGGCGGGACGCGGCTGTGCGCCGGGTCGGAGCTCGC
GAAGCTGGAGATGGCCGTGTTCTTGCACCACCTGGTGCTCAACTTCAGGTGGGAGCTCGCCG
AGCCGGACCAAGCCTTCGTCTTCCCCTTCGTCGACTTCCCCAAGGGCCTTCCCATTAGGGTT
CATAGAATTGCACAGGATGATGAGCAGGAGTAA

SEQ ID NO: 78, seqüência de aminoácidos CYP90B de Oryza sativa

MAAMMASITSELLFFLPFILLALLTFYTTTVAKCHGGHWWRGGTTPAKRKRMNLPPGAAGWP
LVGETFGYLRAHPATSVGRFMEQHIARYGKIYRSSLFGERTVVSADAGLNRYILQNEGRLFE
CSYPRSIGGILGKWSMLVLVGDPHREMRAISLNFLSSVRLRAVLLPEVERHTLLVLRAWPPS
STFSAQHQAKKFTFNLMAKNIMSMDPGEEETERLRREYITFMKGVVSAPLNLPGTPYWKALK
SRAAILGVIERKMEERVEKLSKEDASVEQDDLLGWALKQSNLSKEQILDLLLSLLFAGHETS
SMALALAIFFLEGCPKAVQELREEHLGIARRQRLRGECKLSWEDYKEMVFTQCVINETLRLG
NVVRFLHRKVIKDVHYKGYDIPSGWKILPVLAAVHLDSSLYEDPQRFNPWRWKSSGSSGGLA
QSSSFMPYGGGTRLCAGSELAKLEMAVFLHHLVLNFRWELAEPDQAFVFPFVDFPKGLPIRV
HRIAQDDEQ

SEQ ID NO: 79, seqüência de ácidos nucléicos CYP90B1 de Arabidopsis thaliana NM_114926.2

ATGTTCGAAACAGAGCATCATACTCTCTTACCTCTTCTTCTTCTCCCATCGCTTTTGTCTCT
TCTTCTCTTCTTGATTCTCTTGAAGAGAAGAAATAGAAAAACCAGATTCAATCTACCTCCGG
GTAAATCCGGTTGGCCATTTCTTGGTGAAACCATCGGTTATCTTAAACCGTACACCGCCACA
ACACTCGGTGACTTCATGCAACAACATGTCTCCAAGTATGGTAAGATATATAGATCGAACTT
GTTTGGAGAACCAACGATCGTATCAGCTGATGCTGGACTTAATAGATTCATATTACAAAACG
AAGGAAGGCTCTTTGAATGTAGTTATCCTAGAAGTATAGGTGGGATTCTTGGGAAATGGTCG
ATGCTTGTTCTTGTTGGTGACATGCATAGAGATATGAGAAGTATCTCGCTTAACTTCTTAAG

FIGURA 15

TCACGCACGTCTTAGAACTATTCTACTTAAAGATGTTGAGAGACATACTTTGTTTGTTCTTG
ATTCTTGGCAACAAAACTCTATTTTCTCTGCTCAAGACGAGGCCAAAAAGTTTACGTTTAAT
CTAATGGCGAAGCATATAATGAGTATGGATCCTGGAGAAGAAGAAACAGAGCAATTAAAGAA
AGAGTATGTAACTTTCATGAAAGGAGTTGTCTCTGCTCCTCTAAATCTACCAGGAACTGCTT
ATCATAAAGCTCTTCAGTCACGAGCAACGATATTGAAGTTCATTGAGAGGAAAATGGAAGAG
AGAAAATTGGATATCAAGGAAGAAGATCAAGAAGAAGAAGAAGTGAAAACAGAGGATGAAGC
AGAGATGAGTAAGAGTGATCATGTTAGGAAACAAAGAACAGACGATGATCTTTTGGGATGGG
TTTTGAAACATTCGAATTTATCGACGGAGCAAATTCTCGATCTCATTCTTAGTTTGTTATTT
GCCGGACATGAGACTTCTTCTGTAGCCATTGCTCTCGCTATCTTCTTCTTGCAAGCTTGCCC
TAAAGCCGTTGAAGAGCTTAGGGAAGAGCATCTTGAGATCGCGAGGGCCAAGAAGGAACTAG
GAGAGTCAGAATTAAATTGGGATGATTACAAGAAAATGGACTTTACTCAATGTGTTATAAAT
GAAACTCTTCGATTGGGAAATGTAGTTAGGTTTTTGCATCGCAAAGCACTCAAAGATGTTCG
GTACAAAGGATACGATATCCCTAGTGGGTGGAAAGTGTTACCGGTGATCTCAGCCGTACATT
TGGATAATTCTCGTTATGACCAACCTAATCTCTTTAATCCTTGGAGATGGCAACAGCAAAAC
AACGGAGCGTCATCGTCAGGAAGTGGTAGTTTTTCGACGTGGGGAAACAACTACATGCCGTT
TGGAGGAGGGCCAAGGCTATGTGCTGGTTCAGAGCTAGCCAAGTTAGAAATGGCAGTGTTTA
TTCATCATCTAGTTCTTAAATTCAATTGGGAATTAGCAGAAGATGATCAACCATTTGCTTTT
CCTTTTGTTGATTTTCCTAACGGTTTGCCTATTAGGGTTTCTCGTATTCTGTAA

SEQ ID NO: 80, seqüência de aminoácidos CYP90B1 de Arabidopsis thaliana

MFETEHHTLLPLLLLPSLLSLLLFLILLKRRNRKTRFNLPPGKSGWPFLGETIGYLKPYTAT
TLGDFMQQHVSKYGKIYRSNLFGEPTIVSADAGLNRFILQNEGRLFECSYPRSIGGILGKWS
MLVLVGDMHRDMRSISLNFLSHARLRTILLKDVERHTLFVLDSWQQNSIFSAQDEAKKFTFN
LMAKHIMSMDPGEEETEQLKKEYVTFMKGVVSAPLNLPGTAYHKALQSRATILKFIERKMEE
RKLDIKEEDQEEEEVKTEDEAEMSKSDHVRKQRTDDDLLGWVLKHSNLSTEQILDLILSLLF
AGHETSSVAIALAIFFLQACPKAVEELREEHLEIARAKKELGESELNWDDYKKMDFTQCVIN
ETLRLGNVVRFLHRKALKDVRYKGYDIPSGWKVLPVISAVHLDNSRYDQPNLFNPWRWQQQN
NGASSSGSGSFSTWGNNYMPFGGGPRLCAGSELAKLEMAVFIHHLVLKFNWELAEDDQPFAF
PFVDFPNGLPIRVSRIL

SEQ ID NO: 81, seqüência de ácidos nucléicos CYP90B de Saccharum officinarum ESTs compiladas CA092707.1 CF574030.1 CA217329.1

ATGGGCGCCATGATGGCCTCCATAACCAGCGAGCTCCTCTTCTTCCTTCCCTTCATCCTGCT
GGCCCTCCTCGCCCTGTACACCACCACTGTCGCCAAATGCGACGGCACCCACCAGTGGCGCC
GGCCGAAGAAGAAGCGGCCGAACCTGCCCCCGGGCGCCCTCGGATGGCCTTTCGTCGGCGAG
ACCTTCGGCTACCTCCGCGCCCACCCGGCCACCTCCGTGGGCCTCTTCATGGAGCAGCATGT
CGCACGGTACGGCAAGATATACCGGTCGAGCCTGTTCGGGGAGCGGACGGTGGTGTCGGCGG
ACGCGGGGCTCAACCGCTACATCCTGCAGAACGAGGGGCGGCTGTTCGAGTGCAGCTACCCG
CGCAGCATCGGCGGCATCCTGNGCAAGTGGTCCATGCTGGTGCTGGTGGGCGACGCGCACCG
CGAGATGCGCGCCATCTCGCTCAACTTTCTCAGCTCCGTCCGCCTCCGCGCCGTGCTGCTCC
CGGAGGTGGAGCGCCACACCCTGCTGGTGCTCCGCTCATGGCCGCCCTCCGACGGCACGGTC
TCCGCGCAGCACCAAGCCAAGAAGTTCACGTTTAACCTGATGGCGAAGAACATAATGAGCAT
GGACCCCGGCGAGGAGGAGACGGAGCGGCTGCGGCTGGAGTACATCACCTTCATGAAGGGCG
TCGTGTCAGCGCCGCTCAACTTCCCGGGCACGGCCTACTGGAAGGCGCTCAAGTCACGCGCG
TCCATACTTGGAGTAATAGAGAGGAAGATGGAGGACAGGCTTCAGAAAATGAGTAAGGAGAA
CTCAAGTGTGGAGGAAGACGATCTTCTTGGATGGGCCCTGAAGCAGTCCAATCTGTCAAAGG
AACAGATCCTGGACCTCTTGCTGAGCCTGCTCTTCGCCGGGCACGAGACTTCGTCAATGGCG

FIGURA 15 (Cont.)

CTAGCCCTCGCCATCTTCTTCCTTGAAGGATGCCCCAAGGCTGTTCAAGAACTCCGGGAGGA
GCATCTCGAGATTGCTAGGAGACAAAGGCTAAGAGGGGCGTTCAAATTGAGCTGGGAAGACT
ACAAGGAAATGGTTTTCACGCCATGGTGTATAAACGAGACATTGCGGGTTGGCAACGTGGTC
AGGTTCCTGCACCGGAAGGTCATCCAAGATGTGCACTACAATGGGTACGACATACCACGCGG
GTGGAAAATCCTGCCGGTGTTAGCGGCGGTGCATCTGGATTCGTCGCTGTACAAGGACCCCT
ACCGGTTCAACCCTTGGAGATGGAAGAGCAACGCGCCGAGCAGCTTCATGCCGTACGGCGGC
GGGCCGCGGCTGTGCGCCGGGTCGGAGCTGGCCAAGCTGGAGATCGCCATCTTCCTGCACCA
CCTGGTGCTCAACTTCCGGTGGGAGCTGGCGGAGCCGGACCAAGCCTTCGTCTACCCCTTCG
TCGACTTCCCCAAGGGCCTCCCGATCAGGGTCCAGCGGATCGCCGACGACCAGGGGGCATCG
CAGCGTTTTGACCGAGAAGCACGATGTAGCGGTACGGGTACACAAGCAAGTACAAAATTTTC
GTTGCATTTTGAGGGAATTGGGGCTGGGTGGTTATGGTAA

SEQ ID NO: 82, seqüência de aminoácidos CYP90B de Saccharum officinarum

MGAMMASITSELLFFLPFILLALLALYTTTVAKCDGTHQWRRPKKKRPNLPPGALGWPFVGE
TFGYLRAHPATSVGLFMEQHVARYGKIYRSSLFGERTVVSADAGLNRYILQNEGRLFECSYP
RSIGGILXKWSMLVLVGDAHREMRAISLNFLSSVRLRAVLLPEVERHTLLVLRSWPPSDGTV
SAQHQAKKFTFNLMAKNIMSMDPGEEETERLRLEYITFMKGVVSAPLNFPGTAYWKALKSRA
SILGVIERKMEDRLQKMSKENSSVEEDDLLGWALKQSNLSKEQILDLLLSLLFAGHETSSMA
LALAIFFLEGCPKAVQELREEHLEIARRQRLRGAFKLSWEDYKEMVFTPWCINETLRVGNVV
RFLHRKVIQDVHYNGYDIPRGWKILPVLAAVHLDSSLYKDPYRFNPWRWKSNAPSSFMPYGG
GPRLCAGSELAKLEIAIFLHHLVLNFRWELAEPDQAFVYPFVDFPKGLPIRVQRIADDQGAS
QRFDREARCSGTGTQASTKFSLHFEGIGAGWLW

SEQ ID NO: 83, seqüência de ácidos nucléicos CYP90B de Allium cepa

ATGGAGATCATATTAGTGTCTACATTGATCATATCGCTACTAATATTTCTTGGATTTAGAAG
CAATGGGAAAACGGAGAGAAAATTGCTGCCTACACTACCACCAGGCAATCTTGGAGGTTGGC
CATTCATCGGTGACACCATTCCGTTCATGACACCTCATTCTTCTGCTTTGTTGGGCACTTAC
ATCGATCAAAATATTTCCAAATATGGGAGGATATTTCGAATGAACTTGTTAGGAAAGGCAAC
GATCGTGTCTGTAGACCCTGATTTCAACAGATATATTCTACAGAATGAAGGAAGATTGTTTG
AAAATAGCTGCCCAACGAGCATTAAAGAGATTTTGGGAAAATGGTCTATGCTTGCATTAGCT
GGGGATATACACAGAGAAATGAGATCCATTGCTGTGAATTTCATGAACAGTGTTAAGCTTAG
AACTTATTTTTTAAAGGATATTGATATTCAGGCTGTTAATATTCTTGATGCTTGGAAGGTCA
ACTCTACTTTCTCTGCACAGGATGAAGGAAAGAAGTTTGCATTTAACCTCATGGTGAAGCAT
CTAATGAATATGGATCCTGGAATGCCAGAGACAGAAGAAATCAGAAAAGAGTACATTTTCTT
CATGGAGGGGATGGCTTCCATTCCTTTAAACTTTCCTGGAACAGCCTACAGAAGAGCTTTAC
AGTCAAGGTCCAGGATTCTGGCAATAATGGGGCAAAAGCTTGACGAAAGGATGCAGAAAATA
AAAGAAGGCTGTAAAGGACTGGAAGAAGAGGATCTTCTTGCCTCAGTTGCAAGAAATACTAA
CATAACAAGAGACCAGATTCTTGATTTGATGATCAGCATGCTTTTTGCTGGTCATGAAACTT
CTTCTGCCGCTATTTCTCTTGCCATTTATTTCCTGCAAGCTTCCCCCGATGTTCTTAAAAAG
CTTCGAGAGGAGCACATAAAGATTGCAAAACAAAAGAAAGAAAGGGGCGAAACTGAATTGAA
CTGGGATGATTACAAGCAAATGGAATTCACAAACTGCGTTATTCATGAAACCCTAAGATTAG
GCAACATCGTTAAGTTTTTGCATCGGAAAACCATCAAAGATGTTCAATACAAAGGTTATGAA
ATTCCATGTGGGTGGGAAGTAGTGCCAATCATCTCAGCAGCACATTTGGATTCTTCTATCTT
TGACAACCCAAAAGTTATGAATCCTTCGAGGTGGGAGGCGATATTTTCAGCAGGAGCAAAGA
GCAATATAATGTCATTCAGCGGTGGACCTCGGTTATGTCCAGGAGCAGAGTTGGCGAAACTG
GAGATGGCTATTTTTCTTCATCATCTTGTACAGAGGTTCGACTGGGAATTGGTGGAGAAGGA
TAACCCTGTATCATTCCCCTTCCTTGGATTTCCCAAGAAATTGCCTATCAAAATCACAGCTC
TTAAACATTGA

FIGURA 15 (Cont.)

SEQ ID NO: 84, seqüência de ácidos nucléicos CYP90B de Allium cepa TC2113

MEIILVSTLIISLLIFLGFRSNGKTERKLLPTLPPGNLGGWPFIGDTIPFMTPHSSALLGTY
IDQNISKYGRIFRMNLLGKATIVSVDPDFNRYILQNEGRLFENSCPTSIKEILGKWSMLALA
GDIHREMRSIAVNFMNSVKLRTYFLKDIDIQAVNILDAWKVNSTFSAQDEGKKFAFNLMVKH
LMNMDPGMPETEEIRKEYIFFMEGMASIPLNFPGTAYRRALQSRSRILAIMGQKLDERMQKI
KEGCKGLEEEDLLASVARNTNITRDQILDLMISMLFAGHETSSAAISLAIYFLQASPDVLKK
LREEHIKIAKQKKERGETELNWDDYKQMEFTNCVIHETLRLGNIVKFLHRKTIKDVQYKGYE
IPCGWEVVPIISAAHLDSSIFDNPKVMNPSRWEAIFSAGAKSNIMSFSGGPRLCPGAELAKL
EMAIFLHHLVQRFDWELVEKDNPVSFPFLGFPKKLPIKITALKH

SEQ ID NO: 85, seqüência de ácidos nucléicos CYP90B de Zinnia elegans AB231155

ATGTGTTCAACAACCCTAAATATGTGTGACCTTGAGTTCTTCATCCTTGCATCTTGTCTTGT
CTTGGCTCTTTTTCTCATCTTGAAGCTTGTCAAAAGAAGAACAAACAATGGTTCGACTCGAA
ATCTTCCACCGGGCAACATGGGCTGGCCGTTCATCGGTGAAACCATCGGTTACCTCCAACCG
TATTCGGCTACAACCATCGGCAAGTTTATGGAACAACACATATCCAAGTATGGGAAGATATA
CAAATCTAGTTTGTTTGGTGAGCCAACAATAGTTTCTGCTGATCCAGGGTTGAATAAGTACA
TATTGCAAAATGAAGGGAGGTTATTTGAATGTAGTTATCCAAGAAGCATAGGGGGTATTCTT
GGCAAATGGTCCATGTTGGTTTTGGTTGGTGACATGCATAGAGACATGAGGCAAATTTCACT
CAACTTTTTGTCCAATGCAAGGCTTAAAACTCAACTAGTAAATGAAGTTGAGAAAAATACTT
TGTGGGTATTAGATTCTTGGAAAGAAAACTCACCTTTTTGTGCCCAAGAAGAAGCCAAGAAG
TTTACTTTTAATCTAATGGCAACACATATCATGAGTTTAGACCCGGGTGAACCGGAGACCGA
GCGATTGAAGAAAGAGTATGTAACTTTCATGAAAGGTGTGGTTTCTCCCCCTTTAAACTTCC
CTGGAACTGCATACTGGAAAGCTTTAAAGTCTCGAGCGACGATTCTTAAATTCATCGAAACA
AAAATGGAGGAGCGGATTAGGATGGACGAAGGAAACGGATTAGGGAAACTAGACAATGATCT
TCTTGGATGGTCTATGAAGAACTCAAATCTCACTAAAGAGCAAATACTCGATTTGGTATTGA
GTCTACTCTTTGCTGGTCACGAAACGTCTTCGGTTTCGATATCGTTAGCCGTTTACTTCCTT
GAAGCTTGTCCTACCGCGGTTCGTCAACTTAGAGAAGAACATGAAGAAATTGTGATGAAAAA
AAAGCTATTGGGTGAGAAGTATCTCACTTGGGATGACTACAAAAAGATGGAGTTTACTCAGT
GTGTGATCAATGAGACGCTAAGATTCGGGAATGTGGTGAGATTCCTCCACAGAAAGGCTATT
AAAGATGTGAGGTATAAAGGATATGACATTCCATGTGGTTGGAAAGTGCTGCCAGTGATTGC
AGCCGTGCATTTGGATCCTACACATTTTGACCAACCTTACCTTTTTGATCCATGGAGATGGC
AGAACGCAAGTGTCACGTCATCTACTTGTTCAACCCCGCCATCAGCAAGTAACTTCATGCCA
TTTGGTGGAGGGCCCCGCTTATGCACAGGGTCAGAGCTAGCGAAACTAGAGATGGCGATATT
TATCCACCATTTGGTCCTAAAATACGAGTGGGAATTGGTTGACTCAGATGAAGCATTCGCTT
ATCCATATCTCGACTTTCCAAAAGGTCTGCCAATCAAAATCCGTCACCGAAAACAATCATGT
TAG

SEQ ID NO: 86, seqüência de aminoácidos CYP90B de Zinnia elegans

MCSTTLNMCDLEFFILASCLVLALFLILKLVKRRTNNGSTRNLPPGNMGWPFIGETIGYLQP
YSATTIGKFMEQHISKYGKIYKSSLFGEPTIVSADPGLNKYILQNEGRLFECSYPRSIGGIL
GKWSMLVLVGDMHRDMRQISLNFLSNARLKTQLVNEVEKNTLWVLDSWKENSPFCAQEEAKK
FTFNLMATHIMSLDPGEPETERLKKEYVTFMKGVVSPPLNFPGTAYWKALKSRATILKFIET
KMEERIRMDEGNGLGKLDNDLLGWSMKNSNLTKEQILDLVLSLLFAGHETSSVSISLAVYFL
EACPTAVRQLREEHEEIVMKKKLLGEKYLTWDDYKKMEFTQCVINETLRFGNVVRFLHRKAI
KDVRYKGYDIPCGWKVLPVIAAVHLDPTHFDQPYLFDPWRWQNASVTSSTCSTPPSASNFMP
FGGGPRLCTGSELAKLEMAIFIHHLVLKYEWELVDSDEAFAYPYLDFPKGLPIKIRHRKQSC

FIGURA 15 (Cont.)

SEQ ID NO: 87, Seqüência de ácidos nucléicos CYP90B de Medicago trunculata AC147964.10

ATGTCTAACTCATACTTAACTTGCAGTTTTCTTTCTTCCATCTTTGTTCTTTCTTTGATTTT
CATTTTCATCAAAAGAAAGAAAACAAGGTATAATCTTCCACCTGGAAAAATGGGATGGCCCT
TTATAGGAGAAACCATTGGTTATTTGAAGCCTTACACTGCCACCACAATGGGAGAATTTATG
GAAAATCACATAGCAAGGTATGGGACAATTTACAAGTCAAATTTGTTTGGAGGGCCAGCTAT
TGTATCAGCAGATGCAGAATTGAATAGGTTCATATTACAAAATGATGGAAAATTGTTTGAGT
GTAGCTATCCAAAAAGCATTGGTGGAATACTTGGAAAATGGTCAATGTTGGTTTTAGTAGGT
GACATGCATAGGGAAATGAGGAATATATCACTAAACTTTATGAGCTATGCTAGGCTTAAAAC
ACATTTTTTGAAAGATATGGAGAAGCATACCCTTTTTGTTCTAAGCTCTTGGAAAGAAAATT
GTACATTTTCAGCTCAAGATGAAGCAAAAAAGTTCACCTTCAATTTGATGGCCAAACAAATC
ATGAGTTTGGATCCAGGGAATCTTGAGACAGAACAGTTGAAAAAAGAGTATGTCTGTTTCAT
GAAAGGTGTTGTTTCTGCTCCTTTGAATTTGCCAGGAACTGCATACAGAAAAGCATTAAAGT
CTAGGAACAATATATTGAAGTTCATAGAGGGGAAAATGGAAGAAAGGGTGAAGAGAAACCAA
GAAGGAAAAAAAGGGATGGAGGAAAATGATCTTCTAAATTGGGTTTTAAAGCATTCAAATCT
TTCCACTGAGCAAATTCTTGACTTGATTCTAAGTTTACTTTTTGCTGGCCATGAAACTTCAT
CTGTGGCTATAGCTCTAGCTATTTACTTTTTGCCTAGTTGTCCTCAAGCTATACAACAATTA
AGGGAAGAGCATAGAGAAATAGCTAGATCCAAGAAGAAAGCAGGGGAGGTTGAATTAACTTG
GGATGATTATAAAAGAATGGAATTTACTCATTGTGTTGTGAATGAAACACTAAGGTTGGGTA
ATGTTGTGAGATTCCTTCACAGGAAGGCTATCAAAGATGTTCATTACAAAGGTTATGACATT
CCATGTGGATGGAAAGTCCTTCCGGTGATTTCAGCGGTACATTTGGATCCTTCAAATTTTGA
CCAACCTCAACACTTCAATCCTTGGAGATGGCAGATGAGCAATAACAACTTCATGCCATTTG
GAGGAGGACCAAGGCTATGTGCAGGATTAGAATTAGCCAAACTTGAAATGGCTGTTTTCATT
CACCATATCATCCTCAAATACAACTGGGACATGGTCGATGTTGATCAACCTATTGTATACCC
TTTTGTTGATTTTCCCAAAGGTTTGCCAATTAGAGTCCAAAGCCAAGCCACTCTTTAA

SEQ ID NO: 88, seqüência de aminoácidos CYP90B de Medicago trunculata

MSNSYLTCSFLSSIFVLSLIFIFIKRKKTRYNLPPGKMGWPFIGETIGYLKPYTATTMGEFM
ENHIARYGTIYKSNLFGGPAIVSADAELNRFILQNDGKLFECSYPKSIGGILGKWSMLVLVG
DMHREMRNISLNFMSYARLKTHFLKDMEKHTLFVLSSWKENCTFSAQDEAKKFTFNLMAKQI
MSLDPGNLETEQLKKEYVCFMKGVVSAPLNLPGTAYRKALKSRNNILKFIEGKMEERVKRNQ
EGKKGMEENDLLNWVLKHSNLSTEQILDLILSLLFAGHETSSVAIALAIYFLPSCPQAIQQL
REEHREIARSKKKAGEVELTWDDYKRMEFTHCVVNETLRLGNVVRFLHRKAIKDVHYKGYDI
PCGWKVLPVISAVHLDPSNFDQPQHFNPWRWQQNNDGASGNSNIFLPFGGGPRLCAGLELAK
LEMAVFIHHIILKYNWDMVDVDQPIVYPFVDFPKGLPIRVQSQATL

SEQ ID NO: 89, seqüência de ácidos nucléicos CYP90B de Populus trichocarpa ESTs compiladas CK090847.1 CV280598.1 DT503533.1

ATGTCTCACTCAGAGCTTGTTGTCTTTCTCCTTCCATCGATTTTATCACTACTCTTGCTCTT
CATTCTCGTGCAAAGAAAGCAAGTAAGATTTAATCTCCCACCGGGCAACATGGGGTGGCCAT
TTCTTGGAGAAACCATTGGCTACCTGAAGCCTTACTCTGCTACTTCAATAGGAGAATTCATG
GAACAGCACATATCAAGGTATGGAAAGATTTACAAGTCCAATTTGTTTGGGGAGCCAACAAT
AGTATCCGCAGATGCTGGACTTAGCAGATTTATACTACAGAATGAGGGAAGATTATTTGAAT
GCAGCTATCCAAAAAGTATTGGTGGAATTCTTGGAAAATGGTCCATGATGGTTCTTGTTGGA
GACATGCATAGAGACATGAGGATTATATCTCTCAACTTTTTGAGCCATGCCAGGTTAAGAAC
TCATCTATTGAAAGAAGTGGAGAAGCAAACCCTGCTTGTTCTTAGCTCTTGGAAGGAGAATT
GTACATTTTCAGCTCAAGATGAAGCAAACAAGTTTACCTTCAATTGGATGGCAAAACATATC

FIGURA 15 (Cont.)

ATGAGCTTGGATCCTGGAAAGACAGAGACTGAGCAGCTGAAAAAAGAGTATGTTACTTTCAT
GAAAGGAGTAGTTTCAGGTCCTATAAATTTTCCTGGAACCCCATATAGAAAAGCCTTGAAGT
CTCGATCAATCATCTTGAAATTTATAGAGCGAAAGATGGAGGAGAGAATTGGAGAAACGAAG
GGTGGAGTAGAAAACTTGGAAGACGATGATCTTCTTGGATGGGTCTTGAAGCATTCAAATCT
TTATACGGAGCAAATCCTTGAATTAATCTTAAGCTTGCTCTTTGCTGGCCACGAAACTTCTT
CTGTGTCCATAGCTCTAGCCATATCCTTCTTGCAAGCTTGTCCTGGTTCTATTCAACAGTTA
AAAGAAGAACATATTCAAATCTCCAGAGCCAAGAAACGGTCAGGAGAGACGGAATTGACCTG
GGATGATTACAAAAAAATGGAATTCACTCAATGCGTTATAAGCGAGACAGTGAGGCTTGGAA
ACGTAGTCAGGTTTGTTCACAGAAAAGCTCTAAAAGATGTTCGGTACAAAGGGTANGACATT
CCATGTGGATGGAAAGTGTTACCAGTAATCTCATCCGTCCATTTAGATTCAACTCTTTTCGA
CCAACCTCAACACTTCAATCCATGGAGATGGCAGCAGCACAACAATGCTCGTGGATCTTCTA
CTTGTTCGAGTGCGGCGGCGGCGGCGGCGGCGGCGGTGAGTAGTAATCACTTCATGCCATTT
GGGGGAGGACCGCGACTCTGTGCAGGATCGGAATTGGCAAAACTTGAAATGGCAGTTTTCAT
TCACCATTTGGTTCTGAACTTCCATTGGGAATTGGTCGGTGCCGATCAAGCCTTTGCCTTTC
CTTTTGTTGATTTTCCTAAAGGCTTGCCAATAAGAGTCAAGCACCACACAGTCATATAA

SEQ ID NO: 90, seqüência de aminoácidos CYP90B de Populus trichocarpa

MSHSELVVFLLPSILSLLLLFILVQRKQVRFNLPPGNMGWPFLGETIGYLKPYSATSIGEFM
EQHISRYGKIYKSNLFGEPTIVSADAGLSRFILQNEGRLFECSYPKSIGGILGKWSMMVLVG
DMHRDMRIISLNFLSHARLRTHLLKEVEKQTLLVLSSWKENCTFSAQDEANKFTFNWMAKHI
MSLDPGKTETEQLKKEYVTFMKGVVSGPINFPGTPYRKALKSRSIILKFIERKMEERIGETK
GGVENLEDDDLLGWVLKHSNLYTEQILELILSLLFAGHETSSVSIALAISFLQACPGSIQQL
KEEHIQISRAKKRSGETELTWDDYKKMEFTQCVISETVRLGNVVRFVHRKALKDVRYKGXDI
PCGWKVLPVISSVHLDSTLFDQPQHFNPWRWQQHNNARGSSTCSSAAAAAAAAVSSNHFMPF
GGGPRLCAGSELAKLEMAVFIHHLVLNFHWELVGADQAFAFPFVDFPKGLPIRVKHHTVI

SEQ ID NO: 91, seqüência de ácidos nucléicos (parcial) CYP90B de Aquilegia formosa x Aquilegia pubescens ESTs compiladas DR940523.1 DR940522.1

ATGCCTGAGCTTGTGTTTTTCTTTTCATTAGCTCCAGCAATTTTAGCACTAATACTTCTCTT
GAAACTCTTCAAAAGGAAGAAAAAGTCATATAATCTTCCACCAGGAAACATGGGTTGGCCAT
ATCTAGGCGAAACTCTTGGTTATTTGAAGCCTTATTGTGCTATCACTACTGGAGATTTCATG
GAGCAACATATATCAAGGTATGGGAAGATCTACAAGTCAAATTTATTTGGTTATCCTACAAT
AGTTTCAGTTGATCCTGAATTAAATCGATATGTATTACAAAACGAAGGAAGACTGTTTGAAT
GTAGTTATCCAAGTAGTTTAGGTGGGATTCTTGGCAAATGGTCAATGTTAGTTTTGGTTGGA
GACATGCATAAAAACATGAGGATGATCTCTGTCAACTTCATGAGCAGCGCAAGACTTCGAAC
ACATCTAATTCAAGATGTGGAGACTCAAGCCTTATTGGTGCTGAAATCTTGGCAGGTTGATA
AAAAGATTTTGGCCCAAGATGAAGCAAAGAAGTTTACCTTCAATTTAATTGTAAAAAATATA
ATGAGCATGGAACCTGGAACACCTGAAAGTGAGAAGCTTAGGAGGGAATACATTACATTCAT
GAAAGGAATCATTTCTGCGCCTTTGAATTTGCCTGGAACTGCATATAGAAGAGCCTTAAAGT
CTCGATCGAACATTCTGCAACTTATCGAGCATAACATGAATGAGAGACTCCAAAAGACTAAC
GGAGATGGTAAGAAAGTGGAAGATGATGATCTACTTGGATGGGTCTTGAAGCATTCAAATCT
TACCACTGAACAAATTCTTGATTTGATACTAAGTATGCTTTTCGCGGGTCATGAAACTTCCT
CAGTATCTATATCTCTAGCCATATACCTTTTGCAAGGATGCCTAAAAGCAGTTGAAGAGTTA
AGGGAAGAGCATATTAGAATTGCTAAAGCAAAAGAACAGGCAGGACAGAGATCTGGATTAAA
TTGGGAAGATTACAAACACATGGAATTCACTCAATGTGNNTTTCTTCATAGAAAAACTCTCA
AAGATGTTCAGTACAAAGGGTATGACATTCCATGTGGTTGGAAAGTTCTTCCAGTATTTGCA

FIGURA 15 (Cont.)

GCAGTTCATTTGGACCCTTTAAATTTCGACCAACCTCATGCTTTCAATCCATGGAGATGGCA
GAATGGGAAGACAAGCACGACGACTAACAACTTCATGCCGTTTGGTGGTGGATTACGGTTAT
GTGCTGGTTCAGAGCTAGCCAAGTTGGAAATGGCTATTTTCATTCACCATTTGGTCTTGAAT
TATGATTGGGATATAGCAGAACCAGATCAACCATTTGCCTACCCATTTGTTGAATTTCCAAA
AGGTCTACCAATTAAGGTCTATGACCATCAGTGCTTAACATGA

SEQ ID NO: 92, seqüência de aminoácidos (parcial) CYP90B de Aquilegia formosa x Aquilegia pubescens

MPELVFFFSLAPAILALILLLKLFKRKKKSYNLPPGNMGWPYLGETLGYLKPYCAITTGDFM
EQHISRYGKIYKSNLFGYPTIVSVDPELNRYVLQNEGRLFECSYPSSLGGILGKWSMLVLVG
DMHKNMRMISVNFMSSARLRTHLIQDVETQALLVLKSWQVDKKILAQDEAKKFTFNLIVKNI
MSMEPGTPESEKLRREYITFMKGIISAPLNLPGTAYRRALKSRSNILQLIEHNMNERLQKTN
GDGKKVEDDDLLGWVLKHSNLTTEQILDLILSMLFAGHETSSVSISLAIYLLQGCLKAVEEL
REEHIRIAKAKEQAGQRSGLNWEDYKHMEFTQCXFLHRKTLKDVQYKGYDIPCGWKVLPVFA
AVHLDPLNFDQPHAFNPWRWQNGKTSTTTNNFMPFGGGLRLCAGSELAKLEMAIFIHHLVLN
YDWDIAEPDQPFAYPFVEFPKGLPIKVYDHQCLT

SEQ ID NO: 93, seqüência de ácidos nucléicos (parcial) CYP90B de extremidade 5' de Triticum aestivum BQ620306.1

ATGGCCGCCATCATGGCCTCCATAACCAGCGAGCTCCTTTTCTTCCTCCCCTTCATCCTCCT
GGCCCTGCTCACCTTCTACACCAGCGCCGTGGCTAAATGCCATGGCCTCCACTGGTGGAGCG
GCCGGACGAAGAAGAGGCGGCCGAACCTGCCGCCCGGCGCCGTCGGCTGGCCCTTCATCGGC
GAGACCTTCGGGTACCTCCGCGCCCACCCGGCCACCTCCATCGGCCAGTTCATGGACCAGCA
CATCGCACGGTACGGGAAGATATACCGGTCGAGCCTGTTCGGGGACCGGACGGTGGTGTCGG
CGGACGCGGGGCTGAACCGGTACATCCTGCAGAACGAGGGGCGGCTGTTCGAGTGTAGCTAC
CCGCGGAGCATCGGCGGCATCCTTGGCAAATGGTCGATGCTGGTGCTCGTCGGCGACCCCCA
CCGCGAGATGCGCTTCATCTCCCTCAACTTCCTCAGCTCCGTCCGCCTCCGCGCCGTGCTCC
TCCCGGAGGTGGAGCGCCACACCCTCCTCGTCCTCCGCGACTGGCTGCCTTACTCCTCCTCC
TCCGTCTTCTCCGCGCAGCACGAAGCCAAGAAGTTCACGTTTAACCTGATGGCGAAGAACAT
CATGAGCATGGACCCCGGCGA

SEQ ID NO: 94, seqüência de aminoácidos (parcial) CYP90B N-terminal de Triticum aestivum

MAAIMASITSELLFFLPFILLALLTFYTSAVAKCHGLHWWSGRTKKRRPNLPPGAVGWPFIG
ETFGYLRAHPATSIGQFMDQHIARYGKIYRSSLFGDRTVVSADAGLNRYILQNEGRLFECSY
PRSIGGILGKWSMLVLVGDPHREMRFISLNFLSSVRLRAVLLPEVERHTLLVLRDWLPYSSS
SVFSAQHEAKKFTFNLMAKNIMSMDPG

SEQ ID NO: 95, seqüência de ácidos nucléicos (parcial) CYP90B de extremidade 3' de Triticum aestivum ESTs compiladas BQ619714.1 CA715360.1

GGAAGGCCCTCAAGTCGCGAGCCACCATACTTGGAGTAATAGAGAGGAAAATGGAGGAAAGG
CTTGAGAAAATGAATAAGGAGGCCTCGAGCATGGAAGAAGATGATCTTCTCGGGTGGGCGAT
GAAGCAGTCCAATCTTTCGAAAGAACAAATATTGGACCTCTTGCTGAGCTTGCTGTTTGCCG
GGCATGAGACATCGTCAATGGCGCTCGCCCTCGCCATCTTCTTCCTCGAAGGTTGCCCTAAA
GCCGTTGAAGAACTGCGGGAGGAGCATCTTGAGATTGCTAGGAGACAGAAGCTGAGAGGGGA
GTGCAAACTGAGCTGGGAAGACTACAAAGAGATGGTTTTCACGCAATGCGTTATAAACGAGA
CCTTGCGGCTAGGCAACGTGGTCAGGTTCCTGCACCGGAAGGTCATTCGAGATGTGCACTAC

FIGURA 15 (Cont.)

AATGGGTATGACATCCCGAGCGGATGGAAAATTCTGCCGGTGTTAGCGGCGGTGCATCTCGA
CTCGTCGCTGTACGAGGACCCCAGCAGCTTCAACCCTTGGAGATGGAAGGGCAACGCGTCCG
GCGTGGCGCAGAACAGTAACTTCATGCCCTACGGCGGCGGCACGAGGCTCTGCGCCGGGTCG
GAGCTCGCCAAGCTCGAGATGGCCATCTTCCTGCACCACCTGGTGCTCAACTTCCGGTGGGA
GCTCGCCGAGCCCGACCAGGCGTTCGTGTACCCGTTCGTCGACTTCCCCAAGGGCCTGCCCA
TCAGGGTCCATAGGATTGCACAGGAAGAAGAAGGAGAAGAGGAGTAA

SEQ ID NO: 96, seqüência de aminoácidos (parcial) CYP90B C-terminal de Triticum aestivum

KALKSRATILGVIERKMEERLEKMNKEASSMEEDDLLGWAMKQSNLSKEQILDLLLSLLFAG
HETSSMALALAIFFLEGCPKAVEELREEHLEIARRQKLRGECKLSWEDYKEMVFTQCVINET
LRLGNVVRFLHRKVIRDVHYNGYDIPSGWKILPVLAAVHLDSSLYEDPSSFNPWRWKGNASG
VAQNSNFMPYGGGTRLCAGSELAKLEMAIFLHHLVLNFRWELAEPDQAFVYPFVDFPKGLPI
RVHRIAQEEEGEE

SEQ ID NO: 97, seqüência de ácidos nucléicos (parcial) CYP90B de Euphorbia esula DV141872.1

AGCATGGACCCTGGAAAACCAGAGACTGAAAAGCTCAAGAAAGAATATGTTACTTTCATGAA
AGGAGTTGTTTCTGCTCCTATTAATTTGCCTGGAACTGCTTATAGAAGGGCCTTACAGTCTC
GATCGACAATTTTGAAGTTCATAGAGGAGAAAATGGAGGAAAGAAATGAAAAATTAAAAGAA
GGAAAAGCAGAGGAAGAAGAAGAAGATGATCTTCTAGGATGGGTTTTAAAGCATTCAAATCT
TTCAACTGAGCAAATTCTAGATTTGGTATTAAGTTTAATGTTTGCTGGCCATGAAACTTCTT
CAGTAGCAATTCTTNNAGCTATCTATTTTTTGCAAGATTCTCCTGCTGCTCTTCAACAGCTA
AGGGAAGAACATAAGGAAATTGAAAAAGCCAAGAAGCAGTCAGGAGAGAAGGGATTGAACTG
GGATGTTTACAAAAATATGGAATTCACTCAGTGTGTTATCAATGAAACACTAAGACTTGGAA
ATGTAGTCAGGTTCCTTCATAGGAAGACTATTAAACACGTTCAATACAAAGGATATGACATT
CCACGTGGATGGAAAGTGCTGCCAGTGATTGCAGCCGTGCATTTGGACAGTAGCCATTTTGA
GAAGCCTCAACACTTCAATCCATGGAGATGGTTGCACCAGAATAATGGGATACAAAATATGA
ATAATAATTTCATGCCATTTGGGGGAGGACCAAGATTATGTGCAGGATCAGAATTAGCAAAA
CTAGAAATGGCTATTTTTATTCATCATTTGGTCCTTAATTACCAG

SEQ ID NO: 98, seqüência de aminoácidos (parcial) CYP90B de Euphorbia esula

SMDPGKPETEKLKKEYVTFMKGVVSAPINLPGTAYRRALQSRSTILKFIEEKMEERNEKLKE
GKAEEEEEDDLLGWVLKHSNLSTEQILDLVLSLMFAGHETSSVAILXAIYFLQDSPAALQQL
REEHKEIEKAKKQSGEKGLNWDVYKNMEFTQCVINETLRLGNVVRFLHRKTIKHVQYKGYDI
PRGWKVLPVIAAVHLDSSHFEKPQHFNPWRWLHQNNGIQNMNNNFMPFGGGPRLCAGSELAK
LEMAIFIHHLVLNYQ

SEQ ID NO: 99, seqüência de ácidos nucléicos (parcial) CYP90B de Gossypium hirsutum ESTs compiladas CO125422 DT568185.1

ATGCCTGACTCAGAGCCTATTTTCTTGCTTCTTCCATCTATTTTATCTTTGATTCTGTTCTT
CATTCTCATCAAGAGAAAGCAAAGAAGGTATAATCTTCCACCAGGGAACATGGGGTGGCCTT
TTCTCGGCGAAACAATCGGTTACTTGAGGCCTTACTCTGCTACTTCAGTAGGTGAATTCATG
CACCAGCATATATCAAGGTATGGGAATATCTACAAATCGAATTTGTTTGGTGAGAAGACAAT
AGTGTCTGCAGATGCTGGGTTGAACAAGTTCATATTACAAAACGAAGGGAGATTATTCGAGT
GCAGTTACCCGAGAAGCATTGGTGGCATTCTTGGGAAATGGTCGATGCTGGTTTTGGTTGGG

FIGURA 15 (Cont.)

GATATGCATAGAGACATGAGGATTATATCACTCAACTTCTTGAGCAACGCCAGGCTGAGGAC
TCATCTTTTGAGAGAAGTGGAGAAACATACTTTGCTTGTTCTAAATACTTGGAAAGAGAAGT
GCATATTTTCAGCTCAGGATGAAGCAAAAAAGTTCACTTTCAATTTGATGGCAAAAAATATC
ATGAGCATGGACCCTGGACATCCAGAGACGGAGCAGCTAAAGAAAGAATATGTTACTTTCAT
GAAAGGAGTCGTTTCTGCTCCTTTAAATTTACCTGGAACTGCATACAGAAAAGCCTTACAAT
CTCGATCAACAATCCTGAAATTTATTGAAAAGAAAATGGAAGTAAGGATAAGGAAAATGAAG
GAAGGAAAAGAAAACTCAGAGGAAGATGATCTTCTTGAATGGGTCTTAAAGCATTCTAATCT
TTCCACAGAGCAAATCCTTGACTTGATTTTGAGCTTGCTTTTTGCTGGACATGAGACTTCCT
CGGTAGCCATAACCTTAGCCATCTACTTCTTGCCAGGTTGTCCTTTGGCCATTCAACAGTTG
AGAGAGGAACACCTTGAAGTTGCAG

SEQ ID NO: 100, seqüência de aminoácidos (parcial) CYP90B de Gossypium hirsutum

MPDSEPIFLLLPSILSLILFFILIKRKQRRYNLPPGNMGWPFLGETIGYLRPYSATSVGEFM
HQHISRYGNIYKSNLFGEKTIVSADAGLNKFILQNEGRLFECSYPRSIGGILGKWSMLVLVG
DMHRDMRIISLNFLSNARLRTHLLREVEKHTLLVLNTWKEKCIFSAQDEAKKFTFNLMAKNI
MSMDPGHPETEQLKKEYVTFMKGVVSAPLNLPGTAYRKALQSRSTILKFIEKKMEVRIRKMK
EGKENSEEDDLLEWVLKHSNLSTEQILDLILSLLFAGHETSSVAITLAIYFLPGCPLAIQQL
RENTLKLQSQTNQRDXLNWDDTE

SEQ ID NO: 011, seqüência de ácidos nucléicos (parcial) CYP90B de Lycopersicon esculentum ESTs compiladas BF050501 AW221826.1 BM409833

ATGGGTTGGCCTTTTCTTGGTGAAACTATTGGCTATTTGAGACCTTATTCAGCTACTACTAT
TGGAGATTTCATGCAAGATCATATCTCTAGGTATGGGAAAATTTTCAAGTCAAATTTGTTTG
GAGAGCCAACAATAGTTTCAGCAGATGCAGGGCTAAACAGATACATTCTGCAGAATGAAGGG
AGATTATTTGAGTGTAATTATCCAAGAAGTATAGGTGGGATACTTGGTAAATGGTCTATGTT
AGTTCAAGTTGGACAAATGCATAGAGATATGAGGATGATTTCTCTGAATTTTTTGAGCAATG
CTAGGCTAAGGAATCAACTTTTAAGTGAAGTTGAAAAGCATACTTTGCTTGTTCTTGGCTCT
TGGAAACAGGATTCTGTTGTTTGTGCACAAGATGAAGCAAAGAAGTTAACATTCAACTTTAT
GGCAGAGCATATCATGAGTCTACAACCTGGAAATCCAGAGACAGAGAAGCTGAAAAAAGAGT
ACATCACATTTATGAAAGGAGTGGTTTCTGCTCCATTGAATTTTCCAGGAACAGCTTACAGA
AAGGCCTTACAGTCTCGATCAACAATTCTTGGATTTATTGAGAGAAAAATGGAGGAGAGGCT
TAAGGAAATGAACAGAAACGAAAACGACCTTCTAGGTTGGGTTCTGAAGAATTCAAATCTCT
CAAAAGAGCAAATTCTTGATTTGCTACTGAGTTTGCTCTTTGCTGGCCATGAAACTTCATCA
GTAGCAATAGCTCTGTCTATTTTCTTACTCGAAAGCTGTCCTGCTGCTGTTCAACAATTAAC
AGAAGAGCACTTGGAGATTTCCCGGG

SEQ ID NO: 102, seqüência de aminoácidos (parcial) CYP90B de Lycopersicon esculentum

MGWPFLGETIGYLRPYSATTIGDFMQDHISRYGKIFKSNLFGEPTIVSADAGLNRYILQNEG
RLFECNYPRSIGGILGKWSMLVQVGQMHRDMRMISLNFLSNARLRNQLLSEVEKHTLLVLGS
WKQDSVVCAQDEAKKLTFNFMAEHIMSLQPGNPETEKLKKEYITFMKGVVSAPLNFPGTAYR
KALQSRSTILGFIERKMEERLKEMNRNENDLLGWVLKNSNLSKEQILDLLLSLLFAGHETSS
VAIALSIFLLESCPAAVQQLTEEHLEISR

FIGURA 15 (Cont.)

SE ID NO: 103, seqüência de ácidos nucléicos (parcial) CYP90B de extremidade 5' de Solanum tuberosum ESTs compiladas BQ114367 BQ045917

ATGTCTGACTTAGAGTTTTTTCTTTTTCTTGTTCCTCCAATCTTGGCAGTCCTTATTATTCT
TAATCTATTCAAAAGAAAACACAAATTTCAAAATCTTCCACCAGGGGATATGGGTTGGCCTT
TTCTTGGTGAAACTATTGGTTATTTGAGACCTTACTCAGTTACTACTATTGGAGATTTCATG
CAAGATCATATCTCAAGGTATGGGAAAATTTTCAAGTCAAATTTGTTTGGAGAGCCAACAAT
TGTTTCAGCAGATGCAGGGCTTAACAGATACATTCTGCAGAATGAAGGGAGATTATTTGAGT
GTAATTATCCAAGAAGTATAGGTGGGATACTTGGTAAATGGTCTATGTTGGTTCAAGTTGGA
CAAATGCATAGAGATATGAGGATGATTTCTCTGAATTTTTTGAGCAATGCTAGACTCAGGAA
TCAACTTTTAAGTGAAGTTGAAAAGCATACTGTGCTTGTTCTTGGCTCTTGGAAACAGGATT
CTGTTGTTTGTGCACAAGATGAAGCAAAGAAGTTTACATTCAACTTTATGGCAGAGCATATC
ATGAGTCTACAACCTGGAAATCCAGAGACAGAGAAGCTGAAGAAAGAGTACATCACATTTAT
GAAAGGAGTGGTTTCTGCTCCATTGAATTTTCCAGGAACAGCTTACAGAAAAGCCCTACAGT
CTCGATCAACAATTCTTGGAATTATTGAG

SEQ ID NO: 104, seqüência de aminoácidos (parcial) CYP90B N-terminal de Solanum tuberosum

MSDLEFFLFLVPPILAVLIILNLFKRKHKFQNLPPGDMGWPFLGETIGYLRPYSVTTIGDFM
QDHISRYGKIFKSNLFGEPTIVSADAGLNRYILQNEGRLFECNYPRSIGGILGKWSMLVQVG
QMHRDMRMISLNFLSNARLRNQLLSEVEKHTVLVLGSWKQDSVVCAQDEAKKFTFNFMAEHI
MSLQPGNPETEKLKKEYITFMKGVVSAPLNFPGTAYRKALQSRSTILGIIE

SEQ ID NO: 105, seqüência de ácidos nucléicos (parcial) CYP90B C-terminal de Solanum tuberosum BQ114368.1

ggaaagctgtgaaagatgttcgatacaaaggttatgacattccatgtggatggaaagtattg
ccggtgatttcagcagcgcatttagatccttcacttttgaccgacctcacgactttgatcc
ttggagatggcagaatgcagaagagtcgccttcaggtaaaggaggaagcacaggcacaagca
gcacaacaaaaagtagtaataatttcatgccatttgggggaggtccacgtctatgtgcagga
tctgaactggccaaacttgagatggccattttcattcactatcttgttcttaattttcactg
gaaattagctgcacctgatcaggcttttgcctatccttacgtagattttcccaatgccctac
caatcactatccaacatcgatcgtcaaataaattaaaccaccacacttactacctctaa

SEQ ID NO: 106, seqüência de aminoácidos (parcial) CYP90B C-terminal de Solanum tuberosum

KAVKDVRYKGYDIPCGWKVLPVISAAHLDPSLFDRPHDFDPWRWQNAEESPSGKGGSTGTSS
TTKSSNNFMPFGGGPRLCAGSELAKLEMAIFIHYLVLNFHWKLAAPDQAFAYPYVDFPNALP
ITIQHRSSNKLNHHTYYL

SEQ ID NO: 107, prm06540

GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGGCCGCCATGATGGC

SEQ ID NO: 108, prm06520

GGGGACCACTTTGTACAAGAAAGCTGGGTTTACTCCTGCTCATCATCC

SEQ ID NO: 109, promotor endosperma específico de Oryza sativa

CTTCTACATCGGCTTAGGTGTAGCAACACGACTTTATTATTATTATTATTATTATTATT
ATTTTACAAAAATATAAAATAGATCAGTCCCTCACCACAAGTAGAGCAAGTTGGTGAGTTAT

FIGURA 15 (Cont.)

TGTAAAGTTCTACAAAGCTAATTTAAAAGTTATTGCATTAACTTATTTCATATTACAAACAA
GAGTGTCAATGGAACAATGAAAACCATATGACATACTATAATTTTGTTTTTATTATTGAAAT
TATATAATTCAAAGAGAATAAATCCACATAGCCGTAAAGTTCTACATGTGGTGCATTACCAA
AATATATATAGCTTACAAAACATGACAAGCTTAGTTTGAAAAATTGCAATCCTTATCACATT
GACACATAAAGTGAGTGATGAGTCATAATATTATTTTCTTTGCTACCCATCATGTATATATG
ATAGCCACAAAGTTACTTTGATGATGATATCAAAGAACATTTTTAGGTGCACCTAACAGAAT
ATCCAAATAATATGACTCACTTAGATCATAATAGAGCATCAAGTAAAACTAACACTCTAAAG
CAACCGATGGGAAAGCATCTATAAATAGACAAGCACAATGAAAATCCTCATCATCCTTCACC
ACAATTCAAATATTATAGTTGAAGCATAGTAGTA

SEQ ID NO: 110, promotor embrião/aleurona específico de Oryza sativa

GGTCAGCCAATACATTGATCCGTTGCCAATCATGCAAAGTATTTTGGCTGTGGCCGAGTGCC
GGAATTGATAATTGTGTTCTGACTAAATTAAATGACCAGAAGTCGCTATCTTCCAATGTATC
CGAAACCTGGATTAAACAATCCTGTTCTGTTCTCTAGCCCCTCCTGCATGGCCGGATTGTTT
TTTTGACATGTTTTCTTGACTGAGGCCTGTTTGTTCTAAACTTTTTCTTCAAACTTTTAACT
TTTTCATCACATCAGAACTTTTCTACACATATAAACTTTTAACTTTTCCGTCACATCGTTCC
AATTTCAATCAAACTTTCAATTTTGGCGTGAACTAAACACACCCTGAGTCTTTTATTGCTCC
TCCGTACGGGTTGGCTGGTTGAGAATAGGTATTTTCAGAGAGAAAATCTAGATATTGGGAGG
AACTTGGCATGAATGGCCACTATATTTAGAGCAATTCTACGGTCCTTGAGGAGGTACCATGA
GGTACCAAAATTTTAGTGTAAATTTTAGTATCTCATTATAACTAGGTATTATGAGGTACCAA
ATTTACAATAGAAAAAATAGTACTTCATGGTACTTTCTTAAGTACCGTAAAATTGCTCCTAT
ATTTAAGGGGATGTTTATATCTATCCATATCCATAATTTGATTTTGATAAGAAAAAATGTGA
GCACACCAAGCATGTCCATGACCTTGCACTCTTGGCTCACTCGTCAACTGTGAAGAACCTCA
AAAATGCTCAATATAGCTACAGGTGCCTGAAAAAATAACTTTAAAGTTTTGAACATCGATTT
CACTAAACAACAATTATTATCTCCCTCTGAAAGATGATAGTTTAGAACTCTAGAATCATTGT
CGGCGGAGAAAGTAAATTATTTTCCCCAAATTTCCAGCTATGAAAAAACCCTCACCAAACAC
CATCAAACAAGAGTTCACCAAACCGCCCATGCGGCCATGCTGTCACGCAACGCACCGCATTG
CCTGATGGCCGCTCGATGCATGCATGCTTCCCCGTGCACATATCCGACAGACGCGCCGTGTC
AGCGAGCTCCTCGACCGACCTGTGTAGCCCATGCAAGCATCCACCCCCGCCACGTACACCCC
CTCCTCCTCCCTACGTGTCACCGCTCTCTCCACCTATATATGCCCACCTGGCCCCTCTCCTC
CCATCTCCACTTCACCCGATCGCTTCTTCTTCTTCGTTGCATTCATCTTGCTAGC

SEQ ID NO: 111, seqüência de ácidos nucléicos Glyma_CYP90B de Glycine max agrupamento de BF324723.1, AW759817.1, DX407703, CL885557.1, BG726252.21

ATGTCTGACTCACTCTTAACTTTCTATTCTCTTTCAGCCATTCTTGCTCTTCTCCCAATCTT
CATCTTCATTCTCATCAAAAGAAAGCAAAGCAAACCCAGGCTCAACCTTCCCCCAGGTAACA
TGGGTTGGCCATTTCTTGGTGAAACCATTGGCTATTTGAAGCCTTATTCTGCCACCACAGTA
GGGGAATTCATGGAGCAACACATAGCAAGGTATGGTACAATTTACAAGTCAAAACTGTTTGC
GGGGCCAGCAATAGTGTCAGCAGATGCAGGACTCAACAGGTTCATTCTACAAAACGAAGGGA
AATTGTTCGAGTGCAGCTATCCTAGAAGCATCGGTGGAATACTAGGAAAATGGTCCATGTTG
GTCTTAGTTGGTGACATGCATAGAGACATGCGGGTTATATCACTCAACTTTCTAAGCCACGC
CAGGCTCAGAACACACCTCTTGAAAGAGGTGGAGAAGCAATCCCTCTTGGTTCTGAACTCTT
GGAGCCAAAATTCCATATTCTCAGCCCAAGATGAAGCTAAGAAGTTCACCTTCAATTTAATG
GCTAAGCATATCATGAGCATGGATCCTGGGGATATCGAGACAGAGCAACTAAAGAAAGAGTA
CGTCACTTTCATGAAAGGGGTGGTTTCCGCGCCATTGAATTTACCTGGAACTGCATACCGAA
AGGCATTGAAGTCTCGGTCCATTATACTAAAGTTCATAGAGGGGAAAATGGAAGAGAGAGTT
AGAAGAATCCAAGAGGGGAATGAGAGTTTGGAGGAAGATGATCTTCTAAATTGGGTTTTGAA

FIGURA 15 (Cont.)

GAATTCAAATCTTTCAACCGAGCAAATTCTTGACTTGATTCTCAGCTTGCTCTTCGCTGGCC
ATGAAACTTCGTCGGTAGCCATAGCTCTAGCCATTTACTTCTTACCCGGTTGTCCTCAAGCT
ATACAACAGTTAAAGGAAGAACACAGAGAAATTGCCAGAGCCAAAAAGCAAGCAGGGGAAGT
TGAACTCACTTGGGATGACTACAAACGAATGGAATTTACTCATTGTGTAANNTTGGGAAATG
TTGTGAGGTTTCTCCACAGGAAGGCTGTGAAAGATGTTAACTATAAAGGTTATGACATTCCA
TGTGGGTGGAAAGTCCTCCCGGTGATTGCAGCCGTGCATCTGGATCCTTCACTTTTTGACCA
ACCTCAACACTTCAATCCATGGAGATGGCAGAACAATGGCAGTCATGGAAGTTGTCCAAGCA
AGAACACAGCAAACAACAACTTTCTTCCGTTCGGAGGAGGACCACGATTATGTGCAGGATCA
GAGTTAGCTAAGCTTGAAATGGCTGTTTTCATTCACCATCTCATTCTCAACTACCATTGGGA
ATTGGCTGATACCGATCAAGCTTTTGCCTACCCTTTTGTCGACTTCCCCAAAGGCCTACCCG
TTAGAGTCCAAGCCCATTCTTTACTTTGA

SEQ ID NO: 112, seqüência de polipeptídeo traduzido Glyma_CYP90B de Glycine max

MSDSLLTFYSLSAILALLPIFIFILIKRKQSKPRLNLPPGNMGWPFLGETIGYLKPYSATTV
GEFMEQHIARYGTIYKSKLFAGPAIVSADAGLNRFILQNEGKLFECSYPRSIGGILGKWSML
VLVGDMHRDMRVISLNFLSHARLRTHLLKEVEKQSLLVLNSWSQNSIFSAQDEAKKFTFNLM
AKHIMSMDPGDIETEQLKKEYVTFMKGVVSAPLNLPGTAYRKALKSRSIILKFIEGKMEERV
RRIQEGNESLEEDDLLNWVLKNSNLSTEQILDLILSLLFAGHETSSVAIALAIYFLPGCPQA
IQQLKEEHREIARAKKQAGEVELTWDDYKRMEFTHCVSGMIINELWRSTCLFFYFMYSNCYY
YYLMNMLVVSNIIWDDEAXXVGLGNVVRFLHRKAVKDVNYKGYDIPCGWKVLPVIAAVHLDP
SLFDQPQHFNPWRWQNNGSHGSCPSKNTANNNFLPFGGGPRLCAGSELAKLEMAVFIHHLIL
NYHWELADTDQAFAYPFVDFPKGLPVRVQAHSLL

SEQ ID NO: 113, seqüência de ácidos nucléicos Goshi_CYP90B de Gossypium hirsutum

ATGCCTGACTCAGAGCCTATTTTCTTGCTTCTACCATCTATTTTATCTTTGATTCTGTTCTT
CATTCTCATCAAGAGAAAGCAAAGAAGGTATAATCTTCCACCAGGGAACATGGGGTGGCCTT
TTCTCGGCGAAACAATCGGTTACTTGAGGCCTTACTCTGCTACTTCAGTAGGTGAATTCATG
CACCAGCATATATCAAGGTATGGGAATATCTACAAATCGAATTTGTTTGGTGAGAAGACAAT
AGTGTCTGCAGATGCTGGGTTGAACAAGTTCATATTACAAAACGAAGGGAGATTATTCGAGT
GCAGTTACCCGAGAAGCATTGGTGGCATTCTTGGGAAATGGTCGATGCTGGTTTTGGTTGGG
GATATGCATAGAGACATGAGGATTATATCACTCAACTTCTTGAGCAACGCCAGGCTGAGGAC
TCATCTTTTGAGAGAAGTGGAGAAACATACTTTGCTTGTTCTAAATACTTGGAAAGAGAAGT
GCATATTTTCAGCTCAGGATGAAGCAAAAAAGTTCACTTTCAATTTGGTGGCAAAAAATATC
ATGAGCATGGACCCTGGACATCCAGAGACGGAGCAGCTAAAGAAAGAATATGTTACTTTCAT
GAAAGGAGTCGTTTCTGCTCCTTTAAATTTACCTGGAACTGCATACAGAAAAGCCTTACAAT
CTCGATCAACAATCCTGAAATTTATTGAAAAGAAAATGGAAGTAAGGATAAGGAAAATGAAG
GAAGGAAAAGAAAACTCAGAGGAAGATGATCTTCTTGAATGGGTCTTAAAGCATTCTAATCT
TTCCACAGAGCAAATCCTTGACTTGATTTTGAGCTTGCTTTTTGCTGGACATGAGACCTCCT
CGGTAGCCATAACCTTAGCCATCTACTTCTTGCCAGGTTGTCCTTTGGCCATTCAACAGTTG
AGAGAAGAACACCTTGAAGTTGCCAGAGCCAAGAACCAATCAGGAGAGACTGAACTCAACTG
GGATGATTACAAGAAAATGGAGTTCACTCAATGTGTTATTAACGAGACACTTAGGCTTGGTA
ATGTCGTCAGATTTCTCCACAGAAAAGCTCTCAAAGATATTAGATATAAAGGTTATGATATT
CCATGTGGGTGGAAAGTGCTTCCAGTGATTGCAGCAGTGCACTTGGATCCCTGTCTTTTTGA
CCACCCTCAACTCTTCAATCCATGGCGATGGCAGCAAAATAATGGGAGTCGAGGGGCGGGGA
CGGCAACGTCATCAGCGAGCAGCAGCAATTACTTCATGCCATTCGGGGGAGGACCACGGCTA
TGTGCAGGAACAGAGCTGGCTAAACTGGAAATGGCGGTGTTCATCCACCATTTGGTCCTCAA
CTACCAGTGGGAGTTAGCCGATACGGATGAAGCCTTTGCCTTCCCTTTTGTCGACTTCCCTA
AAGGCCTACCCATCAGAGTCTTCAAATCTTAA

FIGURA 15 (Cont.)

SEQ ID NO: 114, Goshi_CYP90B de Gossypium hirsutum

MPDSEPIFLLLPSILSLILFFILIKRKQRRYNLPPGNMGWPFLGETIGYLRPYSATSVGEFM
HQHISRYGNIYKSNLFGEKTIVSADAGLNKFILQNEGRLFECSYPRSIGGILGKWSMLVLVG
DMHRDMRIISLNFLSNARLRTHLLREVEKHTLLVLNTWKEKCIFSAQDEAKKFTFNLVAKNI
MSMDPGHPETEQLKKEYVTFMKGVVSAPLNLPGTAYRKALQSRSTILKFIEKKMEVRIRKMK
EGKENSEEDDLLEWVLKHSNLSTEQILDLILSLLFAGHETSSVAITLAIYFLPGCPLAIQQL
REEHLEVARAKNQSGETELNWDDYKKMEFTQCVINETLRLGNVVRFLHRKALKDIRYKGYDI
PCGWKVLPVIAAVHLDPCLFDHPQLFNPWRWQQNNGSRGAGTATSSASSSNYFMPFGGGPRL
CAGTELAKLEMAVFIHHLVLNYQWELADTDEAFAFPFVDFPKGLPIRVFKS

SEQ ID NO: 115, seqüência de ácidos nucléicos Zeama_CYP90B_1 de Zea mays agrupamento de KK156535.1 CA452603.1 EE156534.1 CG333964.1

ATGGGCGCCATGATGGCCTCCATAACCAGCGAGCTCCTCTTCTTCCTTCCCTTCATCCTGCT
GGCCCTCCTCGCCTTGTACACCACCACCGTCGCCAAATGCCACGGCACCCACCCGTGGCGCC
GTCAGAAGAAGAAGCGGCCCAACCTGCCCCCGGGCGCCCGCGGATGGCCCTTGGTCGGCGAA
ACTTTCGGCTACCTCCGCGCCCACCCGGCCACCTCCGTGGGCCGCTTCATGGAGCGGCATGT
CGCACGGTACGGGAAGATATACCGGTCGAGCCTGTTCGGGGAGCGGACGGTGGTGTCGGCGG
ACGCGGGGCTGAACCGCTACATCCTGCAGAACGAGGGGCGGCTGTTCGAGTGCAGCTACCCG
CGCAGCATCGGCGGCATCCTGGGCAAGTGGTCCATGCTGGTGCTCGTGGGCGACGCGCACCG
CGAGATGCGCGCTATCTCGCTCAACTTCCTCAGCTCCGTCCGCCTCCGCGCCGTGCTGCTCC
CCGAGGTGGAGCGCCACACCCTGCTGGTCCTCCGCTCGTGGCCGCCCTCCGACGGCACCTTC
TCCGCCCAGCACGAAGCCAAGAAGTTCACGTTTAACCTGATGGCGAAGAACATAATGAGCAT
GGACCCCGGCGAGGAGGAGACGGAGCGGCTGCGGCTGGAGTACATCACCTTCATGAAGGGCG
TCGTGTCAGCGCCGCTCAACTTCCCGGGCACGGCCTACTGGAAGGCGCTCAAGTCACGCGCG
TCCATACTTGGAGTGATAGAGAGGAAGATGGAGGACAGGCTTGAGAAGATGAGCAGGGAGAA
GTCAAGCGTGGAGGAGGACGACCTTCTTGGATGGGCCCTGAAGCAATCCAACCTGTCCAAGG
AACAGATCCTGGACCTCTTGCTGAGCCTGCTCTTCGCGGGGCACGAGACTTCGTCCATGGCG
CTCGCCCTCGCCATCTTCTTCCTCGAAGGGTGCCCTAAGGCCGTGCAAGAACTCCGGGAGGA
GCATCTCCTGATTGCTAGGAGACAAAGGCTAAGAGGGGCGTCTAAATTGAGCTGGGAAGACT
ACAAGGAAATGGTTTTCACGCAGTGTGTTATAAACGAGACATTGCGGCTCGGCAACGTGGTC
AGGTTCCTGCACCGGAAGGTCATCCGAGATGTACACTACAATGGGTACGACATACCGCGGGG
GTGGAAAATCCTGCCGGTTCTAGCGGCGGTGCACCTGGACTCGTCGCTGTACGAGGACCCCA
GCCGGTTCAACCCTTGGAGATGGAAGAGCAACAACGCGCCAAGCAGCTTCATGCCGTACGGC
GGCGGGCCGCGGCTGTGCGCCGGGTCGGAGCTGGCCAAGCTGGAGATGGCCATCTTCCTGCA
CCACCTGGTGCTCAACTTCCGGTGGGAGCTGGCGGAGCCGGACCAGGCCTTCGTTTACCCTT
TCGTCGACTTCCCCAAGGGCCTCCCGATCAGGGTCCAGCGGGTCGCCGACGACCAAGGCCAT
CGTAGCGTTTTGACCGAGAGCACAAGAGGCTGA

SEQ ID NO: 116, polipeptídeo traduzido Zeama_CYP90B_1 de Zea mays

MGAMMASITSELLFFLPFILLALLALYTTTVAKCHGTHPWRRQKKKRPNLPPGARGWPLVGE
TFGYLRAHPATSVGRFMERHVARYGKIYRSSLFGERTVVSADAGLNRYILQNEGRLFECSYP
RSIGGILGKWSMLVLVGDAHREMRAISLNFLSSVRLRAVLLPEVERHTLLVLRSWPPSDGTF
SAQHEAKKFTFNLMAKNIMSMDPGEEETERLRLEYITFMKGVVSAPLNFPGTAYWKALKSRA
SILGVIERKMEDRLEKMSREKSSVEEDDLLGWALKQSNLSKEQILDLLLSLLFAGHETSSMA
LALAIFFLEGCPKAVQELREEHLLIARRQRLRGASKLSWEDYKEMVFTQCVINETLRLGNVV
RFLHRKVIRDVHYNGYDIPRGWKILPVLAAVHLDSSLYEDPSRFNPWRWKSNNAPSSFMPYG
GGPRLCAGSELAKLEMAIFLHHLVLNFRWELAEPDQAFVYPFVDFPKGLPIRVQRVADDQGH
RSVLTESTRG

FIGURA 15 (Cont.)

SEQ ID NO: 117, seqüência de ácidos nucléicos Zeama_CYP90B_2 de Zea mays ZM58195574-58721845

ATGGGCGCCATGATGGCCTCCATAACCAGCGAGCTCCTCTTCTTCCTTCCCTTCATCCTGCT
GGCCCTCCTCGCCTTGTACACCACCACCGTCGCCAAATGCCACGGCACCCACCAGTGGCGCC
GTCAGAAGAAGAAGCGGCCCAACCTGCCCCCGGGCGCCCGCGGATGGCCCTTGGTCGGCGAG
ACTTTCGGCTACCTCCGCGCCCACCCGGCCACCTCCGTGGGCCGCTTCATGGAGCGGCATGT
CGCACGGTACGGGAAGATATACCGGTCGAGCCTGTTCGGGGAGCGGACGGTGGTGTCGGCGG
ACGCGGGGCTGAACCGGTACATCCTGCAGAACGAGGGGCGGCTGTTCGAGTGCAGCTACCCG
CGCAGCATCGGCGGCATCCTGGGCAAGTGGTCCATGCTGGTGCTCGTGGGCGACGCGCACCG
CGAGATGCGCGCTATCTCGCTCAACTTCCTCAGCTCCGTCCGCCTCCGCGCCGTGCTGCTCC
CCGAGGTGGAGCGCCACACCCTGCTGGTCCTCCGCTCCTGGCCGCCTCCGACCGGCACCTTC
TCCGCCCAGCACGAAGCCAAGAAGTTCACGTTTAACCTGATGGCGAAGAACATAATGAGCAT
GGACCCCGGCGAGGAGGAGACGGAGCGGCTGCGGCTGGAGTACATCACCTTCATGAAGGGCG
TCGTGTCAGCGCCGCTCAACTTCCCGGGCACGGCCTACTGGAAGGCGCTCAAGTCGCGCGCG
TCCATACTTGGAGTGATAGAGAGGAAGATGGAGGACAGGCTTGAGAAGATGAGCAGGGAGAA
GTCAAGCGTGGAGGAGGACGACCTTCTTGGATGGGCCCTGAAGCAATCCAACCTGTCCAAGG
AACAGATCCTGGACCTCTTGCTGAGCCTGCTCTTCGCGGGGCACGAGACTTCGTCCATGGCG
CTCGCCCTCGCCATCTTCTTCCTCGAAGGGTGCCCTAAGGCCGTGCAAGAACTCCGGGAGGA
GCATCTCCTGATTGCTAGGAGACAAAGGCTAAGGGGGGCGTCCAAATTGAGCTGGGAAGACT
ACAAGGAAATGGTTTTCACGCAGTGTGTTATAAACGAGACATTGCGGCTCGGCAACGTGGTC
AGGTTCCTGCACCGGAAGGTCATCCGAGATGTACACTACAATGGGTACGACATACCGCGGGG
GTGGAAAATCCTGCCGGTTCTAGCGGCGGTGCACCTGGACTCGTCGCTGTACGAGGATCCCA
GCCGGTTCAACCCTTGGAGATGGAAGGTCAGTGCCCTGCCCCCTCCCCTCGCAAAAGGCACA
AGCCAAGGGACAAGCAAGAGGTTTACACCGTTCGGTGGTGGCCCCCGGCTCTGCCCAGGATC
AGAGCTCGCTAAAGTGGAGACTGCTTTCTTCCTCCATCACCTTGTCCTCAATTATAGATGGA
GAATTGATGGCGATGACATTCCAATGGCATACCCGTATGTGGAGTTTCAGAGAGGTCTGCCA
ATAGAAATCGAGCCAACGTCCCCTGAATTTGACTGTCCTGGAGCTACAGCCATCAGTTATCA
CACCAGAGAGAAATGA

SEQ ID NO: 118, polipeptídeo traduzido Zeama_CYP90B_2 de Zea mays

MGAMMASITSELLFFLPFILLALLALYTTTVAKCHGTHQWRRQKKKRPNLPPGARGWPLVGE
TFGYLRAHPATSVGRFMERHVARYGKIYRSSLFGERTVVSADAGLNRYILQNEGRLFECSYP
RSIGGILGKWSMLVLVGDAHREMRAISLNFLSSVRLRAVLLPEVERHTLLVLRSWPPPTGTF
SAQHEAKKFTFNLMAKNIMSMDPGEEETERLRLEYITFMKGVVSAPLNFPGTAYWKALKSRA
SILGVIERKMEDRLEKMSREKSSVEEDDLLGWALKQSNLSKEQILDLLLSLLFAGHETSSMA
LALAIFFLEGCPKAVQELREEHLLIARRQRLRGASKLSWEDYKEMVFTQCVINETLRLGNVV
RFLHRKVIRDVHYNGYDIPRGWKILPVLAAVHLDSSLYEDPSRFNPWRWKVSALPPPLAKGT
SQGTSKRFTPFGGGPRLCPGSELAKVETAFFLHHLVLNYRWRIDGDDIPMAYPYVEFQRGLP
IEIEPTSPEFDCPGATAISYHTREK

SEQ ID NO: 119, seqüência parcial de ácidos nucléicos Brara_CYP90B de Brassica rapa agrupamento de ED531940.1, DU985218.1, AJ859443.1

ATGTTCGAAACAGAGCATACTCTCGTGCCTCTTCTTCTTCTCCCATCACTTCTATCTCTTCT
CCTCTTCTTGATTCTCTTGAAGAGACGAAGTCGACACAGTTTCAATCTCCCTCCTGGAAAAT
CTGGATGGCCATTTCTAGGCGAAACCATCGGATATCTCAAACCTTACTCTGCCAAAACTCTC
GGTTACTTCATGCAACAACATATCTCCAAGTATGGGAAGATATATAGATCGAATTTGTTTGG
AGAACCAACGATCGTATCAGCTGATGCAGGACTCAACAGGTTCATATTACAAAACGAAGGAA

FIGURA 15 (Cont.)

GACTCTTTGAATGTAGTTATCCTCGAAGTATTGGTGGGATTCTTGGGAAATGGTCGATGCTT
GTTCTTGTTGGAGACATGCATAGAGACATGAGAAGTATCTCGCTAAACTTTCTAAGTCACGC
TCGTCTCAGAACGATTCTTCTTAAGGACGTTGAGAGGCATACTTTGTTCGTTCTTAATTCTT
GGCAACAACATTCTGTTTTCTCTGCTCAAGATGAGGCCAAAAAGTTTACGTTTAATCTAATG
GCGAAGCATATAATGAGTATGGATCCTGGAGAAGAAGAGACAGAGCAGTTAAAGAAAGAGTA
TGTGACTTTCATGAAAGGGGTTGTTTCTGCTCCTCTCAATCTCCCAGGAACTGCTTATCGTA
AAGCTCTCCAGCAGTCACGAGGGACGATATTGAAGTTTATTGAGAAGAAAATGGAAGAGAGA
AAATCAGAGATTCAAGAAGAAGACGAAGAAGATGAAGCAGAGATTAGTAGAAGTGATCATTA
TGAGAGAAAACATAGAGCAGATGATGATCTTTTGGGATGGGTTCTAAAACATTCCAATCTTT
CGACTGAGCAAATTCTCGATCTTATTCTCAGTTTATTATTTGCCGGACATGAGACATCATCT
GTAGCCATCGCTCTCGCTATCTACTTTTTAGCGGGCNNNTTGAAGGAAGAGCATCTTGAAAT
CGCGAGGGTGAAGAAGGAACTTGGAGAGTCAGAATTGAATTGGGATGATTACAAGACAATGG
ACTTTACTCATAGTGTTATAAATGAGACTCTTCGACTAGGAAATGTAGTAAGGTTTTTGCAT
CGTAAAGCACTCAAAAACGTTCGGTATAAAGGATACGATATCCCAAGTGGGTGGAAAGTGTT
ACCAGTGATCTCAGCCGTACATTTGGATAACTCCCGTTACGACGAACCTAATCTCTTTAATC
CTTGGAGATGGCAACAG

SEQ ID NO: 120, polipeptídeo parcial traduzido Brara_CYP90B de Brassica rapa

MFETEHTLVPLLLLPSLLSLLLFLILLKRRSRHSFNLPPGKSGWPFLGETIGYLKPYSAKTL
GYFMQQHISKYGKIYRSNLFGEPTIVSADAGLNRFILQNEGRLFECSYPRSIGGILGKWSML
VLVGDMHRDMRSISLNFLSHARLRTILLKDVERHTLFVLNSWQQHSVFSAQDEAKKFTFNLM
AKHIMSMDPGEEETEQLKKEYVTFMKGVVSAPLNLPGTAYRKALQQSRGTILKFIEKKMEER
KSEIQEEDEEDEAEISRSDHYERKHRADDDLLGWVLKHSNLSTEQILDLILSLLFAGHETSS
VAIALAIYFLAGXLKEEHLEIARVKKELGESELNWDDYKTMDFTHSVINETLRLGNVVRFLH
RKALKNVRYKGYDIPSGWKVLPVISAVHLDNSRYDEPNLFNPWRWQQ

SEQ ID NO: 121, seqüência parcial de ácidos nucléicos Horvu_CYP90B de Hordeum vulgare BJ474344.1, AV835480

ATGGCCGCCATGATGGCATCCATAACCAGCGAGCTCCTTTTCTTCCTCCCCTTCATCCTCCT
GGCCCTGCTCACCTTCTACACCAGCAGCGTGGCCAAATGCCATGGCCTCCACCGGTGGAGCG
GCCGGACGAAGAAGAAGCGGCCGAATCTGCCGCCCGGCGCCGCCGGCTGGCCTTTCGTCGGC
GAGACCTTCGGGTACCTCCGCGCCCACCCGGCCACCTCCATCGGCCAGTTCATGAACCAGCA
CATCGCACGGTACGGGAAGATATACCGGTCGAGCCTGTTCGGGGAGCGGACGGTGGTGTCGG
CGGACGCGGGGCTGAACCGGTACATCCTGCAGAACGAGGGGCGGCTGTTCGAGTGCAGCTAC
CCGCGGAGCATCGGCGGCATCCTGGGCAAATGGTCCATGCTGGTGCTCGTGGGCGACCCCCA
CCGCGAGATGCGCTCCATCTCCCTCAACTTCCTCAGCTCCCTCCGCCTCCGCGCCGTGCTCC
TCCCGGAGGTGGAGCGCCACACCCTCCTCGTCCTCCGCGACTGGCTGCCTTCCTCCTCCTCC
GCCGTCTTCTCCGCCCAGCACGAAGCCAAGAAGTTCACGTTTAACCTGATGGCGAAGAACAT
CATGAGCATGGACCCCGGCGAGGAGGAGACGGAGCGGCTGAGGCTCGAGTACATCACCTTCA
TGAAGGGGGTGGTGTCCGCGCCCCTCAACTTCCCCGGGACGGCCTACTGGAAGGCCCTCAAG
TCTCGAGCCACCATACTTGGGGTAATAGAGAGGAAAATGGAGGATAGGCTCGAGAAAATGAA
CAAGGAGGCCTCGAGCATGAAGAAGATATCTCTCGGGGGGCG

FIGURA 15 (Cont.)

SEQ ID NO: 122, polipeptídeo parcial traduzido Horvu_CYP90B de Hordeum vulgare

MAAMMASITSELLFFLPFILLALLTFYTSSVAKCHGLHRWSGRTKKKRPNLPPGAAGWPFVG
ETFGYLRAHPATSIGQFMNQHIARYGKIYRSSLFGERTVVSADAGLNRYILQNEGRLFECSY
PRSIGGILGKWSMLVLVGDPHREMRSISLNFLSSLRLRAVLLPEVERHTLLVLRDWLPSSSS
AVFSAQHEAKKFTFNLMAKNIMSMDPGEEETERLRLEYITFMKGVVSAPLNFPGTAYWKALK
SRATILGVIERKMEDRLEKMNKEASSMKKISLGG

SEQ ID NO: 123, seqüência parcial de ácidos nucléicos Lotja_CYP90B de Lotus japonicus AP007781.1

ATGTCTGACTCATACCTCACTTTCTGTTTTCTTTCTTCCATTCTTGCTCTCACTGTAATCAT
CATTTTCATGAAAAGAAAGAAGGCAAGGCTTAACCTTCCCCCTGGAAAAATGGGATGGCCCT
TTCTTGGGGAAACCATTGGGTACTTGAAGCCATACTCTGCTACCACAGTAGGAGATTTTATG
GAAAAGCACATAGCAAGGTATGGTACAATTTACAAGTCAAAATTGTTTGGTGAGCCTGCAAT
AGTTTCTGCAGATGCAGAGTTGAACAGGTTCATATTACAGAACGAAGGGAAGCTGTTTGAGT
GCAGCTATCCAAGAAGCATTGGAGGAATACTTGGAAAATGGTCCATGTTGGTCTTAGTTGGG
GACATGCATAGGGACATGAGAAACATTTCACTGAACTTTCTGAGCCTTGCTAGGCTCAAAAC
ACATCTATTGAAAGAAGTGGAGAAGCACTCTCTTCTAGTTCTAGGCTCTTGGAAAGAAAATT
GTACATTCTCAGCTCAAGATGAAGCAAAGAAGTTCACATTCAATTTGATGGCGAAACATATC
ATGAGCTTGGATCCTGGGAATCTAGAGACAGAACAGCTGAAGAAAGAGTATGTCTCTTTCAT
GAATGGTGTGGTGTCTGCACCTTTGAATTTTCCCGGAACTGCATACAGAAAAGCATTAAAGT
CTAGGTCCACCATACTGAAGTTCATAGAGGGAAAAATGGAAGAAAGGATCAAAAGAAACCAA
AATTTGGAGGAAGATCTTCTAAACTGGGTTGTGATGCATTCAAATCTTTCAACTGAGCAAAT
TCTTGACCTGGTTCTGAGCTTGCTCTTTGCAGGCCATGAAACTTCATCTGTGGCTATAGCTT
TAGCTATTTACTTTTTGCCAGGTTGTCCTAAAGCTATACAACAATTAAGGGAAGAACATAGA
GAAATAGCCAGGTCCAAGAAGCAAGCAGGGGAGGTTGAATTAACTTGGGATGATTACAAAAG
AATGGAGTTTACTCAATGTGTAGTTGTGAATGAAACACTGAGGTTGGGAAATGTTGTGAGGT
TCCTTCACAGGAAGGCTCTGAAAGATGTTCGGTACAAAGGTTATGACATTCCACGTGGGTGG
AAAGTCCTCCCTGTGATTTCAGCTATGCATCTGGATCCTGCACTTTTTTACCAACCTCAACA
CTTCAATCCATGGAGATGGAAG

SEQ ID NO: 124, polipeptídeo parcial traduzido Lotja_CYP90B de Lotus japonicus

MSDSYLTFCFLSSILALTVIIIFMKRKKARLNLPPGKMGWPFLGETIGYLKPYSATTVGDFM
EKHIARYGTIYKSKLFGEPAIVSADAELNRFILQNEGKLFECSYPRSIGGILGKWSMLVLVG
DMHRDMRNISLNFLSLARLKTHLLKEVEKHSLLVLGSWKENCTFSAQDEAKKFTFNLMAKHI
MSLDPGNLETEQLKKEYVSFMNGVVSAPLNFPGTAYRKALKSRSTILKFIEGKMEERIKRNQ
NLEEDLLNWVVMHSNLSTEQILDLVLSLLFAGHETSSVAIALAIYFLPGCPKAIQQLREEHR
EIARSKKQAGEVELTWDDYKRMEFTQCVVVNETLRLGNVVRFLHRKALKDVRYKGYDIPRGW
KVLPVISAMHLDPALFYQPQHFNPWRWK

SEQ ID NO: 125, seqüência parcial de ácidos nucléicos Maldo_CYP90B de Malus domestica HG631390

AGAGACATGAGGATGATATCTCTCAACTTCTTGAGCCATGCCAGACTCAGAACCCATCTGAT
GAGAGAAGTTGAAAAGCACACTCTTCTTGTTTTGGGGAGTTGGAAGGAAAACTCTGTTTTTT
CAGCTCAGGATGAAGCTAAGAAGTTCACGTTCAACTTGATGGCCAAACATATCATGAGCTTG
GATCCTGGAAAACCAGAGACTGAGCAGCTGAAGAAATTGTATGTTACTTTCATGAAAGGTGT

FIGURA 15 (Cont.)

GGTTTCTCCTCCTCTGAATTTACCAGGAACAGCTTACAGAAGAGCCTTACAGTCAAGATCAA
CGATTCTGAAGTTTATAGAGTGCAAAATGAAAGAAAGATTGATGGAGGGAACCGAAAACATT
GGGGAAGATGATCTGCTTGGATGGGTTCTGAAGAATTCAAATCTTTCAAAGGAGCAAATTCT
TGACTTAATATTGAGCTTGCTCTTTGCTGGCCATGAAACTTCATCAGTGTCTATAGCCTTAG
CAATTTACTTTTTACCAAGCTGCCCTAATGCAATTCTGCAGTTAAGGGAAGAACACAGTGAA
ATTGCCAAAGCCAAGAAACTGGCAGGCGAGACAGAGTTGAATTGGGAGGACTACAAGAAAAT
GGAGTTCACCGAATGTGTAATCTGTGAGACACTTCGGCTTGGGAATGTGGTGAGATTTTTAC
ACAGAAAGGCTCTGAAGGATGTTCGGTACAAAGGGTATGACATTCCATCTGGGTGGAAAGTG
CTTCCGGTGATTGCAGCCGTGCATTTGGATCCTTTACTTTTTGACCACCCGCAACACTTCAA
TCCATGGAGATGGCAGCAGAATAACAACCACCACCACCGTGGATCATCGTCGTCATCATCAT
GCTACACGAGCATGACAAGTCACAACTTTATGCCATTTGGGGGAGGACCACGACTTTGCGCC
GGTTCAGAATTGGCCAAACTTGAAATGGCCGTGTTCATCCACCACCTTGTCCTCAACTTCCA
CTGGGAGTTAGCCGATCCTTTCGACAAACCTTTTGCTTTCCCCTTCGTCGATTTCCAAAATG
GCCTACCAATCACAGCCCACCGCTACGTACCAAACTAA

SEQ ID NO: 126, polipeptídeo parcial traduzido Maldo_CYP90B de Malus domestica

RDMRMISLNFLSHARLRTHLMREVEKHTLLVLGSWKENSVFSAQDEAKKFTFNLMAKHIMSL
DPGKPETEQLKKLYVTFMKGVVSPPLNLPGTAYRRALQSRSTILKFIECKMKERLMEGTENI
GEDDLLGWVLKNSNLSKEQILDLILSLLFAGHETSSVSIALAIYFLPSCPNAILQLREEHSE
IAKAKKLAGETELNWEDYKKMEFTECVICETLRLGNVVRFLHRKALKDVRYKGYDIPSGWKV
LPVIAAVHLDPLLFDHPQHFNPWRWQQNNNHHHRGSSSSSSCYTSMTSHNFMPFGGGPRLCA
GSELAKLEMAVFIHHLVLNFHWELADPFDKPFAFPFVDFQNGLPITAHRYVPN

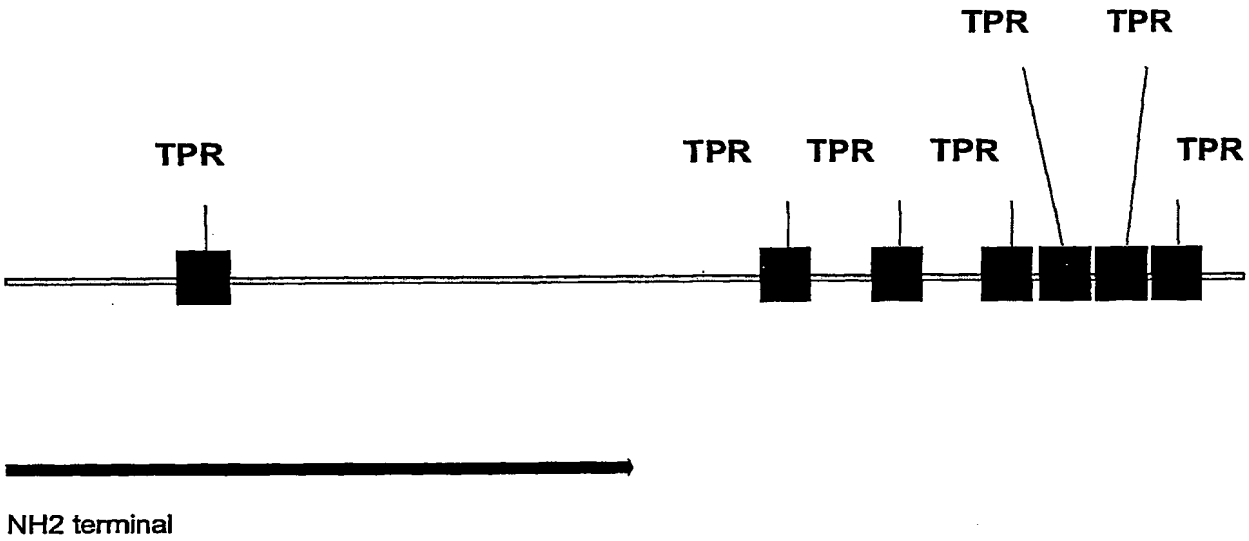
SEQ ID NO: 127, seqüência parcial de ácidos nucléicos Vitvi_CYP90B de Vitis vinifera agrupamento de DT13384, DT011548

AGACTTCATGGAGCAACACATATCAAGGTTCGGAGAAATCTACAAGTCCAATCTGTTTGGCG
AGCCAACCATAGTCTCAGCGGATTCTGGGCTGAACAGATTCATACTACAGAACGAAGGAAAA
TTGTTTGAATGCAGCTATCCCAGAAGCATAGGTGGAATTCTTGGGAAATGGTCCATGCTGGT
TTTAGTTGGAGACATGCATAGAGACATGAGAACCATCTCCCTCAACTTCTTGAGCCATGGCA
GGCTTAGGACTCATCTCCTACCAGAGGTGGTGAAGCACACTTTGCTTGTTCTAAGCTCTTGG
AAGGAGAATTGTACATTTTCTGCTCAAGATGAAGCTAAGAAGTTTACCTTCAATCTGATGGC
AAAACATATCATGAGCTTGGATCCCGGAAAGCCGGAGACTGAGCAGCTTAAGAAAGAGTATG
TTACTTTCATGAAAGGAGTGGTATCTGCTCCTTTGAATTTCCCTGGAACTGCATACAGAAAA
GCCCTACAGTCTCGGTCCACCATCTTAAAATTTATCGAGCTGAAGATGGAAGAGAGGATTCA
GAAACTGAGGGGAGGAGGGTTTGAGAATATGGAGGATGACGATCTTCTTGGATGGGTGCTGA
AGCATTCCAACCTTTCAACTGAGCAAATCCTGGACTTGGTACTGAGCTTGCTCTTTGCTGGC
CATGAAACTTCATCAGTGGCAATAGCTTTAGCTATCTACTTCTTGGAAGGCTGTCCCAACGC
CGTT

SEQ ID NO: 128, polipeptídeo parcial traduzido Vitvi_CYP90B de Vitis vinifera

DFMEQHISRFGEIYKSNLFGEPTIVSADSGLNRFILQNEGKLFECSYPRSIGGILGKWSMLV
LVGDMHRDMRTISLNFLSHGRLRTHLLPEVVKHTLLVLSSWKENCTFSAQDEAKKFTFNLMA
KHIMSLDPGKPETEQLKKEYVTFMKGVVSAPLNFPGTAYRKALQSRSTILKFIELKMEERIQ
KLRGGGFENMEDDDLLGWVLKHSNLSTEQILDLVLSLLFAGHETSSVAIALAIYFLEGCPNA

FIGURA 15 (Cont.)



Comprimento completo
Arath_CDC27B Hobbit

FIGURA 16

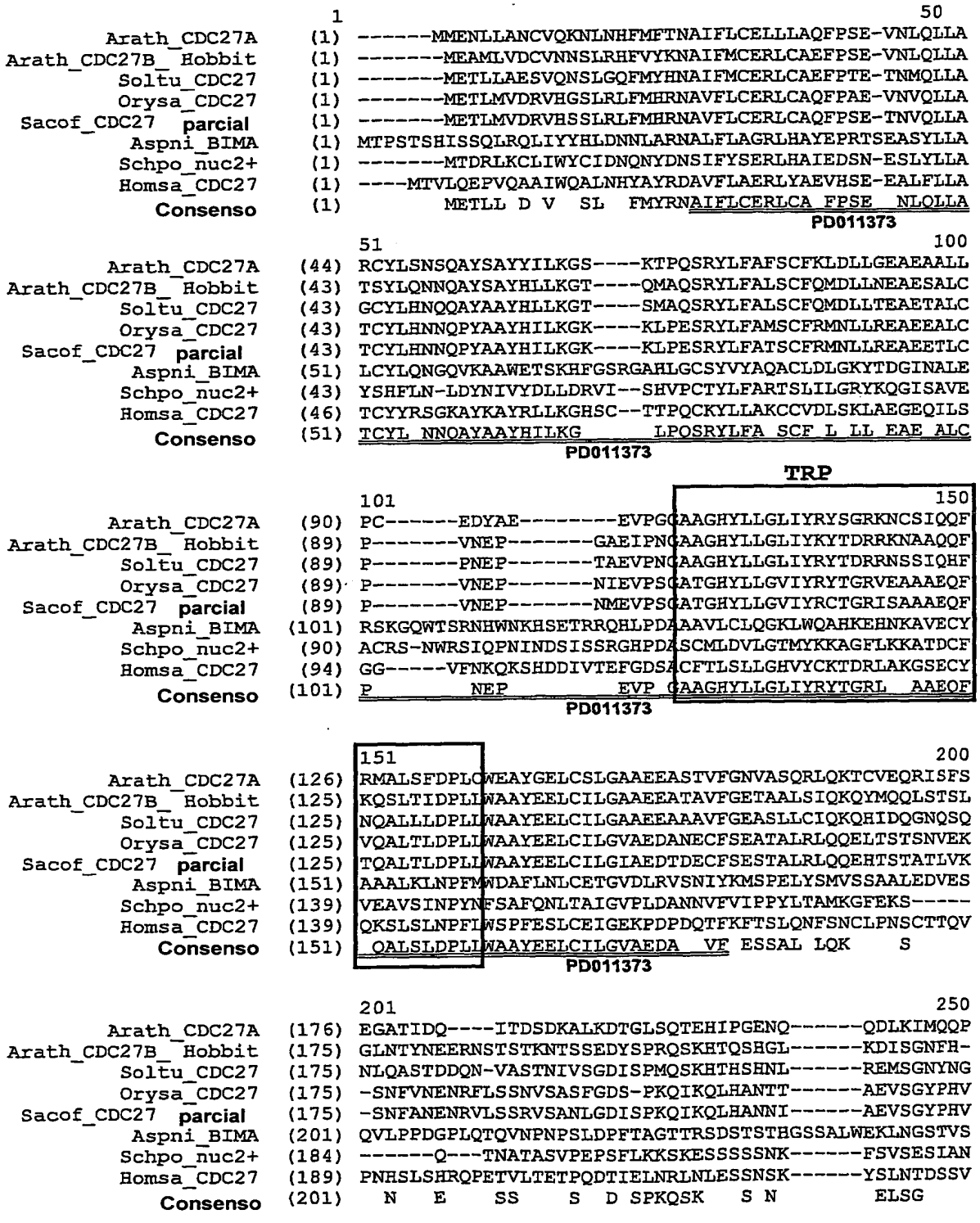

1 50
Arath_CDC27A (1) ------MMENLLANCVQKNLNHFMFTNAIFLCELLLAQFPSE-VNLQLLA
Arath_CDC27B_Hobbit (1) -------MEAMLVDCVNNSLRHFVYKNAIFMCERLCAEFPSE-VNLQLLA
Soltu_CDC27 (1) -------METLLAESVQNSLGQFMYHNAIFMCERLCAEFPTE-TNMQLLA
Orysa_CDC27 (1) -------METLMVDRVHGSLRLFMHRNAVFLCERLCAQFPAE-VNVQLLA
Sacof_CDC27 parcial (1) -------METLMVDRVHSSLRLFMHRNAVFLCERLCAQFPSE-TNVQLLA
Aspni_BIMA (1) MTPSTSHISSQLRQLIYYHLDNNLARNALFLAGRLHAYEPRTSEASYLLA
Schpo_nuc2+ (1) -------MTDRLKCLIWYCIDNQNYDNSIFYSERLHAIEDSN-ESLYLLA
Homsa_CDC27 (1) ----MTVLQEPVQAAIWQALNHYAYRDAVFLAERLYAEVHSE-EALFLLA
Consenso (1) METLL D V SL FMYRNAIFLCERLCA FPSE NLQLLA
PD011373
51 100
Arath_CDC27A (44) RCYLSNSQAYSAYYILKGS----KTPQSRYLFAFSCFKLDLLGEAEAALL
Arath_CDC27B_Hobbit (43) TSYLQNNQAYSAYHLLKGT----QMAQSRYLFALSCFQMDLLNEAESALC
Soltu_CDC27 (43) GCYLHNQQAYAAYHLLKGT----SMAQSRYLFALSCFQMDLLTEAETALC
Orysa_CDC27 (43) TCYLHNNQPYAAYHILKGK----KLPESRYLFAMSCFRMNLLREAEEALC
Sacof_CDC27 parcial (43) TCYLHNNQPYAAYHILKGK----KLPESRYLFATSCFRMNLLREAEETLC
Aspni_BIMA (51) LCYLQNGQVKAAWETSKHFGSRGAHLGCSYVYAQACLDLGKYTDGINALE
Schpo_nuc2+ (43) YSHFLN-LDYNIVYDLLDRVI--SHVPCTYLFARTSLILGRYKQGISAVE
Homsa_CDC27 (46) TCYYRSGKAYKAYRLLKGHSC--TTPQCKYLLAKCCVDLSKLAEGEQILS
Consenso (51) TCYL NNQAYAAYHILKG LPQSRYLFA SCF L LL EAE ALC
PD011373
TRP
101 150
Arath_CDC27A (90) PC------EDYAE--------EVPGCAAGHYLLGLIYRYSGRKNCSIQQF
Arath_CDC27B_Hobbit (89) P-------VNEP-------GAEIPNCAAGHYLLGLIYKYTDRRKNAAQQF
Soltu_CDC27 (89) P-------PNEP-------TAEVPNCAAGHYLLGLIYRYTDRRNSSIQHF
Orysa_CDC27 (89) P-------VNEP-------NIEVPSCATGHYLLGVIYRYTGRVEAAAEQF
Sacof_CDC27 parcial (89) P-------VNEP-------NMEVPSCATGHYLLGVIYRCTGRISAAAEQF
Aspni_BIMA (101) RSKGQWTSRNHWNKHSETRRQHLPDAAAVLCLQGKLWQAHKEHNKAVECY
Schpo_nuc2+ (90) ACRS-NWRSIQPNINDSISSRGHPDASCMLDVLGTMYKKAGFLKKATDCF
Homsa_CDC27 (94) GG-----VFNKQKSHDDIVTEFGDSACFTLSLLGHVYCKTDRLAKGSECY
Consenso (101) P NEP EVP CAAGHYLLGLIYRYTGRL AAEQF
PD011373
151 200
Arath_CDC27A (126) RMALSFDPLCWEAYGELCSLGAAEEASTVFGNVASQRLQKTCVEQRISFS
Arath_CDC27B_Hobbit (125) KQSLTIDPLLWAAYEELCILGAAEEATAVFGETAALSIQKQYMQQLSTSL
Soltu_CDC27 (125) NQALLLDPLLWAAYEELCILGAAEEAAAVFGEASLLCIQKQHIDQGNQSQ
Orysa_CDC27 (125) VQALTLDPLLWAAYEELCILGVAEDANECFSEATALRLQQELTSTSNVEK
Sacof_CDC27 parcial (125) TQALTLDPLLWAAYEELCILGIAEDTDECFSESTALRLQQEHTSTATLVK
Aspni_BIMA (151) AAALKLNPFMWDAFLNLCETGVDLRVSNIYKMSPELYSMVSSAALEDVES
Schpo_nuc2+ (139) VEAVSINPYNFSAFQNLTAIGVPLDANNVFVIPPYLTAMKGFEKS-----
Homsa_CDC27 (139) QKSLSLNPFLWSPFESLCEIGEKPDPDQTFKFTSLQNFSNCLPNSCTTQV
Consenso (151) QALSLDPLLWAAYEELCILGVAEDA VF ESSAL LQK S
PD011373
201 250
Arath_CDC27A (176) EGATIDQ----ITDSDKALKDTGLSQTEHIPGENQ------QDLKIMQQP
Arath_CDC27B_Hobbit (175) GLNTYNEERNSTSTKNTSSEDYSPRQSKHTQSHGL------KDISGNFH-
Soltu_CDC27 (175) NLQASTDDQN-VASTNIVSGDISPMQSKHTHSHNL------REMSGNYNG
Orysa_CDC27 (175) -SNFVNENRFLSSNVSASFGDS-PKQIKQLHANTT------AEVSGYPHV
Sacof_CDC27 parcial (175) -SNFANENRVLSSRVSANLGDISPKQIKQLHANNI------AEVSGYPHV
Aspni_BIMA (201) QVLPPDGPLQTQVNPNPSLDPFTAGTTRSDSTSTHGSSALWEKLNGSTVS
Schpo_nuc2+ (184) ------Q---TNATASVPEPSFLKKSKESSSSSNK------FSVSESIAN
Homsa_CDC27 (189) PNHSLSHRQPETVLTETPQDTIELNRLNLESSNSK------YSLNTDSSV
Consenso (201) N E SS S D SPKQSK S N ELSG

FIGURA 17

```
                          251                                                300
        Arath_CDC27A  (216) GDIP-------PNTDRQLSTNGWDLNTPSPVLLQVMDALPPLLLKNMRRP
Arath_CDC27B_ Hobbit  (218) ----------SHGVNGGVSNMSFYNTPSPVAAQLSGIAPPPLFRNFQPA
         Soltu_CDC27  (218) A-AA-------IQNLGGVSTNMSFYNTPSPMASQLSGVVPPPVCRNFQQN
         Orysa_CDC27  (217) KSTA-------LHMQNGAPSNLSQFDTPSPTSTQVSGIAPPPLFRNMHAY
 Sacof_CDC27 parcial  (218) RPTA-------LHVQNSSTSNVAQFDTPSPTAAQTSSIMPPPLFRNVHA-
          Aspni_BIMA  (251) VASSGTGPHLPREGMETPGGQSSESDDPRVTNGNGTDVFEPPLAPAKKNR
         Schpo_nuc2+  (219) SYS----------NSS--ISAFTKWFDRVDASELPGSEKERHQSLKLQSQ
         Homsa_CDC27  (233) SYIDSAVISPDTVPLGTGTSILSKQVQNKPKTGRSLLGGPAALSPLTPSF
           Consenso   (251)   SA          L   SSNLS F TPSP AAQLS I PPPL RNM A

                          301                                                350
        Arath_CDC27A  (259) -------------------------------AVEGSLMSVHGVRVRRRNFF
Arath_CDC27B_ Hobbit  (257) VANPNSLITDSSP--KSTVNSTLQAPRRKFVDEGKLRKISGRLFSDSG--
         Soltu_CDC27  (260) GTNASVAGADNSP--RATVNSTIQAPRRKFVDEGKLRKISGRLFSDSG--
         Orysa_CDC27  (260) QN--T-AGGNAPS--KPKVNAPNLTLRRKYIDEAGLKKVSGRLFNQSSDS
 Sacof_CDC27 parcial  (261) ----------------------------IQIQIPGVWREWYRLF------
          Aspni_BIMA  (301) TIQTIGGDHPMDPPPKMRPTGIRPRTRTKFESDEGHTERDAGMGHRLG--
         Schpo_nuc2+  (257) ----------------------------SQTSKNLLAFNDAQKADS----
         Homsa_CDC27  (283) GILPLETPSPGDG--SYLQNYTNTPPVIDVPSTGAPSKKSVARIGQTG--
           Consenso   (301)                K   N       R KFI EGGL KISGRLF  SG

                          351                                                400
        Arath_CDC27A  (279) SEELSAEAQEESGRRRSARIAARKKNPMSQSFGK------DSHWLHLSPS
Arath_CDC27B_ Hobbit  (303) -PRRSSRLSADSGANINSSVATVSGN-VNNASKYLG-----GSKLSSLAL
         Soltu_CDC27  (306) -PRRNSRLAGESTGNTNSNVSGASGNGTIHSSKYYG-----SSKLSSMTL
         Orysa_CDC27  (305) VPRRSARLSRDTTINSNSNISQFGGNGTDHSSGKLRVNSSTPSKLCSTAL
 Sacof_CDC27 parcial  (277) ----------------------VR--EIAS------KLVHTIKMVLTTI
          Aspni_BIMA  (349) ---DRKRTVSGQVAHPSVPHSTDQGVGQRRSVRLFNQIKPSTNKISSTAL
         Schpo_nuc2+  (275) ----------------------------NN--R------DTSLKSHFVEP
         Homsa_CDC27  (329) --TKSVFSQSGNSREVTPILAQTQSSGPQTSTTPQVLSPTITSPPNALPR
           Consenso   (351)    RSAR   DS     S IA   GNG   SSK        SSKL ST L

                          401                                                450
        Arath_CDC27A  (323) ESNYAPSLSSMIGKCRIQSSKEVIPDTVTLNDPATTSGQSVSDIGSSVDD
Arath_CDC27B_ Hobbit  (346) RSVTLRKGHSWANENMDEGVR----GEPFDDSRP--NTASTTGSMASNDQ
         Soltu_CDC27  (350) RSMTSRKAQSWATENYGEGTR----NDISNDSRLNMTMSHPSGDARPLEQ
         Orysa_CDC27  (355) RSVQVRKGKPQATENFDEG----DYHFDMDDSVTSTSSSTSIVDGRYPEQ
 Sacof_CDC27 parcial  (296) RSVQVRKGKPRATENFDEGSRYEVIDEMWTDNISGTSSSVSTADGRSFEQ
          Aspni_BIMA  (396) GVKEGREVKKVRTTGNKARTTTSSNVGRVVSGNN-RRHAGEIHDGDSKEY
         Schpo_nuc2+  (289) RTQALRPGARLTYKLREAR----------------SSKRGESTPQSFREE
         Homsa_CDC27  (377) RSSRLFTSDSSTTKENSKKLKMKFPPKIPNRKTKSKTNKGGITQPNINDS
           Consenso   (401) RSV LRKG S ATEN DEGSR      I  D    TS S  S DG S EQ

                          451                                                500
        Arath_CDC27A  (373) EEKSNPSESSPDRFSLIS-------------GISEVLSLLKILGDGHRHL
Arath_CDC27B_ Hobbit  (390) EDETMSIGGIAMSSQTITIG-------------VSEILNLLRTLGEGCRLS
         Soltu_CDC27  (396) ERPRTSASGVNVSSTSIPLS-------------ASEILALFRFLGEGYRLS
         Orysa_CDC27  (401) EK----SERVLSQDSKLAIG-------------IRELMALLRTLGEGYRLS
 Sacof_CDC27 parcial  (346) DK----AERILLQDSKLALG-------------IREILGLVRTLGEGCRLS
          Aspni_BIMA  (445) RGTSSTSNGSQNASSKLAISER----TKSVEALAWILDLFFKIASGYFCL
         Schpo_nuc2+  (323) DN-------------------------------NLMELLKLFGKGVYLL
         Homsa_CDC27  (427) LEITKLDSSIISEGKISTITPQIQAFNLQKAAAAGLMSLLREMGKGYLAL
           Consenso   (451) E  S  S  I    S IAI             ISEILALLR LGEGYRLS
```

FIGURA 17 (Cont.)

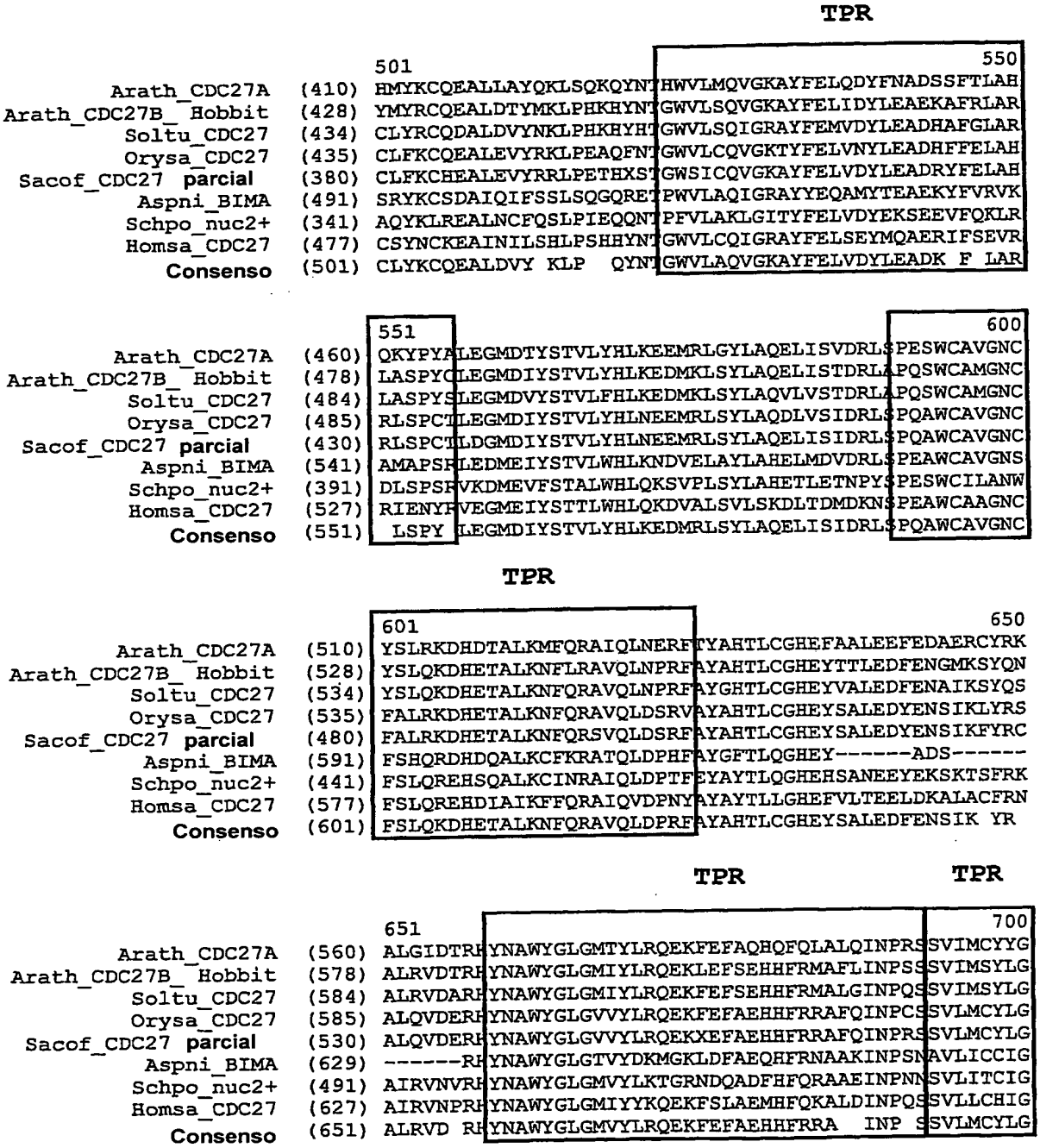


FIGURA 17 (Cont.)

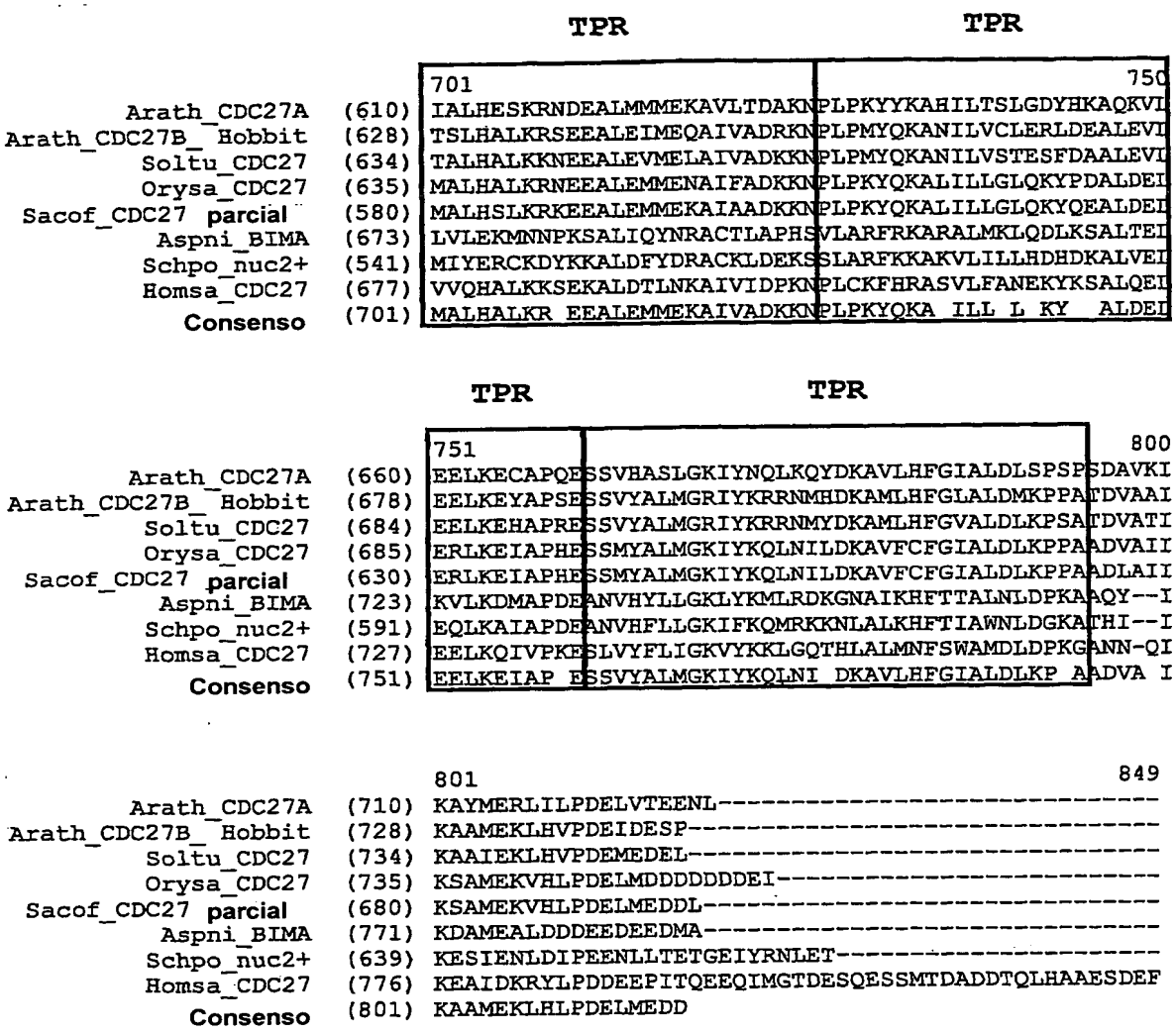


FIGURA 17 (Cont.)

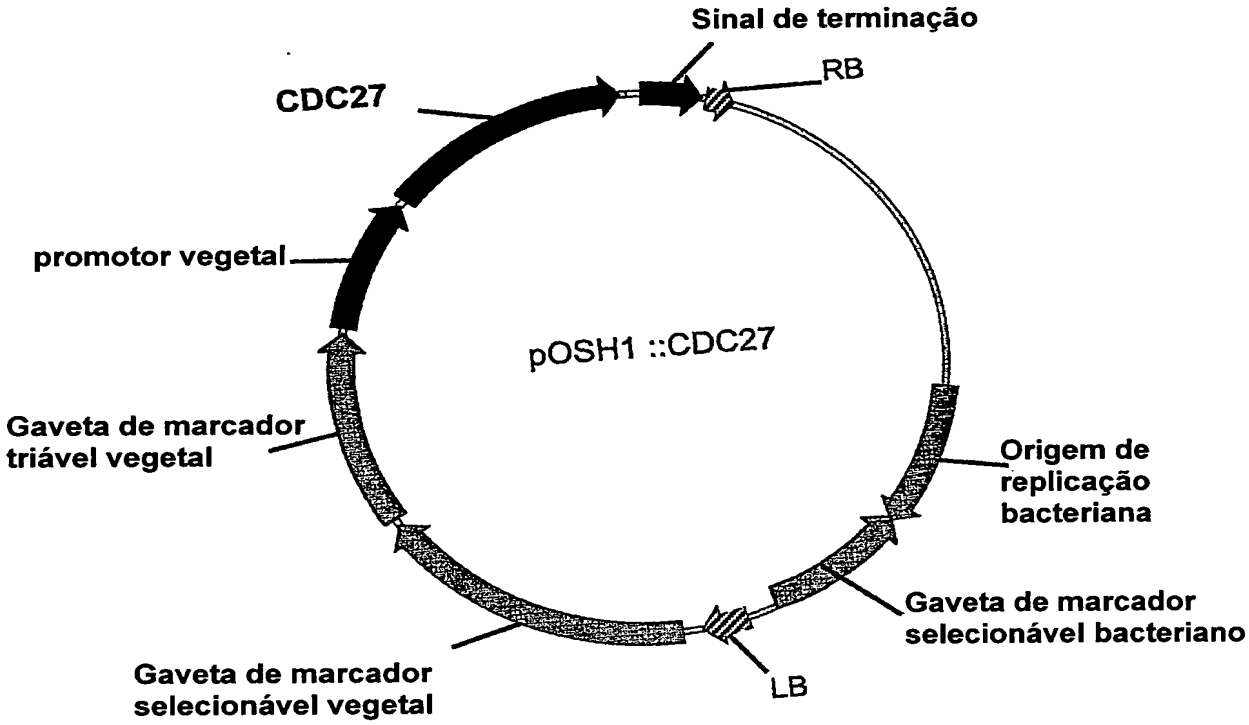

Sinal de terminação
RB
CDC27
promotor vegetal
pOSH1 ::CDC27
Gaveta de marcador triável vegetal
Origem de replicação bacteriana
Gaveta de marcador selecionável bacteriano
LB
Gaveta de marcador selecionável vegetal

FIGURA 18

#	TC 895803	Funcão provável
1	A.nidulans |NP363957	bimA [Aspergillus nidulans]
2	Arabidopsis |TC264528	proteína tipo CDC27/NUC2 [Arabidopsis thaliana] proteína HOBBIT
3	Arabidopsis |NP043602	proteína de ligação de DNA (CDC27SH) isolog
4	C.elegans |TC134668	(Q9N593) proteína 1 defito de transição de metáfase para anáfase
5	C.intestinalis |TC58032	Ciona intestinalis cDNA, clone:cicl059p05, seqüência completa de inserto
6	Drosophila |TC172403	CG8610-PA [Drosophila melanogaster]
7	Frog |TC288783	Xenopus laevis NBRP cDNA clone:XL434l12ex, 5 prime end
8	G.gallus |TC210379	(Q5ZK91) Proteína hipotética
9	Abelha TC10375	(Q7QJW4) EbiP9284 (Fragmento)
10	Humana|THC2241027	proteína CDC27 [Homo sapiens]
11	Humana|THC2235238	H-NUC=[Homo sapiens] [824 aa] homólogo 27 de proteína de ciclo de divisão celular; homólogo de nuc2
12	M.grisea |NP870483	proteína hipotética MG06292.4[Magnaporthe grisea 70-15]
13	Milho |TC251983	(Q8LGU6) proteína HOBBIT
14	Medaka |TC32541	proteína CDC27
15	Medicago |TC104997	(Q8LGU6) proteína HOBBIT
16	Mosquito |NP1265145	ENSANGP00000021543 [Anopheles gambiae str. PEST]
17	camund|TC1480020	unidade completa de repetição de DNA ribossômico humano
18	N.crassa |NP624892	proteína hipotética [Neurospora crassa]
19	porco|TC212778	unidade completa de repetição de DNA ribossômico humano
20	batata |TC119704	proteína da família de ciclo de divisão celular proteína da família CDC baixa similaridade com SP|P30260|
21	R.trout |TC81012	(Q7ZWD5) ciclo de divisão celular 27
22	rato|TC539269	proteína CDC27
23	arroz|NP1231881	proteína tipo aminotransferase [Oriza sativa (grupo de cultivar japonica)
24	S.cerevisiae |TC13353	(P38042) proteína de controle de divisão celular 27
25	S.mansoni |TC11451	proteína CDC27Hs (homólogo 27 de proteína de ciclo de divisão celular) (H-NUC) [Humano]
26	S.officinarum |TC52080	(Q8LGU6) proteína HOBBIT
27	S.pombe |TC6937	proteína tipo de estrutura nuclear p76 [Schizosaccharomyces pombe]
28	soja |TC210267	(Q8LGU6) proteína HOBBIT
29	trigo |TC240260	(Q69XV2) HOBBIT putativo
30	Zebrafish |TC291951	(Q7ZWD5) ciclo de divisão celular 27

Lista experimental de ortólogo/parálogo de TIGR, TC895803 experimental (a partir de http://www.tigr.org/tigr-scripts/tgi/ego/ego_report.pl?ego=895803)

FIGURA 19

SEQ ID NO: 129, seqüência de ácidos nucléicos CDC27B/Hobbit truncada 5' de Arabidopsis thaliana AC006081

ATGCAACAACTGTCAACTTCCCTCGGCTTAAACACTTACAACGAGGAACGTAATTCAACTTC
TACTAAAAACACGAGTTCTGAAGATTATAGTCCAAGGCAGTCTAAACACACACAAAGCCATG
GCCTTAAAGATATCTCCGGAAATTTCCATTCTCATGGAGTTAATGGAGGTGTTTCGAACATG
TCATTCTATAATACGCCTTCGCCAGTGGCTGCACAGCTATCCGGTATAGCTCCACCACCACT
TTTCCGGAATTTTCAGCCAGCTGTTGCAAACCCAAACTCCCTTATTACTGACAGTTCTCCAA
AGTCCACTGTTAACTCTACTCTTCAAGCACCTAGAAGAAAGTTTGTAGATGAAGGAAAGTTA
CGTAAGATTTCTGGCAGACTATTTTCTGATTCTGGTCCACGACGGAGTTCAAGACTGTCTGC
TGATTCAGGGGCAAACATTAATTCAAGTGTTGCAACAGTAAGCGGAAATGTGAACAACGCTT
CCAAGTATTTGGGAGGTTCTAAATTGAGTTCTTTGGCACTTCGTTCTGTAACACTTCGGAAG
GGACACTCCTGGGCAAATGAAAACATGGATGAAGGGGTCCGTGGGGAACCTTTTGATGATTC
AAGGCCTAATACTGCCTCAACGACTGGTTCTATGGCTTCCAATGATCAAGAAGACGAAACAA
TGTCGATTGGTGGCATAGCAATGAGTTCTCAAACAATCACAATTGGTGTTTCGGAAATTTTA
AACCTCCTTAGGACACTCGGAGAAGGGTGTAGACTTTCATACATGTACAGGTGTCAGGAGGC
ACTGGATACGTATATGAAACTTCCACATAAGCATTATAATACAGGATGGGTTCTTTCCCAGG
TCGGGAAAGCATACTTTGAACTAATTGACTATTTAGAGGCTGAAAAGGCATTCCGTCTTGCC
CGTCTGGCTTCTCCTTATTGCTTAGAAGGAATGGATATATACTCTACGGTCCTCTATCATTT
GAAGGAAGACATGAAGCTGAGTTACTTGGCTCAGGAACTAATATCAACCGATCGCTTAGCTC
CTCAATCTTGGTGTGCTATGGGAAATTGCTATAGCTTGCAAAAGGACCATGAGACCGCACTG
AAGAATTTCCTACGAGCTGTTCAACTGAATCCAAGATTTGCATATGCACATACCTTATGTGG
CCACGAATACACAACTCTTGAGGATTTTGAGAACGGAATGAAAAGTTACCAAAACGCACTTC
GTGTAGATACAAGACACTACAACGCATGGTACGGGCTTGGAATGATATATCTACGCCAAGAG
AAGTTAGAGTTCTCAGAGCATCACTTCAGAATGGCTTTCCTAATAAACCCGAGTTCCTCTGT
TATAATGTCTTATTTAGGGACATCTTTGCATGCCTTGAAGAGAAGTGAGGAAGCACTAGAGA
TAATGGAGCAAGCCATAGTAGCAGATAGAAAAAACCCTCTTCCAATGTACCAGAAAGCTAAC
ATACTTGTCTGCTTAGAAAGATTAGATGAAGCTCTAGAAGTTCTTGAGGAGCTCAAAGAGTA
TGCGCCTTCAGAGAGCAGCGTTTACGCTTTAATGGGCAGGATCTATAAGCGGCGAAACATGC
ACGATAAAGCCATGCTTCATTTCGGTCTAGCTTTAGATATGAAACCGCCTGCAACTGACGTT
GCTGCAATAAAGGCTGCAATGGAGAAATTGCATGTTCCAGATGAGATCGATGAGAGCCCGTG
A

SEQ ID NO: 130, seqüência de aminoácidos CDC27B/Hobbit truncada N-terminal de Arabidopsis thaliana

MQQLSTSLGLNTYNEERNSTSTKNTSSEDYSPRQSKHTQSHGLKDISGNFHSHGVNGGVSNM
SFYNTPSPVAAQLSGIAPPPLFRNFQPAVANPNSLITDSSPKSTVNSTLQAPRRKFVDEGKL
RKISGRLFSDSGPRRSSRLSADSGANINSSVATVSGNVNNASKYLGGSKLSSLALRSVTLRK
GHSWANENMDEGVRGEPFDDSRPNTASTTGSMASNDQEDETMSIGGIAMSSQTITIGVSEIL
NLLRTLGEGCRLSYMYRCQEALDTYMKLPHKHYNTGWVLSQVGKAYFELIDYLEAEKAFRLA
RLASPYCLEGMDIYSTVLYHLKEDMKLSYLAQELISTDRLAPQSWCAMGNCYSLQKDHETAL
KNFLRAVQLNPRFAYAHTLCGHEYTTLEDFENGMKSYQNALRVDTRHYNAWYGLGMIYLRQE
KLEFSEHHFRMAFLINPSSSVIMSYLGTSLHALKRSEEALEIMEQAIVADRKNPLPMYQKAN
ILVCLERLDEALEVLEELKEYAPSESSVYALMGRIYKRRNMHDKAMLHFGLALDMKPPATDV
AAIKAAMEKLHVPDEIDESP

FIGURA 20

SEQ ID NO:131 seqüência de aminoácidos aj487669 de ácido nucleico CDC27B/Hobbit de comprimento completo de Arabidopsis thaliana

ATGGAAGCTATGCTTGTGGACTGTGTAAACAACAGTCTTCGTCATTTTGTCTACAAAAATGC
TATTTTCATGTGCGAGCGTCTCTGCGCTGAGTTTCCTTCTGAGGTTAATTTGCAGCTATTAG
CCACCAGCTACCTGCAGAATAATCAAGCTTACAGTGCATATCATCTGCTAAAGGGAACACAA
ATGGCTCAGTCCCGATACTTGTTCGCATTATCATGCTTCCAGATGGACCTTCTCAATGAAGC
TGAATCTGCACTCTGCCCTGTTAATGAACCTGGTGCGGAGATCCCAAATGGTGCAGCAGGCC
ATTACCTTCTTGGACTTATTTACAAGTATACTGATAGAAGGAAGAATGCTGCTCAACAATTT
AAACAGTCCTTGACAATAGACCCTCTACTTTGGGCTGCATATGAGGAATTATGTATATTAGG
TGCTGCTGAGGAAGCAACTGCAGTTTTTGGTGAAACAGCTGCTCTCTCCATTCAAAAGCAGT
ATATGCAACAACTGTCAACTTCCCTCGGCTTAAACACTTACAACGAGGAACGTAATTCAACT
TCTACTAAAAACACGAGTTCTGAAGATTATAGTCCAAGGCAGTCTAAACACACACAAAGCCA
TGGCCTTAAAGATATCTCCGGAAATTTCCATTCTCATGGAGTTAATGGAGGTGTTTCGAACA
TGTCATTCTATAATACGCCTTCGCCAGTGGCTGCACAGCTATCCGGTATAGCTCCACCACCA
CTTTTCCGGAATTTTCAGCCAGCTGTTGCAAACCCAAACTCCCTTATTACTGACAGTTCTCC
AAAGTCCACTGTTAACTCTACTCTTCAAGCACCTAGAAGAAAGTTTGTAGATGAAGGAAAGT
TACGTAAGATTTCTGGCAGACTATTTTCTGATTCTGGTCCACGACGGAGTTCAAGACTGTCT
GCTGATTCAGGGGCAAACATTAATTCAAGTGTTGCAACAGTAAGCGGAAATGTGAACAACGC
TTCCAAGTATTTGGGAGGTTCTAAATTGAGTTCTTTGGCACTTCGTTCTGTAACACTTCGGA
AGGGACACTCCTGGGCAAATGAAAACATGGATGAAGGGGTCCGTGGGGAACCTTTTGATGAT
TCAAGGCCTAATACTGCCTCAACGACTGGTTCTATGGCTTCCAATGATCAAGAAGACGAAAC
AATGTCGATTGGTGGCATAGCAATGAGTTCTCAAACAATCACAATTGGTGTTTCGGAAATTT
TAAACCTCCTTAGGACACTCGGAGAAGGGTGTAGACTTTCATACATGTACAGGTGTCAGGAG
GCACTGGATACGTATATGAAACTTCCACATAAGCATTATAATACAGGATGGGTTCTTTCCCA
GGTCGGGAAAGCATACTTTGAACTAATTGACTATTTAGAGGCTGAAAAGGCATTCCGTCTTG
CCCGTCTGGCTTCTCCTTATTGCTTAGAAGGAATGGATATATACTCTACGGTCCTCTATCAT
TTGAAGGAAGACATGAAGCTGAGTTACTTGGCTCAGGAACTAATATCAACCGATCGCTTAGC
TCCTCAATCTTGGTGTGCTATGGGAAATTGCTATAGCTTGCAAAAGGACCATGAGACCGCAC
TGAAGAATTTCCTACGAGCTGTTCAACTGAATCCAAGATTTGCATATGCACATACCTTATGT
GGCCACGAATACACAACTCTTGAGGATTTTGAGAACGGAATGAAAAGTTACCAAAACGCACT
TCGTGTAGATACAAGACACTACAACGCATGGTACGGGCTTGGAATGATATATCTACGCCAAG
AGAAGTTAGAGTTCTCAGAGCATCACTTCAGAATGGCTTTCCTAATAAACCCGAGTTCCTCT
GTTATAATGTCTTATTTAGGGACATCTTTGCATGCCTTGAAGAGAAGTGAGGAAGCACTAGA
GATAATGGAGCAAGCCATAGTAGCAGATAGAAAAAACCCTCTTCCAATGTACCAGAAAGCTA
ACATACTTGTCTGCTTAGAAAGATTAGATGAAGCTCTAGAAGTTCTTGAGGAGCTCAAAGAG
TATGCGCCTTCAGAGAGCAGCGTTTACGCTTTAATGGGCAGGATCTATAAGCGGCGAAACAT
GCACGATAAAGCCATGCTTCATTTCGGTCTAGCTTTAGATATGAAACCGCCTGCAACTGACG
TTGCTGCAATAAAGGCTGCAATGGAGAAATTGCATGTTCCAGATGAGATCGATGAGAGCCCG
TGA

SEQ ID NO : 132
seqüência de aminoácidos CDC27B/Hobbit de comprimento completo de Arabidopsis thaliana

MEAMLVDCVNNSLRHFVYKNAIFMCERLCAEFPSEVNLQLLATSYLQNNQAYSAYHLLKGTQ
MAQSRYLFALSCFQMDLLNEAESALCPVNEPGAEIPNGAAGHYLLGLIYKYTDRRKNAAQQF
KQSLTIDPLLWAAYEELCILGAAEEATAVFGETAALSIQKQYMQQLSTSLGLNTYNEERNST
STKNTSSEDYSPRQSKHTQSHGLKDISGNFHSHGVNGGVSNMSFYNTPSPVAAQLSGIAPPP
LFRNFQPAVANPNSLITDSSPKSTVNSTLQAPRRKFVDEGKLRKISGRLFSDSGPRRSSRLS

FIGURA 20 (Cont.)

ADSGANINSSVATVSGNVNNASKYLGGSKLSSLALRSVTLRKGHSWANENMDEGVRGEPFDD
SRPNTASTTGSMASNDQEDETMSIGGIAMSSQTITIGVSEILNLLRTLGEGCRLSYMYRCQE
ALDTYMKLPHKHYNTGWVLSQVGKAYFELIDYLEAEKAFRLARLASPYCLEGMDIYSTVLYH
LKEDMKLSYLAQELISTDRLAPQSWCAMGNCYSLQKDHETALKNFLRAVQLNPRFAYAHTLC
GHEYTTLEDFENGMKSYQNALRVDTRHYNAWYGLGMIYLRQEKLEFSEHHFRMAFLINPSSS
VIMSYLGTSLHALKRSEEALEIMEQAIVADRKNPLPMYQKANILVCLERLDEALEVLEELKE
YAPSESSVYALMGRIYKRRNMHDKAMLHFGLALDMKPPATDVAAIKAAMEKLHVPDEIDESP

SEQ ID N: 133 seqüência de ácidos nucléicos CDC27a de Arabidopsis thaliana Sequencia NM_112503.2

ATGATGGAGAATCTACTGGCGAATTGTGTCCAGAAAAACCTTAACCATTTTATGTTCACCAA
TGCTATCTTCCTTTGCGAACTTCTTCTCGCCCAATTTCCATCTGAGGTGAACCTGCAATTGT
TAGCCAGGTGTTACTTGAGTAACAGTCAAGCTTATAGTGCATATTATATCCTTAAAGGTTCA
AAAACGCCTCAGTCTCGGTATTTATTTGCATTCTCATGCTTTAAGTTGGATCTTCTTGGAGA
GGCTGAAGCTGCATTGTTGCCCTGTGAAGATTATGCTGAAGAAGTTCCTGGTGGTGCAGCTG
GGCATTATCTTCTTGGTCTTATATATAGATATTCTGGGAGGAAGAACTGTTCAATACAACAG
TTTAGGATGGCATTGTCATTTGATCCATTGTGTTGGGAAGCATATGGAGAACTTTGTAGTTT
AGGTGCCGCTGAAGAAGCCTCAACAGTTTTCGGGAATGTTGCTTCCCAGCGTCTTAAAACTT
GTGTAGAACAAAGAATAAGCTTCTCAGAAGGAGCAACCATAGACCAGATTACAGATTCTGAT
AAGGCCTTAAAAGATACAGGTTTATCGCAAACAGAACACATTCCAGGAGAGAACCAACAAGA
TCTGAAAATTATGCAGCAGCCTGGAGATATTCCACCAAATACTGACAGGCAACTTAGTACAA
ACGGATGGGACTTGAACACACCTTCTCCAGTGCTTTTACAGGTAATGGATGCTCCACCGCCT
CTGCTTCTTAAGAATATGCGTCGTCCAGCAGTGGAAGGATCTTTGATGTCTGTACATGGAGT
GCGTGTGCGTCGAAGAAACTTTTTTAGTGAAGAATTGTCAGCAGAGGCTCAAGAAGAATCTG
GGCGCCGCCGTAGTGCTAGAATAGCAGCAAGGAAAAAGAATCCTATGTCGCAGTCATTTGGA
AAAGATTCCCATTGGTTACATCTTTCACCTTCCGAGTCAAACTATGCACCTTCTCTTTCCTC
GATGATTGGAAAATGCAGAATCCAAAGCAGCAAAGAAGCGATTCCTGATACCGTTACTCTAA
ATGATCCAGCAACGACGTCAGGCCAGTCTGTAAGTGACACTGGAAGCTCTGTTGATGATGAG
GAAAAGTCAAATCCTAGTGAATCTTCCCCGGATCGTTTCAGCCTTATTTCTGGAATTTCAGA
AGTGCTAGGCATTCTGAAAATTCTTGGAGATGGCCACAGGCATTTACATATGTACAAGTGTC
AGGAAGCTTTGTTGGCATATCAAAAGCTATCTCAGAAACAATACAATACACACTGGGTTCTC
ATGCAGGTTGGAAAAGCATATTTTGAGCTACAAGACTACTTCAACGCTGACTCTTCCTTTAC
TCTTGCTCATCAAAAGTATCCTTATGCTTTGGAAGGAATGGATACATACTCCACTGTTCTTT
ATCACCTGAAAGAAGAGATGAGGTTGGGCTATCTGGCTCAGGAACTGATTTCAGTTGATCGC
CTGTCTCCAGAATCCTGGTGTGCAGTTGGGAACTGTTACAGTTTGCGTAAGGATCATGATAC
TGCTCTCAAAATGTTTCAGAGAGCTATCCAACTGAATGAAAGATTCACATATGCACATACCC
TTTGTGGCCACGAGTTTGCCGCATTGGAAGAATTCGAGGATGCAGAGAGATGCTACCGGAAG
GCTCTGGGCATAGATACGAGACACTATAATGCATGGTACGGTCTTGGAATGACCTATCTTCG
TCAGGAGAAATTCGAGTTTGCGCAGCATCAATTTCAACTGGCTCTCCAAATAAATCCAAGAT
CTTCAGTCATCATGTGTTACTATGGAATTGCTTTGCATGAGTCAAAGAGAAACGATGAGGCG
TTGATGATGATGGAGAAGGCTGTACTCACTGATGCAAAGAATCCGCTCCCCAAGTACTACAA
GGCTCACATATTAACCAGCCTAGGTGATTATCACAAAGCACAGAAAGTTTTAGAAGAGCTCA
AAGAATGTGCTCCTCAAGAAAGCAGTGTCCATGCATCGCTTGGCAAAATATACAATCAGCTA
AAGCAATACGACAAAGCCGTGTTACATTTCGGCATTGCTTTGGATTTAAGCCCTTCTCCATC
TGATGCTGTCAAGATAAAGGCTTACATGGAGAGGTTGATACTACCAGACGAGCTGGTGACGG
AGGAAAATTTG

FIGURA 20 (Cont.)

SEQ ID N: 134 seqüência de aminoácidos CDC27a de Arabidopsis thaliana Sequencia NM_112503.2

MMENLLANCVQKNLNHFMFTNAIFLCELLLAQFPSEVNLQLLARCYLSNSQAYSAYYILKGS
KTPQSRYLFAFSCFKLDLLGEAEAALLPCEDYAEEVPGGAAGHYLLGLIYRYSGRKNCSIQQ
FRMALSFDPLCWEAYGELCSLGAAEEASTVFGNVASQRLQKTCVEQRISFSEGATIDQITDS
DKALKDTGLSQTEHIPGENQQDLKIMQQPGDIPPNTDRQLSTNGWDLNTPSPVLLQVMDALP
PLLLKNMRRPAVEGSLMSVHGVRVRRRNFFSEELSAEAQEESGRRRSARIAARKKNPMSQSF
GKDSHWLHLSPSESNYAPSLSSMIGKCRIQSSKEVIPDTVTLNDPATTSGQSVSDIGSSVDD
EEKSNPSESSPDRFSLISGISEVLSLLKILGDGHRHLHMYKCQEALLAYQKLSQKQYNTHWV
LMQVGKAYFELQDYFNADSSFTLAHQKYPYALEGMDTYSTVLYHLKEEMRLGYLAQELISVD
RLSPESWCAVGNCYSLRKDHDTALKMFQRAIQLNERFTYAHTLCGHEFAALEEFEDAERCYR
KALGIDTRHYNAWYGLGMTYLRQEKFEFAQHQFQLALQINPRSSVIMCYYGIALHESKRNDE
ALMMMEKAVLTDAKNPLPKYYKAHILTSLGDYHKAQKVLEELKECAPQESSVHASLGKIYNQ
LKQYDKAVLHFGIALDLSPSPSDAVKIKAYMERLILPDELVTEENL

SEQ ID N: 135 seqüência de ácidos nucléicos CDC27 de Oryza sativa Sequencia NM_112503.2

ATGGAAACCCTAATGGTGGACCGCGTCCACGGCAGCCTCCGCCTCTTCATGCACCGCAACGC
CGTCTTCCTCTGCGAGCGCCTCTGCGCCCAGTTCCCCGCCGAGACAAATGTCCAGTTGTAGC
AACTTGCTACCTTCACAACAACCAGCCATATGCTGCATACCACATCTTGAAAGGAAAGAAGC
TGCCAGAGTCCCGGTACTTGTTTGCTATGTCATGCTTCCGAATGAACCTCTTACGGGAAGCT
GAAGAAGCCTTGTGTCCTGTCAATGAACCAAATATTGAGGTTCCAAGTGGTGCAACAGGGCA
CTACCTTCTTGGAGTAATTTACAGGTACACTGGCAGAGTGGAAGCTGCAGCTGAGCAATTTG
TACAAGCTCTGACTCTTGATCCTCTTCTATGGGCAGCATACGAGGAATTGTGCATACTAGGT
GTTGCTGAAGATGCAAATGAATGTTTCAGTGAAGCAACAGCTCTACGTCTTCAGCAGGAACT
CACATCCACATCAAATGTGGAAAAGTCAAACTTTGTTAATGAAAATCGGTTTCTATCTTCCA
ATGTGTCAGCAAGTTTTGGTGATAGTCCTAAGCAAATTAAACAGCTGCATGCTAACACCACT
GCAGAAGTATCTGGTTATCCTCATGTAAAGTCAACTGCATTGCATATGCAGAACGGTGCACC
ATCTAATTTATCACAGTTTGACACTCCATCGCCAACTTCAACGCAGNNNNATAATGTAACTT
CAACTTCGTCTTCTACAAGTATAGTTGATGGAAGATATCCCGAGCAAGAGAAATCTGAACGA
GTTCTGTCACAGGACTCCAAATTAGCTATTGGTATCAGGGAGCTAATGGCACTCTTGCGGAC
ACTAGGGGAAGGGTATAGGCTTTCTTGCTTGTTTAAGTGTCAGGAAGCATTGGAAGTATATA
GAAAGCTCCCAGAGGCACAATTTAATACTGGATGGGTTCTTTGCCAGGTTGGGAAGACATAT
TTTGAACTCGTCAATTATTTAGAAGCCGATCATTTTTTTGAGTTAGCGCATCGACTATCACC
ATGCACGTTGGAGGGAATGGACATTTACTCCACTGTTCTTTATCATTTGAATGAGGAAATGC
GGCTAAGTTACCTTGCTCAAGATCTTGTTTCTATTGATCGACTATCTCCCCAAGCATGGTGT
GCTGTGGGAAATTGCTTTGCCTTGAGGAAAGATCATGAGACTGCCTTGAAGAATTTTCAACG
TGCTGTACAGCTTGACTCAAGAGTTGCATACGCTCACACGCTATGCGGTCACGATATAAAAC
TATACCGATCTGCACTTCAGGTAGATGAAAGACACTACAATGCCTGGTATGGCCTTGGAGTG
GTGTACCTTCGCCAGGAAAAGTTTGAGTTTGCTGAGCATCATTTCAGAAGGGCATTCCAGAT
AAATCCTTGCTCTTCTGTTCTTATGTGCTATCTTGGGATGGCCTTGCATGCTTTAAAGAGGA
ATGAGGAAGCCTTGGAAATGATGGAGAAGGCTATATTTGCTGATAAGAAGAATCCACTCCCC
AAGTATCAAAAGGCTTTAATCCTTCTAGGCCTACAAAAATACCCTGATGCTCTGGATGAGTT
GGAACGGCTAAAGGAAATTGCACCTCATGAAAGTAGTATGTATGCACTGATGGGAAAGATTT
ACAAGCAACTTAACATTCTTGACAAGGCTGTATTTTGCTTTGGCATTGCCCTGGATTTGAAA
CCTCCTGCTGCTGACGTTGCTATAATACAATCTGCAATGGAGAAAGTACACCTTCCAGATGA
ACTTATGGATGATGATGATGATGATGAGATTTAA

FIGURA 20 (Cont.)

SEQ ID N: 136 seqüência de ácidos nucléicos CDC27 de Oryza sativa AP003539.3

METLMVDRVHGSLRLFMHRNAVFLCERLCAQFPAEVNVQLLATCYLHNNQPYAAYHILKGKK
LPESRYLFAMSCFRMNLLREAEEALCPVNEPNIEVPSGATGHYLLGVIYRYTGRVEAAAEQF
VQALTLDPLLWAAYEELCILGVAEDANECFSEATALRLQQELTSTSNVEKSNFVNENRFLSS
NVSASFGDSPKQIKQLHANTTAEVSGYPHVKSTALHMQNGAPSNLSQFDTPSPTSTQVSGIA
PPPLFRNMHAYQNTAGGNAPSKPKVNAPNLTLRRKYIDEAGLKKVSGRLFNQSSDSVPRRSA
RLSRDTTINSNSNISQFGGNGTDHSSGKLRVNSSTPSKLCSTALRSVQVRKGKPQATENFDE
GDYHFDMDDSVTSTSSSTSIVDGRYPEQEKSERVLSQDSKLAIGIRELMALLRTLGEGYRLS
CLFKCQEALEVYRKLPEAQFNTGWVLCQVGKTYFELVNYLEADHFFELAHRLSPCTLEGMDI
YSTVLYHLNEEMRLSYLAQDLVSIDRLSPQAWCAVGNCFALRKDHETALKNFQRAVQLDSRV
AYAHTLCGHEYSALEDYENSIKLYRSALQVDERHYNAWYGLGVVYLRQEKFEFAEHHFRRAF
QINPCSSVLMCYLGMALHALKRNEEALEMMENAIFADKKNPLPKYQKALILLGLQKYPDALD
ELERLKEIAPHESSMYALMGKIYKQLNILDKAVFCFGIALDLKPPAADVAIIKSAMEKVHLP
DELMDDDDDDDEI

SEQ ID N: 137 CDC27 de Solanum tuberosum ESTs compiladas BG887406.1 BG590616.1 DN 939130.1 CV470643.1

ATGGAAACCCTACTAGCTGAATCTGTGCAAAACAGCCTTGGCCAATTTATGTACCACAACGC
CATTTTCATGTGTGAACGACTCTGTGCCGAGTTCCCCACTGAGACAAATATGCAGCTTTTAG
CTGGCTGCTACCTGCACAACCAACAGGCTTATGCTGCATATCATCTTCTCAAGGGGACAAGT
ATGGCTCAATCCCGCTACTTGTTTGCACTATCATGCTTTCAGATGGATCTTCTCACTGAAGC
TGAGACAGCACTTTGCCCTCCTAATGAGCCAACTGCAGAGGTTCCAAATGGTGCAGCTGGGC
ATTACCTTCTTGGTCTTATTTACAGGTATACAGATAGAAGAAATAGTTCCATCCAGCATTTC
AATCAGGCATTGTTATTGGATCCATTGCTATGGGCTGCATATGAGGAGTTGTGTATACTAGG
TGCTGCAGAAGAAGCAGCTGCAGTTTTTGGGGAAGCATCTTTGCTTTGCATTCAGAAACAAC
ACATAGACCAAGGGAACCAATCTCAAAATTTACAAGCATCCACTGATGATCAGAATGTAGCT
TCTACGAATATTGTCTCAGGCGACATCAGCCCTATGCAATCAAAACATACACACAGCCATAA
TCTTCGAGAAATGTCTGGAAATTATAATGGAGCAGCTGCTATCCAGAATTTAGGTGGGGTTT
CTACTAACATGTCATTCTACAACACTCCCTCACCAATGGCATCACAGTTGTCAGGAGTGGTT
CCACCTCCAGTTTGTAGAAATTTTCAGCAAAATGGAACTAATGCATCTGTGGCTGGTGCTGA
TAATTCTCCACGAGCAACTGTCAATTCAACCATTCAGGCCCCTCGGAGAAAGTTTGTTGATG
AGGGGAAGTTAAGAAAGATATCTGGGAGGTTATTTTCTGATTCTGGCCCTCGACGAAATTCA
AGGCTTGCTGGAGAATCTACTGGAAACACAAATTCAAATGTATCTGGTGCTTCTGGAAATGG
AACAATTCATTCTTCCAAATATTATGGTAGTTCGAAGCTGAGCTCAATGACTTTACGTTCCA
TGACAAGTCGAAAGGCACAATCTTGGGCTACCGAAAACTATGGTGAAGGGACTCGCAATGAC
ATTTCTAATGATTCTCGGCTAAATATGACTATGTCACACCCTTCTGGAGATGCTAGACCTCT
TGAACAAGAAAGGCCCCGAACTTCTGCTTCTGGGGTTAATGTAAGCAGCACATCTATCCCTC
TCAGTGCTTCAGAGATATTGGCCCTTTTCAGGTTTCTTGGGGAAGGCTATAGACTTTCTTGT
TTATATAGATGTCAGGATGCACTGGATGTTTATAACAAACTCCCACACAAACATTATCACAC
TGGATGGGTTCTTTCTCAGATTGGAAGAGCATACTTCGAAATGGTTGATTACCTAGAAGCAG
ATCATGCATTTGGCCTTGCTCGTCTGGCCTCACCTTATAGTTTAGAAGGAATGGACGTGTAC
TCGACAGTGTTGTTTCATCTCAAGGAGGACATGAAGTTGAGCTATCTGGCGCAGGTGCTGGT
ATCAACTGATAGATTAGCTCCTCAATCTTGGTGTGCTATGGGGAATTGCTATAGTTTACAGA
AAGACCATGAAACTGCTCTTAAAAATTTTCAACGAGCTGTACAACTAAATCCTAGATTTGCA
TACGGGCACACGCTTTGTGGTCATGAATATGTTGCTTTAGAAGATTTTGAAAATGCTATTAA

FIGURA 20 (Cont.)

GAGCTATCAGAGTGCACTTCGTGTGGATGCCAGGCATTACAATGCCTGGTATGGGCTTGGAA
TGATCTATCTCCGACAGGAGAAGTTTGAATTTTCAGAGCATCACTTTCGAATGGCTTTGGGT
ATAAATCCACAGTCTTCTGTTATCATGTCATATCTTGGCACTGCATTACACGCTCTGAAGAA
AAATGAAGAGGCATTGGAAGTGATGGAGCTGGCTATTGTAGCAGACAAGAAAAACCCTCTTC
CAATGTATCAGAAGGCTAACATCCTTGTGAGCACGGAAAGTTTTGATGCCGCTTTAGAAGTC
TTAGAGGAACTTAAAGAGCATGCTCCTCGTGAGAGCAGTGTCTATGCTTTGATGGGTCGGAT
ATACAAGAGGCGTAATATGTACGACAAAGCCATGCTTCATTTTGGAGTGGCATTAGATTTAA
AACCATCTGCAACTGATGTTGCTACCATTAAGGCTGCCATTGAAAAGCTGCATGTACCAGAT
GAGATGGAAGATGAATTATAA

SEQ ID N: 138 seqüência de aminoácidos CDC27 de Solanum tuberosum

METLLAESVQNSLGQFMYHNAIFMCERLCAEFPTETNMQLLAGCYLHNQQAYAAYHLLKGTS
MAQSRYLFALSCFQMDLLTEAETALCPPNEPTAEVPNGAAGHYLLGLIYRYTDRRNSSIQHF
NQALLLDPLLWAAYEELCILGAAEEAAAVFGEASLLCIQKQHIDQGNQSQNLQASTDDQNVA
STNIVSGDISPMQSKHTHSHNLREMSGNYNGAAAIQNLGGVSTNMSFYNTPSPMASQLSGVV
PPPVCRNFQQNGTNASVAGADNSPRATVNSTIQAPRRKFVDEGKLRKISGRLFSDSGPRRNS
RLAGESTGNTNSNVSGASGNGTIHSSKYYGSSKLSSMTLRSMTSRKAQSWATENYGEGTRND
ISNDSRLNMTMSHPSGDARPLEQERPRTSASGVNVSSTSIPLSASEILALFRFLGEGYRLSC
LYRCQDALDVYNKLPHKHYHTGWVLSQIGRAYFEMVDYLEADHAFGLARLASPYSLEGMDVY
STVLFHLKEDMKLSYLAQVLVSTDRLAPQSWCAMGNCYSLQKDHETALKNFQRAVQLNPRFA
YGHTLCGHEYVALEDFENAIKSYQSALRVDARHYNAWYGLGMIYLRQEKFEFSEHHFRMALG
INPQSSVIMSYLGTALHALKKNEEALEVMELAIVADKKNPLPMYQKANILVSTESFDAALEV
LEELKEHAPRESSVYALMGRIYKRRNMYDKAMLHFGVALDLKPSATDVATIKAAIEKLHVPD
EMEDEL

SEQ ID N: 139 seqüência de ácidos nucléicos tipo CDC27 de Schizosaccharomyces pombe nuc2+ CDC27-nm_001020032.1

ATGACAGATCGATTGAAATGTTTAATATGGTATTGCATTGATAATCAGAATTATGATAATTC
AATTTTTTATTCAGAACGTTTACATGCAATTGAAGATTCAAACGAGAGTTTGTATCTTTTGG
CATATTCGCATTTCCTAAACCTCGATTACAATATTGTATACGACTTATTAGATAGAGTAATT
AGTCATGTTCCTTGCACATACTTATTTGCAAGGACCAGCCTTATTTTAGGCAGATATAAACA
AGGAATAAGTGCTGTGGAGGCCTGTCGATCGAATTGGCGCTCCATTCAGCCAAACATAAATG
ACTCAATTAGCAGTCGTGGACATCCAGATGCCTCTTGCATGCTTGATGTTTTGGGTACTATG
TATAAAAAGGCAGGGTTCCTCAAAAAAGCTACAGATTGTTTTGTAGAAGCTGTCTCCATTAA
CCCATATAATTTCTCTGCTTTCCAGAATTTAACTGCAATAGGCGTGCCACTCGATGCTAATA
ATGTATTTGTTATTCCACCCTACCTTACGGCAATGAAGGGTTTTGAAAAATCTCAAACGAAT
GCTACAGCTTCGGTACCAGAACCGTCTTTTTTGAAGAAAAGTAAAGAGTCTTCCTCATCTTC
CAACAAGTTTTCGGTTTCTGAATCGATAGCAAATAGTTATTCAAACTCATCCATTTCAGCAT
TTACTAAGTGGTTTGATAGGGTTGACGCTTCTGAGCTTCCAGGAAGTGAGAAGGAACGACAT
CAAAGCTTGAAATTACAATCTCAATCTCAGACTAGCAAAAACCTTTTGGCTTTCAATGATGC
TCAAAAAGCTGATTCTAACAATAGGGATACGTCTTTAAAATCCCACTTTGTGGAACCTAGAA
CCCAAGCATTAAGACCAGGAGCTCGTTTAACATATAAATTACGCGAAGCGAGAAGTTCTAAA
AGAGGAGAGAGCACACCTCAAAGCTTCCGCGAAGAGGACAATAATTTGATGGAATTACTAAA
GTTATTCGGTAAGGGTGTTTACCTGCTCGCCCAGTATAAGTTACGAGAGGCTTTAAATTGTT
TCCAAAGCTTGCCCATCGAACAGCAAAATACACCTTTTGTTCTTGCAAAGCTTGGAATAACC
TACTTTGAACTGGTTGATTACGAAAAATCTGAAGAAGTGTTTCAAAAATTAAGGGACTTGTC
GCCTTCACGTGTCAAAGATATGGAAGTCTTTTCAACTGCACTTTGGCATTTGCAAAAGTCTG

FIGURA 20 (Cont.)

TTCCTTTATCTTACCTTGCCCATGAAACTTTGGAAACTAATCCTTATTCCCCAGAATCATGG
TGCATTCTTGCTAATTGCTTCTCACTTCAACGTGAACACTCGCAGGCATTAAAATGTATTAA
TAGAGCTATTCAATTGGATCCAACTTTTGAATATGCTTATACGCTTCAAGGGCACGAGCATT
CTGCAAACGAAGAATACGAAAAATCGAAAACATCTTTCCGCAAAGCAATTAGAGTAAATGTT
CGACATTACAATGCATGGTATGGCCTGGGAATGGTTTATTTAAAAACAGGGCGAAATGATCA
AGCAGACTTTCATTTTCAACGTGCTGCAGAGATCAATCCTAACAACTCTGTACTCATCACTT
GTATTGGTATGATTTACGAACGCTGCAAAGATTACAAAAAAGCACTTGATTTTTATGATCGG
GCATGTAAACTGGATGAAAAGTCTTCGCTTGCCAGGTTCAAGAAAGCCAAAGTGCTTATTTT
ATTACATGATCACGATAAAGCACTCGTTGAATTGGAACAATTAAAGGCAATTGCGCCAGATG
AAGCAAATGTTCATTTTTTGCTCGGCAAAATTTTCAAGCAAATGCGGAAAAAAAATTTAGCC
TTAAAGCACTTCACTATAGCATGGAACTTAGACGGCAAGGCTACGCATATTATTAAGGAATC
GATTGAAAATCTGGATATTCCCGAAGAAAATTTGTTAACTGAAACAGGTGAAATTTATAGGA
ATCTGGAAACTTAA

SEQ ID N: 140 seqüência de ácidos nucléicos tipo nuc2+CDC27 de Schizosaccharomyces pombe

MTDRLKCLIWYCIDNQNYDNSIFYSERLHAIEDSNESLYLLAYSHFLNLDYNIVYDLLDRVI
SHVPCTYLFARTSLILGRYKQGISAVEACRSNWRSIQPNINDSISSRGHPDASCMLDVLGTM
YKKAGFLKKATDCFVEAVSINPYNFSAFQNLTAIGVPLDANNVFVIPPYLTAMKGFEKSQTN
ATASVPEPSFLKKSKESSSSSNKFSVSESIANSYSNSSISAFTKWFDRVDASELPGSEKERH
QSLKLQSQSQTSKNLLAFNDAQKADSNNRDTSLKSHFVEPRTQALRPGARLTYKLREARSSK
RGESTPQSFREEDNNLMELLKLFGKGVYLLAQYKLREALNCFQSLPIEQQNTPFVLAKLGIT
YFELVDYEKSEEVFQKLRDLSPSRVKDMEVFSTALWHLQKSVPLSYLAHETLETNPYSPESW
CILANCFSLQREHSQALKCINRAIQLDPTFEYAYTLQGHEHSANEEYEKSKTSFRKAIRVNV
RHYNAWYGLGMVYLKTGRNDQADFHFQRAAEINPNNSVLITCIGMIYERCKDYKKALDFYDR
ACKLDEKSSLARFKKAKVLILLHDHDKALVELEQLKAIAPDEANVHFLLGKIFKQMRKKNLA
LKHFTIAWNLDGKATHIIKESIENLDIPEENLLTETGEIYRNLET

SEQ ID N: 141 seqüência de ácidos nucléicos tipo CDC27 de Aspergillus niger bimA X59269.1

ATGACTCCGTCCACATCACATATTTCGAGCCAGCTAAGGCAGCTGATATATTACCATCTTGA
CAACAACCTTGCTCGGAACGCGCTGTTCCTTGCCGGTCGTTTACACGCCTACGAACCTCGGA
CGTCGGAAGCTTCGTACCTATTAGCTCTGTGTTACCTACAAAATGGTCAGGTGAAAGCAGCA
TGGGAAACTAGCAAGCATTTTGGGTCGAGGGGTGCGCATCTTGGATGTTCTTACGTCTACGC
GCAGGCTTGTCTTGACCTCGGGAAATATACGGACGGTATTAACGCGCTAGAGCGAAGTAAGG
GACAATGGACTTCGCGAAACCACTGGAATAAACACAGTGAGACGCGACGACAACACTTGCCG
GATGCTGCGGCAGTTTTATGTTTGCAAGGAAAATTATGGCAGGCACACAAGGAACACAACAA
GGCTGTGGAGTGTTACGCTGCAGCTTTAAAGCTGAATCCCTTCATGTGGGATGCATTCTTGA
ATCTGTGCGAAACTGGTGTTGATTTGCGTGTTTCAAACATATATAAGATGAGCCCGGAATTG
TACAGCATGGTATCGTCTGCTGCGCTCGAAGATGTTGAATCCCAGGTTCTACCTCCGGACGG
TCCACTCCAGACACAAGTTAACCCAAATCCGAGCTTGGACCCGTTCACTGCCGGTACGACTC
GCAGCGATTCGACCTCTACTCACGGGAGCTCAGCTTTGTGGGAAAAGTTGAATGGAAGCACA
GTTAGTGTGGCATCTTCAGGAACAGGACCGCATCTTCCGCGGGAAGGCATGGAGACTCCTGG
TGGCCAAAGTAGCGAATCTGATGATCCACGTGTAACCAACGGTAATGGCACAGATGTTTTTG
AGCCGCCGCTAGCCCCTGCAAAGAAGAATCGAACGATCCAAACAATAGGCGGCGATCATCCG
ATGGATCCCCCGCCAAAGATGCGGCCAACCGGTATTCGACCGAGGACCCGAACCAAATTCGA
GTCAGATGAAGGTCATACTGAGAGAGACGCGGGCATGGGTCACAGGTTAGGCGATCGGAAAC

FIGURA 20 (Cont.)

GAACGGTCTCTGGACAAGTTGCGCATCCGTCGGTACCGCATTCAACCGACCAAGGTGTCGGG
CAACGGCGAAGTGTACGTCTCTTTAACCAGATAAAGCCATCAACGAACAAAATATCTAGCAC
GGCATTAGGAGTGAAAGAGGGCCGGGAGGTGAAAAAGGTGAGAACTACGGGGAATAAGGCGC
GCACAACAACAAGCTCAAATGTGGGCCGTGTTGTTAGCGGCAACAATCGACGACACGCTGGG
GAGATCCATGACGGAGACAGCAAAGAATACCGTGGAACATCTTCAACATCAAATGGCTCTCA
GAACGCATCGTCTAAGCTCGCTATATCTGAGCGGACCAAATCGGTTGAAGCCTTGGCTTGGA
TTTTGGACTTGTTTTTCAAGATAGCCTCCGGCTATTTCTGTCTCAGCCGGTACAAATGCTCC
GATGCTATCCAGATTTTCAGCTCTCTATCCCAGGGCCAACGGGAGACACCGTGGGTTCTTGC
TCAAATTGGACGAGCTTACTACGAGCAAGCAATGTATACAGAGGCCGAGAAATACTTTGTCC
GGGTGAAGGCCATGGCGCCTTCCCGGTTAGAAGATATGGAGATCTACTCGACGGTTCTTTGG
CATTTGAAGAATGATGTCGAACTTGCTTACCTGGCCCATGAGTTAATGGATGTCGATCGCTT
ATCACCAGAAGCTTGGTGTGCTGTCGGTAACTCGTTCTCACACCAGCGGGACCACGATCAAG
CTCTGAAGTGCTTCAAGCGTGCCACTCAATTGGATCCTCATTTCGCGTATGGGTTCACGCTA
CAGGGCCATGAGTATGTTGCCAACGAGGAATATGACAAGGCATTGGATGCCTATAGAAGCGG
CATCAACGCGGACAGTCGACATTATAATGCCTGGTACGGGCTGGGAACGGTCTACGATAAGA
TGGGCAAACTTGACTTTGCTGAACAGCACTTCCGTAATGCGGCCAAGATTAATCCTTCTAAT
GCTGTTTTAATATGCTGCATTGGCTTGGTCTTAGAGAAGATGAACAATCCGAAGTCAGCGTT
GATCCAGTACAACAGGGCCTGCACCCTCGCGCCTCATTCTGTTCTTGCTCGATTCCGAAAAG
CCCGTGCGTTGATGAAATTGCAGGATCTCAAGTCAGCGCTTACGGAGCTGAAGGTTCTGAAG
GACATGGCGCCGGATGAGGCAAACGTTCATTACCTCCTGGGCAAACTGTATAAGATGCTTCG
TGACAAGGGAAATGCCATTAAGCACTTCACAACTGCATTAAACCTTGATCCCAAGGCGGCGC
AGTACATCAAAGATGCAATGGAAGCCCTTGACGACGATGAAGAGGATGAAGAAGATATGGCG
TGA

SEQ ID N: 142 seqüência de aminoácidos tipo CDC27 de Aspergillus niger

MTPSTSHISSQLRQLIYYHLDNNLARNALFLAGRLHAYEPRTSEASYLLALCYLQNGQVKAA
WETSKHFGSRGAHLGCSYVYAQACLDLGKYTDGINALERSKGQWTSRNHWNKHSETRRQHLP
DAAAVLCLQGKLWQAHKEHNKAVECYAAALKLNPFMWDAFLNLCETGVDLRVSNIYKMSPEL
YSMVSSAALEDVESQVLPPDGPLQTQVNPNPSLDPFTAGTTRSDSTSTHGSSALWEKLNGST
VSVASSGTGPHLPREGMETPGGQSSESDDPRVTNGNGTDVFEPPLAPAKKNRTIQTIGGDHP
MDPPPKMRPTGIRPRTRTKFESDEGHTERDAGMGHRLGDRKRTVSGQVAHPSVPHSTDQGVG
QRRSVRLFNQIKPSTNKISSTALGVKEGREVKKVRTTGNKARTTTSSNVGRVVSGNNRRHAG
EIHDGDSKEYRGTSSTSNGSQNASSKLAISERTKSVEALAWILDLFFKIASGYFCLSRYKCS
DAIQIFSSLSQGQRETPWVLAQIGRAYYEQAMYTEAEKYFVRVKAMAPSRLEDMEIYSTVLW
HLKNDVELAYLAHELMDVDRLSPEAWCAVGNSFSHQRDHDQALKCFKRATQLDPHFAYGFTL
QGHEYVANEEYDKALDAYRSGINADSRHYNAWYGLGTVYDKMGKLDFAEQHFRNAAKINPSN
AVLICCIGLVLEKMNNPKSALIQYNRACTLAPHSVLARFRKARALMKLQDLKSALTELKVLK
DMAPDEANVHYLLGKLYKMLRDKGNAIKHFTTALNLDPKAAQYIKDAMEALDDDEEDEEDMA

SEQ ID N: 143
seqüência de ácidos nucléicos CDC27 de Homo sapiens NM_001256.2

ATGACGGTGCTGCAGGAACCCGTCCAGGCTGCTATATGGCAAGCACTAAACCACTATGCTTA
CCGAGATGCGGTTTTCCTCGCAGAACGCCTTTATGCAGAAGTACACTCAGAAGAAGCCTTGT
TTTTACTGGCAACCTGTTATTACCGCTCAGGAAAGGCATATAAAGCATATAGACTCTTGAAA
GGACACAGTTGTACTACACCGCAATGCAAATACCTGCTTGCAAAATGTTGTGTTGATCTCAG

FIGURA 20 (Cont.)

CAAGCTTGCAGAAGGGGAACAAATCTTATCTGGTGGAGTGTTTAATAAGCAGAAAAGCCATG
ATGATATTGTTACTGAGTTTGGTGATTCAGCTTGCTTTACTCTTTCATTGTTGGGACATGTA
TATTGCAAGACAGATCGGCTTGCCAAAGGATCAGAATGTTACCAAAAGAGCCTTAGTTTAAA
TCCTTTCCTCTGGTCTCCCTTTGAATCATTATGTGAAATAGGTGAAAAGCCAGATCCTGACC
AAACATTTAAATTCACATCTTTACAGAACTTTAGCAACTGTCTGCCCAACTCTTGCACAACA
CAAGTACCTAATCATAGTTTATCTCACAGACAGCCTGAGACAGTTCTTACGGAAACACCCCA
GGACACAATTGAATTAAACAGATTGAATTTAGAATCTTCCAATTCAAAGTACTCCTTGAATA
CAGATTCCTCAGTGTCTTATATTGATTCAGCTGTAATTTCACCTGATACTGTCCCACTGGGA
ACAGGAACTTCCATATTATCTAAACAGGTTCAAAATAAACCAAAAACTGGTCGAAGTTTATT
AGGAGGACCAGCAGCTCTTAGTCCATTAACCCCAAGTTTTGGGATTTTGCCATTAGAAACCC
CAAGTCCTGGAGATGGATCCTATTTACAAAACTACACTAATACACCTCCTGTAATTGATGTG
CCATCCACCGGAGCCCCTTCAAAAAAGTCTGTTGCCAGAATCGGCCAAACTGGAACAAAGTC
TGTCTTCTCACAGAGTGGAAATAGCCGAGAGGTAACTCCAATTCTTGCACAAACACAAAGTT
CTGGTCCACAAACAAGTACAACACCTCAGGTATTGAGCCCCACTATTACATCTCCCCCAAAC
GCACTACCTCGAAGAAGTTCACGACTCTTTACTAGTGACAGCTCCACAACCAAGGAGAATAG
CAAAAAATTAAAAATGAAGTTTCCACCTAAAATCCCAAACAGAAAAACAAAAAGTAAAACTA
ATAAAGGAGGAATAACTCAACCTAACATAAATGATAGCCTGGAAATTACAAAATTGGACTCT
TCCATCATTTCAGAAGGGAAAATATCCACAATCACACCTCAGATTCAGGCCTTTAATCTACA
AAAAGCAGCAGCAGGTTTGATGAGCCTTCTTCGTGAAATGGGGAAAGGTTATTTAGCTTTGT
GTTCATACAACTGCAAAGAAGCTATAAATATTTTGAGCCATCTACCTTCTCACCACTACAAT
ACTGGTTGGGTACTGTGCCAAATTGGAAGGGCCTATTTTGAACTTTCAGAGTACATGCAAGC
TGAAAGAATATTCTCAGAGGTTAGAAGGATTGAGAATTATAGAGTTGAAGGCATGGAGATCT
ACTCTACAACACTTTGGCATCTTCAAAAAGATGTTGCTCTTTCAGTTCTGTCAAAAGACTTA
ACAGACATGGATAAAAATTCGCCAGAGGCCTGGTGTGCTGCAGGGAACTGTTTCAGTCTGCA
ACGGGAACATGATATTGCAATTAAATTCTTCCAGAGAGCTATCCAAGTTGATCCAAATTACG
CTTATGCCTATACTCTATTAGGGCATGAGTTTGTCTTAACTGAAGAATTGGACAAAGCATTA
GCTTGTTTTCGAAATGCTATCAGAGTCAATCCTAGACATTATAATGCATGGTATGGTTTAGG
AATGATTTATTACAAGCAAGAAAAATTCAGCCTTGCAGAAATGCATTTCCAAAAAGCGCTTG
ATATCAACCCTCAAAGTTCAGTTTTACTTTGCCACATTGGAGTAGTTCAACATGCACTGAAA
AAATCAGAGAAGGCTTTGGATACCCTAAACAAAGCCATTGTCATTGATCCCAAGAACCCTCT
ATGCAAATTTCACAGAGCCTCAGTTTTATTTCGAAATGAAAAATATAAGTCTGCTTTACAAG
AACTTGAAGAATTGAAACAAATTGTTCCCAAAGAATCCCTCGTTTACTTCTTAATAGGAAAG
GTTTACAAGAAGTTAGGTCAAACGCACCTCGCCCTGATGAATTTCTCTTGGGCTATGGATTT
AGATCCTAAAGGAGCCAATAACCAGATTAAAGAGGCAATTGATAAGCGTTATCTTCCAGATG
ATGAGGAGCCAATAACCCAAGAAGAACAGATCATGGGAACAGATGAATCCCAGGAGAGCAGC
ATGACAGATGCGGATGACACACAACTTCATGCAGCTGAAAGTGATGAATTTTAA

SEQ ID N: 144
seqüência de ácidos nucléicos CDC27 de Homo sapiens

MTVLQEPVQAAIWQALNHYAYRDAVFLAERLYAEVHSEEALFLLATCYYRSGKAYKAYRLLK
GHSCTTPQCKYLLAKCCVDLSKLAEGEQILSGGVFNKQKSHDDIVTEFGDSACFTLSLLGHV
YCKTDRLAKGSECYQKSLSLNPFLWSPFESLCEIGEKPDPDQTFKFTSLQNFSNCLPNSCTT
QVPNHSLSHRQPETVLTETPQDTIELNRLNLESSNSKYSLNTDSSVSYIDSAVISPDTVPLG
TGTSILSKQVQNKPKTGRSLLGGPAALSPLTPSFGILPLETPSPGDGSYLQNYTNTPPVIDV
PSTGAPSKKSVARIGQTGTKSVFSQSGNSREVTPILAQTQSSGPQTSTTPQVLSPTITSPPN
ALPRRSSRLFTSDSSTTKENSKKLKMKFPPKIPNRKTKSKTNKGGITQPNINDSLEITKLDS
SIISEGKISTITPQIQAFNLQKAAAAGLMSLLREMGKGYLALCSYNCKEAINILSHLPSHHY
NTGWVLCQIGRAYFELSEYMQAERIFSEVRRIENYRVEGMEIYSTTLWHLQKDVALSVLSKD

FIGURA 20 (Cont.)

LTDMDKNSPEAWCAAGNCFSLQREHDIAIKFFQRAIQVDPNYAYAYTLLGHEFVLTEELDKA
LACFRNAIRVNPRHYNAWYGLGMIYYKQEKFSLAEMHFQKALDINPQSSVLLCHIGVVQHAL
KKSEKALDTLNKAIVIDPKNPLCKFHRASVLFANEKYKSALQELEELKQIVPKESLVYFLIG
KVYKKLGQTHLALMNFSWAMDLDPKGANNQIKEAIDKRYLPDDEEPITQEEQIMGTDESQES
SMTDADDTQLHAAESDEF

SEQ ID N: 145
CDC27 de extremidade 5' CDC27 CA102186.1 CA2799358.1 de Saccharum officinarum ESTs compiladadas

ATGGAAACCCTAATGGTGGACCGCGTCCACAGCAGCCTCCGCCTCTTCATGCACCGCAACGC
CGTATTCCTCTGCGAGCGCCTCTGCGCGCAGTTCCCCTCCGAGACCAATGTGCAATTGTTAG
CGACCTGCTACCTCCACAACAATCAGCCATATGCTGCATACCACATTTTGAAAGGGAAGAAG
CTGCCGGAGTCCCGGTACTTGTTTGCTACATCATGCTTTCGAATGAACCTCTTGCGTGAAGC
AGAAGAAACTCTATGTCCAGTCAATGAACCAAACATGGAGGTTCCAAGTGGAGCAACAGGAC
ACTACCTCCTTGGAGTGATTTACAGGTGCACAGGCAGAATTTCAGCTGCAGCTGAACAATTT
ACACAAGCGTTGACTCTAGATCCTCTTTTATGGGCGGCATATGAGGAATTGTGTATATTAGG
TATTGCTGAAGATACTGATGAGTGTTTTAGTGAATCGACTGCTCTACGTCTCCAGCAGGAAC
ACACATCCACGGCCACTCTGGTGAAGTCGAACTTCGCCAATGAAAATCGAGTTCTATCATCC
AGGGTCTCTGCAAATCTTGGGGATATTAGTCCTAAGCAAATCAAACAGCTTCATGCTAACAA
CATAGCAGAAGTATCTGGCTATCCTCATGTAAGACCAACTGCATTGCATGTGCAGAACAGTT
CAACCTCTAATGTAGCACAGTTTGACACCCCATCACCAACTGCAGCACAGACTTCTAGTATC
ATGCCACCACCACTCTTTAGGAATGTCCATGCTTANNNN

SEQ ID N: 146 CDC27 C-terminal de Saccharum officinarum

METLMVDRVHSSLRLFMHRNAVFLCERLCAQFPSETNVQLLATCYLHNNQPYAAYHILKGKK
LPESRYLFATSCFRMNLLREAEETLCPVNEPNMEVPSGATGHYLLGVIYRCTGRISAAAEQF
TQALTLDPLLWAAYEELCILGIAEDTDECFSESTALRLQQEHTSTATLVKSNFANENRVLSS
RVSANLGDISPKQIKQLHANNIAEVSGYPHVRPTALHVQNSSTSNVAQFDTPSPTAAQTSSI
MPPPLFRNVHA

SEQ ID N: 147 CDC27 C-terminal de Saccharum officinarum CA197669.1
CA197670.1 CA203636.1 CA232307.1 Compilado ESTs

ATTCAAATTCAAATACCTGGGGTTTGGAGGGAATGGTACAGGTTATTCGTCAGGGAAATTGC
GAGTAAACTCGTCCACACCATCAAAATGGTGTTAACCACCATACGTTCCGTGCAAGTTAGGA
AAGGAAAACCACGGGCTACAGAAAATTTTGATGAAGGAAGTAGATATGAAGTCATTGATGAA
ATGTGGACAGACAATATATCAGGAACTTCATCTTCTGTAAGTACAGCTGATGGAAGATCCTT
TGAGCAAGATAAAGCTGAACGAATTCTGTTGCAAGACTCCAAATTGGCACTTGGTATTAGGG
AGATATTGGGACTTGTCCGAACACTCGGTGAAGGTTGTAGGCTTTCTTGCTTGTTTAAGTGC
CATGAAGCCTTGGAAGTCTACAGAAGACTCCCTGAGACCCATTNTAGCACTGGATGGAGCAT
ATGCCAGGTTGGTAAGGCATATTTCGAATTAGTTGATTATTTGGAAGCTGATCGTTACTTTG
AATTGGCACACCGACTGTCGCCTTGTACGCTTGATGGAATGGACATCTATTCTACTGTTCTT
TATCATCTGAATGAGGAAATGAGACTAAGCTACCTTGCTCAAGAGCTTATTTCCATTGATCG
ACTATCTCCTCAAGCATGGTGTGCAGTGGGCAATTGCTTTGCCTTGAGGAAAGATCATGAGA
CTGCTTTGAAGAATTTTCAACGTTCGGTACAGCTTGACTCAAGATTTGCATATGCTCACACT
CTATGTGGTCATGAGTATTCTGCATTGGAGGATTATGAGAATAGTATCAAATTCTACCGGTG
TGCACTGCAGGTAGATGAAAGGCACTACAATGCCTGGTATGGCCTTGGGGTGGTGTATCTTC
GCCAGGAAAAGTNTGAGTTTGCTGAGCATCATTTCAGAAGGGCATTTCAGATAAATCCTCGC
TCTTCTGTTCTCATGTGCTATCTTGGGATGGCGTTGCATTCTCTTAAGAGGAAGGAGGAGGC

FIGURA 20 (Cont.)

ATTGGAAATGATGGAGAAAGCTATAGCAGCTGATAAGAAGAATCCACTGCCCAAGTATCAGA
AGGCCTTAATCCTTCTAGGTCTTCAGAAGTATCAAGAAGCTCTGGATGAGTTGGAGCGGCTA
AAGGAGATTGCACCTCATGAGAGCAGTATGTATGCACTGATGGGAAAGATTTACAAGCAACT
CAATATCCTTGACAAAGCTGTTTTCTGCTTTGGCATTGCCCTGGATTTGAAACCTCCTGCTG
CTGATCTTGCTATAATTAAGTCCGCAATGGAGAAAGTACATCTCCCTGATGAACTGATGGAG
GATGACCTGTAA

SEQ ID N: 148 CDC27 C-terminal de Saccharum officinarum

IQIQIPGVWREWYRLFVREIASKLVHTIKMVLTTIRSVQVRKGKPRATENFDEGSRYEVIDE
MWTDNISGTSSSVSTADGRSFEQDKAERILLQDSKLALGIREILGLVRTLGEGCRLSCLFKC
HEALEVYRRLPETHXSTGWSICQVGKAYFELVDYLEADRYFELAHRLSPCTLDGMDIYSTVL
YHLNEEMRLSYLAQELISIDRLSPQAWCAVGNCFALRKDHETALKNFQRSVQLDSRFAYAHT
LCGHEYSALEDYENSIKFYRCALQVDERHYNAWYGLGVVYLRQEKXEFAEHHFRRAFQINPR
SSVLMCYLGMALHSLKRKEEALEMMEKAIAADKKNPLPKYQKALILLGLQKYQEALDELERL
KEIAPHESSMYALMGKIYKQLNILDKAVFCFGIALDLKPPAADLAIIKSAMEKVHLPDELME
DDL

SEQ ID NO : 149, prm00778
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACAATGCAACAACTGTCAACTTC

SEQ ID NO : 150, prm00779
GGGGACCACTTTGTACAAGAAAGCTGGGTTGGAGTAGCTATGGTTTCAC

SEQ ID NO : 151, promotor OSH1 de Oryza sativa
GGGGAGTTAGGAACCTTGACATACAACCAATGATACCATTTTTTTTCAAGCAGCTAGAGGTA
CAGAGGTTCATATTTTTTAGATGGCTCATTAGTGATATTTTAGTCAAAAATTTCAGAAGTAC
GGCCACGGGTGATGGCCTGAACTAATATTTTATTCGAGGTGCCGCTACATCATCGTCTAAAG
TACACGCAAGATTCAACGGAAAAAAGAAACCGATCGATCGAGATCAGTTAGTTAATGAAATA
ACTAGATCAACTCATGTCGTCAAAAACAAAAGATGCTCATCTATGGACAACACACGCTGATG
ATTCGATCATCAAACAAAGGTGGTAGTAGTAGTAAAGCGTATCGTGTTTCTCATCAGAAAGA
AGAATTAAAGAAAAAACTAATCCCGTCTCGCGAGCCAGAGAAAATTCCCTACAAAAGCCACT
CCTTTGATTTGACATGCAAAAGCAAGGCTCCACTCCTCTACTACCCTACAACTACACAACAC
TGTCTCTCTATCTCCAAAGGCAGTAGCTGTATTGGCTTCCAGCTTTTCCTCTCTACCTCTAA
TGATAGCTTGGAGCAAGTTCAATAGTATAGCTAACTACTAGCTCTAATTCATCTATAATCAA
TCTAATAGCTTATTCATACAATAGTTATATACTACATTATTAATATCTGGTCCCATCTATCA
TACACACTGAGTCTGTGCTATAGCTGACTACAAATCTGTAGCCCGCTGCTCTTCTCTCTTCA
TTTATCTTCTTAAAATATATTTGCAGCTGGCTTATGGCTTATAGCTTGATATTGAGAGGGAG
AGGAGTGAGAGCTAGCTCAGCTCAGCTCAAGGTAAACAACAAGGCACACTCTTCCTACCTCT
TCTCCGGTTCTTCCTTTTTCCTCTTTCTCTTCGTCCAAGAACTTCACCTCAATAGCTCGAGC
TACGGCCTAACTTTTGCGTTGCGCAGGAGAGCTCGATCGCTGCACCAATACTTCACTGGAGA
TCGCCTAGCTGCAGCTAGCTAGCTTATCCTGCGTGTCTTGAGCTTCTCTC

FIGURA 20 (Cont.)

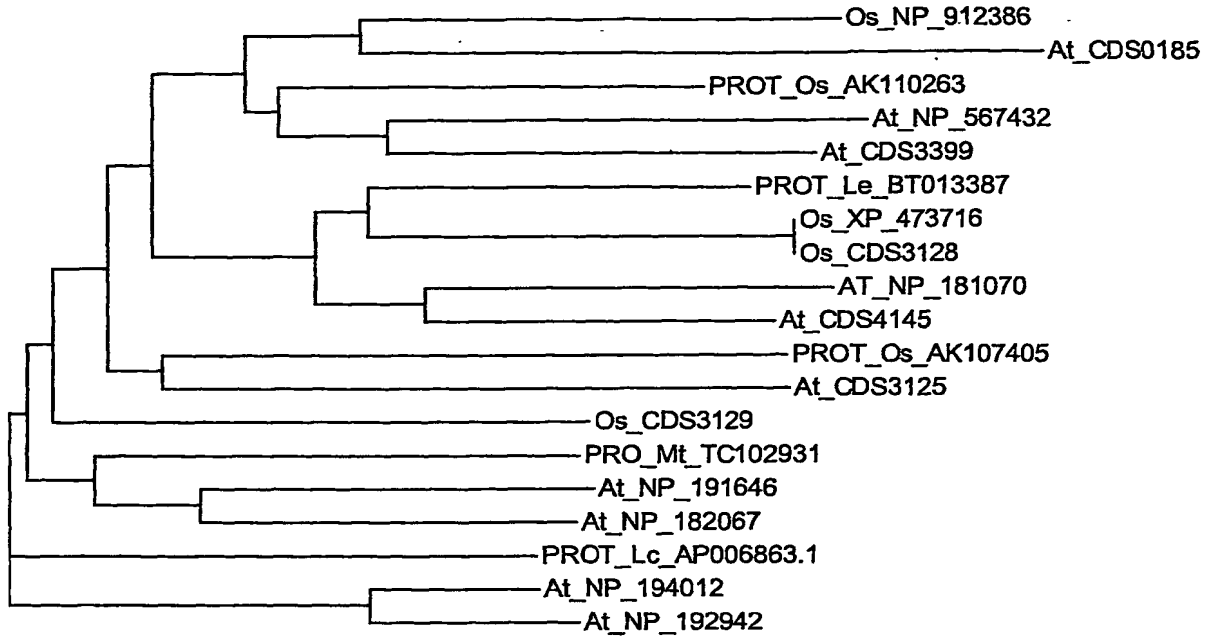


FIGURA 21

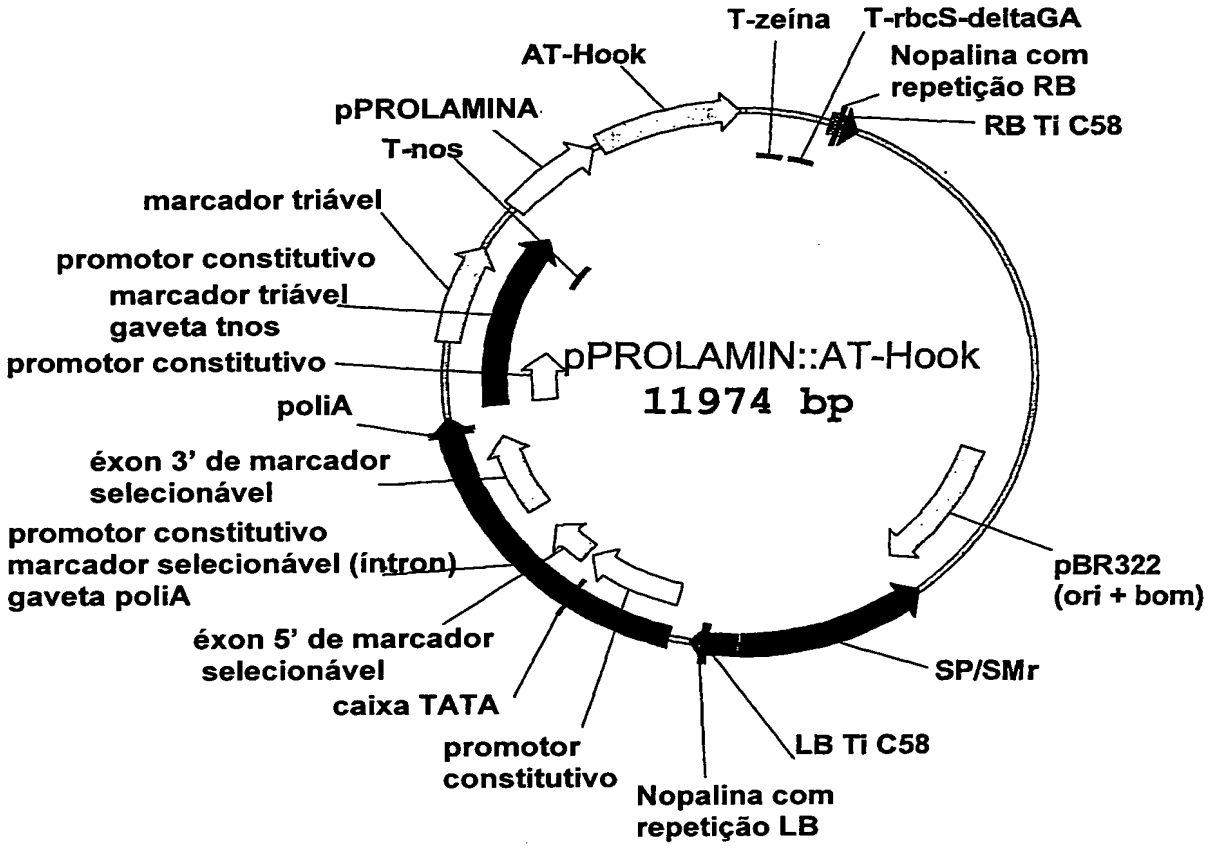


FIGURA 22

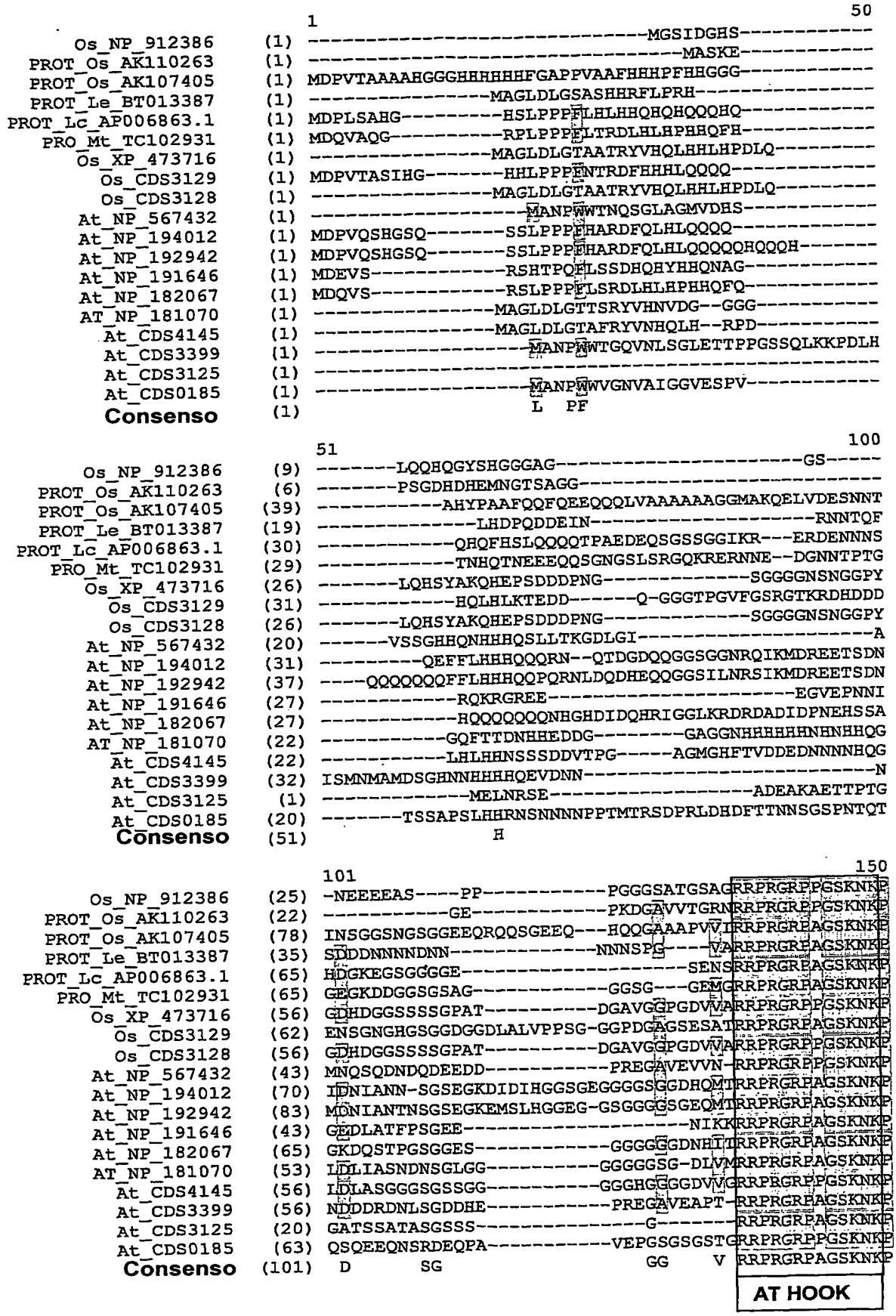
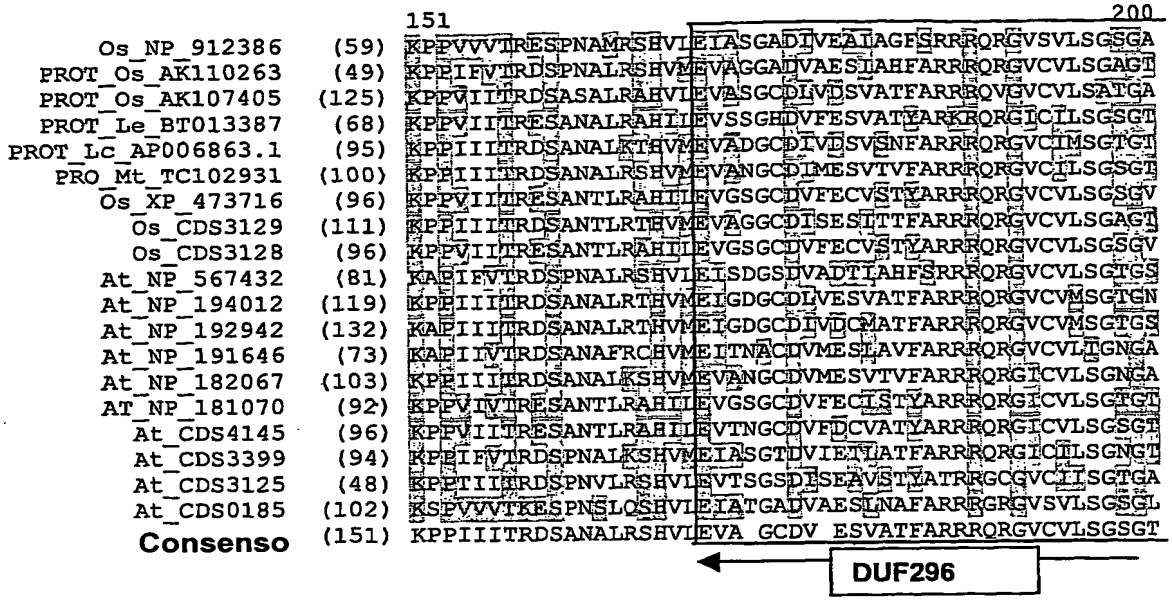


FIGURA 23



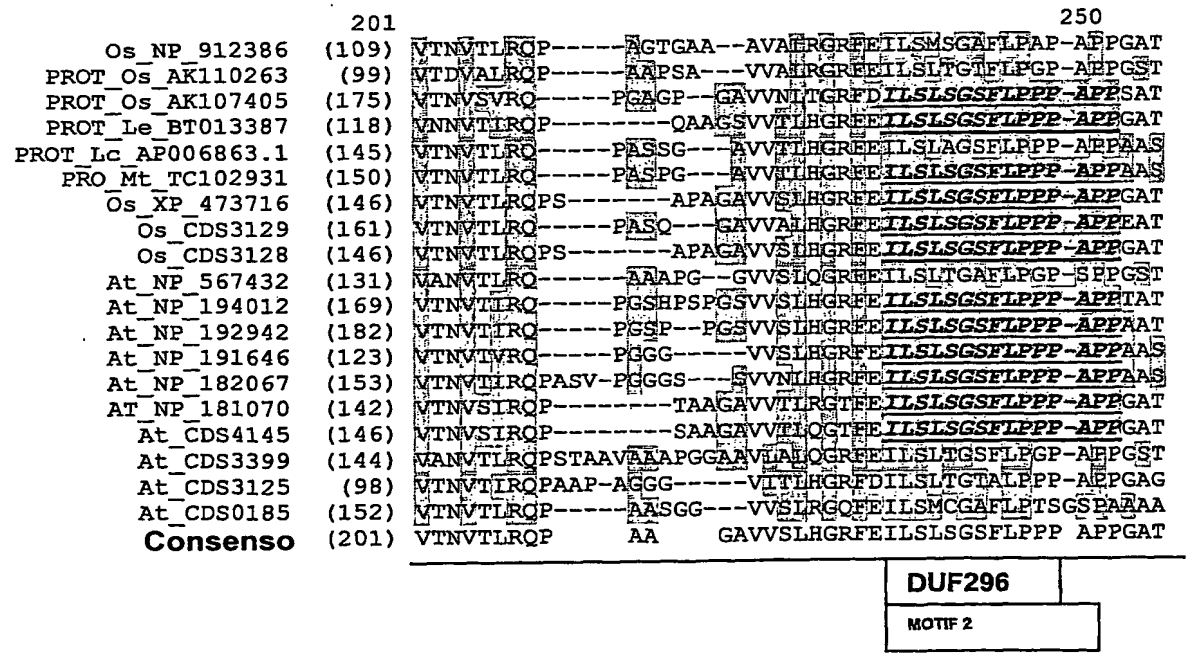


FIGURA 23 cont.

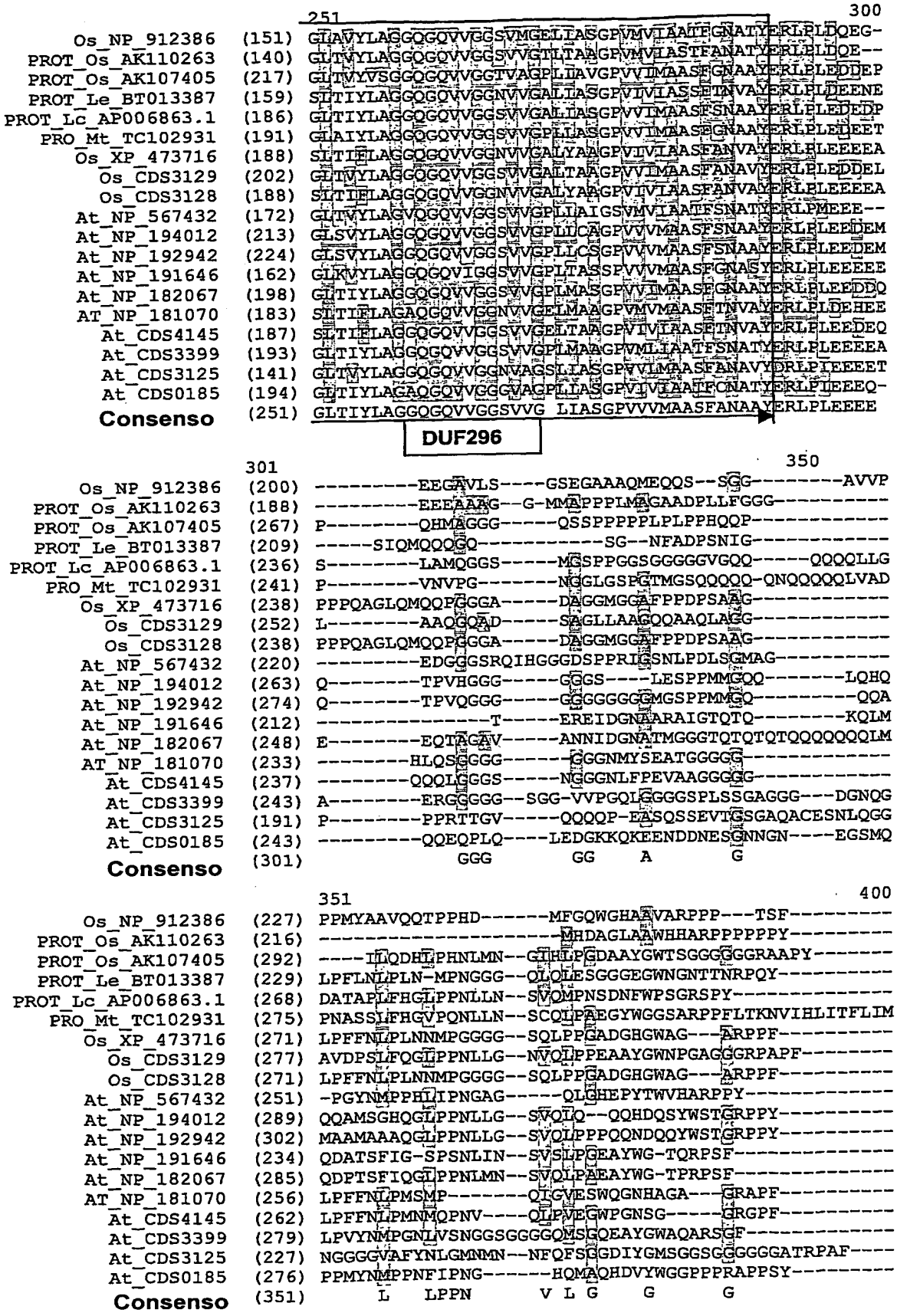


FIGURA 23 cont.

```
                                401
      Os_NP_912386   (259) ---
   PROT_Os_AK110263  (236) ---
   PROT_Os_AK107405  (329) ---
   PROT_Le_BT013387  (266) ---
PROT_Lc_AP006863.1   (302) ---
    PRO_Mt_TC102931  (323) FFV
      Os_XP_473716   (306) ---
        Os_CDS3129   (317) ---
        Os_CDS3128   (306) ---
      At_NP_567432   (282) ---
      At_NP_194012   (325) ---
      At_NP_192942   (340) ---
      At_NP_191646   (266) ---
      At_NP_182067   (318) ---
      AT_NP_181070   (286) ---
        At_CDS4145   (293) ---
        At_CDS3399   (316) ---
        At_CDS3125   (271) ---
        At_CDS0185   (311) ---
          Consenso   (401)
```

FIGURA 23 cont.

SEQ ID NO: 152, Oryza sativa – Sequencia de DNA

ATGGATCCGGTCACGGCATCAATACACGGTCACCATCTTCCTCCACCGTTCAACACCCGCGA
CTTCCATCACCATCTCCAGCAGCAGCAGCACCAGCTGCATCTCAAGACCGAGGATGACCAAG
GCGGCGGCACTCCGGGTGTCTTCGGCAGCCGCGGCACCAAGCGCGACCACGACGACGACGAG
AACAGTGGCAACGGCCATGGAAGCGGTGGTGACGGCGGTGACCTCGCGCTGGTACCCCCCTC
GGGTGGCGGGCCGGACGGCGCCGGGAGCGAGAGCGCCACGCGCCGCCCGAGGGGACGCCCGG
CGGGGTCCAAGAACAAGCCGAAGCCACCGATCATCATCACCAGGGACAGCGCCAACACGCTC
CGGACGCATGTCATGGAGGTGGCCGGCGGCTGCGACATCTCCGAGAGCATCACCACGTTCGC
GCGACGCCGGCAGCGCGGGGTTTGCGTGCTCAGCGGCGCCGGCACCGTCACTAACGTCACGC
TGCGGCAGCCCGCATCGCAGGGAGCGGTCGTTGCGCTCCACGGCCGGTTCGAGATACTCTCC
CTCTCCGGCTCCTTCCTCCCGCCGCCCGCCCCGCCGGAGGCCACGGGGCTCACCGTCTACCT
GGCCGGAGGCCAGGGCCAGGTCGTGGGCGGCAGCGTCGTCGGCGCGCTGACCGCGGCTGGGC
CTGTGGTGATAATGGCGGCGTCTTTTGCGAACGCGGTGTACGAGCGGCTGCCGTTGGAGGAC
GACGAGCTACTGGCGGCTCAAGGGCAAGCCGACAGCGCTGGGTTGCTCGCCGCGGGGCAGCA
AGCGGCGCAGCTCGCCGGCGGGGCCGTCGATCCAAGCCTCTTCCAAGGACTACCACCAAACC
TACTCGGAAACGTGCAGCTGCCGCCGGAAGCCGCCTACGGATGGAACCCTGGCGCCGGCGGT
GGCCGCCCCGGCGCCGTTCTGA

SEQ ID NO: 153, Oryza sativa – Sequencia de proteína

MDPVTASIHGHHLPPPFNTRDFHHHLQQQQHQLHLKTEDDQGGGTPGVFGSRGTKRDHDDDE
NSGNGHGSGGDGGDLALVPPSGGGPDGAGSESATRRPRGRPAGSKNKPKPPIIITRDSANTL
RTHVMEVAGGCDISESITTFARRRQRGVCVLSGAGTVTNVTLRQPASQGAVVALHGRFEILS
LSGSFLPPPAPPEATGLTVYLAGGQGQVVGGSVVGALTAAGPVVIMAASFANAVYERLPLED
DELLAAQGQADSAGLLAAGQQAAQLAGGAVDPSLFQGLPPNLLGNVQLPPEAAYGWNPGAGG
GRPAPF

FIGURA 24

SEQ ID NO: 154, Oryza sativa – Sequencia de DNA

ATGGCAGGTCTCGACCTCGGCACCGCCGCGACGCGCTACGTCCACCAGCTCCACCACCTCCA
CCCCGACCTCCAGCTGCAGCACAGCTACGCCAAGCAGCACGAGCCGTCCGACGACGACCCCA
ACGGCAGCGGCGGCGGCGGCAACAGCAACGGCGGGCCGTACGGGGACCATGACGGCGGGTCC
TCGTCGTCAGGCCCTGCCACCGACGGCGCGGTCGGCGGGCCCGGCGACGTGGTGGCGCGCCG
GCCGCGGGGGCGCCCGCCTGGCTCCAAGAACAAGCCGAAGCCGCCGGTGATCATCACGCGGG
AGAGCGCCAACACGCTGCGCGCCCACATCCTGGAGGTCGGGAGCGGCTGCGACGTGTTCGAG
TGCGTCTCCACGTACGCGCGCCGGCGGCAGCGCGGCGTGTGCGTGCTGAGCGGCAGCGGCGT
GGTCACCAACGTGACGCTGCGTCAGCCGTCGGCGCCCGCGGGCGCCGTCGTGTCGCTGCACG
GGAGGTTCGAGATCCTGTCGCTCTCGGGCTCCTTCCTCCCGCCGCCGGCTCCCCCCGGCGCC
ACCAGCCTCACCATCTTCCTCGCCGGGGGCCAGGGACAGGTCGTCGGCGGCAACGTCGTCGG
CGCGCTCTACGCCGCGGGCCCGGTCATCGTCATCGCGGCGTCCTTCGCCAACGTCGCCTACG
AGCGCCTCCCACTGGAGGAGGAGGAGGCGCCGCCGCCGCAGGCCGGCCTGCAGATGCAGCAG
CCCGGCGGCGGCGCCGATGCTGGTGGCATGGGTGGCGCGTTCCCGCCGGACCCGTCTGCCGC
CGGCCTCCCGTTCTTCAACCTGCCGCTCAACAACATGCCCGGTGGCGGCGGCTCACAGCTCC
CTCCCGGCGCCGACGGCCATGGCTGGGCCGGCGCACGGCCACCGTTCTGA

SEQ ID NO: 155, Oryza sativa – Sequencia de proteína

MAGLDLGTAATRYVHQLHHLHPDLQLQHSYAKQHEPSDDDPNGSGGGGNSNGGPYGDHDGGS
SSSGPATDGAVGGPGDVVARRPRGRPPGSKNKPKPPVIITRESANTLRAHILEVGSGCDVFE
CVSTYARRRQRGVCVLSGSGVVTNVTLRQPSAPAGAVVSLHGRFEILSLSGSFLPPPAPPGA
TSLTIFLAGGQGQVVGGNVVGALYAAGPVIVIAASFANVAYERLPLEEEEAPPPQAGLQMQQ
PGGGADAGGMGGAFPPDPSAAGLPFFNLPLNNMPGGGGSQLPPGADGHGWAGARPPF

SEQ ID NO: 156, Lotus corniculatus – Sequencia de DNA

ATGGATCCATTATCAGCACACGGCCACTCTCTTCCTCCTCCTTTTCTCCATCTGCACCACCA
ACACCAACACCAGCAGCAGCATCAGCAGCATCAGTTCCACTCTTTACAGCAGCAGCAAACCC
CAGCAGAAGATGAACAGAGTGGAAGCAGCGGCGGCATCAAAAGGGAACGCGATGAAAACAAC
AACAGCCATGACGGCAAAGAAGGCTCCGGAGGCGGAGGCGAAAGCGAGAATTCAAGAAGACC
TCGAGGAAGGCCCGCCGGATCGAAGAACAAGCCTAAGCCGCCCATCATCATCACCCGCGACA
GCGCCAACGCACTTAAGACCCACGTCATGGAGGTCGCCGACGGCTGCGACATCGTCGACAGC
GTCTCCAACTTTGCAAGACGCCGCCAGCGCGGCGTCTGCATCATGAGCGGCACTGGAACCGT
CACCAACGTCACTCTCAGGCAGCCAGCTTCTTCCGGCGCTGTTGTCACCCTCCACGGAAGGT
TTGAGATTCTCTCCCTGGCAGGATCGTTCCTGCCGCCGCCTGCTCCACCGGCAGCATCAGGT
TTGACCATTTACTTGGCTGGTGGACAAGGGCAGGTTGTTGGAGGAAGTGTTGTGGGAGCTCT
CATTGCTTCGGGACCTGTGGTTATCATGGCAGCTTCGTTCAGCAACGCTGCGTATGAGAGGC
TTCCTTTGGAAGATGAGGACCCTTCATTGGCAATGCAAGGAGGTTCAATGGGTTCACCACCC
GGTGGTAGTGGTGGTGGTGGTGGAGTTGGTCAGCAACAGCAGCAACAGCTTTTAGGGGATGC
AACTGCCCCACTTTTTCATGGTTTGCCTCCGAATCTTCTCAATTCTGTTCAGATGCCAAACT
CCGATAACTTCTGGCCATCTGGCCGCTCTCCTTACTGA

SEQ ID NO: 157, Lotus corniculatus – Sequencia de Proteína

MDPLSAHGHSLPPPFLHLHHQHQHQQQHQQHQFHSLQQQQTPAEDEQSGSSGGIKRERDENN
NSHDGKEGSGGGGESENSRRPRGRPAGSKNKPKPPIIITRDSANALKTHVMEVADGCDIVDS
VSNFARRRQRGVCIMSGTGTVTNVTLRQPASSGAVVTLHGRFEILSLAGSFLPPPAPPAASG
LTIYLAGGQGQVVGGSVVGALIASGPVVIMAASFSNAAYERLPLEDEDPSLAMQGGSMGSPP
GGSGGGGGVGQQQQQQLLGDATAPLFHGLPPNLLNSVQMPNSDNFWPSGRSPY

FIGURA 24 cont.

SEQ ID NO: 158, Arabidopsis thaliana – Sequencia de DNA

ATGGATCCAGTTCAATCTCATGGATCACAAAGCTCTCTTCCTCCTCCTTTCCATGCTAGAGA
TTTCCAATTACATCTTCAACAACAACAACAACATCAACAACAACATCAACAACAACAACAAC
AACAGTTCTTTCTCCACCATCATCAGCAACCACAAAGAAACCTTGATCAAGATCACGAGCAG
CAAGGAGGGTCAATATTGAATAGATCTATCAAGATGGATCGCGAAGAGACAAGCGATAACAT
GGACAACATCGCTAATACCAACAGCGGTAGCGAAGGTAAAGAGATGAGTTTACACGGAGGAG
AAGGAGGAAGCGGTGGTGGAGGAAGTGGAGAACAGATGACAAGAAGGCCAAGAGGAAGACCA
GCAGGATCCAAGAACAAACCTAAAGCTCCAATAATCATAACAAGAGACAGCGCAAACGCGCT
TCGAACTCACGTCATGGAGATAGGAGACGGATGTGACATAGTTGACTGTATGGCTACGTTCG
CTAGACGCCGCCAAAGAGGCGTTTGCGTTATGAGCGGTACAGGAAGCGTTACTAACGTCACT
ATACGTCAGCCTGGATCGCCACCTGGCTCGGTGGTTAGCCTTCACGGCCGGTTTGAAATCCT
CTCTCTTTCGGGATCTTTCTTGCCTCCGCCTGCGCCGCCTGCAGCCACCGGACTAAGCGTTT
ACCTAGCCGGAGGACAAGGGCAGGTCGTTGGAGGTAGTGTGGTGGGACCTTTGTTGTGTTCG
GGTCCTGTGGTGGTTATGGCGGCTTCTTTTAGCAATGCGGCGTACGAAAGGCTGCCTTTGGA
AGAAGATGAGATGCAGACGCCAGTTCAAGGAGGCGGTGGAGGAGGAGGAGGTGGTGGTGGAA
TGGGATCTCCCCCGATGATGGGACAGCAACAAGCTATGGCAGCTATGGCGGCGGCTCAAGGA
CTACCACCGAATCTTCTTGGTTCGGTTCAGTTGCCACCGCCACAACAGAATGATCAGCAGTA
TTGGTCTACGGGTCGGCCACCGTATTGA

SEQ ID NO: 159, Arabidopsis thaliana – Sequencia de Proteína

MDPVQSHGSQSSLPPPFHARDFQLHLQQQQQHQQQHQQQQQQQFFLHHHQQPQRNLDQDHEQ
QGGSILNRSIKMDREETSDNMDNIANTNSGSEGKEMSLHGGEGGSGGGGSGEQMTRRPRGRP
AGSKNKPKAPIIITRDSANALRTHVMEIGDGCDIVDCMATFARRRQRGVCVMSGTGSVTNVT
IRQPGSPPGSVVSLHGRFEILSLSGSFLPPPAPPAATGLSVYLAGGQGQVVGGSVVGPLLCS
GPVVVMAASFSNAAYERLPLEEDEMQTPVQGGGGGGGGGGGMGSPPMMGQQQAMAAMAAAQG
LPPNLLGSVQLPPPQQNDQQYWSTGRPPY

SEQ ID NO: 160, Arabidopsis thaliana – Sequencia de DNA

ATGGATCCAGTACAATCTCATGGATCACAAAGCTCTCTACCTCCTCCTTTCCACGCAAGAGA
CTTTCAATTACATCTTCAACAACAGCAACAAGAGTTCTTCCTCCACCATCACCAGCAACAAA
GAAACCAAACCGATGGTGACCAACAAGGAGGATCAGGAGGAAACCGACAAATCAAGATGGAT
CGTGAAGAGACAAGCGACAACATAGACAACATAGCTAACAACAGCGGTAGTGAAGGTAAAGA
CATAGATATACACGGTGGTTCAGGAGAAGGAGGTGGTGGCTCCGGAGGAGATCATCAGATGA
CAAGAAGACCAAGAGGAAGACCAGCGGGATCCAAGAACAAACCAAAACCACCGATTATCATC
ACACGGGACAGCGCAAACGCGCTTAGAACCCACGTGATGGAGATCGGAGATGGCTGCGACTT
AGTCGAAAGCGTTGCCACTTTTGCACGAAGACGCCAACGCGGCGTTTGCGTTATGAGCGGTA
CTGGAAATGTTACTAACGTCACTATACGTCAGCCTGGATCTCATCCTTCTCCTGGCTCGGTA
GTTAGTCTTCACGGAAGGTTCGAGATTCTATCTCTCTCAGGATCTTTTCTCCCTCCTCCGGC
TCCTCCTACAGCCACCGGATTGAGTGTTTACCTCGCTGGAGGACAAGGACAGGTGGTTGGAG
GAAGCGTAGTTGGTCCGTTGTTATGTGCTGGTCCTGTCGTTGTCATGGCTGCGTCTTTTAGC
AATGCGGCGTACGAAAGGTTGCCTTTAGAGGAAGATGAGATGCAGACGCCGGTTCATGGCGG
AGGAGGAGGAGGATCATTGGAGTCGCCGCCAATGATGGGACAACAACTGCAACATCAGCAAC
AAGCTATGTCAGGTCATCAAGGGTTACCACCTAATCTTCTTGGTTCGGTTCAGTTGCAGCAG
CAACATGATCAGTCTTATTGGTCAACGGGACGACCACCGTATTGA

FIGURA 24 cont.

SEQ ID NO: 161, Arabidopsis thaliana – Sequencia de Proteína

MDPVQSHGSQSSLPPPFHARDFQLHLQQQQQEFFLHHHQQQRNQTDGDQQGGSGGNRQIKMD
REETSDNIDNIANNSGSEGKDIDIHGGSGEGGGGSGGDHQMTRRPRGRPAGSKNKPKPPIII
TRDSANALRTHVMEIGDGCDLVESVATFARRRQRGVCVMSGTGNVTNVTIRQPGSHPSPGSV
VSLHGRFEILSLSGSFLPPPAPPTATGLSVYLAGGQGQVVGGSVVGPLLCAGPVVVMAASFS
NAAYERLPLEEDEMQTPVHGGGGGGSLESPPMMGQQLQHQQQAMSGHQGLPPNLLGSVQLQQ
QHDQSYWSTGRPPY

SEQ ID NO: 162, Arabidopsis thaliana – Sequencia de DNA

ATGGATCAGGTCTCTCGCTCTCTTCCTCCACCTTTTCTCTCAAGAGATCTCCATCTTCACCC
ACACCATCAATTCCAGCATCAGCAGCAGCAGCAGCAACAGAATCACGGCCACGATATAGACC
AGCACCGAATCGGTGGGCTAAAACGTGACCGAGATGCTGATATCGATCCCAACGAGCACTCT
TCAGCCGGAAAAGATCAAAGTACTCCTGGCTCCGGTGGAGAAAGCGGCGGCGGAGGAGGAGG
AGATAATCACATCACGAGAAGGCCACGTGGCAGACCAGCGGGATCTAAGAACAAACCAAAAC
CGCCAATCATCATCACTCGAGACAGCGCAAACGCTCTCAAATCTCATGTCATGGAAGTAGCA
AACGGATGTGACGTCATGGAAAGTGTCACCGTCTTCGCTCGCCGTCGCCAACGTGGCATCTG
CGTTTTGAGCGGAAACGGCGCCGTTACCAACGTTACCATAAGACAACCAGCTTCAGTACCTG
GTGGTGGCTCATCTGTCGTTAACTTACACGGACGTTTCGAGATTCTTTCTCTCTCGGGATCA
TTCCTTCCTCCTCCGGCTCCACCAGCTGCGTCAGGTCTAACGATTTACTTAGCCGGTGGTCA
GGGACAGGTTGTTGGAGGAAGCGTGGTTGGTCCACTCATGGCTTCAGGACCTGTAGTGATTA
TGGCAGCTTCGTTTGGAAACGCTGCGTATGAGAGACTGCCGTTGGAGGAAGACGATCAAGAA
GAGCAAACAGCTGGAGCGGTTGCTAATAATATCGATGGAAACGCAACAATGGGTGGTGGAAC
GCAAACGCAAACTCAGACGCAGCAGCAACAGCAACAACAGTTGATGCAAGATCCGACGTCGT
TTATACAAGGGTTGCCTCCGAATCTTATGAATTCTGTTCAATTGCCAGCTGAAGCTTATTGG
GGAACTCCGAGACCATCTTTCTAA

SEQ ID NO: 163, Arabidopsis thaliana – Sequencia de Proteína

MDQVSRSLPPPFLSRDLHLHPHHQFQHQQQQQQQNHGHDIDQHRIGGLKRDRDADIDPNEHS
SAGKDQSTPGSGGESGGGGGGDNHITRRPRGRPAGSKNKPKPPIIITRDSANALKSHVMEVA
NGCDVMESVTVFARRRQRGICVLSGNGAVTNVTIRQPASVPGGGSSVVNLHGRFEILSLSGS
FLPPPAPPAASGLTIYLAGGQGQVVGGSVVGPLMASGPVVIMAASFGNAAYERLPLEEDDQE
EQTAGAVANNIDGNATMGGGTQTQTQTQQQQQQQLMQDPTSFIQGLPPNLMNSVQLPAEAYW
GTPRPSF

SEQ ID NO: 164, Arabidopsis thaliana –Sequencia de DNA

ATGGATGAGGTATCTCGTTCTCATACACCGCAATTTCTATCAAGTGATCATCAGCACTATCA
CCATCAAAACGCTGGACGACAAAAACGCGGCAGAGAAGAAGAAGGAGTTGAACCCAACAATA
TAGGGGAAGACCTAGCCACCTTTCCTTCCGGAGAAGAGAATATCAAGAAGAGAAGGCCACGT
GGCAGACCTGCTGGTTCCAAGAACAAACCCAAAGCACCAATCATAGTCACTCGCGACTCCGC
GAACGCCTTCAGATGTCACGTCATGGAGATAACCAACGCCTGCGATGTAATGGAAAGCCTAG
CCGTCTTCGCTAGACGCCGTCAGCGTGGCGTTTGCGTCTTGACCGGAAACGGGGCCGTTACA
AACGTCACCGTTAGACAACCTGGCGGAGGCGTCGTCAGTTTACACGGACGGTTTGAGATTCT
TTCTCTCTCGGGTTCGTTTCTTCCTCCACCGGCACCACCAGCTGCGTCTGGTTTAAAGGTTT
ACTTAGCCGGTGGTCAAGGTCAAGTGATCGGAGGCAGTGTGGTGGGACCGCTTACGGCATCA
AGTCCGGTGGTCGTTATGGCAGCTTCATTTGGAAACGCATCTTACGAGAGGCTGCCACTAGA
GGAGGAGGAGGAAACTGAAAGAGAAATAGATGGAAACGCGGCTAGGGCGATTGGAACGCAAA
CGCAGAAACAGTTAATGCAAGATGCGACATCGTTTATTGGGTCGCCGTCGAATTTAATTAAC
TCTGTTTCGTTGCCAGGTGAAGCTTATTGGGGAACGCAACGACCGTCTTTCTAA

FIGURA 24 cont.

SEQ ID NO: 165, Arabidopsis thaliana – Sequencia de Proteína

MDEVSRSHTPQFLSSDHQHYHHQNAGRQKRGREEEGVEPNNIGEDLATFPSGEENIKKRRPR
GRPAGSKNKPKAPIIVTRDSANAFRCHVMEITNACDVMESLAVFARRRQRGVCVLTGNGAVT
NVTVRQPGGGVVSLHGRFEILSLSGSFLPPPAPPAASGLKVYLAGGQGQVIGGSVVGPLTAS
SPVVVMAASFGNASYERLPLEEEEETEREIDGNAARAIGTQTQKQLMQDATSFIGSPSNLIN
SVSLPGEAYWGTQRPSF

SEQ ID NO: 166, Arabidopsis thaliana – Sequencia de DNA

ATGGCGAATCCTTGGTGGGTAGGGAATGTTGCGATCGGTGGAGTTGAGAGTCCAGTGACGTC
ATCAGCTCCTTCTTTGCACCACAGAAACAGTAACAACAACAACCCACCGACTATGACTCGTT
CGGATCCAAGATTGGACCATGACTTCACCACCAACAACAGTGGAAGCCCTAATACCCAGACT
CAGAGCCAAGAAGAACAGAACAGCAGAGACGAGCAACCAGCTGTTGAACCCGGATCCGGATC
CGGGTCTACGGGTCGTCGTCCTAGAGGTAGACCTCCTGGTTCCAAGAACAAACCAAAGAGTC
CAGTTGTTGTTACCAAAGAAAGCCCTAACTCTCTCCAGAGCCATGTTCTTGAGATTGCTACG
GGAGCTGACGTGGCGGAAAGCTTAAACGCCTTTGCTCGTAGACGCGGCCGGGGCGTTTCGGT
GCTGAGCGGTAGTGGTTTGGTTACTAATGTTACTCTGCGTCAGCCTGCTGCATCCGGTGGAG
TTGTTAGTTTACGTGGTCAGTTTGAGATCTTGTCTATGTGTGGGGCTTTTCTTCCTACGTCT
GGCTCTCCTGCTGCAGCCGCTGGTTTAACCATTTACTTAGCTGGAGCTCAAGGTCAAGTTGT
GGGAGGTGGAGTTGCTGGCCCGCTTATTGCCTCTGGACCCGTTATTGTGATAGCTGCTACGT
TTTGCAATGCCACTTATGAGAGGTTACCGATTGAGGAAGAACAACAGCAAGAGCAGCCGCTT
CAACTAGAAGATGGGAAGAAGCAGAAAGAAGAGAATGATGATAACGAGAGTGGGAATAACGG
AAACGAAGGATCGATGCAGCCGCCGATGTATAATATGCCTCCTAATTTTATCCCAAATGGTC
ATCAAATGGCTCAACACGACGTGTATTGGGGTGGTCCTCCGCCTCGTGCTCCTCCTTCGTAT
TGA

SEQ ID NO: 167, Arabidopsis thaliana – Sequencia de proteína

MANPWWVGNVAIGGVESPVTSSAPSLHHRNSNNNNPPTMTRSDPRLDHDFTTNNSGSPNTQT
QSQEEQNSRDEQPAVEPGSGSGSTGRRPRGRPPGSKNKPKSPVVVTKESPNSLQSHVLEIAT
GADVAESLNAFARRRGRGVSVLSGSGLVTNVTLRQPAASGGVVSLRGQFEILSMCGAFLPTS
GSPAAAAGLTIYLAGAQGQVVGGGVAGPLIASGPVIVIAATFCNATYERLPIEEEQQQEQPL
QLEDGKKQKEENDDNESGNNGNEGSMQPPMYNMPPNFIPNGHQMAQHDVYWGGPPPRAPPSY

SEQ ID NO: 168, Oryza sativa – Sequencia de DNA

ATGGATCCGGTGACGGCGGCGGCGGCGCATGGGGGTGGGCACCACCACCACCACCACTTCGG
AGCGCCACCGGTGGCGGCGTTCCACCACCACCCGTTCCACCACGGCGGCGGGGCGCACTACC
CGGCGGCGTTCCAGCAGTTTCAGGAGGAGCAGCAGCAGCTTGTGGCGGCGGCGGCGGCGGCT
GGTGGGATGGCGAAGCAGGAGCTGGTGGATGAGAGCAACAACACCATCAACAGCGGCGGGAG
CAACGGGAGCGGCGGGGAGGAGCAGAGGCAGCAGTCCGGGGAGGAGCAGCACCAGCAAGGGG
CGGCGGCGCCGGTGGTGATCCGGCGTCCCAGGGGCCGCCCCGCCGGCTCCAAGAACAAGCCC
AAGCCTCCGGTCATCATCACGCGCGACAGCGCCAGCGCGCTGCGGGCGCACGTCCTCGAGGT
CGCCTCCGGGTGCGACCTCGTCGACAGCGTCGCCACGTTCGCGCGCCGCCGCCAGGTCGGTG
TCTGCGTGCTCAGCGCCACCGGCGCCGTCACCAACGTCTCCGTCCGGCAGCCCGGCGCGGGC
CCCGGCGCCGTCGTCAACCTCACCGGCCGCTTCGACATCCTCTCGCTGTCCGGCTCCTTCCT
CCCGCCGCCGGCGCCTCCCTCCGCCACCGGCCTCACCGTCTACGTCTCCGGCGGCCAGGGGC
AGGTCGTGGGCGGCACGGTCGCCGGACCGCTCATCGCCGTCGGCCCCGTCGTCATCATGGCC
GCCTCGTTCGGGAACGCCGCCTACGAGCGCCTCCCGCTCGAGGACGACGAGCCGCCGCAGCA
CATGGCGGGCGGCGGCCAGTCCTCGCCGCCGCCGCCGCCGCTGCCATTACCACCACACCAGC
AGCCGATTCTTCAAGACCATCTGCCACACAACCTGATGAACGGAATCCACCTCCCCGGCGAC
GCCGCCTACGGCTGGACCAGCGGCGGCGGCGGCGGCGGCCGCGCGGCGCCGTACTGA

FIGURA 24 cont.

SEQ ID NO: 169, Oryza sativa – Sequencia de Proteína

MDPVTAAAAHGGGHHHHHHFGAPPVAAFHHHPFHHGGGAHYPAAFQQFQEEQQQLVAAAAAA
GGMAKQELVDESNNTINSGGSNGSGGEEQRQQSGEEQHQQGAAAPVVIRRPRGRPAGSKNKP
KPPVIITRDSASALRAHVLEVASGCDLVDSVATFARRRQVGVCVLSATGAVTNVSVRQPGAG
PGAVVNLTGRFDILSLSGSFLPPPAPPSATGLTVYVSGGQGQVVGGTVAGPLIAVGPVVIMA
ASFGNAAYERLPLEDDEPPQHMAGGGQSSPPPPPLPLPPHQQPILQDHLPHNLMNGIHLPGD
AAYGWTSGGGGGGRAAPY

SEQ ID NO: 170, Oryza sativa – Sequencia de DNA

ATGGGGAGCATCGACGGCCACTCGCTGCAGCAGCATCAGGGGTACTCCCACGGCGGCGGCGC
GGGAGGGAGCAACGAGGAGGAGGAGGCGTCGCCGCCGCCCGGCGGTGGCTCGGCTACGGGGT
CGGCGGGCCGCCGGCCGAGGGGGAGGCCGCCGGGCTCCAAGAACAAGCCGAAGCCGCCCGTC
GTGGTGACGCGGGAGAGCCCCAACGCGATGCGTTCCCACGTGCTGGAGATCGCCAGCGGCGC
CGACATCGTCGAGGCCATCGCGGGCTTCTCCCGCCGCAGGCAGCGCGGCGTCTCCGTGCTCA
GCGGGAGCGGCGCCGTCACCAACGTCACGCTCCGGCAGCCCGCGGGGACTGGGGCCGCCGCC
GTCGCGCTGCGGGGGAGGTTCGAGATATTGTCCATGTCTGGCGCCTTCCTCCCGGCGCCGGC
GCCGCCAGGGGCCACGGGGCTCGCCGTGTACCTCGCCGGCGGGCAGGGGCAGGTGGTGGGTG
GGAGCGTCATGGGGGAGCTGATCGCGTCGGGCCCCGTCATGGTGATCGCGGCCACGTTCGGC
AACGCCACGTACGAGAGGCTGCCGCTGGACCAGGAAGGCGAGGAGGGCGCCGTGCTGTCCGG
GTCGGAGGGCGCCGCCGCGCAGATGGAGCAGCAGAGCAGCGGAGGCGCCGTCGTGCCCCCGC
CGATGTACGCCGCCGTCCAGCAGACGCCGCCGCACGACATGTTCGGGCAGTGGGGGCATGCA
GCGGTGGCTCGGCCGCCGCCGACATCGTTCTAG

SEQ ID NO: 171, Oryza sativa – Sequencia de Proteína

MGSIDGHSLQQHQGYSHGGGAGGSNEEEEASPPPGGGSATGSAGRRPRGRPPGSKNKPKPPV
VVTRESPNAMRSHVLEIASGADIVEAIAGFSRRRQRGVSVLSGSGAVTNVTLRQPAGTGAAA
VALRGRFEILSMSGAFLPAPAPPGATGLAVYLAGGQGQVVGGSVMGELIASGPVMVIAATFG
NATYERLPLDQEGEEGAVLSGSEGAAAQMEQQSSGGAVVPPPMYAAVQQTPPHDMFGQWGHA
AVARPPPTSF

SEQ ID NO: 172, Lycopersicon esculentum – Sequencia de DNA

ATGGCTGGTTTGGACTTAGGTTCCGCTTCTCATCATCGTTTTCTCCCTCGTCATCTCCACGA
TCCTCAAGACGATGAAATCAATCGCAACAACACTCAATTTTCCGATGATGACAACAACAACA
ACGATAACAATAACAATAATAGCCCCGGGGTAGCACGTAGACCTAGAGGTCGTCCCGCGGGG
TCCAAAAATAAGCCTAAGCCTCCCGTGATCATAACACGGGAGAGCGCTAACGCGCTCCGTGC
GCATATTTTAGAAGTGAGTAGCGGACATGATGTCTTTGAATCAGTTGCTACTTATGCTAGAA
AAAGACAAAGAGGAATTTGCATACTGAGCGGGAGCGGTACGGTGAATAACGTCACCATACGG
CAGCCACAGGCTGCCGGTTCTGTGGTGACGTTACACGGAAGATTCGAGATATTATCTTTATC
CGGATCTTTCCTACCACCACCCGCTCCCCCTGGGGCCACCAGCTTAACGATTTATTTAGCGG
GTGGTCAAGGTCAAGTTGTTGGTGGAAACGTTGTGGGTGCGCTAATTGCATCAGGACCGGTT
ATTGTTATTGCTTCGTCATTTACTAATGTTGCTTATGAGAGATTGCCTTTGGATGAAGAAAA
TGAGTCAATTCAGATGCAGCAACAAGGACAAAGTGGTAATTTTGCTGATCCATCTAATATTG
GATTACCTTTTCTTAATTTGCCATTAAACATGCCAAATGGTGGTGGTCAACTACAATTGGAA
AGTGGAGGAGGTGAAGGTTGGAATGGAAACACAACAAACAGGCCACAATATTAG

FIGURA 24 cont.

SEQ ID NO: 173, Lycopersicon esculentum – Sequencia de Proteína

MAGLDLGSASHHRFLPRHLHDPQDDEINRNNTQFSDDDNNNNDNNNNNSPGVARRPRGRPAG
SKNKPKPPVIITRESANALRAHILEVSSGHDVFESVATYARKRQRGICILSGSGTVNNVTIR
QPQAAGSVVTLHGRFEILSLSGSFLPPPAPPGATSLTIYLAGGQGQVVGGNVVGALIASGPV
IVIASSFTNVAYERLPLDEENESIQMQQQGQSGNFADPSNIGLPFLNLPLNMPNGGGQLQLE
SGGGEGWNGNTTNRPQY

SEQ ID NO: 174, Arabidopsis thaliana – Sequencia de DNA

ATGGAACTTAACAGATCTGAAGCAGACGAAGCAAAGGCCGAGACCACTCCCACCGGTGGAGC
CACCAGCTCAGCCACAGCCTCTGGCTCTTCCTCCGGACGTCGTCCACGTGGTCGTCCTGCAG
GTTCCAAAAACAAACCCAAACCTCCGACGATTATAACTAGAGATAGTCCTAACGTCCTTAGA
TCACACGTTCTTGAAGTCACCTCCGGTTCGGACATATCCGAGGCAGTCTCCACCTACGCCAC
TCGTCGCGGCTGCGGCGTTTGCATTATAAGCGGCACGGGTGCGGTCACTAACGTCACGATAC
GGCAACCTGCGGCTCCGGCTGGTGGAGGTGTGATTACCCTGCATGGTCGGTTTGACATTTTG
TCTTTGACCGGTACTGCGCTTCCACCGCCTGCACCACCGGGAGCAGGAGGTTTGACGGTGTA
TCTAGCCGGAGGTCAAGGACAAGTTGTAGGAGGGAATGTGGCTGGTTCGTTAATTGCTTCGG
GACCGGTAGTGTTGATGGCTGCTTCTTTTGCAAACGCAGTTTATGATAGGTTACCGATTGAA
GAGGAAGAAACCCCACCGCCGAGAACCACCGGGGTGCAGCAGCAGCAGCCGGAGGCGTCTCA
GTCGTCGGAGGTTACGGGGAGTGGGGCCCAGGCGTGTGAGTCAAACCTCCAAGGTGGAAATG
GTGGAGGAGGTGTTGCTTTCTACAATCTTGGAATGAATATGAACAATTTTCAATTCTCCGGG
GGAGATATTTACGGTATGAGCGGCGGTAGCGGAGGAGGTGGTGGCGGTGCGACTAGACCCGC
GTTTTAG

SEQ ID NO: 175, Arabidopsis thaliana – Sequencia de Proteína

MELNRSEADEAKAETTPTGGATSSATASGSSSGRRPRGRPAGSKNKPKPPTIITRDSPNVLR
SHVLEVTSGSDISEAVSTYATRRGCGVCIISGTGAVTNVTIRQPAAPAGGGVITLHGRFDIL
SLTGTALPPPAPPGAGGLTVYLAGGQGQVVGGNVAGSLIASGPVVLMAASFANAVYDRLPIE
EEETPPPRTTGVQQQQPEASQSSEVTGSGAQACESNLQGGNGGGGVAFYNLGMNMNNFQFSG
GDIYGMSGGSGGGGGGATRPAF

SEQ ID NO: 176, Arabidopsis thaliana – Sequencia de DNA

ATGGCGAATCCATGGTGGACAGGACAAGTGAACCTATCCGGCCTCGAAACGACGCCGCCTGG
TTCCTCTCAGTTAAAGAAACCAGATCTCCACATCTCCATGAACATGGCCATGGACTCAGGTC
ACAATAATCATCACCATCACCAAGAAGTCGATAACAACAACAACGACGACGATAGAGACAAC
TTGAGTGGAGACGACCACGAGCCACGTGAAGGAGCCGTAGAAGCCCCCACGCGCCGTCCACG
TGGACGTCCTGCTGGTTCCAAGAACAAACCAAAGCCACCGATCTTCGTCACTCGCGATTCTC
CAAATGCTCTCAAGAGCCATGTCATGGAGATCGCTAGTGGGACTGACGTCATCGAAACCCTA
GCTACTTTTGCTAGGCGGCGTCAACGTGGCATCTGCATCTTGAGCGGAAATGGCACAGTGGC
TAACGTCACCCTCCGTCAACCCTCGACCGCTGCCGTTGCGGCGGCTCCTGGTGGTGCGGCTG
TTTTGGCTTTACAAGGGAGGTTTGAGATTCTTTCTTTAACCGGTTCTTTCTTGCCAGGACCG
GCTCCACCTGGTTCCACCGGTTTAACGATTTACTTAGCCGGTGGTCAAGGTCAGGTTGTTGG
AGGAAGCGTGGTGGGCCCATTGATGGCAGCAGGTCCGGTGATGCTGATCGCCGCCACGTTCT
CTAACGCGACTTACGAGAGATTGCCATTGGAGGAGGAAGAGGCAGCAGAGAGAGGCGGTGGT
GGAGGCAGCGGAGGAGTGGTTCCGGGGCAGCTCGGAGGCGGAGGTTCGCCACTAAGCAGCGG
TGCTGGTGGAGGCGACGGTAACCAAGGACTTCCGGTGTATAATATGCCGGGAAATCTTGTTT
CTAATGGTGGCAGTGGTGGAGGAGGACAGATGAGCGGCCAAGAAGCTTATGGTTGGGCTCAA
GCTAGGTCAGGATTTTAA

FIGURA 24 cont.

SEQ ID NO: 177, Arabidopsis thaliana – Sequencia de Proteína

MANPWWTGQVNLSGLETTPPGSSQLKKPDLHISMNMAMDSGHNNHHHHQEVDNNNNDDDRDN
LSGDDHEPREGAVEAPTRRPRGRPAGSKNKPKPPIFVTRDSPNALKSHVMEIASGTDVIETL
ATFARRRQRGICILSGNGTVANVTLRQPSTAAVAAAPGGAAVLALQGRFEILSLTGSFLPGP
APPGSTGLTIYLAGGQGQVVGGSVVGPLMAAGPVMLIAATFSNATYERLPLEEEEAAERGGG
GGSGGVVPGQLGGGGSPLSSGAGGGDGNQGLPVYNMPGNLVSNGGSGGGGQMSGQEAYGWAQ
ARSGF

SEQ ID NO: 178, Oryza sativa – Sequencia de DNA

ATGGCGTCCAAGGAGCCAAGCGGCGACCACGACCACGAGATGAACGGGACCAGCGCCGGGGG
CGGCGAGCCCAAGGACGGCGCGGTGGTGACCGGCCGCAACCGGCGCCCCCGCGGACGGCCGC
CGGGCTCCAAGAACAAGCCCAAGCCGCCCATCTTCGTGACGCGGGACAGCCCGAACGCGCTG
CGCAGCCACGTCATGGAGGTGGCCGGCGGCGCCGATGTCGCCGAGTCCATCGCGCACTTCGC
GCGGCGGCGGCAGCGCGGCGTCTGCGTGCTCAGCGGGGCCGGCACCGTGACCGACGTGGCCC
TGCGCCAGCCGGCCGCGCCGAGCGCCGTGGTGGCGCTCCGTGGGCGGTTCGAGATCCTGTCC
CTGACGGGGACGTTCCTGCCGGGGCCGGCGCCGCCGGGCTCCACCGGGCTGACCGTGTACCT
CGCCGGCGGGCAGGGGCAGGTGGTGGGCGGCAGCGTGGTGGGGACGCTCACCGCGGCGGGGC
CGGTCATGGTGATCGCCTCCACCTTCGCCAACGCCACCTACGAGAGGCTGCCGCTGGATCAG
GAGGAGGAGGAAGCAGCGGCAGGCGGCATGATGGCGCCGCCGCCACTCATGGCCGGCGCCGC
CGATCCACTACTTTTCGGCGGGGGAATGCACGACGCCGGGCTTGCTGCATGGCACCATGCCC
GCCCTCCGCCGCCGCCGCCCTACTAG

SEQ ID NO: 179, Oryza sativa –Sequencia de Proteína

MASKEPSGDHDHEMNGTSAGGGEPKDGAVVTGRNRRPRGRPPGSKNKPKPPIFVTRDSPNAL
RSHVMEVAGGADVAESIAHFARRRQRGVCVLSGAGTVTDVALRQPAAPSAVVALRGRFEILS
LTGTFLPGPAPPGSTGLTVYLAGGQGQVVGGSVVGTLTAAGPVMVIASTFANATYERLPLDQ
EEEEAAAGGMMAPPPLMAGAADPLLFGGGMHDAGLAAWHHARPPPPPPY

SEQ ID NO: 180, Arabidopsis thaliana – Sequencia de DNA

ATGGCAAACCCTTGGTGGACGAACCAGAGTGGTTTAGCGGGCATGGTGGACCATTCGGTCTC
CTCAGGCCATCACCAAAACCATCACCACCAAAGTCTTCTTACCAAAGGAGATCTTGGAATAG
CCATGAATCAGAGCCAAGACAACGACCAAGACGAAGAAGATGATCCTAGAGAAGGAGCCGTT
GAGGTGGTCAACCGTAGACCAAGAGGTAGACCACCAGGATCCAAAAACAAACCCAAAGCTCC
AATCTTTGTGACAAGAGACAGCCCCAACGCACTCCGTAGCCATGTCTTGGAGATCTCCGACG
GCAGTGACGTCGCCGACACAATCGCTCACTTCTCAAGACGCAGGCAACGCGGCGTTTGCGTT
CTCAGCGGGACAGGCTCAGTCGCTAACGTCACCCTCCGCCAAGCCGCCGCACCAGGAGGTGT
GGTCTCTCTCCAAGGCAGGTTTGAAATCTTATCTTTAACCGGTGCTTTCCTCCCTGGACCTT
CCCCACCCGGGTCAACCGGTTTAACGGTTTACTTAGCCGGGGTCCAGGGTCAGGTCGTTGGA
GGTAGCGTTGTAGGCCCACTCTTAGCCATAGGGTCGGTCATGGTGATTGCTGCTACTTTCTC
TAACGCTACTTATGAGAGATTGCCCATGGAAGAAGAGGAAGACGGTGGCGGCTCAAGACAGA
TTCACGGAGGCGGTGACTCACCGCCCAGAATCGGTAGTAACCTGCCTGATCTATCAGGGATG
GCCGGGCCAGGCTACAATATGCCGCCGCATCTGATTCCAAATGGGGCTGGTCAGCTAGGGCA
CGAACCATATACATGGGTCCACGCAAGACCACCTTACTGA

SEQ ID NO: 181, Arabidopsis thaliana – Sequencia de Proteína

MANPWWTNQSGLAGMVDHSVSSGHHQNHHHQSLLTKGDLGIAMNQSQDNDQDEEDDPREGAV
EVVNRRPRGRPPGSKNKPKAPIFVTRDSPNALRSHVLEISDGSDVADTIAHFSRRRQRGVCV
LSGTGSVANVTLRQAAAPGGVVSLQGRFEILSLTGAFLPGPSPPGSTGLTVYLAGVQGQVVG
GSVVGPLLAIGSVMVIAATFSNATYERLPMEEEEDGGGSRQIHGGGDSPPRIGSNLPDLSGM
AGPGYNMPPHLIPNGAGQLGHEPYTWVHARPPY

FIGURA 24 cont.

SEQ ID NO: 182, Arabidopsis thaliana – Sequencia de DNA

ATGGCTGGTCTTGATCTAGGCACAGCTTTTCGTTACGTTAATCACCAGCTCCATCGTCCCGA
TCTCCACCTTCACCACAATTCCTCCTCCGATGACGTCACTCCCGGAGCCGGGATGGGTCATT
TCACCGTCGACGACGAAGACAACAACAACAACCATCAAGGTCTTGACTTAGCCTCTGGTGGA
GGATCAGGAAGCTCTGGAGGAGGAGGAGGTCACGGCGGGGGAGGAGACGTCGTTGGTCGTCG
TCCACGTGGCAGACCACCGGGATCCAAGAACAAACCGAAACCTCCGGTAATTATCACGCGCG
AGAGCGCAAACACTCTAAGAGCTCACATTCTTGAAGTAACAAACGGCTGCGATGTTTTCGAC
TGCGTTGCGACTTATGCTCGTCGGAGACAGCGAGGGATCTGCGTTCTGAGCGGTAGCGGAAC
GGTCACGAACGTCAGCATACGTCAGCCATCTGCGGCTGGAGCGGTTGTGACGCTACAAGGAA
CGTTCGAGATTCTTTCTCTCTCCGGATCGTTTCTTCCTCCTCCGGCACCTCCCGGAGCAACG
AGTTTGACAATTTTCTTAGCCGGAGGACAAGGTCAGGTGGTTGGAGGAAGCGTTGTGGGTGA
GCTTACGGCGGCTGGACCGGTGATTGTGATTGCAGCTTCGTTTACTAATGTTGCTTATGAGA
GACTTCCTTTAGAAGAAGATGAGCAGCAGCAACAGCTTGGAGGAGGATCTAACGGCGGAGGT
AATTTGTTTCCGGAGGTGGCAGCTGGAGGAGGAGGAGGACTTCCGTTCTTTAATTTACCGAT
GAATATGCAACCAAATGTGCAACTTCCGGTGGAAGGTTGGCCGGGGAATTCCGGTGGAAGAG
GTCCTTTCTGA

SEQ ID NO: 183, Arabidopsis thaliana – Sequencia de Proteína

MAGLDLGTAFRYVNHQLHRPDLHLHHNSSSDDVTPGAGMGHFTVDDEDNNNNHQGLDLASGG
GSGSSGGGGGHGGGGDVVGRRPRGRPPGSKNKPKPPVIITRESANTLRAHILEVTNGCDVFD
CVATYARRRQRGICVLSGSGTVTNVSIRQPSAAGAVVTLQGTFEILSLSGSFLPPPAPPGAT
SLTIFLAGGQGQVVGGSVVGELTAAGPVIVIAASFTNVAYERLPLEEDEQQQQLGGGSNGGG
NLFPEVAAGGGGGLPFFNLPMNMQPNVQLPVEGWPGNSGGRGPF

SEQ ID NO: 184, Oryza sativa – Sequencia de DNA

ATGGCAGGTCTCGACCTCGGCACCGCCGCGACGCGCTACGTCCACCAGCTCCACCACCTCCA
CCCCGACCTCCAGCTGCAGCACAGCTACGCCAAGCAGCACGAGCCGTCCGACGACGACCCCA
ACGGCAGCGGCGGCGGCGGCAACAGCAACGGCGGGCCGTACGGGGACCATGACGGCGGGTCC
TCGTCGTCAGGTCCTGCCACCGACGGCGCGGTCGGCGGGCCCGGCGACGTGGTGGCGCGCCG
GCCGCGGGGGCGCCCGCCTGGCTCCAAGAACAAGCCGAAGCCGCCGGTGATCATCACGCGGG
AGAGCGCCAACACGCTGCGCGCCCACATCCTGGAGGTCGGGAGCGGCTGCGACGTGTTCGAG
TGCGTCTCCACGTACGCGCGCCGGCGGCAGCGCGGCGTGTGCGTGCTGAGCGGCAGCGGCGT
GGTCACCAACGTGACGCTGCGTCAGCCGTCGGCGCCCGCGGGCGCCGTCGTGTCGCTGCACG
GGAGGTTCGAGATCCTGTCGCTCTCGGGCTCCTTCCTCCCGCCGCCGGCTCCCCCCGGCGCC
ACCAGCCTCACCATCTTCCTCGCCGGGGGCCAGGGACAGGTCGTCGGCGGCAACGTCGTCGG
CGCGCTCTACGCCGCGGGCCCGGTCATCGTCATCGCGGCGTCCTTCGCCAACGTCGCCTACG
AGCGCCTCCCACTGGAGGAGGAGGAGGCGCCGCCGCCGCAGGCCGGCCTGCAGATGCAGCAG
CCCGGCGGCGGCGCCGATGCTGGTGGCATGGGTGGCGCGTTCCCGCCGGACCCGTCTGCCGC
CGGCCTCCCGTTCTTCAACCTGCCGCTCAACAACATGCCCGGTGGCGGCGGCTCACAGCTCC
CTCCCGGCGCCGACGGCCATGGCTGGGCCGGCGCACGGCCACCGTTCTGA

SEQ ID NO: 185, Oryza sativa – Sequencia de Proteína

MAGLDLGTAATRYVHQLHHLHPDLQLQHSYAKQHEPSDDDPNGSGGGGNSNGGPYGDHDGGS
SSSGPATDGAVGGPGDVVARRPRGRPPGSKNKPKPPVIITRESANTLRAHILEVGSGCDVFE
CVSTYARRRQRGVCVLSGSGVVTNVTLRQPSAPAGAVVSLHGRFEILSLSGSFLPPPAPPGA
TSLTIFLAGGQGQVVGGNVVGALYAAGPVIVIAASFANVAYERLPLEEEEAPPPQAGLQMQQ
PGGGADAGGMGGAFPPDPSAAGLPFFNLPLNNMPGGGGSQLPPGADGHGWAGARPPF

FIGURA 24 cont.

SEQ ID NO: 186, Arabidopsis thaliana – Sequencia de DNA

ATGGCTGGTCTCGATCTAGGCACAACTTCTCGCTACGTCCACAACGTCGATGGTGGCGGCGG
CGGACAGTTCACCACCGACAACCACCACGAAGATGACGGTGGCGCTGGAGGAAACCACCATC
ATCACCATCATAATCATAATCACCATCAAGGTTTAGATTTAATAGCTTCTAATGATAACTCT
GGACTAGGCGGCGGTGGAGGAGGAGGGAGCGGTGACCTCGTCATGCGTCGGCCACGTGGCCG
TCCAGCTGGATCGAAGAACAAACCGAAGCCGCCGGTGATTGTCACGCGCGAGAGCGCAAACA
CTCTTAGGGCTCACATTCTTGAAGTTGGAAGTGGCTGCGACGTTTTCGAATGTATCTCCACT
TACGCTCGTCGGAGACAGCGCGGGATTTGCGTTTTATCCGGGACGGGAACCGTCACTAACGT
CAGCATCCGTCAGCCTACGGCGGCCGGAGCTGTTGTGACTCTGCGGGGTACTTTTGAGATTC
TTTCCCTCTCCGGATCTTTTCTTCCGCCACCTGCTCCTCCAGGGGCGACTAGCTTGACGATA
TTCCTCGCTGGAGCTCAAGGACAGGTCGTCGGAGGTAACGTAGTTGGTGAGTTAATGGCGGC
GGGGCCGGTAATGGTCATGGCAGCGTCTTTTACAAACGTGGCTTACGAAAGGTTGCCTTTGG
ACGAGCATGAGGAGCACTTGCAAAGTGGCGGCGGCGGAGGTGGAGGGAATATGTACTCGGAA
GCCACTGGCGGTGGCGGAGGGTTGCCTTTCTTTAATTTGCCGATGAGTATGCCTCAGATTGG
AGTTGAAAGTTGGCAGGGGAATCACGCCGGCGCCGGTAGGGCTCCGTTTTAG

SEQ ID NO: 187, Arabidopsis thaliana – Sequencia de Proteína

MAGLDLGTTSRYVHNVDGGGGGQFTTDNHHEDDGGAGGNHHHHHHNHNHHQGLDLIASNDNS
GLGGGGGGGSGDLVMRRPRGRPAGSKNKPKPPVIVTRESANTLRAHILEVGSGCDVFECIST
YARRRQRGICVLSGTGTVTNVSIRQPTAAGAVVTLRGTFEILSLSGSFLPPPAPPGATSLTI
FLAGAQGQVVGGNVVGELMAAGPVMVMAASFTNVAYERLPLDEHEEHLQSGGGGGGGNMYSE
ATGGGGGLPFFNLPMSMPQIGVESWQGNHAGAGRAPF

SEQ ID NO: 188, Medicago truncatula - Sequencia de DNA

ATGGATCAAGTAGCACAAGGTCGTCCTCTTCCACCTCCATTTCTCACTAGAGACCTTCATCT
TCATCCTCACCATCAATTTCACACCAACCACCAAACCAATGAAGAGGAACAACAAAGTGGCA
ATGGGAGCTTAAGCCGAGGCCAAAAAAGAGAACGAAACAACGAAGACGGCAACAACACTCCC
ACCGGAGGAGAAGGAAAAGACGACGGTGGTAGCGGAAGTGCTGGTGGAGGAAGTGGTGGTGA
GATGGGAAGAAGACCAAGAGGAAGACCAGCAGGTTCGAAAAACAAACCAAAACCACCTATCA
TCATCACGAGGGACAGCGCGAACGCACTCCGATCCCACGTGATGGAAGTTGCAAATGGATGT
GACATCATGGAAAGTGTGACGGTCTTTGCGCGAAGGAGGCAGCGTGGTGTCTGCATCCTTAG
CGGAAGTGGGACCGTCACAAACGTGACTCTCCGTCAACCAGCATCGCCTGGTGCGGTAGTCA
CACTTCATGGAAGATTTGAGATATTATCATTATCTGGCTCTTTCCTGCCGCCGCCTGCTCCA
CCAGCGGCGTCAGGATTAGCCATATATCTAGCTGGTGGACAAGGACAGGTCGTCGGTGGTAG
CGTGGTGGGACCGTTGTTAGCTTCCGGTCCGGTTGTTATCATGGCAGCTTCCTTTGGAAATG
CTGCTTATGAAAGGCTACCTTTAGAAGATGAAGAAACACCAGTGAATGTGCCAGGAAATGGA
GGGTTAGGGTCACCGGGAACCATGGGAAGTCAACAACAACAGCAGCAGAACCAGCAACAGCA
ACAACTTGTAGCAGATCCTAATGCTTCATCACTTTTCCATGGAGTTCCTCAAAATCTTCTCA
ATTCATGCCAATTACCAGCTGAAGGTTATTGGGGTGGAAGTGCTCGTCCTCCTTTTTTAACC
AAAAATGTTATTCACTTAATCACTTTTCTCATCATGTTTTTCGTTTAA

SEQ ID NO: 189, Medicago truncatula – Sequencia de Proteína

MDQVAQGRPLPPPFLTRDLHLHPHHQFHTNHQTNEEEQQSGNGSLSRGQKRERNNEDGNNTP
TGGEGKDDGGSGSAGGGSGGEMGRRPRGRPAGSKNKPKPPIIITRDSANALRSHVMEVANGC
DIMESVTVFARRRQRGVCILSGSGTVTNVTLRQPASPGAVVTLHGRFEILSLSGSFLPPPAP
PAASGLAIYLAGGQGQVVGGSVVGPLLASGPVVIMAASFGNAAYERLPLEDEETPVNVPGNG
GLGSPGTMGSQQQQQQNQQQQQLVADPNASSLFHGVPQNLLNSCQLPAEGYWGGSARPPFLT
KNVIHLITFLIMFFV

FIGURA 24 cont.

SEQ ID NO: 190, Seqüência artificial – motivo 1
QGQ V/I GG

SEQ ID NO: 191, Seqüência artificial – motivo 2
ILSLSGSFLPPPAPP

SEQ ID NO: 192, Seqüência artificial – motivo 3
NATYERLP

SEQ ID NO: 193, Seqüência artificial – motivo 4
SFTNVAYERLPL

SEQ ID NO: 194, Seqüência artificial – motivo 5
GRFEILSLTGSFLPGPAPPGSTGLTIYLAGGQGQVVGGSVVG

SEQ ID NO: 195, Oryza sativa – promotor de prolamina
CTTCTACATCGGCTTAGGTGTAGCAACACGACTTTATTATTATTATTATTATTATTATT
ATTTTACAAAAATATAAAATAGATCAGTCCCTCACCACAAGTAGAGCAAGTTGGTGAGTTAT
TGTAAAGTTCTACAAAGCTAATTTAAAAGTTATTGCATTAACTTATTTCATATTACAAACAA
GAGTGTCAATGGAACAATGAAAACCATATGACATACTATAATTTTGTTTTTATTATTGAAAT
TATATAATTCAAAGAGAATAAATCCACATAGCCGTAAAGTTCTACATGTGGTGCATTACCAA
AATATATATAGCTTACAAAACATGACAAGCTTAGTTTGAAAAATTGCAATCCTTATCACATT
GACACATAAAGTGAGTGATGAGTCATAATATTATTTTCTTTGCTACCCATCATGTATATATG
ATAGCCACAAAGTTACTTTGATGATGATATCAAAGAACATTTTTAGGTGCACCTAACAGAAT
ATCCAAATAATATGACTCACTTAGATCATAATAGAGCATCAAGTAAAACTAACACTCTAAAG
CAACCGATGGGAAAGCATCTATAAATAGACAAGCACAATGAAAATCCTCATCATCCTTCACC
ACAATTCAAATATTATAGTTGAAGCATAGTAGTA

SEQ ID NO: 196, Seqüência artificial – iniciador
GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGGATCCGGTCACGG

SEQ ID NO: 197, Seqüência artificial – iniciador
GGGGACCACTTTGTACAAGAAAGCTGGGTGGAATCGATCCATCTCAGAA

FIGURA 24 cont.

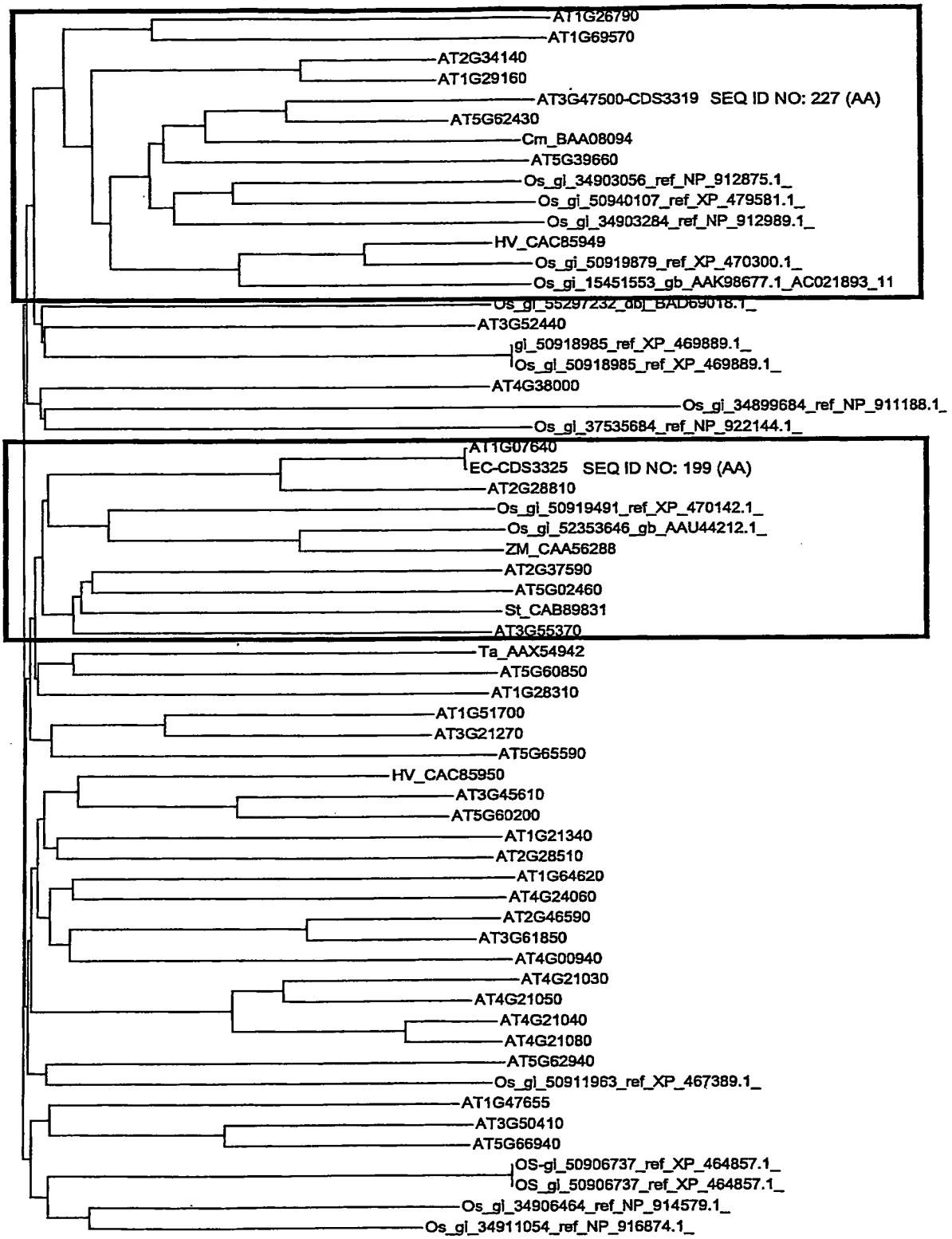

AT1G26790
AT1G69570
AT2G34140
AT1G29160
AT3G47500-CDS3319 SEQ ID NO: 227 (AA)
AT5G62430
Cm_BAA08094
AT5G39660
Os_gi_34903056_ref_NP_912875.1_
Os_gi_50940107_ref_XP_479581.1_
Os_gi_34903284_ref_NP_912989.1_
HV_CAC85949
Os_gi_50919879_ref_XP_470300.1_
Os_gi_15451553_gb_AAK98677.1_AC021893_11
Os_gi_55297232_db|_BAD69018.1_
AT3G52440
gi_50918985_ref_XP_469889.1_
Os_gi_50918985_ref_XP_469889.1_
AT4G38000
Os_gi_34899684_ref_NP_911188.1_
Os_gi_37535684_ref_NP_922144.1_
AT1G07640
EC-CDS3325 SEQ ID NO: 199 (AA)
AT2G28810
Os_gi_50919491_ref_XP_470142.1_
Os_gi_52353646_gb_AAU44212.1_
ZM_CAA56288
AT2G37590
AT5G02460
St_CAB89831
AT3G55370
Ta_AAX54942
AT5G60850
AT1G28310
AT1G51700
AT3G21270
AT5G65590
HV_CAC85950
AT3G45610
AT5G60200
AT1G21340
AT2G28510
AT1G64620
AT4G24060
AT2G46590
AT3G61850
AT4G00940
AT4G21030
AT4G21050
AT4G21040
AT4G21080
AT5G62940
Os_gi_50911963_ref_XP_467389.1_
AT1G47655
AT3G50410
AT5G66940
OS-gi_50906737_ref_XP_464857.1_
OS_gi_50906737_ref_XP_464857.1_
Os_gi_34906464_ref_NP_914579.1_
Os_gi_34911054_ref_NP_916874.1_

FIGURA 25

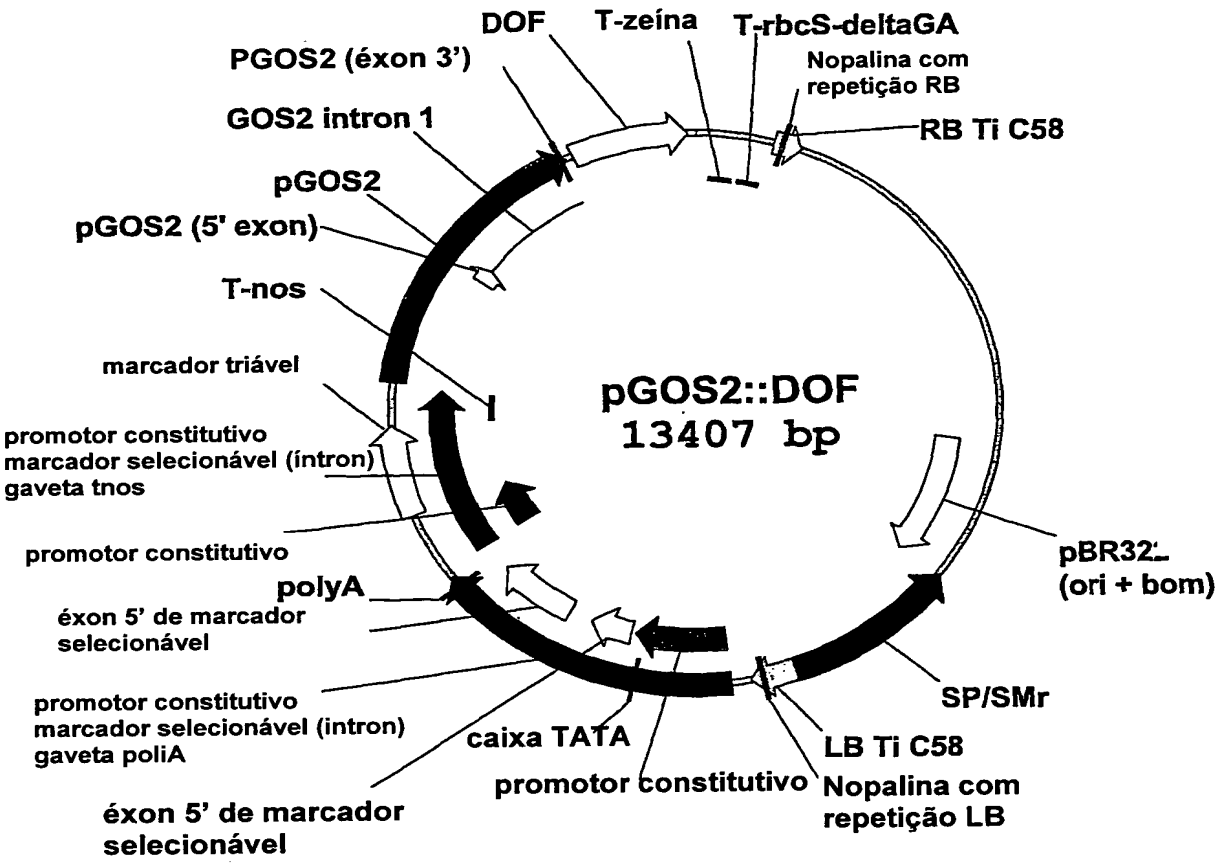

DOF
T-zeína
T-rbcS-deltaGA
Nopalina com repetição RB
RB Ti C58
PGOS2 (éxon 3')
GOS2 intron 1
pGOS2
pGOS2 (5' exon)
T-nos
marcador triável
promotor constitutivo
marcador selecionável (íntron)
gaveta tnos
promotor constitutivo
polyA
éxon 5' de marcador selecionável
promotor constitutivo
marcador selecionável (intron)
gaveta poliA
éxon 5' de marcador selecionável
pGOS2::DOF
13407 bp
pBR32
(ori + bom)
SP/SMr
LB Ti C58
Nopalina com repetição LB
promotor constitutivo
caixa TATA

FIGURA 26

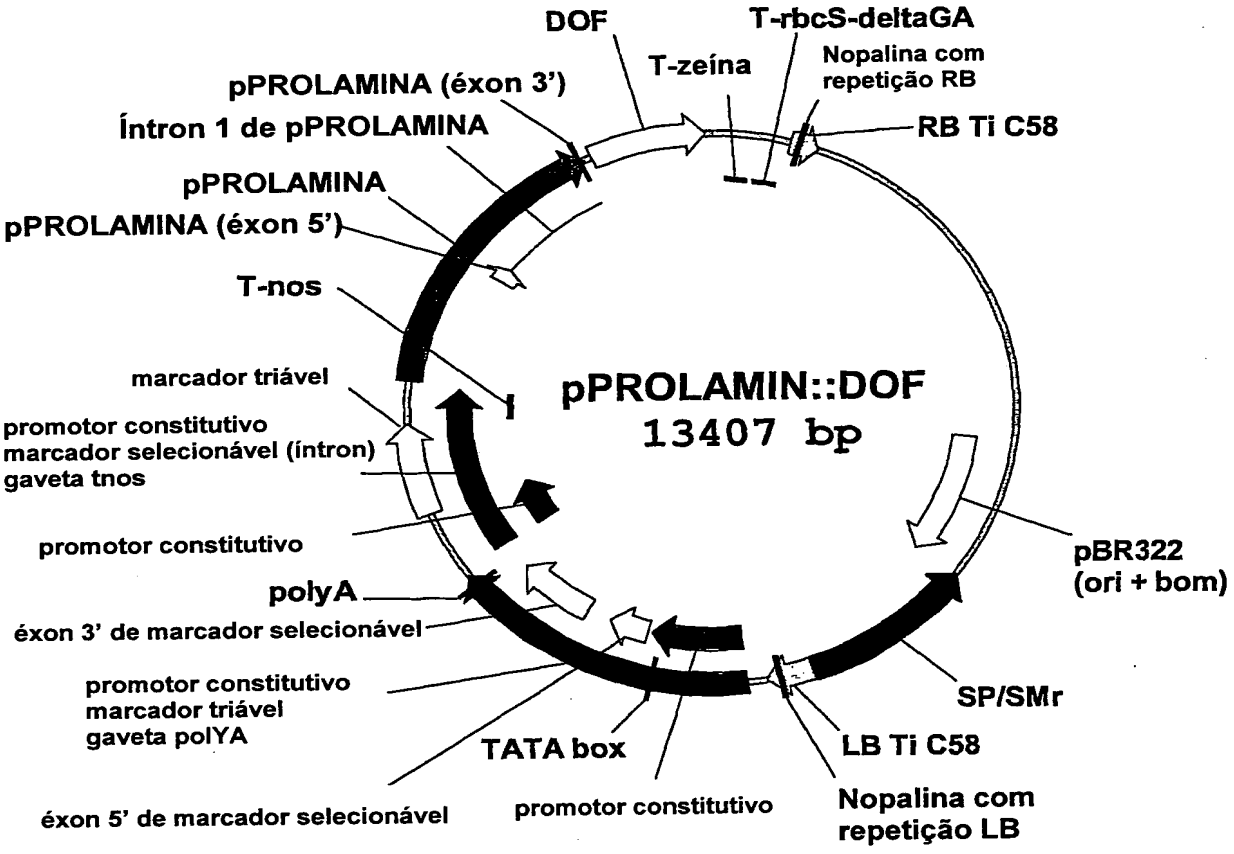

DOF
T-rbcS-deltaGA
T-zeína
Nopalina com repetição RB
RB Ti C58
pPROLAMINA (éxon 3')
Íntron 1 de pPROLAMINA
pPROLAMINA
pPROLAMINA (éxon 5')
T-nos
marcador triável
promotor constitutivo
marcador selecionável (íntron)
gaveta tnos
promotor constitutivo
polyA
éxon 3' de marcador selecionável
promotor constitutivo
marcador triável
gaveta polYA
éxon 5' de marcador selecionável
TATA box
promotor constitutivo
pPROLAMIN::DOF
13407 bp
pBR322
(ori + bom)
SP/SMr
LB Ti C58
Nopalina com repetição LB

FIGURA 27

SEQ ID NO: 198, Arabidopsis thaliana – DNA

ATGGGTGGATCGATGTCGGAGAGAGCAAGGCAGGCCAACATTCCTCCACTAGCGGGACCCCT
AAAGTGTCCTCGATGCGACTCCAGCAACACTAAGTTCTGTTACTACAACAACTATAACCTCA
CTCAGCCTCGTCACTTCTGCAAAGGTTGCCGTCGCTACTGGACACAAGGGGGCGCCCTGAGA
AACGTCCCTGTAGGTGGAGGCTGCCGGAGGAATAACAAGAAGGGCAAAAATGGAAATTTAAA
ATCTTCTTCTTCTTCGTCCAAACAGTCTTCCTCGGTCAACGCTCAAAGTCCTAGCTCAGGAC
AGCTAAGGACAAATCATCAGTTCCCTTTTTCACCAACTCTTTACAATCTCACTCAACTCGGA
GGTATTGGTTTGAACTTAGCCGCTACTAATGGCAACAACCAAGCTCACCAGATCGGTTCCAG
TTTGATGATGAGCGATCTAGGGTTTCTCCATGGACGAAATACTTCAACTCCGATGACGGGAA
ACATTCATGAAAACAACAACAATAATAACAATGAAAACAACCTAATGGCATCCGTTGGATCT
TTGAGCCCCTTTGCTCTCTTCGATCCAACGACGGGGCTATACGCTTTCCAGAACGACGGTAA
TATCGGGAACAACGTTGGGATATCTGGCTCTTCTACTTCCATGGTTGATTCTAGGGTTTATC
AGACGCCTCCGGTGAAGATGGAAGAACAACCTAATTTGGCTAACTTGTCTAGACCGGTCTCC
GGTTTGACGTCTCCTGGGAATCAAACAAATCAGTACTTTTGGCCTGGTTCGGATTTCTCGGG
TCCTTCTAATGATCTCTTGTGA

SEQ ID NO: 199, Arabidopsis thaliana – Proteína

MGGSMSERARQANIPPLAGPLKCPRCDSSNTKFCYYNNYNLTQPRHFCKGCRRYWTQGGALR
NVPVGGGCRRNNKKGKNGNLKSSSSSSKQSSSVNAQSPSSGQLRTNHQFPFSPTLYNLTQLG
GIGLNLAATNGNNQAHQIGSSLMMSDLGFLHGRNTSTPMTGNIHENNNNNNNENNLMASVGS
LSPFALFDPTTGLYAFQNDGNIGNNVGISGSSTSMVDSRVYQTPPVKMEEQPNLANLSRPVS
GLTSPGNQTNQYFWPGSDFSGPSNDLL

SEQ ID NO: 200, Sequencia artificial proteína – domínio DOF

PLAGPLKCPRCDSSNTKFCYYNNYNLTQPRHFCKGCRRYWTQGGALRNVPVGGGCRRNNKKG
K

SEQ ID NO: 201, Arabidopsis thaliana – DNA – ARF155816

ATGGCGGAGAGAGCAAGGCAGGCCAACATTCCTCCACTAGCGGGACCCCTAAAGTGTCCTCG
ATGCGACTCCAGCAACACTAAGTTCTGTTACTACAACAACTATAACCTCACTCAGCCTCGTC
ACTTCTGCAAAGGTTGCCGTCGCTACTGGACACAAGGGGGCGCCCTGAGAAACGTCCCTGTA
GGTGGAGGCTGCCGGAGGAATAACAAGAAGGGCAAAAATGGAAATTTAAAATCTTCTTCTTC
TTCGTCCAAACAGTCTTCCTCGGTCAACGCTCAAAGTCCTAGCTCAGGACAGCTAAGGACAA
ATCATCAGTTCCCTTTTTCACCAACTCTTTACAATCTCACTCAACTCGGAGGTATTGGTTTG
AACTTAGCCGCTACTAATGGCAACAACCAAGCTCACCAGATCGGTTCCAGTTTGATGATGAG
CGATCTAGGGTTTCTCCATGGACGAAATACTTCAACTCCGATGACGGGAAACATTCATGAAA
ACAACAACAATAATAACAATGAAAACAACCTAATGGCATCCGTTGGATCTTTGAGCCCCTTT
GCTCTCTTCGATCCAACGACGGGGCTATACGCTTTCCAGAACGACGGTAATATCGGGAACAA
CGTTGGGATATCTGGTTCTTCTACTTCCATGGTTGATTCTAGGGTTTATCAGACGCTCCGGT
GA

SEQ ID NO: 202, Arabidopsis thaliana – Proteína

MAERARQANIPPLAGPLKCPRCDSSNTKFCYYNNYNLTQPRHFCKGCRRYWTQGGALRNVPV
GGGCRRNNKKGKNGNLKSSSSSSKQSSSVNAQSPSSGQLRTNHQFPFSPTLYNLTQLGGIGL
NLAATNGNNQAHQIGSSLMMSDLGFLHGRNTSTPMTGNIHENNNNNNNENNLMASVGSLSPF
ALFDPTTGLYAFQNDGNIGNNVGISGSSTSMVDSRVYQTLR

FIGURA 28

SEQ ID NO: 203, Arabidopsis thaliana - DNA - AY088198

ATGCCTACGAATTCGAATCATCAGCATCATCTTCAACACCAGCTTAACGAAAATGGAAGTAT
AATAAGTGGCCACGGACTAGTACTCTCTCACCAACTTCCACCTCTCCAAGCAAACCCTAACC
CTAACCACCACCATGTCGCTACCTCTGCTGGTCTTCCGTCAAGGATGGGTGGATCGATGGCG
GAGAGAGCAAGGCAGGCCAACATTCCTCCACTAGCGGGACCCCTAAAGTGTCCTCGATGCGA
CTCCAGCAACACTAAGTTCTGTTACTACAACAACTATAACCTCACTCAGCCTCGTTACTTCT
GCAAAGGTTGCCGTCGCTACTGGACACAAGGGGGCGCCCTGAGAAACGTCCCTGTAGGTGGA
GGCTGCCGGAGGAATAACAAGAAGGGCAAAAATGGAAATTTAAAATCTTCTTCTTCTTCGTC
CAAACAGTCTTCCTCGGTCAACGCTCAAAGTCCTAGCTCAGGACAGCTAAGGACAAATCATC
AGTTCCCTTTTTCACCAACTCTTTACAATCTCACTCAACTCGGAGGTATTGGTTTGAACTTA
GCCGCTACTAATGGCAACAACCAAGCTCACCAGATCGGTTCCAGTTTGATGATGAGCGATCT
AGGGTTTCTCCATGGACGAAATACTTCAACTCCGATGACGGGAAACATTCATGAAAACAACA
ACAATAATAACAATGAAAACAACCTAATGGCATCCGTTGGATCTTTGAGCCCCTTTGCTCTC
TTCGATCCAACGACGGGGCTATACGCTTTCCAGAACGACGGTAATATCGGGAACAACGTTGG
GATATCTGGTTCTTCTACTTCCATGGTTGATTCTAGGGTTTATCAGACGCCTCCGGTGAAGA
TGGAAGAACAACCTAATTTGGCTAACTTGTCTAGACCGGTCTCCGGTTTGACGTCTCCTGGG
AATCAAACAAATCAATACTTTTGGCCTGGTTCGGATTTCTCGGGTCCTTCTAATGATATCTT
G

SEQ ID NO: 204, Arabidopsis thaliana - Proteína

MPTNSNHQHHLQHQLNENGSIISGHGLVLSHQLPPLQANPNPNHHHVATSAGLPSRMGGSMA
ERARQANIPPLAGPLKCPRCDSSNTKFCYYNNYNLTQPRYFCKGCRRYWTQGGALRNVPVGG
GCRRNNKKGKNGNLKSSSSSSKQSSSVNAQSPSSGQLRTNHQFPFSPTLYNLTQLGGIGLNL
AATNGNNQAHQIGSSLMMSDLGFLHGRNTSTPMTGNIHENNNNNNNENNLMASVGSLSPFAL
FDPTTGLYAFQNDGNIGNNVGISGSSTSMVDSRVYQTPPVKMEEQPNLANLSRPVSGLTSPG
NQTNQYFWPGSDFSGPSNDIL

SEQ ID NO: 205, Oryza sativa - DNA

ATGATTCCGGGCACGCTTGCCGATGGCGGCGGCGGAGGCGGCGCGGTGGGGCCGGCGAAGCC
GATGTCGATGTCGGAGAGGGCGCGGCTGGCGAGGATCCCGCTGCCGGAGCCGGGGCTCAAGT
GCCCGCGCTGCGACTCCACCAACACCAAGTTCTGCTACTTCAACAACTACTCCCTCTCCCAG
CCCCGCCACTTCTGCCGCGCCTGCCGCCGCTACTGGACCCGCGGCGGCGCGCTCCGCAACGT
CCCCGTCGGCGGCGGCTACCGCCGCCACGCCAAGCGCGCCAAGCCCAAGCCGGCGTCCGCGG
CGGGCTCCGCCTCAGCCGCCACCACCACCGCCGGCTCGACGCCAGCGGGGTCGACGACGACG
ACGACGACGTCCTCCACCTGCGCCACGCCCAACGCGCCCGCCCTCCCGGCGATGCTGGGCGG
CAACCTCTCCATCCTGCCGCCGCTGCTCCGCCTCGCCGACTTCGACGCCATGAGCCTCGGCT
CCACCTTCTCTGGCATGGCGGCGGCCGCCGGCAAGCCACCGCCCGTCGACGCGGCCGGCTGC
TACTCCGTCGGCGCCGCCACCGGCCTCGAGCAATGGAGACTACAGCAGATGCAGAGCTTCCC
GTTCTTCCACGCCATGGATCACCAGGCGGCGATGGCGGCGCCACCGCCGGCAATGGCAATGC
CGGGGATGTTCCAGCTAGGCCTAGACGGCGACGGCCATGGCAGCGGCGGCGGCGAAGACGGT
GGAGAGCTCCACCATGCGATGCCATCATCGAAGAGAGAAGGCTACCCAAGGGGCATGTATGG
CGATCATCACCTCGCTGGAGGATACACCTCCTACTCCAGTGCAACCACAGGTAACCATCTCT
TGTAA

SEQ ID NO: 206, Oryza sativa - Proteína - AAU44212.1

MIPGTLADGGGGGGAVGPAKPMSMSERARLARIPLPEPGLKCPRCDSTNTKFCYFNNYSLSQ
PRHFCRACRRYWTRGGALRNVPVGGGYRRHAKRAKPKPASAAGSASAATTTAGSTPAGSTTT
TTTSSTCATPNAPALPAMLGGNLSILPPLLRLADFDAMSLGSTFSGMAAAAGKPPPVDAAGC
YSVGAATGLEQWRLQQMQSFPFFHAMDHQAAMAAPPPAMAMPGMFQLGLDGDGHGSGGGEDG
GELHHAMPSSKREGYPRGMYGDHHLAGGYTSYSSATTGNHLL

FIGURA 28 cont.

SEQ ID NO: 207, Oryza sativa - DNA

ATGCCGCCGCATCACGGCGGCCTCATGGCGCCTCGGCCTGACATGGTAGCAGCGGCCGTCGC
GGCGAGCGGCGGCGGCGGTGGCGGCGGCGGCCCGACCGGCGGCACGGCGGTGCGGCCGGGCT
CGATGACGGAACGGGCTCGGCTGGCGAAGATCCCGCAGCCGGAGCCGGGGCTCAAGTGCCCG
CGCTGCGAGTCCACCAACACCAAGTTCTGCTACTTCAACAACTACTCGCTCTCGCAGCCGCG
CCACTTCTGCAAGACGTGCCGCCGCTACTGGACGCGCGGCGGAGCGCTCCGCAACGTCCCCG
TCGGCGGCGGGTGCCGCCGCAACAAGCGCACCAAGTCGTCCAAGTCGTCCTCGTCGACGTCG
GCCGCCGGCTCGGCCTCCGCCACCGGCGGCACGTCGTCGTCCACATCGTCGACCGCCACGGG
TGGCAGCAGCAGCGCCGCGGCGGCCGCGGCGATGATGCCGCCGCAGGCGCAGCTGCCGTTCC
TGGCCTCGTTGCACCACCCGCTCGGCGGCGGCGATCACTACAGCTCCGGTGCGTCCAGGCTA
GGGTTTCCCGGATTGAGCTCGCTGGATCCCGTCGACTACCAGCTCGGCGGCGGCGCCGCCGC
CGCCGCCGCCATCGGGCTAGAGCAGTGGCGCCTCCCGCAGATACAGCAATTCCCCTTCTTGA
GCCGCAACGACGCCATGCCGCCGCCAATGTCCGGCATTTACCCGTTCGACGCGGAGGCCGCC
GCCGACGCCGCCGGCTTCGCCGGCCAGTTGCTGGCCGGCACCAAGGTGCCCGGCTCGTCGGG
CCTGATCACGCAGCTCGCATCCGTCAAGATGGAGGACAGCAACGCTCAGTCCGCGGCGATGA
ACAGCTCGCCGAGGGAGTTCTTGGGCCTCCCCGGCAACCTCCAATTCTGGGGCGGTGGCAAC
GGCGCGGGACCCGGCGGCAATGGAGACGGCGCCACCGGCGGCAGCGGCGCCGGTGTCGCTCC
GGGAGGCGGCGGCAGCGGCGGCGGATGGGCTGATCTCTCCGGATTCAACTCGTCGTCGTCGG
GGAACATACTGTGA

SEQ ID NO: 208, Oryza sativa – Proteína - XP_470142.1

MPPHHGGLMAPRPDMVAAAVAASGGGGGGGGPTGGTAVRPGSMTERARLAKIPQPEPGLKCP
RCESTNTKFCYFNNYSLSQPRHFCKTCRRYWTRGGALRNVPVGGGCRRNKRTKSSKSSSSTS
AAGSASATGGTSSSTSSTATGGSSSAAAAAAMMPPQAQLPFLASLHHPLGGGDHYSSGASRL
GFPGLSSLDPVDYQLGGGAAAAAAIGLEQWRLPQIQQFPFLSRNDAMPPPMSGIYPFDAEAA
ADAAGFAGQLLAGTKVPGSSGLITQLASVKMEDSNAQSAAMNSSPREFLGLPGNLQFWGGGN
GAGPGGNGDGATGGSGAGVAPGGGGSGGGWADLSGFNSSSSGNIL

SEQ ID NO: 209, Triticum aestivum - DNA

CTCCTAGCTCGTCGACAAGCATGCCCACAAAGCCACTGATAACAAGCGCATAGCAGCGCACG
CTCTCTTATAGTAGAATGTTCTGAACTCCACACCAGCCCATGCAAGACTTCCAGTCCATCCC
GGGCCTCGCCGGGCGGCTGTTCGGCGGCGCGGCCGCGGCAGACATCCGGCGCGTGCAGGGCC
CGGCGTCCCGGTGCGGCGTGTTCTCGCAGGCGGCGTCCGCGCAGCCGGAGGCGGCCGTCAAG
TGCCCGCGGTGCGAGTCCACCAACACCAAGTTCTGCTACTACAACAACTACAACCTGTCGCA
GCCGCGCCACTTCTGCAAGAGCTGCCGCCGGTACTGGACCAAGGGCGGCGTCCTCCGCAACG
TCCCCGTCGGCGGCGGCTGCCGCAAGGCCAAGCGCAGCTCCTCGTCGGCGTCCGCACCGTCG
ACGCCCGCGGCCACGGACGCCAAGAGCCAGCGGCGCGCGTCCGCGTCGTCCTCCTCCCGCTC
CAACAGCGGCAGCGGCAGCGCCAGCCCCACGGCCGCTGCGGAAGAGACGACGACAACGGAGA
CCGAGCCCCCTCCTCCGCCCACGCCGTCGTCCAACTCCAACTCCAACGCGGTCTCCTTCGCC
AACCGCATGACGAACTACCCCTTCGCGGCAGACGTGCCACCTCTGGCGCCGATATTCGCCGA
CCAGGCCGCCGCGCTCGCGTCCCTCTTCGCGCCGCCTCCTCCACCGCCTCTCCCGGTGTTCA
GCTTCTCGGCGGAGCCCAAGATGGAGGAGGCGATCGGGTCACTGCTGCTCCCGGGGCAGGAG
GCGTCGCAGGAGCCAGAGGAGCCCACCTGCACCTCCACCGTCGCGGACATGGCGCCGTTCAT
GTCGCTGGACGCGGGGATCTTCGAGCTCGGCGACGCGTCGCCGGCCGATTACTGGAACGGCG
GGAGCTGCTGGACGGACGTCCAGGACCCGTCCGTCTACCTACCCTAGTTTGGCTTAGTTCCT
CATCACGA

FIGURA 28 cont.

SEQ ID NO: 210, Triticum aestivum – Proteína – AAX54942

MQDFQSIPGLAGRLFGGAAAADIRRVQGPASRCGVFSQAASAQPEAAVKCPRCESTNTKFCY
YNNYNLSQPRHFCKSCRRYWTKGGVLRNVPVGGGCRKAKRSSSSASAPSTPAATDAKSQRRA
SASSSSRSNSGSGSASPTAAAEETTTTETEPPPPPTPSSNSNSNAVSFANRMTNYPFAADVP
PLAPIFADQAAALASLFAPPPPPPLPVFSFSAEPKMEEAIGSLLLPGQEASQEPEEPTCTST
VADMAPFMSLDAGIFELGDASPADYWNGGSCWTDVQDPSVYLP

SEQ ID NO: 211, Zea mays - DNA

AGGTGGTGGCGGGGGACAGGTTGGGGGCCCCGCCAAGCCCATGTCCATGGCGGAGCGCGCGC
GCCTCGCGAGGATCCCACTGCCGGAGCCGGGACTCAAGTGCCCGCGCTGCGACTCAACCAAC
ACCAAGTTCTGCTACTTCAACAACTACTCCCTCACGCAGCCGCGCCACTTCTGCCGGGCCTG
CCGCCGCTACTGGACGCGTGGCGGCGCGCTCCGCAACGTGCCGGTCGGTGGCGGATACCGCC
GCCACGCCAAGCGCGCCAAGCCCAAGCAGCAGCAGCAGCACGCGGCCGCCGGGACCGGAGCT
GCCAACGGCGCGATGCAGCAGCCTCCCGCTGGGTCTATGGCGTCGTCGGCCGCCGCCTGCAC
CGCGACCACGACGATGACGACCAACGCGCTGGACGCCGGGCCCGGCGGCATGCTGCCCATGC
TGCCGCCGCTCGTCCGCCTAGCAGACTTCGACGCAATGAGCCTCGGCTCCAGCTTCTCCGGG
ATATCGTCCATGGGGAAGCCCGGATCCATCGGCGCGGCCTGCTACCCGCACTCTGTCGGCGG
GCTGGAGCAGTGGAGGGTGCAGCAGATGCAGAGCTTCCCGTTCTTGCATGCGATGGACCAGG
GCCCGCTGGGGCCACCTCTGGCCATGGCGATGGCGGCGCCAGGAGGGATGTTCCAGCTAGGT
CTAGACACCACCAGTGATAACAGCCGTGGCGGTGGCGGCGGCGGCTGCGGCGAAGACGGGTC
GTCTGCGGGAGAGGCGCTCCATATGATGCAAGCAGCCACCAAGAGGGAGAGCTACCCGGCAC
CACCAAGAGCCATGTACGGCGACCAACACCACAATCACCTCGCTGCTGCTGGTGGCTACACT
TCCTATTCCACCAATGCTGCTGCAGGTAACCATCTCTTGTAATGGCCGGCCGATCGATCGAT
CGAGAGCTCAACAATTCAAGTGTCCTGCTATAGCTAGCTACTACGACGTCGTGCTACGATCG
TATCGGTTTCGGTTCGTTCCTACAAATATAGCCTAGAGATAGAGTGTCTCTGTCTGTGTGAT
CGATGGTATTGTTATGATCATATAGAAAAGACCAGTGTAGCATGCATGCACTTGTTGCAATG
TTTGCTTTCAAGAAGAAACTGGAGGGAGGAGAGGCTCGGTTTGGATGCTGATCATGCACAAA
TATTACTAGTGTCTACAGCTGCTACTTCATTATAAAAAAAAAAAAAAAAAA

SEQ ID NO: 212, Zea mays - Proteína – CAA54288

GGGGGQVGGPAKPMSMAERARLARIPLPEPGLKCPRCDSTNTKFCYFNNYSLTQPRHFCRAC
RRYWTRGGALRNVPVGGGYRRHAKRAKPKQQQQHAAAGTGAANGAMQQPPAGSMASSAAACT
ATTTMTTNALDAGPGGMLPMLPPLVRLADFDAMSLGSSFSGISSMGKPGSIGAACYPHSVGG
LEQWRVQQMQSFPFLHAMDQGPLGPPLAMAMAAPGGMFQLGLDTTSDNSRGGGGGGCGEDGS
SAGEALHMMQAATKRESYPAPPRAMYGDQHHNHLAAAGGYTSYSTNAAAGNHLL

SEQ ID NO: 213, Solanum tuberosum - DNA

GAGGACTTATCAACCTTTTTATATAAAAGAGATAAAGATCAAAGAGATCAAAGAAAGAAAAT
CATGGTTTTCTCATCTTTTCCTGTGTATCTAGATCATCCCAATTTGCATCAGTTACAACAGC
CAGACGGCCATCAACAAGGAAATACTGGGCTGGAGAATCCAACTCTGCAGCCCCCACCTATG
CAGGTGGGGGCTAGTCCGGGCTCAATCAGACCAGGTTCTATGGTGGATCGAGCCCGGTTAGC
TAAGATTCCACTACCGGAAGCTGGATTAAAGTGTCCAAGGTGTGATTCAACAAACACAAAGT
TCTGCTACTTCAACAACTACAACCTTTCCCAACCAAGACACTTCTGCAAGACCTGTCGCCGG
TACTGGACTAGAGGGGGCGCCTTGAGAAGCGTGCCGGTAGGAGGAGGATGCCGGAGGAACAA
GAGAAGCAAAAGTAATAACAACAACAACAGCTCAAAGACAGCTGGCAGTAATGTCAATACTA
ATACTATTGCTTCCGGTACTTCAACAAGTGCAAGTCCTTCAAGCTGCAGCACGGAAATAATG
AATGGTCGCCATCACTTCTCTCATGAGCAACCACCGCAGTTAACTCCTCTCATGGCTGCTTT

FIGURA 28 cont.

CCAAAATCTTAATCATCACTATGGCGGATTTCAACCTCCTCCTTTGGTTTCAACACACCATG
GGAATGGAACCGGCGCGCTTGGTCATCATCATGAAATGGGATTTCAAATAGGGAGTAGTACT
AATACTAACAATTTACCTGTACCCCCAGGAGGAGGATCTGATCATCAGTGGAGATTACCATC
TTTGGCAGCAAACACAAATTTGTACCCTTTTCAACACGGTACTGATCAAGGAATTCATGAAT
CATCATCTGTTAACAATAATAATATTAATGCTCATGATGACCAAGGGTTAAATTCGACCAAA
CAGTTTTTGGGAACAATGGAAAATAATACTAATCAATATTGGGGTGGAAACGCATGGACAGG
GTTTTCTGGATTAAACTCATCTTCTTCAGCAAGCCATCTACTTTGATTTCATTCATGTAAGT
CAATTTGTGTACTTGACTCTCAACGATTAAAATGTTGTATGTGTTGGGAGGGGGGTCTAAAT
TAGATATATAGTATACTGGGGGCATGTTTAAGTCTTTGTTATTCCATGGTCATTATCACCAC
TTGTAATTTTACTGGATGTTTTTTTTTTATAACTTAGGTGTGTGAAGTTGTATTGTGAGTTT
TAA

SEQ ID NO: 214, Solanum tuberosum – Proteína – CAB89831

MVFSSFPVYLDHPNLHQLQQPDGHQQGNTGLENPTLQPPPMQVGASPGSIRPGSMVDRARLA
KIPLPEAGLKCPRCDSTNTKFCYFNNYNLSQPRHFCKTCRRYWTRGGALRSVPVGGGCRRNK
RSKSNNNNNSSKTAGSNVNTNTIASGTSTSASPSSCSTEIMNGRHHFSHEQPPQLTPLMAAF
QNLNHHYGGFQPPPLVSTHHGNGTGALGHHHEMGFQIGSSTNTNNLPVPPGGGSDHQWRLPS
LAANTNLYPFQHGTDQGIHESSSVNNNNINAHDDQGLNSTKQFLGTMENNTNQYWGGNAWTG
FSGLNSSSSASHLL

SEQ ID NO: 215, Arabidopsis thaliana - DNA

ATGCCTACGAATTCGAATCATCAGCATCATCTTCAACACCAGCTTAACGAAAATGGAAGTAT
AATAAGTGGCCACGGACTAGTACTCTCTCACCAACTTCCACCTCTCCAAGCAAACCCTAACC
CTAACCACCACCATGTCGCTACCTCTGCTGGTCTTCCGTCAAGGATGGGTGGATCGATGGCG
GAGAGAGCAAGGCAGGCCAACATTCCTCCACTAGCGGGACCCCTAAAGTGTCCTCGATGCGA
CTCCAGCAACACTAAGTTCTGTTACTACAACAACTATAACCTCACTCAGCCTCGTCACTTCT
GCAAAGGTTGCCGTCGCTACTGGACACAAGGGGGCGCCCTGAGAAACGTCCCTGTAGGTGGA
GGCTGCCGGAGGAATAACAAGAAGGGCAAAAATGGAAATTTAAAATCTTCTTCTTCTTCGTC
CAAACAGTCTTCCTCGGTCAACGCTCAAAGTCCTAGCTCAGGACAGCTAAGGACAAATCATC
AGTTCCCTTTTTCACCAACTCTTTACAATCTCACTCAACTCGGAGGTATTGGTTTGAACTTA
GCCGCTACTAATGGCAACAACCAAGCTCACCAGATCGGTTCCAGTTTGATGATGAGCGATCT
AGGGTTTCTCCATGGACGAAATACTTCAACTCCGATGACGGGAAACATTCATGAAAACAACA
ACAATAATAACAATGAAAACAACCTAATGGCATCCGTTGGATCTTTGAGCCCCTTTGCTCTC
TTCGATCCAACGACGGGGCTATACGCTTTCCAGAACGACGGTAATATCGGGAACAACGTTGG
GATATCTGGTTCTTCTACTTCCATGGTTGATTCTAGGGTTTATCAGACGCCTCCGGTGAAGA
TGGAAGAACAACCTAATTTGGCTAACTTGTCTAGACCGGTCTCCGGTTTGACGTCTCCTGGG
AATCAAACAAATCAGTACTTTTGGCCTGGTTCGGATTTCTCGGGTCCTTCTAATGATCTCTT
GTGA

SEQ ID NO: 216, Arabidopsis thaliana – Proteína – At1g07640

MGGSMAERARQANIPPLAGPLKCPRCDSSNTKFCYYNNYNLTQPRHFCKGCRRYWTQGGALR
NVPVGGGCRRNNKKGKNGNLKSSSSSSKQSSSVNAQSPSSGQLRTNHQFPFSPTLYNLTQLG
GIGLNLAATNGNNQAHQIGSSLMMSDLGFLHGRNTSTPMTGNIHENNNNNNNENNLMASVGS
LSPFALFDPTTGLYAFQNDGNIGNNVGISGSSTSMVDSRVYQTPPVKMEEQPNLANLSRPVS
GLTSPGNQTNQYFWPGSDFSGPSNDLL

FIGURA 28 cont.

SEQ ID NO: 217, Arabidopsis thaliana - DNA

ATGGTTTTCTCCTCCATCCAAGCCTATCTTGATTCATCCAACTGGCAACAGGCTCCTCCGAG
CAATTATAATCATGACGGAACAGGCGCCTCAGCAAATGGAGGTCATGTTCTTCGTCCTCAGC
TGCAGCCACAGCAGCAGCCACAGCAGCAGCCGCATCCTAATGGGAGCGGGGGCGGAGGTGGA
GGTGGAGGCGGCTCGATCCGAGCAGGATCAATGGTGGACAGAGCAAGACAAGCAAACGTAGC
CTTGCCAGAAGCAGCACTGAAATGTCCGAGATGCGAATCCACCAACACCAAGTTCTGCTACT
TCAACAACTACAGCCTCACTCAACCACGCCACTTCTGCAAGACCTGCCGGAGATACTGGACA
CGTGGCGGAGCTCTCCGCAACGTCCCAGTCGGCGGTGGCTGCCGGAGAAACAGGCGTACCAA
AAGCAACAGCAACAACAACAATAACAGCACTGCTACTAGCAATAACACCAGTTTCTCCTCCG
GGAATGCATCCACCATCAGCACGATTCTCTCCTCCCACTATGGAGGAAACCAAGAGAGTATC
TTAAGCCAGATTTTGTCTCCGGCGAGGCTAATGAATCCTACTTACAATCATCTCGGAGATCT
CACAAGTAATACAAAAACAGACAACAACATGAGCTTGTTGAACTATGGAGGATTGAGTCAAG
ACTTGAGGTCAATCCACATGGGAGCTTCTGGTGGCTCGCTTATGAGCTGTGTTGATGAATGG
AGATCGGCGTCTTATCATCAGCAGTCAAGTATGGGCGGTGGGAACTTGGAGGATTCTTCTAA
TCCTAATCCATCCGCAAATGGGTTTTACTCTTTTGAGTCGCCGAGGATAACTTCAGCGTCAA
TCTCTTCTGCTTTAGCATCGCAGTTTTCTTCGGTTAAAGTTGAAGATAATCCTTACAAATGG
GTTAATGTCAATGGTAATTGCTCTTCCTGGAATGATCTTTCTGCTTTCGGCTCTTCTCGTTG
A

SEQ ID NO: 218, Arabidopsis thaliana – Proteína – At2g37590

MVFSSIQAYLDSSNWQQAPPSNYNHDGTGASANGGHVLRPQLQPQQQPQQQPHPNGSGGGGG
GGGGSIRAGSMVDRARQANVALPEAALKCPRCESTNTKFCYFNNYSLTQPRHFCKTCRRYWT
RGGALRNVPVGGGCRRNRRTKSNSNNNNNSTATSNNTSFSSGNASTISTILSSHYGGNQESI
LSQILSPARLMNPTYNHLGDLTSNTKTDNNMSLLNYGGLSQDLRSIHMGASGGSLMSCVDEW
RSASYHQQSSMGGGNLEDSSNPNPSANGFYSFESPRITSASISSALASQFSSVKVEDNPYKW
VNVNGNCSSWNDLSAFGSSR

SEQ ID NO: 219, Arabidopsis thaliana - DNA

ATGGTTTTCTCATCTCTTCCAGTGAATCAGTTCGATTCCCAAAATTGGCAGCAGCAAGGGAA
CCAACATCAGCTAGAATGTGTCACAACTGACCAGAACCCTAATAATTACTTACGGCAGCTCT
CATCACCACCGACTTCTCAGGTTGCAGGTTCGAGTCAAGCTAGAGTGAATTCAATGGTGGAA
CGTGCTCGGATCGCAAAAGTCCCATTGCCTGAAGCAGCTCTAAATTGCCCTAGATGTGACTC
AACCAATACTAAGTTCTGTTACTTCAATAACTATAGCCTTACTCAACCTCGCCATTTCTGCA
AAACATGTCGTCGCTATTGGACACGTGGCGGTTCCTTGAGGAATGTTCCTGTTGGAGGAGGC
TTTAGGAGGAACAAGAGAAGCAAATCCAGATCGAAATCTACGGTCGTGGTCTCGACTGATAA
TACTACTAGTACTTCATCACTTACTTCTCGCCCAAGTTACTCAAACCCTAGCAAGTTTCATA
GCTACGGTCAAATCCCGGAGTTTAATTCCAACTTGCCCATCTTGCCTCCTCTCCAAAGCCTT
GGAGATTACAATTCAAGCAACACTGGATTAGATTTTGGTGGAACTCAAATAAGCAACATGAT
AAGTGGTATGAGTTCTAGTGGTGGGATCTTGGATGCATGGAGAATACCTCCATCACAACAAG
CTCAGCAATTCCCTTTCTTGATCAACACTACCGGATTGGTGCAATCTTCAAACGCGTTATAT
CCATTACTAGAAGGCGGGGTTAGCGCCACGCAAACAAGAAATGTGAAGGCGGAAGAGAATGA
TCAGGATCGGGGTAGGGATGGGGATGGAGTGAATAACTTATCAAGAAACTTTTTGGGTAATA
TCAACATAAACTCAGGCAGGAACGAGGAATACACATCATGGGGAGGTAACAGTTCTTGGACC
GGTTTCACCTCCAACAACTCAACAGGCCATCTCTCATTCTAA

FIGURA 28 cont.

SEQ ID NO: 220, Arabidopsis thaliana – Proteína – At3g55370
MVFSSLPVNQFDSQNWQQQGNQHQLECVTTDQNPNNYLRQLSSPPTSQVAGSSQARVNSMVE
RARIAKVPLPEAALNCPRCDSTNTKFCYFNNYSLTQPRHFCKTCRRYWTRGGSLRNVPVGGG
FRRNKRSKSRSKSTVVVSTDNTTSTSSLTSRPSYSNPSKFHSYGQIPEFNSNLPILPPLQSL
GDYNSSNTGLDFGGTQISNMISGMSSSGGILDAWRIPPSQQAQQFPFLINTTGLVQSSNALY
PLLEGGVSATQTRNVKAEENDQDRGRDGDGVNNLSRNFLGNININSGRNEEYTSWGGNSSWT
GFTSNNSTGHLSF

SEQ ID NO: 221, Arabidopsis thaliana – DNA
ATGGTTTTTTCTTCATTTCCTACTTATCCTGATCATTCATCAAACTGGCAACAACAACATCA
ACCAATCACAACCACCGTTGGATTCACGGGAAATAACATCAACCAACAGTTTCTTCCTCACC
ATCCCCTCCCACCGCAACAGCAACAAACGCCTCCGCAGCTTCACCACAACAACGGTAACGGC
GGAGTCGCTGTTCCCGGTGGACCTGGCGGGTTAATCCGACCAGGTTCGATGGCGGAAAGAGC
AAGGCTAGCCAACATACCATTACCTGAAACAGCCTTGAAGTGTCCAAGATGTGACTCAACTA
ACACCAAATTCTGTTACTTCAACAACTACAGTCTCACTCAACCTCGCCACTTCTGCAAAGCA
TGCCGTCGTTACTGGACACGTGGCGGTGCTCTAAGGAGCGTTCCCGTCGGTGGCGGTTGCCG
TAGAAACAAAAGAACCAAAAACAGCAGCGGTGGAGGTGGCGGTAGCACCAGTAGCGGTAACA
GCAAGTCACAAGACAGCGCCACGAGCAACGACCAATACCACCACCGAGCCATGGCTAACAAT
CAGATGGGACCACCTTCTTCGTCATCGTCTCTAAGCTCGTTGCTGTCTTCTTACAACGCAGG
GTTAATCCCCGGACATGATCATAACAGCAATAACAACAACATACTTGGACTTGGATCATCTT
TGCCTCCTCTTAAGCTTATGCCTCCTTTAGACTTCACAGACAACTTCACCTTACAATACGGT
GCCGTTTCAGCTCCTTCTTATCATATAGGCGGTGGAAGCAGTGGAGGAGCGGCGGCTCTTTT
AAACGGTTTTGACCAGTGGAGATTCCCGGCAACAAACCAACTTCCTTTAGGCGGTTTAGACC
CGTTTGATCAACAACATCAAATGGAGCAGCAGAATCCAGGTTACGGATTGGTTACCGGGTCG
GGTCAGTATCGACCTAAGAACATTTTCCATAACCTTATCTCCTCTTCTTCGTCTGCTTCATC
AGCTATGGTTACAGCCACCGCGTCGCAATTAGCTTCAGTGAAAATGGAAGATAGTAACAATC
AGCTCAACTTGTCTAGACAACTTTTTGGAGACGAACAACAGCTCTGGAATATTCATGGCGCT
GCTGCAGCATCCACCGCAGCTGCAACAAGTTCGTGGAGTGAAGTCTCTAATAATTTCAGTTC
TTCTTCTACTAGCAATATATAA

SEQ ID NO: 222, Arabidopsis thaliana – Proteína – At5g02460
MVFSSFPTYPDHSSNWQQQHQPITTTVGFTGNNINQQFLPHHPLPPQQQQTPPQLHHNNGNG
GVAVPGGPGGLIRPGSMAERARLANIPLPETALKCPRCDSTNTKFCYFNNYSLTQPRHFCKA
CRRYWTRGGALRSVPVGGGCRRNKRTKNSSGGGGGSTSSGNSKSQDSATSNDQYHHRAMANN
QMGPPSSSSSLSSLLSSYNAGLIPGHDHNSNNNNILGLGSSLPPLKLMPPLDFTDNFTLQYG
AVSAPSYHIGGGSSGGAAALLNGFDQWRFPATNQLPLGGLDPFDQQHQMEQQNPGYGLVTGS
GQYRPKNIFHNLISSSSSASSAMVTATASQLASVKMEDSNNQLNLSRQLFGDEQQLWNIHGA
AAASTAAATSSWSEVSNNFSSSSTSNI

SEQ ID NO: 223, Seqüência artificial – iniciador prm07315
GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGGGTGGATCGATGGC

SEQ ID NO: 224, Seqüência artificial – iniciador prm07316
GGGGACCACTTTGTACAAGAAAGCTGGGTCGTTAATGATCCGACAAAACA

FIGURA 28 cont.

SEQ ID NO: 225, Oryza sativa – DNA - promotor GOS2

AATCCGAAAAGTTTCTGCACCGTTTTCACCCCCTAACTAACAATATAGGGAACGTGTGCTAA
ATATAAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAAACTCATC
CACCTACTTTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCCT
TAGTAATTAAGTGGGAAAATGAAATCATTATTGCTTAGAATATACGTTCACATCTCTGTCAT
GAAGTTAAATTATTCGAGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTT
TCTAGCTGAACTCAATGGGTAAAGAGAGAGATTTTTTTTAAAAAAATAGAATGAAGATATTC
TGAACGTATTGGCAAAGATTTAAACATATAATTATATAATTTTATAGTTTGTGCATTCGTCA
TATCGCACATCATTAAGGACATGTCTTACTCCATCCCAATTTTTATTTAGTAATTAAAGACA
ATTGACTTATTTTTATTATTTATCTTTTTTCGATTAGATGCAAGGTACTTACGCACACACTT
TGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAATACACGTTCAACTAGCAACACATCTCT
AATATCACTCGCCTATTTAATACATTTAGGTAGCAATATCTGAATTCAAGCACTCCACCATC
ACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTTACAGAATAGCATGAAAAG
TATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAAAAAAGAATTTTGCTCGTGCGCGA
GCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCCACAGAACAACCCA
CAAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAGGCTTTG
CGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGCATCCTCCTCCTCCCATCTATAAAT
TCCTCCCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCA
AGGACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTTCGATCCATATCTTCCGGTCGAGT
TCTTGGTCGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCCTTCGGTTGTTC
TTGGATTTATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATC
TGTGATGATTCCTGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTT
CGGTTTGATTAGTAGTATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTA
GGGTACGGAATCTTGCGATTTTGTGAGTACCTTTTGTTTGAGGTAAAATCAGAGCACCGGTG
ATTTTGCTTGGTGTAATAAAAGTACGGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTC
GATTTGACGAAGCTATCCTTTGTTTATTCCCTATTGAACAAAAATAATCCAACTTTGAAGAC
GGTCCCGTTGATGAGATTGAATGATTGATTCTTAAGCCTGTCCAAAATTTCGCAGCTGGCTT
GTTTAGATACAGTAGTCCCCATCACGAAATTCATGGAAACAGTTATAATCCTCAGGAACAGG
GGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTTTTTTCCCAAATATCTTAAAAA
GTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATGCTTTTATAGCGTTATCC
TAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGAAGAACTTATCCGA
TTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATTCATTTGGAT
TATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAACTGTC
CTCAATTTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGTA
GAAGTTTCTTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAA
TCGGGATAGTTATACTGCTTGTTCTTATGATTCATTTCCTTTGTGCAGTTCTTGGTGTAGCT
TGCCACTTTCACCAGCAAAGTTC

SEQ ID NO: 226, Arabidopsis thaliana – DNA

ATGATGATGGAGACTAGAGATCCAGCTATTAAGCTTTTCGGTATGAAAATCCCTTTTCCGTC
GGTTTTTGAATCGGCAGTTACGGTGGAGGATGACGAAGAAGATGACTGGAGCGGCGGAGATG
ACAAATCACCAGAGAAGGTAACTCCAGAGTTATCAGATAAGAACAACAACAACTGTAACGAC
AACAGTTTTNACAATTCGAAACCCGAAACCTTGGACAAAGAGGAAGCGACATCAACTGATCA
GATAGAGAGTAGTGACACGCCTGAGGATAATCAGCAGACGACACCTGATGGTAAAACCCTAA
AGAAACCGACTAAGATTCTACCGTGTCCGAGATGCAAAAGCATGGAGACCAAGTTCTGTTAT
TACAACAACTACAACATAAACCAGCCTCGTCATTTCTGCAAGGCTTGTCAGAGATATTGGAC
TGCTGGAGGGACTATGAGGAATGTTCNTGTGGGGGCAGGACGTCGTAAGAACAAAAGCTTAT

FIGURA 28 cont.

CTTCTCATTACCGTCACATCACTATTTCCGAGGCTCTTGAGGCTGCGAGGCTTGACCCGGGC
TTACAGGCAAACACAAGGGTCTTGAGTTTTGGTCTCGAAGCTCAGCAGCAGCACGTTGCTGC
TCCCATGACACCTGTGATGAAGCTACAAGAAGATCAAAAGGTCTCAAACGGTGCTAGGAACA
GGTTTCACGGGTTAGCGGATCAACGGCTTGTAGCTCGGGTAGAGAATGGAGATGATTGCTCA
AGCGGATCCTCTGTGACCACCTCTAACAATCACTCAGTGGATGAATCAAGAGCACAAAGCGG
CAGTGTTGTTGAAGCACAAATGAACAACAACAACAATAACATGAATGGTTATGCTTGCATCC
CAGGTGTTCCATGGCCTTACACGTGGAATCCAGCGATGCCTCCACCAGGTTTTTACCCGCCT
CCAGGGTATCCAATGCCGTTTTACCCTTACTGGACCATCCCAATGCTACCACCGCATCAATC
CTCATCGCCTATAAGCCAAAAGTGTTCAAATACAAACTCTCCGACTCTCGGAAAGCATCCGA
GAGATGAAGGATCATCGAAAAAGGACAACGAGACAGAGCGAAAACAGAAGGCCGGGTGCGTT
CTGGTCCCGAAAACGTTGAGAATAGATGATCCTAACGAAGCAGCAAAGAGCTCGATATGGAC
AACATTGGGAATCAAGAACGAGGCGATGTGCAAAGCCGGTGGTATGTTCAAAGGGTTTGATC
ATAAGACAAAGATGTATAACAACGACAAAGCTGAGAACTCCCCTGTTCTTTCTGCTAACCCT
GCTGCTCTATCAAGATCACACAATTTCCATGAACAGATTTAG

SEQ ID NO: 227, Arabidopsis thaliana - proteína
MMMETRDPAIKLFGMKIPFPSVFESAVTVEDDEEDDWSGGDDKSPEKVTPELSDKNNNNCND
NSFXNSKPETLDKEEATSTDQIESSDTPEDNQQTTPDGKTLKKPTKILPCPRCKSMETKFCY
YNNYNINQPRHFCKACQRYWTAGGTMRNVXVGAGRRKNKSLSSHYRHITISEALEAARLDPG
LQANTRVLSFGLEAQQQHVAAPMTPVMKLQEDQKVSNGARNRFHGLADQRLVARVENGDDCS
SGSSVTTSNNHSVDESRAQSGSVVEAQMNNNNNNNMNGYACIPGVPWPYTWNPAMPPPGFYPP
PGYPMPFYPYWTIPMLPPHQSSSPISQKCSNTNSPTLGKHPRDEGSSKKDNETERKQKAGCV
LVPKTLRIDDPNEAAKSSIWTTLGIKNEAMCKAGGMFKGFDHKTKMYNNDKAENSPVLSANP
AALSRSHNFHEQI

SEQ ID NO: 228, Seqüência artificial – proteína – motivo 1
KPTKILPCPRCKSMETKFCYYNNYNINQPRHFCKACQRYWTAGGTMRNVXVGAGRRKNKSLS
S

SEQ ID NO: 229, Seqüência artificial – proteína – motivo 1
KALKKPDKILP

SEQ ID NO: 230, Seqüência artificial – proteína – motivo 2
DDPGIKLFGKTIPF

SEQ ID NO: 231, Seqüência artificial – proteína – motivo 3
SPTLGKHSRDE

SEQ ID NO: 232, Seqüência artificial – proteína – motivo 4
LQANPAALSRSQNFQE

SEQ ID NO: 233, Seqüência artificial – proteína – motivo 5
KGEGCLWVPKTLRIDDPDEAAKSSIWTTLGIK

FIGURA 28 cont.

SEQ ID NO: 234, Oryza sativa - DNA

ATGAACAACGTTGAAGAGAAAGCAGCATCAGACTCAAAAGATGAAAATGAAAAGACAGCAAA
TGATGAATCAGGCCAGGACAAAGTACTTAAGAAGCCAGATAAGATTCTCCCTTGCCCTCGGT
GCAACAGTATGGACACAAAGTTTTGTTATTACAACAACTACAATGTTAATCAACCCAGGCAC
TTCTGTAAGAACTGCCAAAGGTATTGGACTGCCGGGGGAACCATGAGAAATGTACCTGTTGG
TGCTGGGAGGCGCAAAAGCAAGAGCTCATCGTTGCACTACCGTCACTTACTGATGGCCCCTG
ATTGCATGATGGGGTCTAGAGTGGAAATATCCAAGTCAATGAACCCTGAAGCTTTCGCATCT
GCGCATTCGACCCCTATACAACCAATTGGCAGAAACGAAACAGTTCTCAAATTTGGGCCTGA
GGTGCCACTCTGTGAATCGATGGCATCAGTGCTGAACATTCAGGAGCAGAATGGAACCAATG
CTGCAGCAGTACCAACGGGTGAAAATCAGGAAGATAACTCTTGCATCTCTTCAATCACATCA
CACAACGTGTTACCTGAAAATGCAGCCCAAGTTGACAAGAACAGCACGCCGGTGTATTGCAA
CGGAGTCGGCCCAGTGCCGCAGTACTACCTTGGAGCTCCTTACATGTACCCATGGAACATAG
GATGGAACAACGTTCCTATGATGGTGCCAGGTACAAGCATGCCAGAGTCTGCTTCCCAATCT
GAGAGCTGCAGCACCAGTTCAGCTCCATGGATGAACATGAACTCCCCCATGATGCCGGTTGC
CTCGAGGCTTTCTGCACCACCATTTCCATACCCTCTAGTGCCACCTGCACTATGGGGTTGCT
TATCCAGCTGGCCGGCCACGGCATGGAACATACCGTGGATCAGAACGAATGGCGGCTGCATG
TCTCCATCGTCGTCGAGCAACAGCAGCTGTTCAGGCAATGGCTCCCCTCTGGGGAAGCATTC
CAGGGACTCCTCTCTCCCACTGAAGGAGGACAAGGAGGAGAAATCACTGTGGGTTCCCAAGA
CGCTCCGCATCGACGATCCCGACGAGGCGGCGAAGAGCTCCATCTGGGCCACCCTGGGGATC
AAGCCTGGAGACCCTGGCATCTTCAAGCCGTTCCAGTCCAAAGGTGAGAGCAAAGGCCAAGC
AGCATCAGAGACTCGTCCTGCTCGTGCTCTTAAGGCAAACCCAGCTGCATTGTCGCGGTCGC
AGTCGTTCCAGGAGACTTCTTGA

SEQ ID NO: 235, Oryza sativa – proteína - NP_912875.1

MNNVEEKAASDSKDENEKTANDESGQDKVLKKPDKILPCPRCNSMDTKFCYYNNYNVNQPRH
FCKNCQRYWTAGGTMRNVPVGAGRRKSKSSSLHYRHLLMAPDCMMGSRVEISKSMNPEAFAS
AHSTPIQPIGRNETVLKFGPEVPLCESMASVLNIQEQNGTNAAAVPTGENQEDNSCISSITS
HNVLPENAAQVDKNSTPVYCNGVGPVPQYYLGAPYMYPWNIGWNNVPMMVPGTSMPESASQS
ESCSTSSAPWMNMNSPMMPVASRLSAPPFPYPLVPPALWGCLSSWPATAWNIPWIRTNGGCM
SPSSSSNSSCSGNGSPLGKHSRDSSLPLKEDKEEKSLWVPKTLRIDDPDEAAKSSIWATLGI
KPGDPGIFKPFQSKGESKGQAASETRPARALKANPAALSRSQSFQETS

SEQ ID NO: 236, Oryza sativa - DNA

ATGCCTAATTTAGGAAATGGTGTAAAAACCAATAATGACTTACCTTTAGTCTCAGACAAGCT
TTTGATTGTCAAAGGTATTCCTTTTTGTCCCAACAATAGTAAGAAAAATGATTTACAAGGCA
TCAGCAGACCGGATGGAAGAATAGAAATCGATTCCATGACTGAGGATGTTAAGACTGAGCCT
GATGGATCTGTCCCTGAGAAGATACTCAAGAAGCCAGATAAGATTCTGCCATGTCCACGCTG
CAATAGCATGGAAACAAAGTTCTGCTACTTCAACAACTACAATGTTCACCAGCCCAGGCACT
TCTGCAGGAACTGCCAAAGATATTGGACCGCTGGTGGAGCTATGAGGAATGTCCCAGTTGGT
GCTGGAAGACGCAGAAATAAGCATGTGTCAAAATACTGTCAGGCGATGATGACGTGCAATAA
TACTGTAGCTCCTGGAGATGTTTCTGATGTGGTTCACCATCAGGTTATTACACATGGATCTT
CTCTCCTTCCAGCAACACTGAAGGAAAATGAAACACCTACAGAATTCATATCAGAAGTACCA
CCATGTAAGTCTTCAGCTTCAATCCTTGATATTGGAGAGCCGAATGATACTGACCTTGTTCC
CTTGGCCTCTGGTGATAACAAGGAAGAAAAATCATGTGCATCTTCTGTGGTAGTATCCAGCT
GTTCAGAGAATCTGATGCCAGATAATGCAATTATGAAAGAGCCAAACAACAGGTCAGGATGT
TGTAATGGTGTGGCATTGCCTTTCCCTACTGGACCTGCTTTGGTGCTCCCCTGGAGTCTTGG
ATGGAACAGTGTTGCTCTCATGCCAGCTACCCAGTGCTCGATGCAGCCCGTTCTTGGGTTAA

FIGURA 28 cont.

AAGATGGGATACCCTGCCCGCCTTCATGGCCACCGCAACTGATGGTGCCGGCCCCAGGAATC
TGTACTCCTGTTGTTCCAATCCCTCTTGTGCCACCTCTGTGGAGCTGCTTCCCTGGCTGGCC
TAATGGAATGTGGAATGCACAATGCCCTGGAGGTAATACTACTGTGCTGCCGTCAACCGCTC
CAAACAAAATTTCTTGTTCAGGAAGCAGTTCTCTGGTGTTGGGAAAGCATTCAAGAGAAGAA
AGCTTGCAGGAAGAAGAGAAGACGAGAAATTACTTATGGGTACCTAAAACTCTTAGGATTGA
TGATCCAGCTGAGGCTGCAAAGAGTTCAATCTGGGCGACCCTTGGCATCAAGCCTGACGATA
AAGGCATATTCAAGTCTTTCCAGCCAAATGTTGCGAAAAATGGCACGGCACCAGAATCGCCT
CAGGCTCTGCAGGCCAATCCAGCAGCATTTTCGCGTTCTCAATCGTTTCAAGAGACGACTTG
A

SEQ ID NO: 237, Oryza sativa – proteína – NP_912989.1

MPNLGNGVKTNNDLPLVSDKLLIVKGIPFCPNNSKKNDLQGISRPDGRIEIDSMTEDVKTEP
DGSVPEKILKKPDKILPCPRCNSMETKFCYFNNYNVHQPRHFCRNCQRYWTAGGAMRNVPVG
AGRRRNKHVSKYCQAMMTCNNTVAPGDVSDVVHHQVITHGSSLLPATLKENETPTEFISEVP
PCKSSASILDIGEPNDTDLVPLASGDNKEEKSCASSVVVSSCSENLMPDNAIMKEPNNRSGC
CNGVALPFPTGPALVLPWSLGWNSVALMPATQCSMQPVLGLKDGIPCPPSWPPQLMVPAPGI
CTPVVPIPLVPPLWSCFPGWPNGMWNAQCPGGNTTVLPSTAPNKISCSGSSSLVLGKHSREE
SLQEEEKTRNYLWVPKTLRIDDPAEAAKSSIWATLGIKPDDKGIFKSFQPNVAKNGTAPESP
QALQANPAAFSRSQSFQETT

SEQ ID NO: 238, Oryza sativa – DNA

ATGGGAGGGGGTGGTAAGTGCAAAAGAGGGAAAAGAGGAAAGATTGCAGCCAAGCGAAGAAG
GGGCCAGTGTTGCAGCAGCAGCAGCAGCAGCGGAGGCGCAAGAAGACGAGCTAGAGAGAGAG
AATCAGAGGAGAGCTTGGGTGAGATGGGTGAGTGCAGAGGAGGAGGAGGAGGAGGAGATGGG
CTGATCAAGCTGTTCGGGAAGACGATCCCGGTGCAGCCGGATGCCAAGGATGTTCAGCAGCA
TAGTGGCAGTAGCAGCAGCTCAACCGAATCCGACGTCCAAGAAACCGCCGCTGTCGCCGTCG
CCGACCCCTCCCCGCGGTCGGAGGTCGTCGACGGCGAGAGCCCGCCGCAGCCGGGCGGCGAG
GCGGCGAGCCATCAGCAGCAGCAGAAGGAGATGAAGCTGAAGAAGCCGGACAAGATCCTGCC
ATGCCCGCGGTGCAGCAGCATGGACACCAAGTTCTGCTACTTCAACAACTACAACGTCAACC
AGCCTCGCCACTTCTGCAAGCACTGCCAGCGCTACTGGACCGCCGGCGGCGCCATGCGCAAC
GTCCCCGTCGGCGCCGGCCGCCGCAAGAACAAGAACGCCACCGCCGCCGCCCACTTCCTCCA
CCGCGTCCGCGCCTGCGCCGCCGCCGCCGCCATGCCCGCGGCGCCCCACGACGCCACCAACG
CCACCGTGCTCAGCTTCGGCGGCGGCGGAGGCGGACACGACGCGCCGCCGGTCACCCTGGAC
CTCGCCGACAAGATGACGCGCCTCGGCAAGGAGGGGCTCGTCGCCCATGCCCGGAACGCCGA
CGCCGCCGCCGCGTGCAGCGAGGTGTCGAGCAACAGGGACGACGAGCAGATCGGCAACACTG
TAGCAAAACCTGCAAACGGGTTGCAGCAGCATCCTCCTCCTCCTCATCATCATCATCATTCA
GCCATGAACGGTGGCGGCATCTGGCCCTACTACACCTCGGGGATCGCGATCCCGATATACCC
GGCGGCGCCGGCGTACTGGGGCTGCATGATTCCACCTCCTGGAGCTTGGAGCCTCCCATGGC
CGGCCACAGTCCAGTCTCAGGCCATCTCATCATCATCACCACCTACAAGTGCTACACCTTCA
GTCTCCTCCTTCACACTAGGCAAGCATCCCAGAGAGGGTGGTGATCATGAGGCAAGAGATCA
CCATGGCAATGGTAAAGTGTGGGTGCCGAAGACGATCCGGATCGACAACGCCGACGAGGTTG
CCCGGAGCTCAATCCGGTCACTCTTCGCCTTCAGAGGCGGCGACAAGGCGGATGATAACAAC
GACGACGATGGCACCGGCGTGCACAAGCTCGCCACCACGGTGTTCGAGCCAAAGAGGGACAG
CAAGACGGCGAAACATCCGGCGATCACGAGCTTGCCGCTCTTGCACACCAACCCCGTCGCGC
TTACCCGATCCGCGACCTTCCAGGAGGGATCTTGA

FIGURA 28 cont.

SEQ ID NO: 239, Oryza sativa – proteína – AAK98677.1

MGGGGKCKRGKRGKIAAKRRRGQCCSSSSSSGGARRRARERESEESLGEMGECRGGGGGGDG
LIKLFGKTIPVQPDAKDVQQHSGSSSSSTESDVQETAAVAVADPSPRSEVVDGESPPQPGGE
AASHQQQQKEMKLKKPDKILPCPRCSSMDTKFCYFNNYNVNQPRHFCKHCQRYWTAGGAMRN
VPVGAGRRKNKNATAAAHFLHRVRACAAAAAMPAAPHDATNATVLSFGGGGGGHDAPPVTLD
LADKMTRLGKEGLVAHARNADAAAACSEVSSNRDDEQIGNTVAKPANGLQQHPPPPHHHHHS
AMNGGGIWPYYTSGIAIPIYPAAPAYWGCMIPPPGAWSLPWPATVQSQAISSSSPPTSATPS
VSSFTLGKHPREGGDHEARDHHGNGKVWVPKTIRIDNADEVARSSIRSLFAFRGGDKADDNN
DDDGTGVHKLATTVFEPKRDSKTAKHPAITSLPLLHTNPVALTRSATFQEGS

SEQ ID NO: 240, Oryza sativa – DNA

ATGGGAGGGGGTGGTAAGTGCAAAAGAGGGAAAAGAGGAAAGATTGCAGCCAAGCGAAGAAG
GGGCCAGTGTTGCAGCAGCAGCAGCAGCAGCGGAGGCGCAAGAAGACGAGCTAGAGAGAGAG
AATCAGAGGAGAGCTTGGGTGAGATGGGTGAGTGCAGAGGAGGAGGAGGAGGAGGAGATGGG
CTGATCAAGCTGTTCGGGAAGACGATCCCGGTGCAGCCGGATGCCAAGGATGTTCAGCAGCA
TAGTGGCAGTAGCAGCAGCTCAACCGAATCCGACGTCCAAGAAACCGCCGCTGTCGCCGTCG
CCGACCCCTCCCCGCGGTCGGAGGTCGTCGACGGCGAGAGCCCGCCGCAGCCGGGCGGCGAG
GCGGCGAGCCATCAGCAGCAGCAGAAGGAGATGAAGCTGAAGAAGCCGGACAAGATCCTGCC
ATGCCCGCGGTGCAGCAGCATGGACACCAAGTTCTGCTACTTCAACAACTACAACGTCAACC
AGCCTCGCCACTTCTGCAAGCACTGCCAGCGCTACTGGACCGCCGGCGGCGCCATGCGCAAC
GTCCCCGTCGGCGCCGGCCGCCGCAAGAACAAGAACGCCACCGCCGCCGCCCACTTCCTCCA
CCGCGTCCGCGCCTGCGCCGCCGCCGCCGCCATGCCCGCGGCGCCCCACGACGCCACCAACG
CCACCGTGCTCAGCTTCGGCGGCGGCGGAGGCGGACACGACGCGCCGCCGGTCACCCTGGAC
CTCGCCGACAAGATGACGCGCCTCGGCAAGGAGGGGCTCGTCGCCCATGCCCGGAACGCCGA
CGCCGCCGCCGCGTGCAGCGAGGTGTCGAGCAACAGGGACGACGAGCAGATCGGCAACACTG
TAGCAAAACCTGCAAACGGGTTGCAGCAGCATCCTCCTCCTCCTCATCATCATCATCATTCA
GCCATGAACGGTGGCGGCATCTGGCCCTACTACACCTCGGGGATCGCGATCCCGATATACCC
GGCGGCGCCGGCGTACTGGGGCTGCATGATTCCACCTCCTGGAGCTTGGAGCCTCCCATGGC
CGGCCACAGTCCAGTCTCAGGCCATCTCATCATCATCACCACCTACAAGTGCTACACCTTCA
GTCTCCTCCTTCACACTAGGCAAGCATCCCAGAGAGGGTGGTGATCATGAGGCAAGAGATCA
CCATGGCAATGGTAAAGTGTGGGTGCCGAAGACGATCCGGATCGACAACGCCGACGAGGTTG
CCCGGAGCTCAATCCGGTCACTCTTCGCCTTCAGAGGCGGCGACAAGGCGGATGATAACAAC
GACGACGATGGCACCGGCGTGCACAAGCTCGCCACCACGGTGTTCGAGCCAAAGAGGGACAG
CAAGACGGCGAAACATCCGGCGATCACGAGCTTGCCGCTCTTGCACACCAACCCCGTCGCGC
TTACCCGATCCGCGACCTTCCAGGAGGGATCTTGA

SEQ ID NO: 241, Oryza sativa – proteína – XP_470300.1

MGECKVGGGGGGGDCLIKLFGKTIPVPEPGACAAGDVDKDLQHSGSSTTEPKTQENTVQDST
SPPPQPEVVDTEDSSADKNSSENQQQQGDTANQKEKLKKPDKILPCPRCSSMDTKFCYYNNY
NINQPRHFCKNCQRYWTAGGAMRNVPVGAGRRKSKSVSAASHFLQRVRAALPGDPPLYAPVK
TNGTVLSFGSDLSTLDLTEQMKHLKDKFIPTTGIKNTDEMPVGLCAEGLSKTEESNQTNLKE
KVSADRSPNVAQHPCMNGGAMWPFGVAPPPAYYTSSIAIPFYPAAAAAVAAYWGCMVPGAWN
APWPPQSQSQSVSSSSAASPVSTMTNCFRLGKHPRDGDEELDSKGNGKVWVPKTVRIDDVDE
VARSSIWSLIGIKGDKVGADHGRGCKLAKVFESKDEAKASTHTAISSLPFMQGNPAALTRSV
TFQEGS

FIGURA 28 cont.

SEQ ID NO: 242, Oryza sativa - DNA

ACTTGTCGGATCTCTCTCTCTCTCTCTCTCCCTCTCTCTCTCGCCCATGTACCATCTCCA
TGGACGACCTCGCCGCCGCCTCCCCTCCTCACCCGCCGCCGCCGCCGCCGGAATCCCACGTG
CCTCCGCCCCCGCAGACGCCGGAGAAGGATTCATGTGAGGATACAGGAGACATGAGGATCAG
TGAAGAAAAGCCATGCACAGATCAGGAGTTAGATGCTGATCAGATGAATAGTTCTAGCTTTA
ATAGTTCCAGCGAGTGTGAGAATCAAACACCTAGCAATGATGAAATGACTGGATCAGAGTCC
AAATCTGAGGCAGCTCAAACAGAGGGTGGTGGATCGAGTGAAGAGAAGGTCCTGAAGAAGCC
AGACAAGATCCTGCCTTGTCCTCGTTGCAACAGCATGGATACAAAGTTCTGCTACTACAACA
ACTACAACATTAACCAGCCAAGACATTTTTGCAAGAGTTGTCAGAGATATTGGACGGCAGGT
GGAAGCATGAGGAACCTCCCTGTTGGTGCTGGTAGGCGCAAGAGTAAGAGCTCCACTGCAAA
TTACCGCAGTATATTAATCACGGGCAGCAATCTAGCTGCTCCTGCTGGAGATGCTCCCCTCT
ATCAACTCTCTATAAAAGGAGATCAAACAGCAACGGCAGTTAAATTTGCACCTGATTCCCCA
CTCTGTAATTCCATGGCCTCTGTGCTGAAAATTGGAGAGCAGAGTAAGAATGCCAAGCCTAC
CTCAACAGCACAACCCAGAAATGGAGAAACCCAGACCTGCCCGGCTTCAGGAACAACTTCAG
ATAGTCCCCGGAATGAACCAGTTAATGGAGCAGTTAGTGGGCATCAAAATGGAATTGTTGGG
CATAGTGGAGTCCCTCCCATGCATCCCATACCATGCTTCCCTGGTCCTCCTTTTGTGTACCC
ATGGAGTCCAGCATGGAATGGCATTCCTGCCATGGCACCACCGGTATGCACAGCACCAGCTG
AACCAGCAAATTCTTCAGACAATGGAAGCACAGCTAGTGTTCAGTGGAGCATGCCACCAGTG
ATGCCGGTACCAGGATACTTTCCGGTAATTCCATCTTCAGTTTGGCCCTTCATTTCTCCCTG
GCCAAATGGTGCATGGAGCTCGCCATGGATTCAACCTAATTGCAGCGTGTCAGCTTCATCTC
CTACAAGCACTAGTACATGTTCAGACAACGGCTCTCCTGTCCTAGGAAAGCACTCCAGGGAC
TCCAAACCGCAAGGAGATGACAAGGCAGAGAAAAACTTGTGGATTCCGAAAACGCTTCGGAT
CGATGATCCTGACGAAGCTGCAAAGAGCTCAATCTGGACAACCCTTGGCATTGAACCTGGTG
ACCGTAGCATGTTCAGATCATTCCAGTCGAAACCTGAAAGCAGGGAGCAGATATCCGGTGCT
GCACGAGTCCTGCAGGCGAATCCAGCAGCTCTATCTCGATCTCAATCTTTCCAGGAGACAAC
GTGATGTATATTGAAGAAATCGTGTGACAATTGTAGAACATGTTACTACTATAATTTAAGCA
AGTGCTACAGCCTGAAGGATGATTGGCGATGTAGGCGCTGCTGCAAACATGGAGGCAGCGTG
ATCTGTACTATTGAAACCTGAAGAGTACTATTCTTGACATTTTTCTATAATTTGCCTGTGGA
GCATGCACCAGTCTACTGGGTTCATGATCAGGCTGTCAGCATATGTTTTGTTGTACAATAAT
AATAATTGTTAATAGCTTTGCCTAAAGGATACTGATAGTATTCTGCTGATCCTTCAGCTCTG
TGATC

SEQ ID NO: 243, Oryza sativa - XP_479581.1

MDDLAAASPPHPPPPPPESHVPPPPQTPEKDSCEDTGDMRISEEKPCTDQELDADQMNSSSF
NSSSECENQTPSNDEMTGSESKSEAAQTEGGGSSEEKVLKKPDKILPCPRCNSMDTKFCYYN
NYNINQPRHFCKSCQRYWTAGGSMRNLPVGAGRRKSKSSTANYRSILITGSNLAAPAGDAPL
YQLSIKGDQTATAVKFAPDSPLCNSMASVLKIGEQSKNAKPTSTAQPRNGETQTCPASGTTS
DSPRNEPVNGAVSGHQNGIVGHSGVPPMHPIPCFPGPPFVYPWSPAWNGIPAMAPPVCTAPA
EPANSSDNGSTASVQWSMPPVMPVPGYFPVIPSSVWPFISPWPNGAWSSPWIQPNCSVSASS
PTSTSTCSDNGSPVLGKHSRDSKPQGDDKAEKNLWIPKTLRIDDPDEAAKSSIWTTLGIEPG
DRSMFRSFQSKPESREQISGAARVLQANPAALSRSQSFQETT

SEQ ID NO: 244, Hordeum vulgare - DNA

CTAGACATGATTGCCAAGCTCGAAAATAACCCTCACAAAGGGAACAAAACTGGTACCGGGCC
CCCCCTCGAGGTCGACAAGGTCGGTGTCCATGCCCCCTCCCCCTCCTCCTTTGTGTTCTTGT
GTATGAAATCCCGGGAATTTTGATTGGTTTCTAATAGAGCCTGTGAAAAAAAAAAAGGCCAT
TTCGAGGCCTGCTTTCTTTGGCTTTGGAGCCGATCATGGCGACGAGTCGTGTTTGGTCTAGT

FIGURA 28 cont.

CATCAGGTCTTAAGCATGGAGTTCCCTGCCTGTTTTGGGAGGTACTAGTAAGCGCTAAGCAG
GCAGGTTTAGCTTCCATTAGCAAGGGAATTTCGCCTTTGCTGGTCTCTAATAATAATAGCAG
TTTCAACAGTTTTAGTATTGCATGATCTAGTCGTCTGAACATCTGGACTTCTGAAAGACAGG
CCTTTCTAGGTTATTATCTTGCCTCTTTTAATAGGTCGTTTTCGCGTGCTCCCAAGGATTTT
GCTGAAAGTGGCTACTGATATGTGTTGCCGTAACCTACTGGACTTCAGAAGTTCAGATGCGT
CCGACTCTGAGTACATCACCCACTGCAAATCAAAGCACCATAGCTTCTTTTTTTCATTAACA
AGATATAAGAACTGGGAATATGAACAAACTTTGTCAAATTTTACACTTGGGTCAGATGGGGG
AAAGCATGTTGTAATCACCGGTCAATAACCCGAATATTACACCGTGTTTGTGGACCTCTAAT
TCACAACAAAATTATTTCAAAGTTAGACTATATATATATACAAGGAACTCATCGAGATTTTC
TGAATGTTTACAGGACCTTCAGCACAGAGGAAGCACCACGGCTGAACCGAAAGTACAAGAAA
GCGCCCCACAGGACTCCACGGGCTCGCCTCCGCAGCCGGAGGTTGTGGACGTCGAGGATCCA
TCTGCTGTCAAGAACTCAGCAGCAGATCAACAAGAAGAAGAAGAAGAACAGGGTGACACGGC
CAACCAGAAGGAGAAGCTCAAGAAGCCTGACAAGATCCTGCCCTGCCCACGCTGTAGCAGCA
TGGACACCAAGTTCTGCTACTACAACAACTACAACATCAACCAGCCGCGCCACTTCTGCAAG
AACTGCCAGAGGTACTGGACGGCCGGCGGGGCCATGCGCAACGTGCCCGTGGGCGCGGGCCG
CCGCAAGAGCAAGAGCATATCGGCAGCGTCCCACTTCCTTCAGAGGATAAGGGCCGCTCTGC
CCGGTGATCCTCTCTGCACCCCGGTCAACACTAACGGCACGGTGCTCAGCTTCGGCTCCGAC
GCATCCACCTTGGACGTCGTCTCAGAACAGATGAAGCACATGAAGGAGCTCAGCTCAGTAAC
CCGGACCGAGAACACCGATGCCCCGTCAGTAGGGTCTTGTGCTGAAGGATGGGCAAAGGGAG
AAGAGTCGAGCCAGATGAACTCGAGGGAGAGAGTTGCAGCAGATAGATCGCCAAATTTTGCG
CAGCACCCGTGCATGAACGGGGCAGCCATGTGGCCATTCAGTTGTGCACCATCGCCTGCCTA
TTTCACCCCAAACGTAGCAATTCCATTCTATCCAGCTGCTGCTGCTGCCTACTGGGGCTGCA
TGGTTCCGGGAGCCTGGAACACTCCATGGCAGCCGCAGCCGCAGCCGCAGTCTCAGTGTCAA
TCTAGCTCACCACCTAGTGCCGCTTCTCCGGTATCAACAATGTCCAGTTGCTTCCAATCCCG
GAAGCATCCTAGAGATGGTGATGAGGAAAGAGATACCAAGGGTAATGGCAAGGTGTGGGTGC
CCAAGACGATCGAGTCGAC

SEQ ID NO: 245, Hordeum vulgare - proteína - CAC85949 parcial

LKKPDKILPCPRCSSMDTKFCYYNNYNINQPRHFCKNCQRYWTAGGAMRNVPVGAGRRKSKS
ISAASHFLQRIRAALPGDPLCTPVNTNGTVLSFGSDASTLDVVSEQMKHMKELSSVTRTENT
DAPSVGSCAEGWAKGEESSQMNSRERVAADRSPNFAQHPCMNGAAMWPFSCAPSPAYFTPNV
AIPFYPAAAAAYWGCMVPGAWNTPWQPQPQPQSQCQSSSPPSAASPVSTMSSCFQSRKHPRD
GDEERDTKGNGKVWVPKTIEST

SEQ ID NO: 246, Cucurbita maxima - DNA

CAGCTGCAAAAGAGGAAACTCTTCATGATTCAGAGGATTATGCATGTGTAAAAACAGCAAAT
GAGGCGCACATGAATCCTGAAGTGTTATCAACAGATGAAAATGATAAGCTCGCAACAAGGAA
AACTGAGAAAGAACAGAATGATGCCCCCAACTCGAAAGAGAAACTGAAGAAACCAGATAAGA
TACTTCCATGTCCCCGCTGCAATAGCATGGAAACCAAGTTTTGTTATTATAATAATTATAAT
GTCAATCAACCACGCCATTTTTGCAAAGCCTGTCAAAGATATTGGACTGAAGGCGGTACCAT
CAGGAATGTCCCTGTTGGAGCTGGCCGCCGAAAAAACAAGAACTCAGCCTCACACTACCGAC
ACATTACAATCTCAGAGGCTCTCCGAGCTGCACAAATTGATGTTCCTATTGAGGTCAACCAC
CTAGCATCAAAAGGCAATGGACGAGTCCTCAATTTCAGTGTAAGCCCACCTGTATGTGAATC
TATGGTCAATGTATCACATCCTGCAGAAAGAAAGGTCTTGAATGGAACAAGGAATGAATTTG
AGGGAGCCAAAGGACCTTGTGAGGGTGGGGAAACTGGTGATGATTGTTCTAGTGCATCTTCA
GTAACAATGTCAAGCTCAATGAAGAATGGAGCCAGGAGGTTCCCTCAAGAACCACATATGCA
GAACATCAATGGTTTTCCTTCTCAAATCCCATGTCTTCCTGGTGTTCCTTGGCCTTGTTCAT

FIGURA 28 cont.

GGACTGCACCAATACCTCCACCAGCCTTGTGCCCTCCTGGAGTTCCTTTATCATTCTACCCT
GCAACATATTGGAGTTGCAGTGCTTCAGGTTCTTGGAATATTCCTTGGGTCACTCCACAACC
CTGTCCTCCAATCCCTGGTCCAAATTCTCCGACGCTAGGCAAGCATTCGAGAGATGGCGATG
AACTCCAGGCTGATAATTCTGAGATGAAAGATCCTCCAAAACAGAAAAATGGATCTGTTTTG
GTTCCCAAAACTTTAAGGATTGATGATCCAAACGAAGCTGCTAAGAGTTCAATATGGGAGAC
ACTTGGTATTAAGAATGATTCAATCAAAGCTGTTGATCTGTCTAATGTTTTCCAATCAAAGG
GCGACCTAAAGAGTAACGTTTCTGAAGTGTTGTCTCCAGTTTTGCAAGCCAACCCTGCAGCC
TTGTCAAGATCTCTTACTTTCCACGAGCGGTCGTGAATGGTATTTTTGATTTTCCAGGTTCA
TTGTAAAAGAAATCTTTAAGTTAAGAAGCTTAAGCATGAAGTGGATACAAGCATCAATCTCA
AGCTTCTTCTGCTCAAGAAGCAATCCAGTTTTTCAAGAATCCCGGAATTGTTCTCTCTATAG
AAAGAGTAGCTGCCATTATCAAGCCATCCTTTTGTACATAACTTTACATATGTGAATAACCG
ATGTAGAGTGGCAGCTAAAGATTCTCAGATGTATAAAAATGCAGCCAAAAGTTTGCTTC

SEQ ID NO: 247, Cucurbita maxima - BAA08094

MNPEVLSTDENDKLATRKTEKEQNDAPNSKEKLKKPDKILPCPRCNSMETKFCYYNNYNVNQ
PRHFCKACQRYWTEGGTIRNVPVGAGRRKNKNSASHYRHITISEALRAAQIDVPIEVNHLAS
KGNGRVLNFSVSPPVCESMVNVSHPAERKVLNGTRNEFEGAKGPCEGGETGDDCSSASSVTM
SSSMKNGARRFPQEPHMQNINGFPSQIPCLPGVPWPCSWTAPIPPPALCPPGVPLSFYPATY
WSCSASGSWNIPWVTPQPCPPIPGPNSPTLGKHSRDGDELQADNSEMKDPPKQKNGSVLVPK
TLRIDDPNEAAKSSIWETLGIKNDSIKAVDLSNVFQSKGDLKSNVSEVLSPVLQANPAALSR
SLTFHERS

SEQ ID NO: 248, Arabidopsis thaliana - DNA

ATGTGGCTCTCTCATCTCTTCATGTCCCTCTCTAAGCTGACGTGTAATTTCTCCATTTTCTC
TGTCTTTATGGCTTGTGGCTCTATTGGTATGTCTCAAGTTAGAGATACTCCGGTTAAATTGT
TTGGCTGGACAATTACACCGGTTTCTCATGATCCATACTCTTCTTCGTCCCATGTTCTTCCT
GATTCTTCCTCGTCTTCCTCGTCTTCTTCTCTATCACTTCGACCACACATGATGAATAACCA
ATCTGTTACTGACAATACAAGTCTTAAGCTGTCATCTAATCTTAACAACGAGTCAAAAGAAA
CATCTGAGAACAGTGATGACCAACACAGCGAGATCACAACAATTACATCGGAAGAAGAGAAA
ACAACTGAACTGAAGAAACCAGACAAGATTCTTCCATGTCCGAGATGCAACAGCGCAGACAC
CAAATTCTGTTACTACAACAACTACAACGTTAACCAGCCACGTCACTTCTGTAGAAAATGCC
AGAGGTATTGGACCGCTGGTGGATCCATGAGGATCGTCCCGGTTGGCTCAGGCCGTCGCAAG
AACAAGGGATGGGTTTCTTCAGACCAGTACCTGCACATCACTTCCGAGGATACTGACAATTA
CAATAGCTCCTCAACAAAGATTCTAAGCTTCGAGTCTTCGGACTCTTTGGTAACTGAGAGGC
CTAAGCATCAATCAAACGAAGTGAAGATAAACGCTGAACCTGTTTCACAAGAACCCAACAAC
TTCCAAGGGTTACTTCCTCCCCAAGCATCCCCTGTTTCGCCTCCTTGGCCTTACCAATACCC
TCCAAACCCTAGTTTCTACCACATGCCCGTCTACTGGGGCTGCGCGATACCGGTTTGGTCTA
CCCTCGACACTTCTACATGTCTTGGGAAAAGGACAAGAGACGAAACTTCTCATGAAACTGTT
AAAGAGAGTAAAAATGCTTTTGAGAGAACAAGCTTGCTTTTGGAATCTCAGAGCATCAAAAA
TGAAACAAGTATGGCTACAAATAACCATGTGTGGTATCCAGTACCGATGACCCGCGAGAAGA
CACAAGAATTCAGCTTTTTCAGTAATGGAGCTGAAACAAAGAGCAGCAACAACAGATTCGTC
CCTGAAACGTATCTTAACCTGCAAGCAAACCCTGCAGCCATGGCAAGATCTATGAACTTCAG
AGAGAGCATATAA

FIGURA 28 cont.

SEQ ID NO: 249, Arabidopsis thaliana - proteína - At1g26790

MWLSHLFMSLSKLTCNFSIFSVFMACGSIGMSQVRDTPVKLFGWTITPVSHDPYSSSSHVLP
DSSSSSSSSSLSLRPHMMNNQSVTDNTSLKLSSNLNNESKETSENSDDQHSEITTITSEEEK
TTELKKPDKILPCPRCNSADTKFCYYNNYNVNQPRHFCRKCQRYWTAGGSMRIVPVGSGRRK
NKGWVSSDQYLHITSEDTDNYNSSSTKILSFESSDSLVTERPKHQSNEVKINAEPVSQEPNN
FQGLLPPQASPVSPPWPYQYPPNPSFYHMPVYWGCAIPVWSTLDTSTCLGKRTRDETSHETV
KESKNAFERTSLLLESQSIKNETSMATNNHVWYPVPMTREKTQEFSFFSNGAETKSSNNRFV
PETYLNLQANPAAMARSMNFRESI

SEQ ID NO: 250, Arabidopsis thaliana - DNA

ATGTCTAAATCTAGAGATACGGAGATAAAGTTGTTTGGGAGGACAATCACATCTCTTTTAGA
TGTGAATTGTTATGATCCGTCGTCGTTGTCCCCTGTTCACGATGTTTCTTCTGATCCAAGCA
AGGAGGATTCGTCTTCTTCTTCATCTTCTTGTTCTCCAACTATTGGACCAATCAGGGTTCCG
GTTAAAAAAAGTGAGCAAGAGAGTAACAAATTCAAAGATCCATATATATTATCCGATCTAAA
CGAACCACCAAAAGCAGTATCTGAGATTTCATCACCAAGAAGTTCCAAGAACAACTGTGATC
AACAGAGCGAGATCACAACAACAACTACCACAAGTACTACATCAGGAGAGAAATCAACGGCT
CTCAAGAAACCGGACAAGCTTATTCCATGTCCTAGATGTGAAAGCGCAAACACCAAATTCTG
TTATTACAACAACTACAACGTGAACCAGCCACGTTACTTCTGCAGGAACTGTCAGAGGTATT
GGACAGCTGGTGGATCTATGAGGAACGTTCCTGTTGGCTCAGGTCGTCGCAAGAACAAAGGA
TGGCCTTCTTCAAACCATTACTTGCAAGTCACTTCTGAGGATTGTGATAATAATAACTCGGG
GACGATCCTTAGTTTCGGTTCTTCGGAGTCTTCGGTTACAGAGACTGGTAAGCATCAGTCAG
GTGATACAGCAAAGATAAGTGCTGATTCAGTTTCTCAAGAAAATAAAAGCTACCAAGGGTTT
CTTCCTCCGCAAGTAATGTTACCTAATAATTCTTCTCCTTGGCCTTACCAATGGAGTCCAAC
GGGTCCTAACGCTAGTTTCTACCCTGTCCCCTTCTACTGGGGATGCACGGTTCCGATATACC
CTACCTCAGAGACTTCATCATGTTTAGGAAAACGGTCAAGAGATCAAACTGAAGGAAGAATC
AATGATACTAATACAACAATAACTACTACAAGAGCAAGATTGGTCTCAGAATCTCTTAGAAT
GAATATCGAAGCTAGTAAGAGCGCTGTGTGGTCTAAGTTACCGACAAAACCCGAGAAAAAAA
CGCAAGGATTCAGTTTGTTCAATGGATTTGACACAAAGGGAAACAGCAACAGAAGTAGCTTG
GTCTCCGAAACTTCTCACAGTCTACAAGCAAACCCTGCAGCGATGTCTAGAGCTATGAACTT
CAGGGAGAGCATGCAACAATAA

SEQ ID NO: 251, Arabidopsis thaliana - proteína - At1g69570

MSKSRDTEIKLFGRTITSLLDVNCYDPSSLSPVHDVSSDPSKEDSSSSSSSCSPTIGPIRVP
VKKSEQESNKFKDPYILSDLNEPPKAVSEISSPRSSKNNCDQQSEITTTTTTSTTSGEKSTA
LKKPDKLIPCPRCESANTKFCYYNNYNVNQPRYFCRNCQRYWTAGGSMRNVPVGSGRRKNKG
WPSSNHYLQVTSEDCDNNNSGTILSFGSSESSVTETGKHQSGDTAKISADSVSQENKSYQGF
LPPQVMLPNNSSPWPYQWSPTGPNASFYPVPFYWGCTVPIYPTSETSSCLGKRSRDQTEGRI
NDTNTTITTTRARLVSESLRMNIEASKSAVWSKLPTKPEKKTQGFSLFNGFDTKGNSNRSSL
VSETSHSLQANPAAMSRAMNFRESMQQ

SEQ ID NO: 252, Arabidopsis thaliana - DNA

ATGATGATGGAGACTAGAGATCCAGCTATTAAGCTTTTCGGTATGAAAATCCCTTTTCCGTC
GGTTTTTGAATCGGCAGTTACGGTGGAGGATGACGAAGAAGATGACTGGAGCGGCGGAGATG
ACAAATCACCAGAGAAGGTAACTCCAGAGTTATCAGATAAGAACAACAACAACTGTAACGAC
AACAGTTTTAACAATTCGAAACCCGAAACCTTGGACAAAGAGGAAGCGACATCAACTGATCA
GATAGAGAGTAGTGACACGCCTGAGGATAATCAGCAGACGACACCTGATGGTAAAACCCTAA
AGAAACCGACTAAGATTCTACCGTGTCCGAGATGCAAAAGCATGGAGACCAAGTTCTGTTAT

FIGURA 28 cont.

TACAACAACTACAACATAAACCAGCCTCGTCATTTCTGCAAGGCTTGTCAGAGATATTGGAC
TGCTGGAGGGACTATGAGGAATGTTCCTGTGGGGGCAGGACGTCGTAAGAACAAAAGCTCAT
CTTCTCATTACCGTCACATCACTATTTCCGAGGCTCTTGAGGCTGCGAGGCTTGACCCGGGC
TTACAGGCAAACACAAGGGTCTTGAGTTTTGGTCTCGAAGCTCAGCAGCAGCACGTTGCTGC
TCCCATGACACCTGTGATGAAGCTACAAGAAGATCAAAAGGTCTCAAACGGTGCTAGGAACA
GGTTTCACGGGTTAGCGGATCAACGGCTTGTAGCTCGGGTAGAGAATGGAGATGATTGCTCA
AGCGGATCCTCTGTGACCACCTCTAACAATCACTCAGTGGATGAATCAAGAGCACAAAGCGG
CAGTGTTGTTGAAGCACAAATGAACAACAACAACAACAATAACATGAATGGTTATGCTTGCA
TCCCAGGTGTTCCATGGCCTTACACGTGGAATCCAGCGATGCCTCCACCAGGTTTTTACCCG
CCTCCAGGGTATCCAATGCCGTTTTACCCTTACTGGACCATCCCAATGCTACCACCGCATCA
ATCCTCATCGCCTATAAGCCAAAAGTGTTCAAATACAAACTCTCCGACTCTCGGAAAGCATC
CGAGAGATGAAGGATCATCGAAAAAGGACAACGAGACAGAGCGAAAACAGAAGGCCGGGTGC
GTTCTGGTCCCGAAAACGTTGAGAATAGATGATCCTAACGAAGCAGCAAAGAGCTCGATATG
GACAACATTGGGAATCAAGAACGAGGCGATGTGCAAAGCCGGTGGTATGTTCAAAGGGTTTG
ATCATAAGACAAAGATGTATAACAACGACAAAGCTGAGAACTCCCCTGTTCTTTCTGCTAAC
CCTGCTGCTCTATCAAGATCACACAATTTCCATGAACAGATTTAG

SEQ ID NO: 253, Arabidopsis thaliana - proteína - At3g4750

MMMETRDPAIKLFGMKIPFPSVFESAVTVEDDEEDDWSGGDDKSPEKVTPELSDKNNNNCND
NSFNNSKPETLDKEEATSTDQIESSDTPEDNQQTTPDGKTLKKPTKILPCPRCKSMETKFCY
YNNYNINQPRHFCKACQRYWTAGGTMRNVPVGAGRRKNKSSSSHYRHITISEALEAARLDPG
LQANTRVLSFGLEAQQQHVAAPMTPVMKLQEDQKVSNGARNRFHGLADQRLVARVENGDDCS
SGSSVTTSNNHSVDESRAQSGSVVEAQMNNNNNNNNMNGYACIPGVPWPYTWNPAMPPPGFYP
PPGYPMPFYPYWTIPMLPPHQSSSPISQKCSNTNSPTLGKHPRDEGSSKKDNETERKQKAGC
VLVPKTLRIDDPNEAAKSSIWTTLGIKNEAMCKAGGMFKGFDHKTKMYNNDKAENSPVLSAN
PAALSRSHNFHEQI

SEQ ID NO: 254, Arabidopsis thaliana - DNA

ATGGCTGATCCGGCGATTAAGCTCTTTGGAAAGACGATTCCTTTACCTGAGCTTGGTGTTGT
TGATTCTTCTTCTAGCTATACCGGATTTTTAACCGAAACTCAGATTCCTGTTCGGTTATCAG
ATTCGTGTACCGGCGATGATGATGATGAAGAGATGGGTGATTCCGGTTTAGGACGAGAAGAA
GGTGATGATGTTGGTGATGGTGGAGGAGAGAGCGAGACTGATAAAAAGGAAGAAAAAGATAG
TGAGTGTCAGGAAGAGTCATTGAGGAATGAATCTAATGATGTTACTACTACTACATCGGGTA
TAACTGAAAAAACGGAAACAACAAAAGCTGCAAAGACGAATGAAGAGTCAGGTGGTACTGCT
TGCTCTCAAGAGGGGAAGTTAAAGAAACCTGATAAGATTCTACCGTGTCCGCGATGTAACAG
CATGGAAACCAAGTTCTGTTACTACAACAACTATAATGTTAACCAACCTCGCCATTTCTGCA
AGAAATGTCAGAGATATTGGACAGCTGGTGGAACGATGAGGAATGTTCCGGTTGGTGCTGGG
AGACGTAAGAATAAGAGTCCAGCTTCTCATTATAACCGTCATGTAAGTATAACATCTGCGGA
AGCTATGCAGAAGGTGGCGAGAACTGATCTTCAACATCCTAATGGTGCAAATCTTCTCACTT
TTGGCTCTGATTCTGTGCTTTGTGAATCTATGGCTTCTGGATTGAATCTTGTTGAGAAGTCA
TTGTTGAAGACACAAACTGTATTGCAAGAACCCAATGAAGGCTTGAAGATTACGGTTCCGTT
AAACCAGACAAACGAAGAAGCTGGAACAGTCAGCCCGTTACCAAAAGTTCCATGCTTTCCAG
GACCACCACCAACTTGGCCTTACGCTTGGAACGGAGTTTCGTGGACGATTTTACCGTTTTAC
CCTCCACCGGCTTACTGGAGCTGCCCGGGGGTTTCACCGGGGGCATGGAACAGCTTCACATG
GATGCCACAACCCAATTCACCATCTGGTTCCAATCCAAATTCTCCTACACTAGGTAAACATT

FIGURA 28 cont.

CACGTGACGAGAACGCTGCTGAACCAGGAACCGCTTTTGATGAAACCGAGTCACTTGGTAGG
GAGAAAAGCAAACCCGAGAGATGCTTGTGGGTTCCCAAGACGCTGAGGATTGATGATCCAGA
GGAAGCTGCTAAAAGTTCCATCTGGGAAACATTAGGGATCAAAAAAGACGAAAATGCGGATA
CTTTCGGAGCTTTCAGATCATCAACCAAAGAAAAAAGCAGTCTTTCTGAAGGAAGACTTCCG
GGAAGAAGACCGGAGTTGCAAGCGAATCCTGCTGCTCTTTCTAGGTCAGCAAACTTCCATGA
GAGCTCATAG

SEQ ID NO: 255, Arabidopsis thaliana – proteína – At5g39660
MADPAIKLFGKTIPLPELGVVDSSSSYTGFLTETQIPVRLSDSCTGDDDDEEMGDSGLGREE
GDDVGDGGGESETDKKEEKDSECQEESLRNESNDVTTTTSGITEKTETTKAAKTNEESGGTA
CSQEGKLKKPDKILPCPRCNSMETKFCYYNNYNVNQPRHFCKKCQRYWTAGGTMRNVPVGAG
RRKNKSPASHYNRHVSITSAEAMQKVARTDLQHPNGANLLTFGSDSVLCESMASGLNLVEKS
LLKTQTVLQEPNEGLKITVPLNQTNEEAGTVSPLPKVPCFPGPPPTWPYAWNGVSWTILPFY
PPPAYWSCPGVSPGAWNSFTWMPQPNSPSGSNPNSPTLGKHSRDENAAEPGTAFDETESLGR
EKSKPERCLWVPKTLRIDDPEEAAKSSIWETLGIKKDENADTFGAFRSSTKEKSSLSEGRLP
GRRPELQANPAALSRSANFHESS

SEQ ID NO: 256, Seqüência artificial – iniciador prm07303
GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGATGATGGAGACTAGAGATC

SEQ ID NO: 257, Seqüência artificial – iniciador prm07304
GGGGACCACTTTGTACAAGAAAGCTGGGTCATATGTAACTCTAAATCTGTTCA

SEQ ID NO: 258, Oryza sativa – DNA – promotor de prolamina
CTTCTACATCGGCTTAGGTGTAGCAACACGACTTTATTATTATTATTATTATTATTATTATT
ATTTTACAAAAATATAAAATAGATCAGTCCCTCACCACAAGTAGAGCAAGTTGGTGAGTTAT
TGTAAAGTTCTACAAAGCTAATTTAAAAGTTATTGCATTAACTTATTTCATATTACAAACAA
GAGTGTCAATGGAACAATGAAAACCATATGACATACTATAATTTTGTTTTTATTATTGAAAT
TATATAATTCAAAGAGAATAAATCCACATAGCCGTAAAGTTCTACATGTGGTGCATTACCAA
AATATATATAGCTTACAAAACATGACAAGCTTAGTTTGAAAAATTGCAATCCTTATCACATT
GACACATAAAGTGAGTGATGAGTCATAATATTATTTTCTTTGCTACCCATCATGTATATATG
ATAGCCACAAAGTTACTTTGATGATGATATCAAAGAACATTTTTAGGTGCACCTAACAGAAT
ATCCAAATAATATGACTCACTTAGATCATAATAGAGCATCAAGTAAAACTAACACTCTAAAG
CAACCGATGGGAAAGCATCTATAAATAGACAAGCACAATGAAAATCCTCATCATCCTTCACC
ACAATTCAAATATTATAGTTGAAGCATAGTAGTA

SEQ ID NO: 259, Oryza sativa - DNA – promotor wsi18
GCTTGAGTCATAGGGAGAAAACAAATCGATCATATTTGACTCTTTTCCCTCCATCTCTCTTA
CCGGCAAAAAAAGTAGTACTGGTTTATATGTAAAGTAAGATTCTTTAATTATGTGAGATCCG
GCTTAATGCTTTTCTTTTGTCACATATACTGCATTGCAACAATTGCCATATATTCACTTCTG
CCATCCCATTATATAGCAACTCAAGAATGGATTGATATATCCCCTATTACTAATCTAGACAT
GTTAAGGCTGAGTTGGGCAGTCCATCTTCCCAACCCACCACCTTCGTTTTTCGCGCACATAC
TTTTCAAACTACTAAATGGTGTGTTTTTTAAAAATATTTTCAATACAAAAGTTGCTTTAAAA
AATTATATTGATCCATTTTTTTAAAAAAAATAGCTAATACTTAATTAATCACGTGTTAAAAG
ACCGCTCCGTTTTGCGTGCAGGAGGGATAGGTTCACATCCTGCATTACCGAACACAGCCTAA
ATCTTGTTGTCTAGATTCGTAGTACTGGATATATTAAATCATGTTCTAAGTTACTATATACT
GAGATGAATAGAATAAGTAAAATTAGACCCACCTTAAGTCTTGATGAAGTTACTACTAGCTG

FIGURA 28 cont.

CGTTTGGGAGGACTTCCCAAAAAAAAAAGTATTAGCCATTAGCACGTGATTAATTAAGTACT
AGTTTAAAAAACTTAAAAAATAAATTAATATGATTCTCTTAAGTAACTCTCCTATAGAAAAC
TTTTACAAAATTACACCGTTTAATAGTTTGGAAAATATGTCAGTAAAAAATAAGAGAGTAGA
AGTTATGAAAGTTAGAAAAAGAATTGTTTTAGTAGTATACAGTTATAAACTATTCCCTCTGT
TCTAAAACATAAGGGATTATGGATGGATTCGACATGTACCAGTACCATGAATCGAATCCAGA
CAAGTTTTTTATGCATATTTATTCTACTATAATATATCACATCTGCTCTAAATATCTTATAT
TTCGAGGTGGAGACTGTCGCTATGTTTTTCTGCCCGTTGCTAAGCACACGCCACCCCCGATG
CGGGGACGCCTCTGGCCTTCTTGCCACGATAATTGAATGGAACTTCCACATTCAGATTCGAT
AGGTGACCGTCGACTCCAAGTGCTTTGCACAAAACAACTCCGGCCTCCCGGCCACCAGTCAC
ACGACTCACGGCACTACCACCCCTGACTCCCTGAGGCGGACCTGCCACTGTTCTGCATGCGA
AGCTATCTAAAATTCTGAAGCAAAGAAAGCACAGCACATGCTCCGGGACACGCGCCACCCGG
CGGAAAAGGGCTCGGTGTGGCGATCTCACAGCCGCATATCGCATTTCACAAGCCGCCCATCT
CCACCGGCTTCACGAGGCTCATCGCGGCACGACCGCGCACGGAACGCACGCGGCCGACCCGC
GCGCCTCGATGCGCGAGCCCATCCGCCGCGTCCTCCCTTTGCCTTTGCCGCTATCCTCTCGG
TCGTATCCCGTTTCTCTGTCTTTTGCTCCCCGGCGCGCGCCAGTTCGGAGTACCAGCGAAAC
CCGGACACCTGGTACACCTCCGCCGGCCACAACGCGTGTCCCCCCTACGTGGCCGCGCAGCA
CATGCCCATGCGCGACACGTGCACCTCCTCATCCAAACTCTCAAGTCTCAACGGTCCTATAA
ATGCACGGATAGCCTCAAGCTGCTCGTCACAAGGCAAGAGGCAAGAGGCAAGAGCATCCGTA
TTAACCAGCCTTTTGAGACTTGAGAGTGTGTGTGACTCGATCCAGCGTAGTTTCAGTTCGTG
TGTTGGTGAGTGATTCCAGCCAAGTTTGCG

SEQ ID NO: 260, Oryza sativa – DNA – promotor de expansina

AAAACCACCGAGGGACCTGATCTGCACCGGTTTTGATAGTTGAGGGACCCGTTGTGTCTGGT
TTTCCGATCGAGGGACGAAAATCGGATTCGGTGTAAAGTTAAGGGACCTCAGATGAACTTAT
TCCGGAGCATGATTGGGAAGGGAGGACATAAGGCCCATGTCGCATGTGTTTGGACGGTCCAG
ATCTCCAGATCACTCAGCAGGATCGGCCGCGTTCGCGTAGCACCCGCGGTTTGATTCGGCTT
CCCGCAAGGCGGCGGCCGGTGGCCGTGCCGCCGTAGCTTCCGCCGGAAGCGAGCACGCCGCC
GCCGCCGACCCGGCTCTGCGTTTGCACCGCCTTGCACGCGATACATCGGGATAGATAGCTAC
TACTCTCTCCGTTTCACAATGTAAATCATTCTACTATTTTCCACATTCATATTGATGTTAAT
GAATATAGACATATATATCTATTTAGATTCATTAACATCAATATGAATGTAGGAAATGCTAG
AATGACTTACATTGTGAATTGTGAAATGGACGAAGTACCTACGATGGATGGATGCAGGATCA
TGAAAGAATTAATGCAAGATCGTATCTGCCGCATGCAAAATCTTACTAATTGCGCTGCATAT
ATGCATGACAGCCTGCATGCGGGCGTGTAAGCGTGTTCATCCATTAGGAAGTAACCTTGTCA
TTACTTATACCAGTACTACATACTATATAGTATTGATTTCATGAGCAAATCTACAAAACTGG
AAAGCAATAAGAAATACGGGACTGGAAAAGACTCAACATTAATCACCAAATATTTCGCCTTC
TCCAGCAGAATATATATCTCTCCATCTTGATCACTGTACACACTGACAGTGTACGCATAAAC
GCAGCAGCCAGCTTAACTGTCGTCTCACCGTCGCACACTGGCCTTCCATCTCAGGCTAGCTT
TCTCAGCCACCCATCGTACATGTCAACTCGGCGCGCGCACAGGCACAAATTACGTACAAAAC
GCATGACCAAATCAAAACCACCGGAGAAGAATCGCTCCCGCGCGCGGCGGCGACGCGCACGT
ACGAACGCACGCACGCACGCCCAACCCCACGACACGATCGCGCGCGACGCCGGCGACACCGG
CCGTCCACCCGCGCCCTCACCTCGCCGACTATAAATACGTAGGCATCTGCTTGATCTTGTCA
TCCATCTCACCACCAAAAAAAAAAGGAAAAAAAAACAAAACACACCAAGCCAAATAAAAGCG
ACA

FIGURA 28 cont.



FIGURA 29

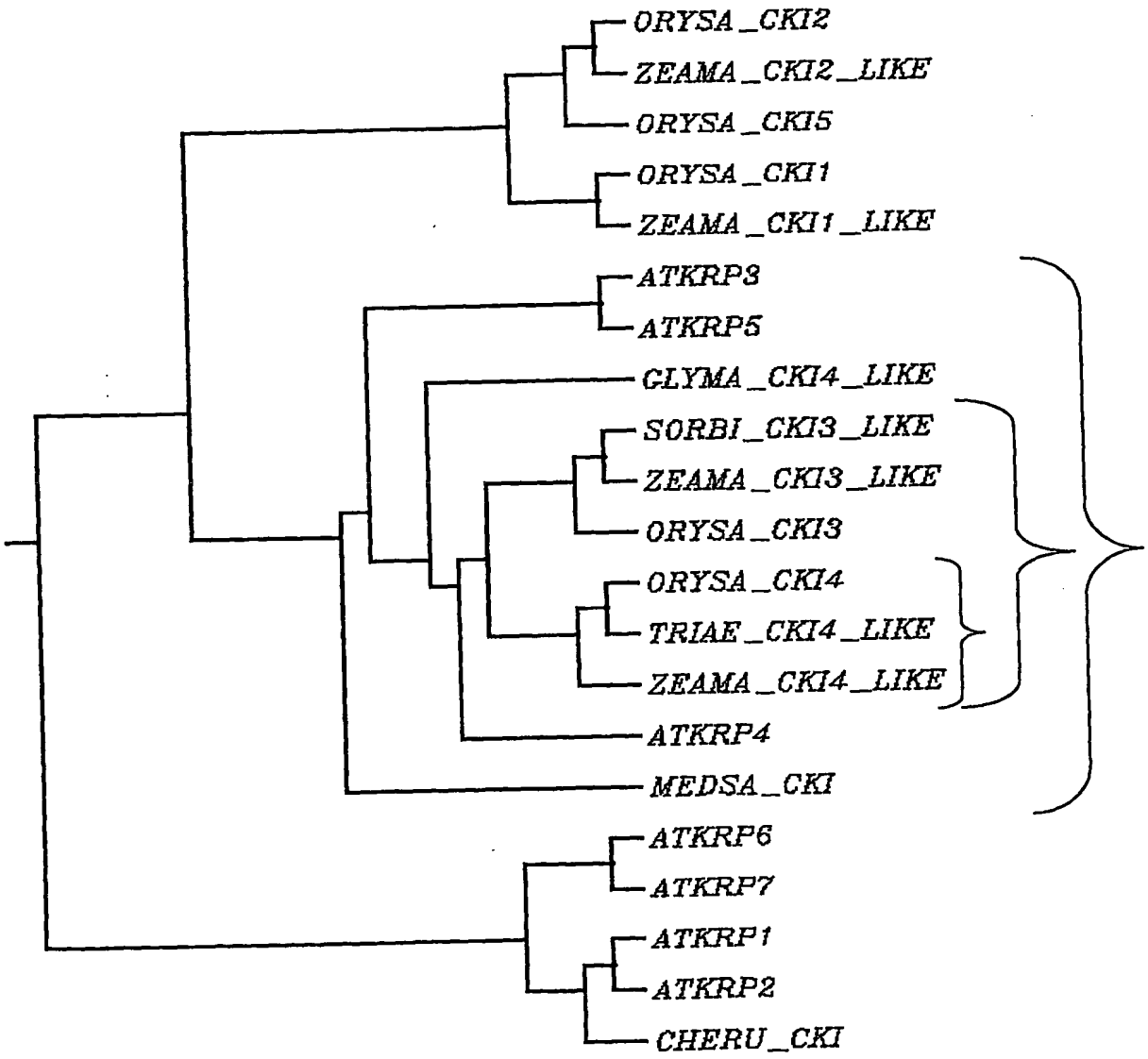

ORYSA_CKI2
ZEAMA_CKI2_LIKE
ORYSA_CKI5
ORYSA_CKI1
ZEAMA_CKI1_LIKE
ATKRP3
ATKRP5
GLYMA_CKI4_LIKE
SORBI_CKI3_LIKE
ZEAMA_CKI3_LIKE
ORYSA_CKI3
ORYSA_CKI4
TRIAE_CKI4_LIKE
ZEAMA_CKI4_LIKE
ATKRP4
MEDSA_CKI
ATKRP6
ATKRP7
ATKRP1
ATKRP2
CHERU_CKI

FIGURA 30

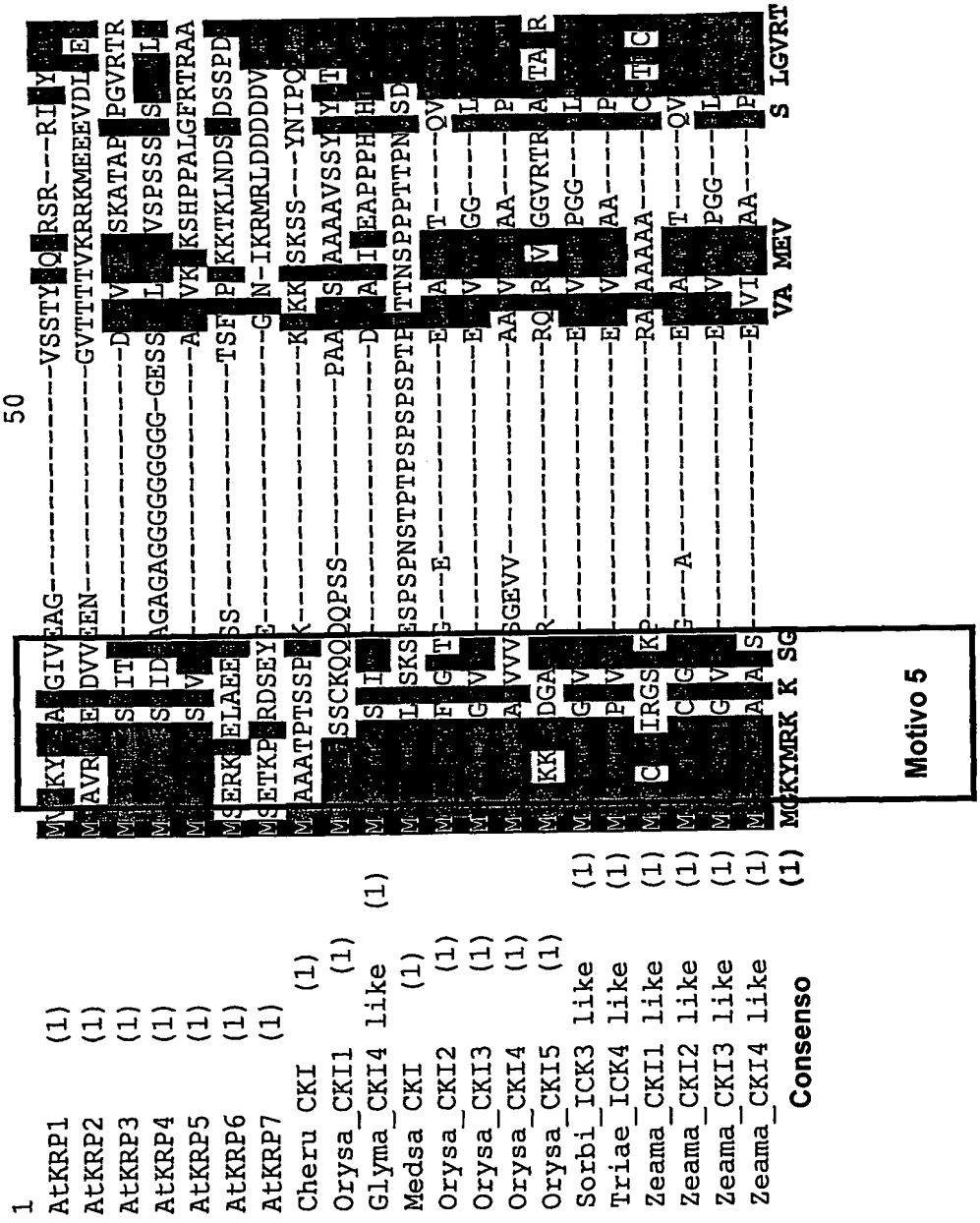

1
50
AtKRP1 (1)
AtKRP2 (1)
AtKRP3 (1)
AtKRP4 (1)
AtKRP5 (1)
AtKRP6 (1)
AtKRP7 (1)
Cheru CKI (1)
Orysa_CKI1 (1)
Glyma_CKI4 like (1)
Medsa CKI (1)
Orysa_CKI2 (1)
Orysa_CKI3 (1)
Orysa_CKI4 (1)
Orysa_CKI5 (1)
Sorbi_ICK3 like (1)
Triae_ICK4 like (1)
Zeama_CKI1 like (1)
Zeama_CKI2 like (1)
Zeama_CKI3 like (1)
Zeama_CKI4 like (1)
Consenso
Motivo 5

FIGURA 31

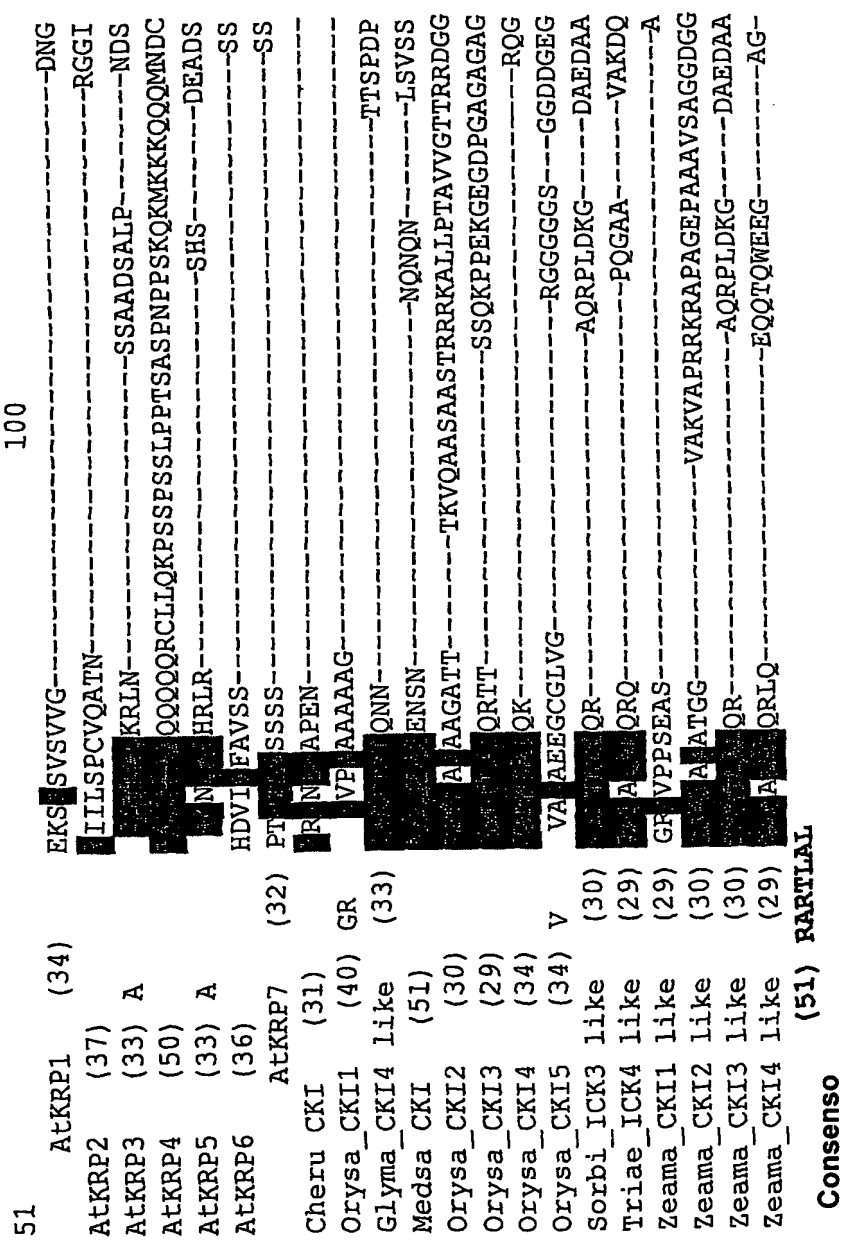


FIGURA 31 cont.

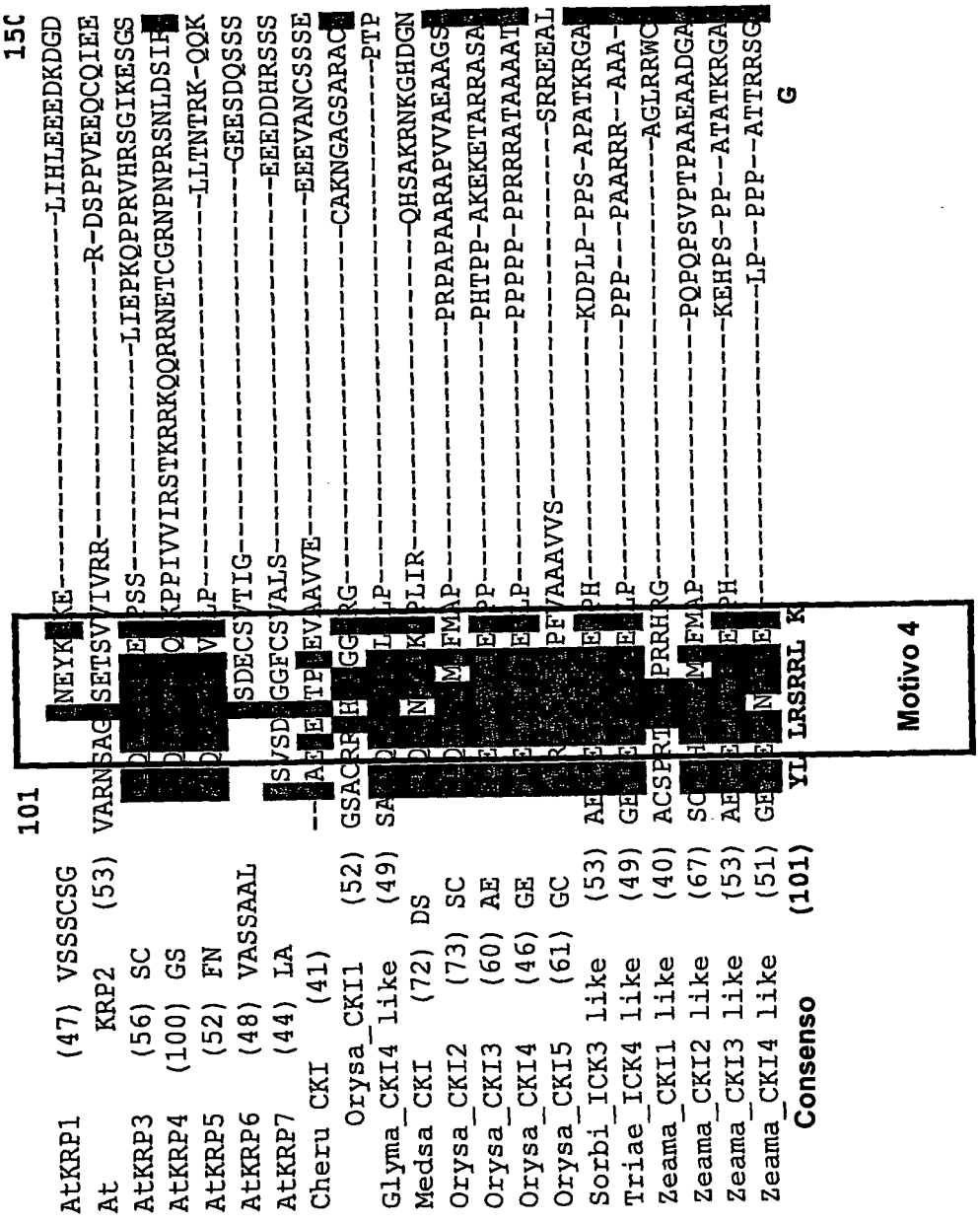


FIGURA 31 cont.

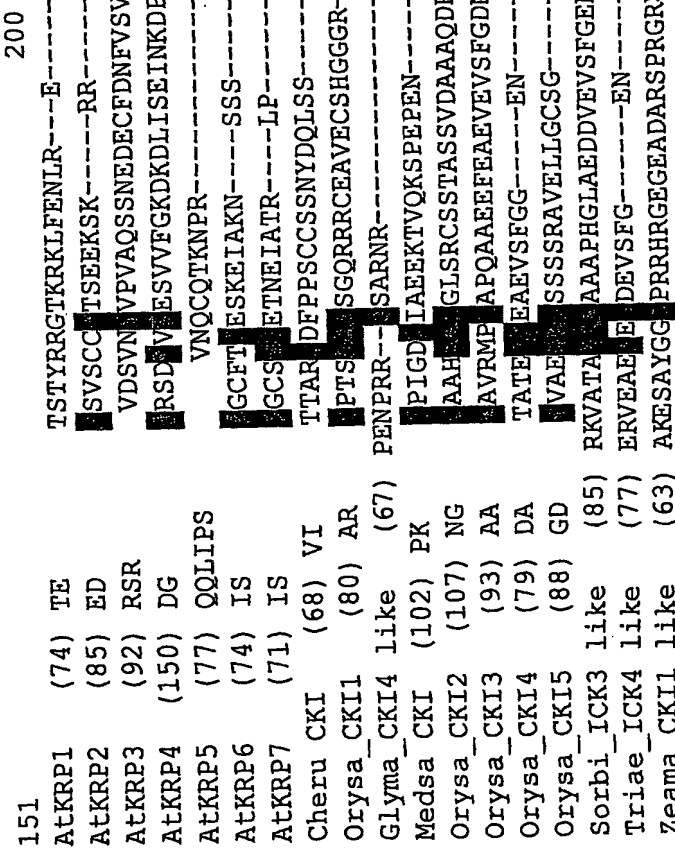


FIGURA 31 cont.

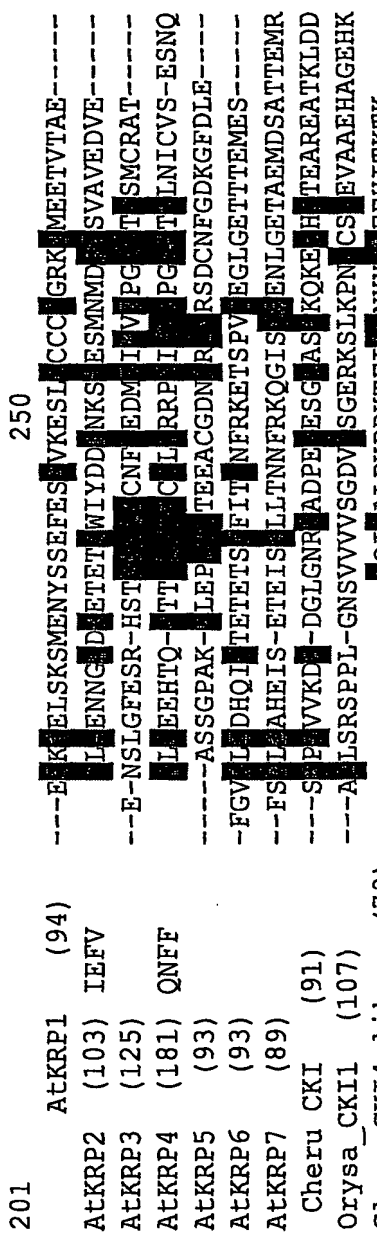


FIGURA 31 cont.

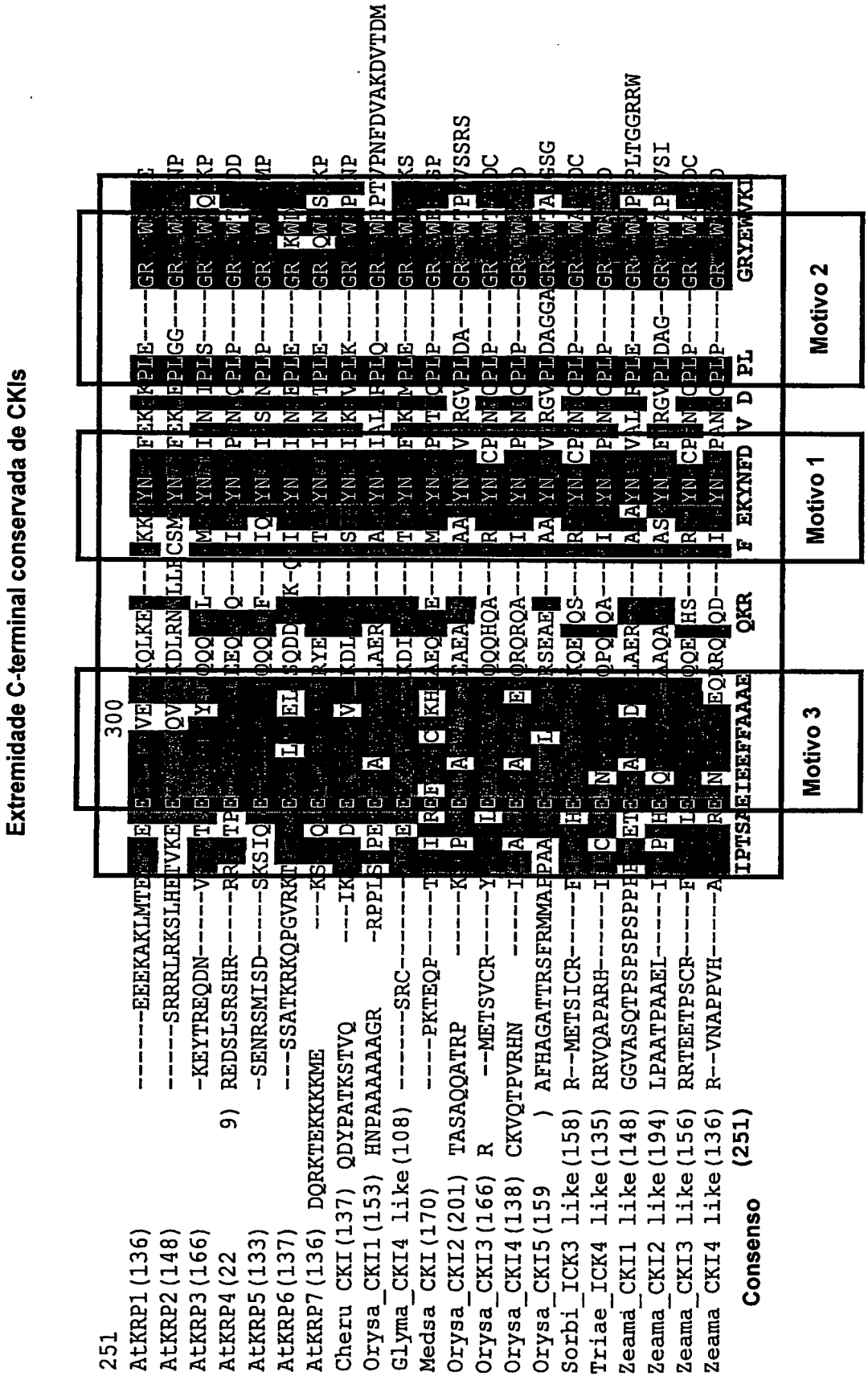


FIGURA 31 cont.

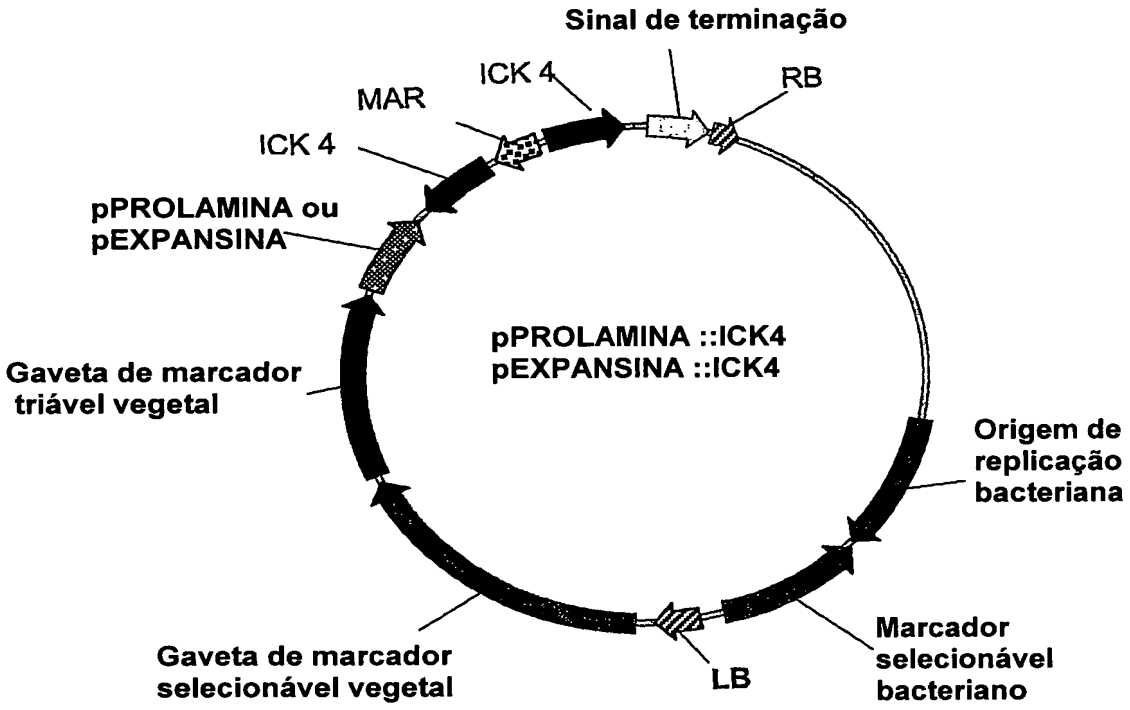

Sinal de terminação
ICK 4
RB
MAR
ICK 4
pPROLAMINA ou
pEXPANSINA
pPROLAMINA ::ICK4
pEXPANSINA ::ICK4
Gaveta de marcador
triável vegetal
Origem de
replicação
bacteriana
Marcador
selecionável
bacteriano
LB
Gaveta de marcador
selecionável vegetal

FIGURA 32

SEQ ID NO: 261, MOTIVO 1
FXXKYNFD **caracterizado pelo fato de que X é qualquer aminoácido**

SEQ ID NO: 262, MOTIVO 2
[P/L]LXGRYEW

caracterizado pelo fato de que X é qualquer aminoácido e [P/L] significa que ou uma prolina ou uma leucina aparece na posição indicada

SEQ ID NO: 263, MOTIVO 3
EXE[D/E]FFXXXE (SEQ ID NO: 3)

caracterizado pelo fato de que X é qualquer aminoácido e [D/E] significa que ou um aspartato ou um glutamato aparece na posição indicada

SEQ ID NO: 264, MOTIVO 4
YXQLRSRR

caracterizado pelo fato de que X é qualquer aminoácido

SEQ ID NO: 265, MOTIVO 5
MGKY[M/I][K/R]KX[K/R]

caracterizado pelo fato de que X é qualquer aminoácido e [M/I] significa que ou uma metionina ou uma isoleucina aparece na posição indicada, e [K/R] significa que ou uma lisina ou uma arginina aparece na posição indicada

SEQ ID NO: 266, MOTIVO 6
SXGVRTRA

caracterizado pelo fato de que X é qualquer aminoácido

SEQ ID NO: 267, cDNA de comprimento completo CKI4 de Oryza sativa AK073804.1
ATGGGCAAGTACATGCGCAAGGCCAAGGTGGTGGTCTCCGGCGAGGTGGTGGCCGCCGCCGT
CATGGAGCTCGCCGCGGCGCCGCTCGGGGTGCGCACCCGCGCCCGCTCCCTCGCGCTGCAGA
AGAGGCAGGGCGGGGAGTACCTCGAGCTCAGGAGCCGCAGGCTCGAGAAGCTCCCTCCTCCC
CCGCCGCCGCCGCCGAGGAGGAGGGCGACGGCTGCGGCTGCGACTGCTGATGCGACGGCGAC
GGAGAGCGCGGAGGCGGAGGTGTCGTTCGGGGGGGAGAACGTCCTCGAGCTGGAGGCCATGG
AAAGGAATACCAGGGAGACGACACCTTGCAGCTTGATCAGGGACCCCGATACGATTAGCACC
CCTGGATCTACCACAAGGCGCAGCCACTCGAGTTCTCATTGCAAGGTGCAAACACCCGTGCG
CCACAACATTATTCCAGCATCAGCAGAGCTGGAAGCGTTCTTCGCCGCCGAAGAGCAACGGC
AACGACAGGCTTTCATCGACAAGTATAACTTTGATCCTGTGAATGACTGCCCTCTTCCCGGC
CGATTTGAATGGGTCAAGCTAGACTGA

FIGURA 33

SEQ ID NO: 268, polipeptídeo deduzido CKI4 de Oryza sativa

MGKYMRKAKVVVSGEVVAAAVMELAAAPLGVRTRARSLALQKRQGGEYLELRSRRLEKLPPP
PPPPPRRRATAAAATADATATESAEAEVSFGGENVLELEAMERNTRETTPCSLIRDPDTIST
PGSTTRRSHSSSHCKVQTPVRHNIIPASAELEAFFAAEEQRQRQAFIDKYNFDPVNDCPLPG
RFEWVKLD

SEQ ID NO: 269, cDNA de comprimento completo tipo CKI4 de Zea mays AY 986792

ATGGGCAAGTACATGCGCAAGGCCAAGGCTTCCAGCGAGGTTGTCATCATGGATGTCGCCGC
CGCTCCGCTCGGAGTCCGCACCCGAGCGCGCGCCCTCGCGCTGCAGCGTCTGCAGGAGCAAC
AGACGCAGTGGGAAGAAGGTGCTGGCGGCGAGTACCTGGAGCTAAGGAACCGGAGGCTCGAG
AAGCTGCCGCCGCCGCCGGCGACCACTAGGAGGTCGGGCGGGAGGAAAGCGGCAGCCGAGGC
CGCCGCAACTAAGGAGGCTGAGGCGTCGTACGGGGAGAACATGCTCGAGTTGGAGGCCATGG
AGAGGATTACCAGGGAGACGACGCCTTGCAGCTTGATTAACACCCAGATGACTAGCACTCCT
GGGTCCACGAGATCCAGCCACTCTTGCCACCGCAGGGTGAACGCTCCTCCGGTGCACGCCGT
CCCAAGTTCTAGGGAGATGAATGAGTACTTCGCTGCCGAACAGCGACGGCAACAGCAGGATT
TCATTGACAAGTACAACTTCGATCCTGCAAACGACTGCCCTCTCCCAGGCAGGTTTGAGTGG
GTGAAGCTAGACTGA

SEQ ID NO: 270, polipeptídeo deduzido tipo CKI4 de Zea mays

MGKYMRKAKASSEVVIMDVAAAPLGVRTRARALALQRLQEQQTQWEEGAGGEYLELRNRRLE
KLPPPPATTRRSGGRKAAAEAAATKEAEASYGENMLELEAMERITRETTPCSLINTQMTSTP
GSTRSGHSCHRRVNAPPVHAVPSSREMNEYFAAEQRRQQQDFIDKYNFDPANDCPLPGRFEW
VKLD

SEQ ID NO: 271, cDNA de comprimento completo tipo CKI4 de Triticum aestivum agrupamento de bg 908519.1 e CA640135.1

ATGGGCAAGTACATGCGCAAGCCCAAGGTCTCCGGCGAGGTGGCCGTCATGGAGGTCGCCGC
CGCGCCGCTAGGGGTCCGCACCCGCGCACGAGCGCTCGCGATGCAGAGGCAGCCGCAGGGGG
CGGCGGTGGCCAAGGACCAGGGGGAGTACCTGGAGCTCAGGAGTCGGAAGCTCGAGAAGCTG
CCCCCGCCGCCGCCGGCGGCGAGGAGGAGGGCGGCCGCGGCGGAGCGTGTCGAGGCCGAGGC
CGAGGCCGACGAGGTGTCCTTCGGTGAGAACGTGCTCGAGTCGGAGGCCATGGGGAGGGGTA
CCAGGGAGACGACGCCCTGCAGCTTGATTAGGGACTCGGGAACGATAAGCACTCCTGGATCC
ACAACAAGACCGAGCCACTCGAATTCCCATCGCAGGGTGCAAGCTCCAGCGCGCCATATTAT
TCCATGTTCAGCAGAGATGAATGAGTTCTTCTCTGCTGCGGAGCAACCGCAACAGCAAGCCT
TCATTGACAAGTACAACTTTGATCCTGTGAACGACTGTCCTCTCCCAGGCCGATACGAGTGG
GTGAAGCTAGACTGATAATTCTCCAGGAAGGAGAGCACCATGTACTTCTCCGCTCCCTCCAC
CTTAGCGTCGTGGTAAAGGCCGCCCCGTCGTGTAGCTTTGTTTCCGTTGTAAAAAGAATAGT
TAGCTGTAGTAGCCTCAATGGCGTTACATACAGAGTAATGCTGATTACACCTAATCCTCAAA
CCATGTACGTT

SEQ ID NO: 272, polipeptídeo deduzido tipo CKI4 de Triticum aestivum

MGKYMRKPKVSGEVAVMEVAAAPLGVRTRARALAMQRQPQGAAVAKDQGEYLELRSRKLEKL
PPPPPAARRRAAAAERVEAEAEADEVSFGENVLESEAMGRGTRETTPCSLIRDSGTISTPGS
TTRPSHSNSHRRVQAPARHIIPCSAEMNEFFSAAEQPQQQAFIDKYNFDPVNDCPLPGRYEW
VKLD

FIGURA 33 CONT.

SEQ ID NO: 273, **cDNA de comprimento completo CKI3 de Oryza sativa AK064723.1**

ATGGGGAAGTACATGCGGAAGGGGAAGGTGTCGGGGGAGGTGGCGGTGATGGAGGTGGGCGG
GGCGCTGCTCGGCGTCCGCACCCGCTCCCGCACGCTCGCGCTGCAGCGGACGACCTCGTCGC
AGAAGCCGCCGGAGAAGGGGGAGGGGGACCCCGGTGCGGGCGCGGGCGCGGGGGCGGAGTAC
CTCGAGCTCAGGAGCCGGAGGCTCGAGAAGCCGCCTCCGCACACGCCGCCGGCCAAGGAGAA
GGAGACCGCCAGGAGGGCTTCCGCCGCCGCCGCCGCCGCCGCCGTGAGGATGCCGGCGGCGCCGC
AAGCGGCCGAGGAGTTCGAGGCGGAGGTCGAGGTGTCCTTCGGCGACAACGTTCTTGACCTC
GACGGCGACGCCATGGAGAGGAGTACCAGGGAGACAACGCCTTGCAGTTTAATTAGGAGCTC
AGAAATGATAAGCACCCCTGGCTCCACAACTAAAACCAACACCTCGATCAGTTCCCGGCGCA
GAATGGAGACCTCTGTTTGTCGTTACGTTCCGAGTTCTCTTGAGATGGAAGAGTTCTTTGCA
GCTGCTGAACAACAGCAACATCAGGCTTTCAGAGAGAGGTATAACTTCTGTCCTGTGAACGA
CTGCCCACTTCCTGGACGGTACGAATGGACAAGGCTAGACTGCTAG

SEQ ID NO: 274, **polipeptídeo deduzido CKI3 de Oryza sativa**

MGKYMRKGKVSGEVAVMEVGGALLGVRTRSRTLALQRTTSSQKPPEKGEGDPGAGAGAGAEY
LELRSRRLEKPPPHTPPAKEKETARRASAAAAAAVRMPAAPQAAEEFEAEVEVSFGDNVLDL
DGDAMERSTRETTPCSLIRSSEMISTPGSTTKTNTSISSRRRMETSVCRYVPSSLEMEEFFA
AAEQQQHQAFRERYNFCPVNDCPLPGRYEWTRLDC

SEQ ID NO: 275, **cDNA de comprimento completo tipo CKI3 de Zea mays DV174570.1**

ATGGGGAAGTACATGCGCAAGGGCAAGGTGTCCGGGGAGGTCGCCGTCATGGAGGTACCCGG
CGGCGCGCTGCTCGGCGTCCGCACCCGCTCCCGCACGCTCGCGCTGCAGCGCGCGCAGAGGC
CGCTCGACAAGGGCGACGCGGAGGACGCCGCCGCGGAGTACCTCGAGCTCAGGAGCCGGAGG
CTCGAGAAGCCGCACAAGGAGCATCCGTCGCCGCCCGCGACCGCGACCAAGAGGGGCGCCGG
GAGGAAGGCCGCCGCCGCCGCCGCGGTGCAGCACGTGCTGATGCAGGACGAGGTCGAGGTCG
AGGTCTCGTTCGGGGACAACGTGCTTGACTTGGACACCATGGAAAGGAGTACCAGAGAGACA
ACACCGTGCAGCCTGATTAGGAACCCAGAGATGATAAGCACCCCAGGATCCACAACTAAAAG
CAAAACCAGCAGCAACTCGACGACTTCCCGCCGCAGAACGGAGGAAACCCCGAGCTGCCGGT
TCATACCGAGCTCGCTCGAGATGGAGGAGTTCTTCTCGGCGGCCGAGCAACAGGAGCAGCAT
AGCTTCAGGGAGAAGTACAACTTCTGTCCCGTGAACGACTGTCCTCTCCCTGGCCGGTACGA
ATGGGCGAGGCTAGACTGCTAG

SEQ ID NO: 276, **polipeptídeo deduzido tipo CKI3 de Zea mays**

MGKYMRKGKVSGEVAVMEVPGGALLGVRTRSRTLALQRAQRPLDKGDAEDAAAEYLELRSRR
LEKPHKEHPSPPATATKRGAGRKAAAAAAVQHVLMQDEVEVEVSFGDNVLDLDTMERSTRET
TPCSLIRNPEMISTPGSTTKSKTSSNSTTSRRRTEETPSCRFIPSSLEMEEFFSAAEQQEQH
SFREKYNFCPVNDCPLPGRYEWARLDC

SEQ ID NO: 277, **cDNA de comprimento completo tipo CKI3 de Sorghum bicolor agrupamento de CN152732.1 E CD224882.1 DE CN152732**

ATGGGGAAGTACATGCGCAAGGGCAAGGTGTCCGGGGAGGTCGCCGTCATGGAGGTCCCCGG
CGGCGCGCTGCTCGGCGTCCGCACCCGCTCCCGCACGCTCGCGCTGCAGCGCGCGCAGAGGC
CGCTCGACAAGGGCGACGCCGAGGACGCCGCTGCGGAGTACCTCGAGCTCAGGAGCCGGAGG
CTCGAGAAGCCGCACAAGGACCCGTTACCGCCGCCGTCTGCGCCTGCGACCAAGAGGGGCGC
CGGGAGGAAGGTCGCCACCGCCGCCGCCGCCGCCGCGGCGCCGCACGGGCTGGCGGAGGACG
ACGTCGAGGTCTCCTTCGGCGAGAACGTGCTCGACTTCGACGCCATGGAAAGGAGTACCAGA

FIGURA 33 CONT.

GAGACAACACCGTGCAGTTTGATTAGGAACCCAGAGATGATAAGCACCCCAGGATCCACAAC
TAAGAGTAAAACCAGCAACTCGATGACCTCCCGTCGCAGAATGGAAACCTCAATCTGCCGTT
TCATACCAAGTTCGCATGAGATGGAAGAGTTCTTCTCAGCAGCTGAAAAACAGGAGCAGCAA
AGCTTCAGGGAGAAGTATAACTTCTGTCCTGTGAACGACTGTCCTCTTCCGGGTCGGTATGA
ATGGGCGAGGCTAGACTGCTAG

SEQ ID NO: 278, polipeptídeo deduzido tipo CKI3 de Sorghum bicolor
MGKYMRKGKVSGEVAVMEVPGGALLGVRTRSRTLALQRAQRPLDKGDAEDAAAEYLELRSRR
LEKPHKDPLPPPSAPATKRGAGRKVATAAAAAAAPHGLAEDDVEVSFGENVLDFDAMERSTR
ETTPCSLIRNPEMISTPGSTTKSKTSNSMTSRRRMETSICRFIPSSHEMEEFFSAAEKQEQQ
SFREKYNFCPVNDCPLPGRYEWARLDC

SEQ ID NO: 279, cDNA parcial tipo CKI4 de Saccharum officinarum
CO373621.1
GTCGACCCACGCGTCCGGTACTTCGCTGCTGAACAGCGACGCCAACAACAGGCTTTCATTGA
CAAGTACAACTTTGATCCTGTAAATGACTGCCCTCTCCCAGGCAGGTTTGAATGGGTGAAGC
TAGACTGATTGATTCAGAGGACAAGAGAGCAGCAGCATGGAACTCACCTCCGCTCCCTCCAC
CACCGCAGTGTTGTGGCAGAGGCGCATACCGTCGTGTTAGCTTTGTTTCTGTTGTAAAAACT
TAGTGTTAGCCTGTAGCCTTAATTGTCGTGTGTTACAGTACAAAACTGATGCTGAGTTACAA
CACCCTGATCTGATCTGATCCCTCAACTCAGNNGTAACCCTTAACAGCTTATTCTGTAAGGA
ACCTTACCCACCCTTGTTACCAGTT

SEQ ID NO: 280,
polipeptídeo deduzido parcial tipo CKI4 de Saccharum officinarum
YFAAEQRRQQQAFIDKYNFDPVNDCPLPGRFEWVKLD

SEQ ID NO: 281, promotor de prolamina RP6 de Oryza sativa
CTTCTACATCGGCTTAGGTGTAGCAACACGACTTTATTATTATTATTATTATTATTATT
ATTTTACAAAAATATAAAATAGATCAGTCCCTCACCACAAGTAGAGCAAGTTGGTGAGTTAT
TGTAAAGTTCTACAAAGCTAATTTAAAAGTTATTGCATTAACTTATTTCATATTACAAACAA
GAGTGTCAATGGAACAATGAAAACCATATGACATACTATAATTTTGTTTTTATTATTGAAAT
TATATAATTCAAAGAGAATAAATCCACATAGCCGTAAAGTTCTACATGTGGTGCATTACCAA
AATATATATAGCTTACAAAACATGACAAGCTTAGTTTGAAAAATTGCAATCCTTATCACATT
GACACATAAAGTGAGTGATGAGTCATAATATTATTTTCTTTGCTACCCATCATGTATATATG
ATAGCCACAAAGTTACTTTGATGATGATATCAAAGAACATTTTTAGGTGCACCTAACAGAAT
ATCCAAATAATATGACTCACTTAGATCATAATAGAGCATCAAGTAAAACTAACACTCTAAAG
CAACCGATGGGAAAGCATCTATAAATAGACAAGCACAATGAAAATCCTCATCATCCTTCACC
ACAATTCAAATATTATAGTTGAAGCATAGTAGTA

SEQ ID NO: 282, promotor de Beta-expansina de Oryza sativa
AAAACCACCGAGGGACCTGATCTGCACCGGTTTTGATAGTTGAGGGACCCGTTGTGTCTGGT
TTTCCGATCGAGGGACGAAAATCGGATTCGGTGTAAAGTTAAGGGACCTCAGATGAACTTAT
TCCGGAGCATGATTGGGAAGGGAGGACATAAGGCCCATGTCGCATGTGTTTGGACGGTCCAG
ATCTCCAGATCACTCAGCAGGATCGGCCGCGTTCGCGTAGCACCCGCGGTTTGATTCGGCTT
CCCGCAAGGCGGCGGCCGGTGGCCGTGCCGCCGTAGCTTCCGCCGGAAGCGAGCACGCCGCC
GCCGCCGACCCGGCTCTGCGTTTGCACCGCCTTGCACGCGATACATCGGGATAGATAGCTAC
TACTCTCTCCGTTTCACAATGTAAATCATTCTACTATTTTCCACATTCATATTGATGTTAAT

FIGURA 33 CONT.

GAATATAGACATATATATCTATTTAGATTCATTAACATCAATATGAATGTAGGAAATGCTAG
AATGACTTACATTGTGAATTGTGAAATGGACGAAGTACCTACGATGGATGGATGCAGGATCA
TGAAAGAATTAATGCAAGATCGTATCTGCCGCATGCAAAATCTTACTAATTGCGCTGCATAT
ATGCATGACAGCCTGCATGCGGGCGTGTAAGCGTGTTCATCCATTAGGAAGTAACCTTGTCA
TTACTTATACCAGTACTACATACTATATAGTATTGATTTCATGAGCAAATCTACAAAACTGG
AAAGCAATAAGAAATACGGGACTGGAAAAGACTCAACATTAATCACCAAATATTTCGCCTTC
TCCAGCAGAATATATATCTCTCCATCTTGATCACTGTACACACTGACAGTGTACGCATAAAC
GCAGCAGCCAGCTTAACTGTCGTCTCACCGTCGCACACTGGCCTTCCATCTCAGGCTAGCTT
TCTCAGCCACCCATCGTACATGTCAACTCGGCGCGCGCACAGGCACAAATTACGTACAAAAC
GCATGACCAAATCAAAACCACCGGAGAAGAATCGCTCCCGCGCGCGGCGGCGACGCGCACGT
ACGAACGCACGCACGCACGCCCAACCCCACGACACGATCGCGCGCGACGCCGGCGACACCGG
CCGTCCACCCGCGCCCTCACCTCGCCGACTATAAATACGTAGGCATCTGCTTGATCTTGTCA
TCCATCTCACCACCAAAAAAAAAAGGAAAAAAAAACAAAACACACCAAGCCAAATAAAAGCG
ACA

SEQ ID NO: 283, região de acoplamento de matriz/seqüência de região de acoplamento de armação de Nicotiana tabacum

CTGCCATTCTTTAGAGGGGATGCTTGTTTAAGAACAAAAAATATATCACTTTCTTTTGTTCC
AAGTCATTGCGTATTTTTTTAAAAATATTTGTTCCTTCGTATATTTCGAGCTTCAATCACTT
TATGGTTCTTTGTATTCTGGCTTTGCTGTAAATCGTAGCTAACCTTCTTCCTAGCAGAAATT
ATTAATACTTGGGATATTTTTTTAGAATCAAGTAAATTACATATTACCACCACATCGAGCTG
CTTTTAAATTCATATTACAGCCATATAGGCTTGATTCATTTTGCAAAATTTCCAGGATATTG
ACAAC

SEQ ID NO: 284, prm00472 iniciador direto
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACAATGGGCAAGTACATGCGCA

SEQ ID NO: 285, prm00473 iniciador reverso
GGGGACCACTTTGTACAAGAAAGCTGGGTGGAGCAGAGAGGTCCATGGTGCCC

FIGURA 33 CONT.

PI0619242-4

RESUMO

"PROTEÍNA, MÉTODOS PARA AUMENTAR PRODUÇÃO DE SEMENTE E/OU AUMENTAR TAXA DE CRESCIMENTO DE PLANTAS E TAXA DE CRESCIMENTO, CÉLULA VEGETAL, CONSTRUÇÃO, E, MÉTODOS PARA PRODUZIR UMA PLANTA TRANSGÊNICA, PARA MELHORAR CARACTERÍSTICAS DE CRESCIMENTO DE PLANTAS E CARACTERÍSTICAS DE CRESCIMENTO EM RELAÇÃO A CORRESPONDENTES PLANTAS DO TIPO SELVAGEM, PARA AUMENTAR PRODUÇÃO VEGETAL, NÚMERO DE SEMENTES EM PLANTAS E PRODUÇÃO DE SEMENTE EM PLANTAS"

A presente invenção geralmente se refere ao campo de biologia molecular e se relaciona a um método para melhorar várias características de crescimento vegetal modulando expressão em uma planta de um ácido nucleico codificando uma GRP (Proteína Relacionada a Crescimento). A presente invenção também se relaciona a plantas tendo expressão modulada de um ácido nucleico codificando uma GRP, cujas plantas têm características de crescimento melhoradas em relação a correspondentes plantas do tipo selvagem ou outras plantas de controle. A invenção também fornece construções úteis nos métodos da invenção. A GRP pode ser uma das seguintes: Regulador de Produção de Sementes (SYR), FG-GAP, CYP90B, CDC27, fatores de transcrição AT-hook, fatores de transcrição DOF e Inibidores de Quinase Dependentes de Ciclina (CKIs).