

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

(21) **PI 0712304-3 A2**

(22) Data de Depósito: 31/05/2007
(43) Data da Publicação: 17/01/2012
(RPI 2141)



(51) *Int.Cl.:*
C12N 9/12
C12N 15/82

(54) Título: MÉTODO E PROCESSO PARA MELHORAR CARACTERÍSTICAS DE CRESCIMENTO DE PLANTA EM RELAÇÃO ÀS PLANTAS DE CONTROLE CORRESPONDENTES, PLANTA, PARTE DE PLANTA OU CÉLULA DE PLANTA, MOLÉCULA DE ÁCIDO NUCLEICO ISOLADA, CONSTRUTO DE GENE, VETOR, PLANTA TRANSGÊNICA, E, USO DE UMA MOLÉCULA DE ÁCIDO NUCLEICO OU DO CONSTRUTO DE GENE OU DO VETOR

(30) Prioridade Unionista: 08/06/2006 EP 06115168.4

(73) Titular(es): Basf Plant Science GMBH

(72) Inventor(es): Christian Dammann, Stefan Henkes

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

(86) Pedido Internacional: PCT EP2007055336 de 31/05/2007

(87) Publicação Internacional: WO 2007/141189 de 13/12/2007

(57) Resumo: MÉTODO E PROCESSO PARA MELHORAR CARACTERÍSTICAS DE CRESCIMENTO DE PLANTA EM RELAÇÃO ÀS PLANTAS DE CONTROLE CORRESPONDENTES, PLANTA, PARTE DE PLANTA OU CÉLULA DE PLANTA, MOLÉCULA DE ÁCIDO NUCLEICO ISOLADA, CONSTRUTO DE GENE, VETOR, PLANTA TRANSGÊNICA, E, USO DE UMA MOLÉCULA DE ÁCIDO NUCLEICO OU DO CONSTRUTO DE GENE OU DO VETOR A presente invenção refere-se em geral ao campo de biologia molecular e concerne a um método para melhorar as características de crescimento de planta. Mais especificamente, a presente invenção concerne a um método para melhorar as características de crescimento de planta pela modulação da expressão em uma planta de um ácido nucleico de planta codificador de uma quinase dependente de ciclina (CDK) e/ou pela modulação da atividade em uma planta de uma proteína CDK de planta, cuja proteína CDK compreende motivos diferentes ou cujo ácido nucleico CDK codifica tal proteína. A presente invenção também concerne às plantas tendo expressão modulada de um ácido nucleico CDK de planta e/ou atividade modulada de uma proteína CDK de planta, cuja proteína CDK compreende diferentes motivos de sequência ou cujo ácido nucleico codifica tal proteína e cujas plantas têm características de crescimento melhoradas em relação às plantas de tipo selvagem correspondentes. A invenção adicionalmente se refere às sequências de ácido nucleico específicas codificadoras das proteínas anteriormente mencionadas tendo atividade melhorada de crescimento de planta, aos construtos de ácido nucleico, aos vetores e às plantas contendo ditas sequências de ácido nucleico.



“MÉTODO E PROCESSO PARA MELHORAR CARACTERÍSTICAS DE CRESCIMENTO DE PLANTA EM RELAÇÃO ÀS PLANTAS DE CONTROLE CORRESPONDENTES, PLANTA, PARTE DE PLANTA OU CÉLULA DE PLANTA, MOLÉCULA DE ÁCIDO NUCLEICO ISOLADA, CONSTRUTO DE GENE, VETOR, PLANTA TRANSGÊNICA, E, USO DE UMA MOLÉCULA DE ÁCIDO NUCLEICO OU DO CONSTRUTO DE GENE OU DO VETOR”

A presente invenção refere-se em geral ao campo da biologia molecular e concerne a um método para melhorar as características de crescimento de planta. Mais especificamente, a presente invenção concerne a um método para melhorar as características de crescimento de planta pela modulação de expressão em uma planta de um ácido nucleico de planta codificador de uma quinase dependente de ciclina (CDK) e/ou pela modulação da atividade em uma planta de uma proteína CDK de planta, cuja proteína CDK compreende motivos diferentes ou cujo ácido nucleico CDK codifica tal proteína. A presente invenção também concerne às plantas tendo expressão modulada de um ácido nucleico CDK de planta e/ou atividade modulada de uma proteína CDK de planta, cuja proteína CDK compreende motivos de sequência diferentes ou cujo ácido nucleico codifica tal proteína e cujas plantas possuem características de crescimento melhoradas em relação às plantas de tipo selvagem correspondentes.

A invenção adicionalmente se refere às seqüências de ácido nucleico específicas codificadoras das proteínas anteriormente mencionadas tendo a atividade melhoradora de crescimento de planta acima citada, aos construtos de ácido nucleico, aos vetores e às plantas contendo as ditas seqüências de ácido nucleico.

A população mundial sempre crescente e o fornecimento decrescente de terra arável disponível para agricultura estimulam a pesquisa na direção do melhoramento da eficiência de agricultura. Meios

convencionais para melhoramentos de colheita e horticulturais utilizam técnicas de geração seletivas para identificar plantas tendo características desejáveis. Contudo, tais técnicas de geração seletivas têm várias desvantagens, isto é estas técnicas são tipicamente laboriosamente intensas e resultam em plantas que muitas vezes contêm componentes genéticos heterogêneos que nem sempre podem resultar na feição desejável sendo transferidas das plantas parentais. Avanços em biologia molecular têm permitido que a humanidade modifique o plasma germinativo de animais e plantas. Engenharia genética de plantas exige o isolamento e a manipulação de material genético (tipicamente na forma de DNA ou RNA) e a subsequente introdução daquele material genético em uma planta. Tal tecnologia tem a capacidade para dar colheitas ou plantas tendo várias feições econômicas, agronômicas ou horticulturais melhoradas. Uma feição de um interesse econômico particular é o rendimento. O rendimento é normalmente definido como o produto mensurável de valor econômico de uma colheita. Isto pode ser definido em termos de quantidade e/ou qualidade. Rendimento de colheita é influenciado pelos estresses atípicos aos quais as plantas ou colheitas estão sujeitas. Tais estresses incluem estresses ambientais (abióticos) (tais como estresses de temperatura causados por temperaturas altas ou baixas atípicas; estresses causados por deficiência de nutriente; estresses causados pela falta de água (estiagem)) e estresses bióticos (que podem ser infligidos às plantas por outras plantas (ervas daninhas), pestes animais e patógenos). Rendimento de colheita pode não apenas ser aumentado pelo combate de um ou mais dos estresses aos quais a colheita ou planta é submetida, mas também pode ser aumentado pela modificação dos mecanismos de crescimento inerentes de uma planta. Biomassa de planta é o rendimento para colheitas de forragem como alfafa, feno e milho de silagem. Muitos substitutos para rendimento têm sido usados em colheitas de grãos. Mais importantes dentre estes são as estimativas do tamanho de planta. O tamanho de planta pode ser medido em

muitas maneiras dependendo da espécie e do estágio desenvolvvente, mas incluem peso seco de planta total, peso seco acima do solo, peso fresco acima do solo, área foliar, volume de caule, altura de planta, diâmetro de invólucro de flor composta, comprimento foliar, comprimento de raiz, massa de raiz, número de brotos e número de folhas. Muitas espécies mantêm uma razão conservativa entre o tamanho de partes diferentes da planta em um dado estágio desenvolvvente. Estas relações alométricas são usadas para extrapolar de uma destas medições de tamanho para outra (e.g. Tittone et al 2005 *Agric Ecosys & Environ* 105: 213). Tamanho de planta em um estágio desenvolvvente inicial tipicamente se correlacionará com o tamanho de planta mais tarde no desenvolvimento. Uma planta maior com uma área foliar maior pode tipicamente absorver mais luz e dióxido de carbono do que uma planta menor e portanto provavelmente ganhará um peso maior durante o mesmo período (Fasoula & Tollenaar 2005 *Maydica* 50:39). Isto está além da continuação potencial da vantagem genética ou microambiental que a planta tinha para alcançar o tamanho maior inicialmente. Há um componente genético forte para a taxa de crescimento e o tamanho da planta (e.g. ver Steege et al 2005 *Plant Physiology* 139:1078), e assim para uma variedade de genótipos diversos tamanho de planta sob condição ambiental provavelmente está correlacionado com o tamanho sob outra (Hittalmani et al 2003 *Theoretical Applied Genetics* 107:679). Nesta maneira um ambiente padrão é usado como um substituto dos ambientes diversos e dinâmicos encontrados em localizações e tempos diferentes pelas colheitas no campo.

Índice de colheita, a razão de rendimento de sementes para peso seco acima do solo, é relativamente estável sob muitas condições ambientais e assim pode ser muitas vezes obtida uma correlação robusta entre tamanho de planta e rendimento de grãos (e.g. Rebetzke et al 2002 *Crop Science* 42:739). Estes processos estão intrinsecamente ligados porque a maioria da biomassa de grão é dependente da produtividade fotossintética

armazenada ou corrente pelas folhas e caule da planta (Gardener et al 1985 Physiology of Crop Plants. Iowa State University Press, pp68-73) Portanto, seleção do tamanho de planta, mesmo em estágios de desenvolvimentos iniciais, tem sido usada como um indicador para rendimento potencial futuro (e.g. Tiftonell et al 2005 Agric Ecosys & Environ 105: 213). Quando se testa o impacto de diferenças genéticas sobre tolerância ao estresse, a capacidade para padronizar propriedades de solo, temperatura, disponibilidade de água e nutriente e intensidade de luz é uma vantagem intrínseca de ambientes de câmara de crescimento de planta ou estufa comparado com o campo.

Contudo, limitações artificiais sobre rendimento devido à polinização insatisfatória por causa da ausência de vento ou insetos, ou espaço insuficiente para maturação de raiz ou crescimento de copa, podem restringir o uso destes ambientes controlados para testar diferenças de rendimento. Por isso, medições de tamanho de planta em desenvolvimento inicial, sob condições padronizadas em uma câmara de crescimento ou estufa, são práticas padrão para proporcionar indicação de vantagens de rendimento genético potencial.

Os mecanismos de crescimento inerentes de uma planta residem em uma seqüência elevadamente ordenada de eventos coletivamente conhecida como o 'ciclo celular'. A capacidade para influenciar o ciclo celular em uma planta (quer usando tecnologia de DNA recombinante quer usando meio não-recombinante), e para deste modo modificar as características de crescimento de uma planta, teria muitas aplicações em áreas tais como aumento de colheita, geração de planta, produção de plantas ornamentais, arboricultura, horticultura, silvicultura, a produção de algas ou plantas (por exemplo para uso como biorreatores, para a produção de substâncias tais como fármacos, anticorpos, ou vacinas, ou para a bioconversão de resíduo orgânico ou para uso como combustível no caso de plantas e algas de rendimento alto).

Avanço através do ciclo celular é fundamental para o crescimento e o desenvolvimento de organismos multicelulares e é crucial para proliferação de células. Os componentes maiores do ciclo celular estão elevadamente conservados em levedura, mamíferos, e plantas. O ciclo celular é tipicamente dividido nas seguintes fases sequenciais: G0-G1-S-G2-M. Síntese ou replicação 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 intervalo (*gap*) interpostas, G1 (durante a qual as células crescem antes da replicação de DNA) e G2 (um período após a replicação de DNA durante o qual a célula se prepara para divisão). Divisão celular é completada após a citocinese, a última etapa da fase M. Células que têm saído do ciclo celular e que têm se tornado quiescentes são ditas em estarem na fase G0. Células nesta fase podem ser estimuladas para reentrarem no ciclo celular na fase G1. O “G” em G1, G2 e G0 representa “*gap*” (intervalo). Completitude do processo de ciclo celular permite que cada célula filha durante a divisão celular receba uma cópia completa do genoma parental.

Divisão celular é controlada por dois eventos de ciclo celular principais, a saber iniciação de síntese de DNA e iniciação de mitose. Cada transição para cada um destes eventos chave é controlada por um ponto de checagem representado por complexos de proteína específicos (envolvidos em replicação de DNA e divisão). A expressão de genes necessários para síntese de DNA no limite G1/S é regulada pela família E2F de fatores de transcrição em células de planta e mamíferos (WO 96/25494; Muller et al., Genes and Development 15, 267-285, 2001; De Veylder et al., EMBO J. 21, 13602-1368, 2002). Entrada no ciclo celular é regulada/ativada por um complexo E2F/Rb que integra sinais e permite ativação de transcrição de genes de ciclo celular. A transição entre as fases diferentes do ciclo celular, e portanto o avanço através do ciclo celular, é conduzida pela formação e ativação de

serina/treonina proteína quinases heterodiméricas diferentes, geralmente chamadas de quinases dependentes de ciclina (CDK). Um pré-requisito para atividade destas quinases é a associação física com uma ciclina específica, a regulação do tempo de ativação sendo grandemente dependente da expressão de ciclina. Ligação de ciclina induz mudanças conformacionais no lobo N-terminal da CDK associada e contribui para a localização e a especificidade de substrato do complexo. CDKs monoméricas são ativadas quando estão associadas com ciclinas e assim têm atividade de quinase. Níveis de proteína ciclina flutuam no ciclo celular e portanto representam um fator maior na determinação da regulação do tempo de ativação de CDK. A ativação periódica destes complexos contendo ciclinas e CDK durante o ciclo celular medeia a regulação temporal de transições do ciclo celular (pontos de checagem). Outros fatores reguladores da atividade de CDK incluem inibidores de CDK (CKIs or ICKs, KIPs, CIPs, INKs), quinase ativadora de CDK (CAK), CDK fosfatase (Cdc25) e subunidade de CDK (CKS) (Mironov et al. *Plant Cell* 11, 509-522, 1999; Reed, S.I. *Progress in Cell Cycle Research* 2, 5-27, 1996).

Em plantas, duas classes maiores de CDKs, conhecidas como CDKs de tipo-A e de tipo-B, têm sido estudadas até o presente. As CDKs de tipo-A regulam ambas as transições de G1-para-S e de G2-para-M, enquanto que as CDKs de tipo-B parecem controlar apenas o ponto de checagem de G2-para-M (Hemerly et al., 1995; Magyar et al., 1997; Porceddu et al., 2001). Em adição, a presença de CDKs de tipo-C e de quinases ativadoras de CDK (CAKs) tem sido relatada (Magyar et al., 1997; Umeda et al., 1998; Joubès et al., 2001), do mesmo modo a presença de CDKs de tipo-D, tipo-E e tipo-F (Vandepoele et al. *Plant Cell* 14, 903-916, 2002).

É sabido que as CDKs de tipo-A têm uma estrutura terciária conservada (Goldsmith e Cobb, *Curr. Opin. Struct. Biol.* 4, 833-840), incluindo um motivo PSTAIRE elevadamente conservado que está envolvido

em ligação de ciclina. O núcleo catalítico de uma CDK é composto de um lobo N-terminal e um lobo C-terminal. O lobo C-terminal inclui uma fenda catalítica (responsável pela ligação de substrato e ATP) e adicionalmente compreende uma denominada alça-T, chamada segundo um resíduo de treonina que está conservado em várias famílias de quinase. Em CDK2 de humano, este resíduo de treonina está na posição 161, enquanto que em *Saccharomyces cerevisiae cdc28* e em *Schizosaccharomyces pombe cdc2* está localizado em posição 169 e 167 respectivamente. Fosforilação deste resíduo de treonina é relatada em causar uma mudança de conformação estrutural na alça-T que é necessária para mudar a quinase para um estado ativo (Gu et al., EMBO J. 11, 3995-4005). Vários estudos descrevem mutações da treonina conservada na alça-T (Ducommun et al. EMBO J. 10, 3311-3319, 1991; Gould et al. EMBO J. 10, 3297-3309; Marcote et al. Mol. Cell. Biol. 13, 5122-5131, 1993; Ducommun et al. Mol Cell. Biol. 11, 6177-6184, 1991; Coleman et al. J. Biol. Chem. 272, 18869-18874, 1997; Martinez et al. EMBO J. 16, 343-354, 1997; Gould et al. Mol. Gen. Genet. 259, 437-448, 1998; Booher et al. Mol. Cell. Biol. 6, 3523-3530, 1986; Solomon et al. Mol. Biol. Cell 3, 13-27, 1992; Lim et al. Mol. Cell. Biol. 16, 4573-4583, 1996), foi mostrado que todas as mutações testadas têm um impacto sério sobre a ligação de ligantes (tal como ciclina ou Suc1/ICK) e/ou sobre atividade de quinase, resultando em fenótipos letais ou defectivos em experimentos de complementação de levedura (de acordo com a numeração de levedura *cdc28*), e por analogia também a mutação T169D, imita uma fosforilação, foi demonstrado que nenhuma das CDKs com tais mutações foram capazes de totalmente complementar levedura.

Outros resíduos que desempenham um papel importante em atividade de proteína CDK de tipo-A são treonina na posição 14 e tirosina na posição 15. Sob fosforilação de pelo menos um destes aminoácidos, a CDK se torna inativada. WO 99/54489 descreve o uso de uma CDK com treonina 14 e

tirosina 15 substituídas por alanina e fenilalanina respectivamente para aumentar a tolerância de plantas ao estresse salino. WO 00/52171 descreve um método de modificar uma ou mais características ou propriedades fisiológicas, bioquímicas e morfológicas mediadas por citocinina compreendendo expressão de uma Cdc25 fosfoproteína fosfatase em uma planta.

Como mencionado acima CDKs são pontos de checagem de ciclo celular, que estão envolvidos em cascatas de transdução de sinal que garantem integridade genômica durante a divisão celular. Como pontos de checagem em mitose CDKs são reguladas por ciclina A ou ciclina B. As CDKs são apenas ativas durante o ciclo celular em conexão com sua respectiva ciclina. Embora seus papéis mitóticos essenciais estejam claros, os mecanismos moleculares pelos quais estas proteína quinases atuam na célula viva precisam ser clarificados. Em particular, as funções de diferentes isoformas de CDK e subunidades de CDK permanecem não esclarecidas. Análises genéticas e bioquímicas em vários microorganismos têm mostrado que as CDKs elevadamente conservadas são exigidas para entrada e saída mitóticas. Estudos bioquímicos e estruturais predizem que CDKs coordenam reconhecimento de substrato específico, mas atualmente os efetores a jusantes diretos de CDKs são desconhecidos. A situação é ainda mais difícil porque a maioria das CDKs existe em isoformas diferente cada uma tendo mais provavelmente uma função diferente. Isto significa que outros estudos bioquímicos são necessários para clarificar as rotas moleculares pelas quais CDKs atuam.

Portanto ainda há uma grande demanda por genes novos e mais adequados, que codificam CDKs, que participam na diferenciação de plantas. Conseqüentemente ditos genes novos devem ter tanto quanto possível as seguintes características:

- participação no ciclo celular e/ou na divisão celular;

- participação na organogênese;
 - participação na morfogênese;
 - influência na anatomia das plantas;
 - aumento da atividade metabólica;
- 5
- aumento do tamanho de órgãos diferentes das plantas, de preferência de sementes ou grãos; e/ou
 - uma ampla atividade em diferentes órgãos e/ou compartimentos celulares.

Portanto um objetivo foi obter outros genes CDK, que são adequados para aumento de rendimento em plantas. Este objetivo foi alcançado pelo processo de acordo com a invenção para a produção de compostos de fórmula I.

10

Portanto, de acordo com uma modalidade da presente invenção é proporcionado um método para melhorar as características de crescimento de planta em relação às plantas de tipo selvagem correspondentes, compreendendo modular a atividade em uma planta de um gene CDK de preferência de uma CDK de tipo-A e/ou modular a expressão de um ácido nucleico codificador de tal CDK de preferência de CDK de tipo-A, e opcionalmente selecionar plantas tendo características de crescimento melhoradas.

15

20

Vantajosamente, desempenho do método de acordo com a presente invenção resulta em plantas tendo uma variedade de características de crescimento melhoradas em relação às plantas de tipo selvagem correspondentes e cujas características de crescimento melhoradas compreendem pelo menos rendimento aumentado em relação às plantas de tipo selvagem correspondentes.

25

O termo “rendimento aumentado” como aqui definido é adotado para significar um aumento em qualquer um ou mais dos seguintes, cada um em relação às plantas de tipo selvagem correspondentes:

- (i) biomassa (peso) aumentada de uma ou mais partes de uma planta, particularmente partes acima do solo (capazes de serem colhidas), biomassa de raiz aumentada ou biomassa aumentada de qualquer outra parte capaz de ser colhida;
- 5 (ii) rendimento de sementes total aumentado, que inclui um aumento em biomassa de sementes (peso de sementes) e que pode ser um aumento no peso de sementes por planta ou em uma base de semente individual;
- (iii) número aumentado de flores (“flósculos”) por panícula
- 10 (iv) número aumentado de sementes (cheias);
- (v) tamanho de semente aumentado, que também pode influenciar a composição das sementes;
- (vi) volume de semente aumentado, que também pode influenciar a composição das sementes (incluindo composição e teor total de
- 15 óleo, proteína e carboidrato);
- (vii) área de semente individual aumentada;
- (viii) largura e/ou área de semente individual aumentada;
- (ix) índice de colheita aumentado, que é expressado como uma razão do rendimento de partes capazes de serem colhidas, tais como
- 20 sementes, sobre a biomassa total; e
- (x) peso de mil grãos aumentado (TKW), que é extrapolado do número de sementes cheias contadas e seu peso total. Um TKW aumentado pode resultar de um aumento em tamanho de embrião e/ou tamanho de endosperma.

25 Tomando-se milho como um exemplo, um aumento de rendimento pode ser manifestado como um ou mais dos seguintes: aumento no número de plantas por hectare ou acre, um aumento no número de espigas por planta, um aumento no número de filas, número de grãos por fila, peso de grãos, TKW, comprimento/diâmetro de espiga, dentre outros. Tomando-se

arroz como um exemplo, um aumento de rendimento pode ser manifestado por um aumento em um ou mais dos seguintes: número de plantas por hectare ou acre, número de panículas por planta, número de espiguetas por panícula, número de flores por panícula, aumento na taxa de enchimento de semente, expressada (em %) como a proporção do número de sementes cheias sobre o número de flósculos (número total de sementes), aumento em TKW, dentre outros. Um aumento em rendimento também pode resultar em arquitetura modificada, ou pode ocorrer como resultado de arquitetura modificada.

De acordo com um aspecto preferido, o desempenho dos métodos de acordo com a presente invenção resulta em plantas tendo rendimento aumentado e mais particularmente, biomassa aumentada e/ou rendimento de sementes aumentado. De preferência, o rendimento de sementes aumentado compreende um aumento em um ou mais dos seguintes: número de sementes (cheias), peso de semente total, tamanho de semente, volume de semente, peso de mil grãos e índice de colheita, cada um em relação às plantas de controle.

Portanto, de acordo com a presente invenção, é proporcionado um método para aumentar rendimento de planta em relação às plantas de controle correspondentes, cujo método compreende modular atividade de uma CDK ou de um seu homólogo em uma planta, cuja CDK ou homólogo possui um dos motivos aqui mencionados, e/ou modular a expressão de um ácido nucleico codificador de uma tal CDK ou de seu homólogo.

Visto que as plantas de acordo com a presente invenção têm rendimento aumentado, é provável que estas plantas exibam uma taxa de crescimento aumentada (durante pelo menos parte de seu ciclo de vida), relativo à taxa de crescimento de plantas de 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 de uma planta ou tipos de célula, incluindo sementes, de uma planta, ou pode ser

completamente substancialmente a planta inteira. Plantas tendo uma taxa de crescimento aumentada podem ter um ciclo de vida mais curto. O ciclo de vida de uma planta pode ser adotado para significar o tempo necessário para crescer de uma semente madura seca até o estágio onde a planta tem produzido sementes maduras secas, similar ao material inicial. Este ciclo de vida pode ser influenciado por fatores tais como vigor inicial, taxa de crescimento, tempo de floração e velocidade de maturação de semente. Um aumento em taxa de crescimento pode ocorrer em um ou mais estágios no ciclo de vida de uma planta ou durante substancialmente o ciclo de vida inteiro das planta. Taxa de crescimento aumentada durante os estágios iniciais no ciclo de vida de uma planta pode refletir vigor aumentado. O aumento em taxa de crescimento pode alterar o ciclo de safra de uma planta permitindo que plantas sejam semeadas mais tarde e/ou colhidas mais cedo do que de outra forma seria possível. Se a taxa de crescimento for suficientemente aumentada, ela pode permitir a semeadura de outras sementes de mesma espécie de planta (por exemplo semeadura e colheita de plantas de arroz seguidas por semeadura e colheita de outras plantas de arroz tudo dentro de um período de crescimento convencional). Similarmente, se a taxa de crescimento for suficientemente aumentada, ela pode permitir outra semeadura de sementes de espécies de planta diferentes (por exemplo a semeadura e a colheita de plantas de arroz seguidas por, por exemplo, a semeadura e a colheita opcional de feijão-soja, batatas ou qualquer outra planta adequada). Tempos adicionais de colheita do mesmo rizoma no caso de algumas plantas também pode ser possível. Alteração do 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 (a dizer 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 mais ampla do que a de suas contra-partes de tipo selvagem,

porque as limitações territoriais para crescimento de uma colheita são muitas vezes determinadas pelas condições ambientais adversas quer no momento do plantio (estação inicial) quer no momento da colheita (estação tardia). Tais condições adversas podem ser evitadas se o ciclo de colheita for encurtado. A taxa de crescimento pode ser determinada pela derivação de vários parâmetros de curvas de crescimento plotando experimentos de crescimento, tais parâmetros podem ser: T-Mid (o tempo demorado para as plantas alcançarem 50% de seu tamanho máximo) e T-90 (tempo demorado para as plantas alcançarem 90% de seu tamanho máximo), dentre outros.

Desempenho dos métodos da invenção dá plantas tendo uma taxa de crescimento aumentada. Portanto, de acordo com a presente invenção, é proporcionado um método para aumentar a taxa de crescimento de plantas, cujo método compreende modular atividade de uma CDK, de suas isoformas ou de um seu homólogo em uma planta, cuja CDK ou homólogo possui um motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE ou outro motivo aqui mencionado, e/ou modular a expressão de um ácido nucleico codificador de uma tal CDK ou de seu homólogo.

O termo “isoforma” como aqui usado deve significar versões diferentes de uma proteína com algumas diferenças pequenas, que também conhecido como uma isoenzima se a proteína for uma enzima. Isoformas podem ser normalmente separadas por eletroforese ou alguma outra técnica de separação. Existem por mecanismos múltiplos: diferentes loci de gene, alelos múltiplos (também chamado de alelomorfos, alelozimas, ou alozimas), interação de subunidades diferentes, formas de união diferentes, ou modificação de pós-tradução diferente.

Um aumento em rendimento e/ou taxa de crescimento ocorre se a planta estiver sob condições de não-estresse ou se a planta for exposta a vários estresses comparada com as plantas de controle. Plantas tipicamente respondem à exposição ao estresse pelo crescimento mais lento. Em

condições de estresse severo, a planta pode até mesmo parar inteiramente de crescer. Estresse suave por outro lado é aqui definido como sendo qualquer estresse ao qual uma planta é exposta que não resulta na cessação da planta em crescer inteiramente sem a capacidade para continuar a crescer. Devido às

5 vantagens em práticas agrícolas (irrigação, fertilização, tratamentos com pesticida) estresses severos não são muitas vezes encontrados em plantas de colheita cultivadas. Como uma consequência, o crescimento comprometido induzido por estresse suave é muitas vezes um aspecto indesejável para agricultura. Estresses suaves são os estresses típicos aos quais uma planta

10 pode ser exposta. Estes estresses podem ser estresses bióticos e/ou abióticos (ambientais) diários aos quais uma planta é exposta. Estresses abióticos ou ambientais típicos incluem estresses de temperatura causados por temperaturas congelantes / frias ou quentes atípicas; estresse salino, estresse aquoso (estiagem ou excesso de água). Estresses abióticos também podem ser

15 causados por compostos químicos. Estresses bióticos são tipicamente aqueles causados por patógenos, tais como bactérias, vírus, fungos e insetos. O termo “condições não-estressantes” como aqui usado são aquelas condições ambientais que não significativamente vão além das condições climáticas ou outras condições abióticas diárias que as plantas podem encontrar. Pessoas

20 experientes na técnica estão cientes das condições de solo normais e das condições climáticas para uma dada localização geográfica.

Os termos “aumentar”, “melhorar” ou “melhora” são intercambiáveis e devem significar no sentido do pedido pelo menos 10%, 20%, 30%, 40% ou 50%, de preferência pelo menos 60%, 70%, 80%, 90% ou

25 100%, mais preferivelmente 150%, 200%, 300%, 400% ou 500% mais crescimento em comparação com a planta de tipo selvagem como aqui definido, e.g. que significa em comparação com uma planta sem a introdução da sequência de ácido nucleico codificadora de CDK de acordo com a invenção.

O aumento referido à atividade do polipeptídeo totaliza em uma célula, um tecido, uma organela, um órgão ou um organismo ou uma sua parte preferivelmente pelo menos 5%, preferivelmente pelo menos 20% ou pelo menos 50%, especialmente preferivelmente pelo menos 70%, 80%, 90% ou mais, muito especialmente preferivelmente pelo menos 200%, 300% ou 400%, mais preferivelmente pelo menos 500% ou mais em comparação com o controle, a referência ou tipo selvagem.

O termo “modular a atividade” deve significar qualquer mudança da expressão das seqüências de ácido nucleico ou proteínas codificadas da invenção, que leva a um aumento no crescimento da plantas.

As características de crescimento acima mencionadas podem ser vantajosamente modificadas em qualquer planta.

O termo “planta” como aqui usado inclui plantas inteiras, ancestrais e progênie das plantas e partes de planta, incluindo sementes, rebentos, caules, folhas, raízes (incluindo tubérculos), frutas, pedúnculo, mudas, tubérculos, flores, e tecidos e órgãos, sendo que cada um dos acima mencionados compreendem o gene / ácido nucleico de interesse ou a modificação específica no gene / ácido nucleico de interesse. O termo “planta” também inclui células de planta, culturas em suspensão, tecido de calo, embriões, regiões meristemáticas, gametófitos, esporófitos, pólen, e microesporos, de novo sendo que cada um dos acima mencionados compreendem o gene / ácido nucleico de interesse.

Um(a) “referência”, “controle” ou “tipo selvagem” é em particular uma célula, um tecido, um órgão, um organismo, ou uma sua parte, que não era produzida de acordo com o processo da invenção. Conseqüentemente, o termos “tipo selvagem”, “controle” ou “referência” são intercambiáveis e podem ser uma célula ou uma parte da planta tal como uma organela ou um tecido, ou uma planta, que não foi modificada ou tratada de acordo com o processo aqui descrito conforme a invenção.

Conseqüentemente, a célula ou uma parte da planta tal como uma organela ou uma planta usada como tipo selvagem, controle ou referência corresponde à célula, planta ou parte da mesma tanto quanto possível e é em qualquer outra propriedade exceto no resultado do processo da invenção tão idêntica quanto possível ao tema da invenção. Assim, o tipo selvagem, controle ou a referência é tratado(a) identicamente ou tão identicamente quanto possível, dizendo que podem ser diferentes apenas condições ou propriedades que não influenciam a qualidade da propriedade testada. Isto significa que em outras palavras que o tipo selvagem denota (1) uma planta, que transporta a forma inalterada de um gene ou alelo ou (2) o material/planta inicial do(a) qual as plantas produzidas pelo processo ou método da invenção são derivadas.

Preferivelmente, qualquer comparação entre as planta de tipo selvagem e as plantas produzidas pelo processo da invenção é realizada sob condições análogas. O termo “condições análogas” significa que todas as condições tais como, por exemplo, condições de crescimento ou cultura, condições de ensaio (tais como composição de tampão, temperatura, substratos, cepa patogênica, concentrações e semelhante) são mantidas idênticas entre os experimentos a serem comparados.

A “referência”, “controle”, ou “tipo selvagem” é preferivelmente um tema, e.g. uma organela, uma célula, um tecido, uma planta, que não foi modificado ou tratado de acordo com o processo aqui descrito da invenção e é em qualquer outra propriedade tão similar quanto possível ao tema da invenção. A referência, controle ou tipo selvagem é em seu genoma, transcriptoma, proteoma ou metaboloma tão similar quanto possível ao tema da presente invenção. Preferivelmente, o termo organela, célula, tecido ou planta de “referência”, “controle” ou “tipo selvagem”, refere-se a uma organela, célula, tecido ou planta, que é quase geneticamente idêntica(o) à organela, célula, tecido ou planta, da presente invenção ou uma sua parte preferivelmente em 95%, mais preferido são em 98%, ainda mais

preferido são em 99,00%, em particular em 99,10%, 99,30%, 99,50%, 99,70%, 99,90%, 99,99%, 99,999% ou mais. Mais preferível a “referência”, “controle”, ou “tipo selvagem” é preferivelmente um tema, e.g. uma organela, uma célula, um tecido, uma planta, que é geneticamente idêntica(o) à planta, célula, organela usada de acordo com o processo da invenção exceto que as moléculas de ácido nucleico ou o produto de gene codificado por elas são modificadas ou mudadas de acordo com o processo da invenção.

Se, um controle, uma referência ou um tipo selvagem diferindo do tema da presente invenção apenas por não ser tema do processo da invenção não puder ser obtido(a), um controle, uma referência ou um tipo selvagem pode ser um organismo no qual a causa para a modulação da atividade conferindo o aumento de composto de química fina como aqui descrito tem sido retornada à situação anterior ou desligada, e.g. por complementação de produto de gene reduzido confiável, e.g. por (sobre)expressão transiente ou estável, por ativação de um ativador ou agonista, por inativação de um inibidor ou antagonista, por adição de compostos ativos como e.g. hormônios, por introdução de intensificadores etc.

Plantas que são particularmente úteis nos métodos ou processos da invenção incluem algas, samambaias, e todas as plantas que pertencem à superfamília Viridiplantae, em particular plantas monocotiledôneas e dicotiledôneas, incluindo leguminosas para forragem ou ração de animais, plantas ornamentais, colheitas alimentícias, árvores, ou arbustos selecionados da lista compreendendo *Abelmoschus* spp., *Acer* spp., *Actinidia* spp., *Agropyron* spp., *Allium* spp., *Amaranthus* spp., *Ananas comosus*, *Annona* spp., *Apium graveolens*, *Arabidopsis thaliana*, *Arachis* spp., *Artocarpus* spp., *Asparagus officinalis*, *Avena sativa*, *Averrhoa carambola*, *Benincasa hispida*, *Bertholletia excelsea*, *Beta vulgaris*, *Brassica* spp., *Cadaba farinosa*, *Camellia sinensis*, *Canna indica*, *Capsicum* spp., *Carica papaya*,

Carissa macrocarpa, Carthamus tinctorius, Carya spp., Castanea spp.,
 Cichorium endivia, Cinnamomum spp., Citrullus lanatus, Citrus spp., Cocos
 spp., Coffea spp., Cola spp., Colocasia esculenta, Corylus spp., Crataegus
 spp., Cucumis spp., Cucurbita spp., Cynara spp., Daucus carota, Desmodium
 5 spp., Dimocarpus longan, Dioscorea spp., Diospyros spp., Echinochloa spp.,
 Eleusine coracana, Eriobotrya japonica, Eugenia uniflora, Fagopyrum spp.,
 Fagus spp., Ficus carica, Fortunella spp., Fragaria spp., Ginkgo biloba,
 Glycine spp., Gossypium hirsutum, Helianthus spp., Hibiscus spp., Hordeum
 spp., Ipomoea batatas, Juglans spp., Lactuca sativa, Lathyrus spp., Lemna
 10 spp., Lens culinaris, Linum usitatissimum, Litchi chinensis, Lotus spp., Luffa
 acutangula, Lupinus spp., Macrotyloma spp., Malpighia emarginata, Malus
 spp., 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
 15 spp., Panicum miliaceum, Passiflora edulis, Pastinaca sativa, Persea spp.,
 Petroselinum crispum, Phaseolus spp., Phoenix spp., Physalis spp., Pinus
 spp., Pistacia vera, Pisum spp., Poa spp., Populus spp., Prosopis spp., Prunus
 spp., Psidium spp., Punica granatum, Pyrus communis, Quercus spp.,
 Raphanus sativus, Rheum rhabarbarum, Ribes spp., Rubus spp., Saccharum
 20 spp., Sambucus spp., Secale cereale, Sesamum spp., Solanum spp., Sorghum
 bicolor, Spinacia spp., Syzygium spp., Tamarindus indica, Theobroma cacao,
 Trifolium spp., Triticosecale rimpai, Triticum spp., Vaccinium spp., Vicia
 spp., Vigna spp., Vitis spp., Zea mays, Zizania palustris, Ziziphus spp., dentre
 outras.

25 De acordo com um aspecto preferido da presente invenção, a
 planta é uma planta de colheita compreendendo feijão-soja, girassol, canola,
 alfafa, colza, algodão, tomate, batata ou tabaco. Mais preferivelmente, a
 planta de acordo com a presente invenção é uma planta monocotiledônea tal
 como cana-de-açúcar, mais preferivelmente um cereal, tal como arroz, milho,

trigo, painço, cevada, centeio, aveia ou sorgo.

Plantas preferidas particulares são plantas selecionadas do grupo consistindo de Asteraceae tais como dos gêneros *Helianthus*, *Tagetes* e.g. as espécies *Helianthus annuus* [girassol], *Tagetes lucida*, *Tagetes erecta* ou *Tagetes tenuifolia* [cravo-de-defunto], Brassicaceae tais como dos gêneros *Brassica*, *Arabidopsis* e.g. as espécies *Brassica napus*, *Brassica rapa* ssp. [canola, colza de semente oleosa, nabo silvestre] ou *Arabidopsis thaliana*. Fabaceae tal como o gênero *Glycine* e.g. as espécies *Glycine max*, Soja hispida ou Soja max [feijão-soja] (como eu não estou seguro, se , Soja max de fato existe, que realmente se chama *Glycine max*). Linaceae tal como o gênero *Linum* e.g. a espécie *Linum usitatissimum*, [linho, linhaça]; Poaceae tais como os gêneros *Hordeum*, *Secale*, *Avena*, *Sorgo*, *Oryza*, *Zea*, *Triticum* e.g. as espécies *Hordeum vulgare* [cevada]; *Secale cereale* [centeio], *Avena sativa*, *Avena fatua*, *Avena byzantina*, *Avena fatua* var. *sativa*, *Avena hybrida* [aveia], *Sorgo bicolor* [Sorgo, painço], *Oryza sativa*, *Oryza latifolia* [arroz], *Zea mays* [milho] *Triticum aestivum*, *Triticum durum*, *Triticum turgidum*, *Triticum hybernun*, *Triticum macha*, *Triticum sativum* ou *Triticum vulgare* [trigo, trigo de pão, trigo comum]; Solanaceae tais como os gêneros *Solanum*, *Lycopersicon* e.g. as espécies *Solanum tuberosum* [batateira], *Lycopersicon esculentum*, *Lycopersicon lycopersicum*., *Lycopersicon pyriforme*, *Solanum integrifolium* ou *Solanum lycopersicum* [tomateiro].

A atividade de uma proteína CDKA pode ser modulada por modulação dos níveis da proteína CDKA. Alternativamente, atividade também pode ser modulada quando não houver mudança nos níveis de uma proteína CDKA, isto pode ocorrer quando as propriedades intrínsecas do polipeptídeo são alteradas, por exemplo por preparação de um mutante. De acordo com um aspecto preferido da invenção, atividade modulada da proteína CDKA e/ou expressão moduladas de um ácido nucleico codificador desta CDKA é atividade introduzida e/ou aumentada de uma proteína CDKA

e/ou expressão aumentada de um ácido nucleico codificador desta CDKA.

Os termos “CDK de tipo-A” ou “CDKA” como aqui definidos podem ser usados intercambiavelmente e incluem qualquer seqüência de aminoácidos tendo atividade de quinase dependente de ciclina e cuja
5 seqüência quando usada na construção de uma árvore filogenética de CDK, tal como aquela mostrada no protocolo de seqüência de preferência de SEQ ID NO: 45, 47, 49, 51, 53 e/ou SEQ ID NO: 55, agrupamentos ao redor de CDKs de tipo-A em vez de qualquer dos outros grupos de CDK e cuja seqüência de aminoácidos compreende uma seqüência de aminoácidos
10 (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE são preferidos. Uma pessoa experiente na técnica poderia prontamente determinar se qualquer seqüência de aminoácidos em questão cai dentro da definição de uma “CDK de tipo-A” usando técnicas e programas de computador conhecidos para a preparação de uma tal árvore filogenética, tais como o pacote GCG, EBI ou CLUSTAL,
15 usando parâmetros pré-estabelecidos (veja por exemplo Vandepoele et al. 2002). Sob construção de uma tal árvore filogenética, seqüências agrupando no grupo de CDK de tipo-A serão consideradas caindo dentro da definição de uma “CDK de tipo-A” ou “CDKA”, e portanto serão úteis na realização dos métodos da invenção. Preferivelmente a CDK adicionalmente compreende em
20 ordem crescente de preferência pelo menos 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% ou mais de identidade de seqüência total ao aminoácido mostrado em SEQ ID NO: 2. Portanto programas baseados nos algoritmos acima mencionados são preferidos. Vantajosamente as comparações de seqüências podem ser feitas com o programa PileUp (J. Mol. Evolution., 25,
25 351-360, 1987, Higgins et al., CABIOS, 5 1989: 151-153) ou preferivelmente com os programas Gap e BestFit, que são respectivamente baseados nos algoritmos de Needleman e Wunsch [J. Mol. Biol. 48; 443-453 (1970)] e Smith e Waterman [Adv. Appl. Math. 2; 482-489 (1981)]. Ambos os programas são parte do pacote de programas de computador GCG [Genetics

Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711 (1991); Altschul et al. (1997) *Nucleic Acids Res.* 25:3389 et seq.]. Portanto preferivelmente os cálculos para determinar as percentagens de homologia de seqüência são feitos com o programa Gap sobre a extensão inteira das seqüências. Os seguintes ajustes padrão para a comparação de seqüências de ácido nucleico foram usados: peso de intervalo: 50, peso de comprimento: 3, combinação média: 10.000, não-combinação média: 0.000.

Homologia entre dois polipeptídeos é entendida com o significado de a identidade da seqüência de aminoácidos sobre, em cada caso, o comprimento inteiro da seqüência que é calculada por comparação com o auxílio do algoritmo de programa GAP (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA), ajustando os seguintes parâmetros:

Peso de intervalo:	8	Peso de comprimento:	2
Combinação média:	2,912	Não-combinação média:	-2,003.

Em ambos os casos (comparação de seqüência de ácido nucleico ou de seqüência de aminoácidos) dos parâmetros mencionados Combinação Média e Não-Combinação Média os números dados acima são os resultados do cálculo.

Os vários domínios estruturais em uma proteína CDKA são bem conhecidos (De Bondt et al., *Nature* 363, 595-602, 1993) e podem ser identificados usando banco 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 e Bairoch (1994), Uma sintaxe de perfil generalizada para motivos de seqüências biomoleculares e sua função em interpretação de seqüência automática. (Em) 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/>).

O domínio quinase de CDK é de um TKc de tipo-S (Número de Acesso em SMART de SM00220, Número de Acesso em InterPro de IPR002290), e tem atividade de Ser/Thr quinase. O sítio ativo predito (VLHRDLKPQNLLI, sendo que D é o resíduo catalítico predito) corresponde à assinatura PS00108 de PROSITE. Em posição 1 do sítio ativo em vez de uma Valina pode existir uma Fenilalanina. Em posição 6 uma troca de Leucina por Metionina pode ocorrer e em posição 9 Gln pode ser trocada por Asn. O sítio de ligação de ATP (IGEGTYGVVYRARDKVTNETIALK) corresponde à assinatura PS00107 de PROSITE. Também no sítio de ligação de ATP podem ocorrer algumas mutações. Elas são as seguintes: posição 11 Arg → Lys; posição 12 Ala → Gly, posição 13 Arg → Leu, posição 15 Lys → Arg e posição 16 Val → Leu, Ala, Ser, Thr ou Asn.

Métodos para a pesquisa e a identificação de homólogos de CDK de tipo-A estariam dentro do domínio de pessoas experientes na técnica. Tais métodos compreendem comparação das seqüências representadas por SEQ ID NO 1 ou 2, ou pelo Número de Acesso em GenBank de CAA42922, em um formato legível por computador, com seqüências que estão disponíveis em bancos de dados públicos tais como MIPS (<http://mips.gsf.de/>), GenBank (<http://www.ncbi.nlm.nih.gov/Genbank/index.html>) ou EMBL Nucleotide Sequence Database (<http://www.ebi.ac.uk/embl/index.html>), usando algoritmos bem conhecidos na técnica para o alinhamento ou a comparação de seqüências, tais como GAP (Needleman e Wunsch, J. Mol. Biol. 48; 443-453 (1970)), BESTFIT (usando o algoritmo de homologia local de Smith e 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 e D. J. Lipman Proc.Natl.Acad.Sci. USA 85:2444- 2448 (1988)). O programa de computador para realizar análise BLAST está disponível para o público
5 através do National Centre for Biotechnology Information (NCBI). Os homólogos mencionados abaixo foram identificados usando parâmetros pré-estabelecidos de BLAST (matriz BLOSUM62, penalidade de abertura de intervalo 11 e penalidade de extensão de intervalo 1) e são preferivelmente usadas para análise as seqüências de comprimento total. Estes métodos de
10 alinhamento também facilmente permitem a identificação da treonina conservada que corresponde à treonina 161 em CDC2 de humano ou CDKA de arroz; 1 (SEQ ID NO: 8).

É para ser entendido que o termo “CDK ou de preferência CDK de tipo-A ou um seu homólogo” não é para ser limitado às seqüências
15 como representadas no protocolo de seqüência, mas que qualquer polipeptídeo atendendo aos critérios de ter atividade de quinase dependente de ciclina, ter um domínio (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE ou outro domínio como aqui descrito, e ter pelo menos 80%, 85% or 90%, preferivelmente 91%, 92%, 93%, 94% ou 95%, mais preferivelmente 96%,
20 97%, 98%, 99% ou 100% de identidade de seqüência em relação às seqüências descritas no protocolo de seqüência preferivelmente em relação às seqüências de SEQ ID NO: 45, 47, 49, 51, 53 e/ou SEQ ID NO: 55, pode ser adequado para uso nos métodos da invenção, desde que as CDKs tenham a propriedade de aumento de rendimento.

25 Para determinar a atividade de quinase de CDKs de tipo-A, vários ensaios estão disponíveis e são bem conhecidos na técnica (por exemplo Current Protocols in Molecular Biology, Volumes 1 e 2, Ausubel et al. (1994), Current Protocols; ou *online*, tal como <http://www.protocol-online.org>).

Em resumo, o ensaio de quinase geralmente envolve: (1) contatar a proteína quinase com um polipeptídeo substrato contendo o sítio alvo a ser fosforilado; (2) permitir fosforilação do sítio alvo em um tampão quinase apropriado sob condições adequadas; (3) separar produtos fosforilados de substrato não-fosforilado após um período de reação adequado. A presença ou ausência de atividade de quinase é determinada pela presença ou ausência do alvo fosforilado. Também, medições quantitativas podem ser realizadas. Proteína CDK purificada, ou extratos de célula contendo ou enriquecidos com a proteína CDK podem ser usados como uma fonte de proteína quinase. Histona H1 ou peptídeos pequenos são particularmente bem adequados como um substrato. O peptídeo precisa compreender um ou mais resíduos de serina, treonina, ou tirosina em um motivo de sítio de fosforilação. Uma compilação de sítios de fosforilação pode ser encontrada em *Biochimica et Biophysica Acta* 1314, 191-225, (1996). Também, os substratos de peptídeo podem ter vantajosamente uma carga positiva efetiva para facilitar ligação em filtros de fosfocelulose, (permitindo separação dos peptídeos fosforilados dos não-fosforilados e detecção dos peptídeos fosforilados). Se um motivo de sítio de fosforilação não é conhecido, um substrato geral de Ser/Thr quinase pode ser usado. Por exemplo, o peptídeo “ADAQHATPPKKRKVEDPKDF” (Marshak et al. *J. Cell. Biochem.* 45, 391, 1991) é um substrato específico para CDK de tipo-A. Para determinar os parâmetros cinéticos para fosforilação do peptídeo sintético, uma faixa de concentrações de peptídeo é exigida. Para reações iniciais, uma concentração de peptídeo de 0,7-1,5 mM pode ser usada. Para cada enzima quinase, é importante determinar os ótimos tampão, força iônica, e pH para atividade. Um Tampão Quinase 5x padrão geralmente contém BSA 5 mg/ml (Albumina de Soro Bovino prevenindo a adsorção de quinase no tubo de ensaio), Tris-Cl 150 mM (pH 7,5), $MgCl_2$ 100 mM. As concentrações ótimas de cátions divalentes precisam ser determinadas empiricamente para

cada proteína quinase. Tampões adequados para ensaios de CDK são conhecidos na técnica (por exemplo John et al., Protoplasma 161, 70-74, 1991). Um doador comumente usado do grupo fosforila é [gama-³²P]ATP radiomarcado (normalmente em concentração final de 0,2 mM). A quantidade de ³²P incorporado nos peptídeos pode ser determinada pela medição da atividade sobre as almofadas secas de nitrocelulose em um contador de cintilação.

Adicionalmente, tal “CDK ou seu homólogo ou derivado”, quando expressada sob o controle de um promotor específico para broto em *Oryza sativa*, aumenta o rendimento de sementes comparado com as plantas de tipo selvagem. Este aumento em rendimento de semente pode ser medido em várias maneiras, por exemplo como um aumento no peso total de sementes, como um aumento no número de sementes cheias colhidas de uma planta ou como um aumentado índice de colheita.

A atividade biológica e/ou funcional de uma CDK ou de um seu homólogo de acordo com a presente invenção inclui pelo menos uma tendo atividade de quinase dependente de ciclina ou tendo atividade aumentadora de rendimento em plantas como descrito acima.

“Fragmentos ativos” de uma CDK de preferência de uma proteína CDK de tipo-A incluem pelo menos 100, 110, 120, 130, 140 ou 150, de preferência 160, 170, 180, 190 ou 200 resíduos de aminoácido de uma proteína CDK, cujos resíduos contíguos retêm atividade biológica e/ou funcional em relação à proteína naturalmente ocorrente.

Uma CDK ou um seu homólogo como definido aqui acima é codificada por uma molécula de ácido nucleico CDK. O ácido nucleico codificador de uma CDK ou de um seu homólogo pode ser qualquer ácido nucleico natural ou sintético. Portanto o termo “molécula de ácido nucleico CDK” ou “gene CDK” como aqui definido é qualquer molécula de ácido nucleico (incluindo aquelas como um resultado da degeneração do código

genético) codificadora de um polipeptídeo CDK ou de um seu homólogo como descrito aqui acima. Exemplos de moléculas de ácido nucleico CDK incluem aquelas no protocolo de seqüência, e aquelas codificadoras dos homólogos mencionados acima. Ácidos nucleicos CDK e suas variantes

5 funcionais podem ser adequados na prática dos métodos da invenção. Tais ácidos nucleicos CDK variantes funcionais incluem porções de uma molécula de ácido nucleico CDK, variantes alélicas e/ou ácidos nucleicos capazes de hibridizarem com uma molécula de ácido nucleico CDK. O termo “funcional” no contexto de uma variante funcional refere-se a uma variante (i.e. uma

10 porção ou uma seqüência de hibridização), que codifica um polipeptídeo tendo atividade de quinase dependente de ciclina.

Uma outra modalidade da invenção é uma molécula de ácido nucleico isolada compreendendo uma molécula de ácido nucleico selecionada do grupo consistindo de:

- 15 a) uma molécula de ácido nucleico isolada como representada em SEQ ID NO: 45, 47, 49, 51, 53 ou 55;
- b) uma molécula de ácido nucleico isolada codificadora da seqüência de aminoácidos como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56;
- 20 c) uma molécula de ácido nucleico isolada cuja seqüência pode ser deduzida de uma seqüência de polipeptídeos como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56 como um resultado da degeneração do código genético;
- d) uma molécula de ácido nucleico isolada que codifica um
- 25 polipeptídeo que tem pelo menos 80% de identidade com a seqüência de aminoácidos do polipeptídeo codificado pela molécula de ácido nucleico de (i) a (iii);
- e) uma molécula de ácido nucleico isolada codificadora de um homólogo, derivado ou fragmento ativo da molécula de aminoácido

como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56, cujo homólogo, derivado ou fragmento é de uma origem de planta ou compreende vantajosamente um motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I);

- 5 f) uma molécula de ácido nucleico isolada codificadora de uma proteína compreendendo uma seqüência de aminoácidos selecionada do grupo consistindo de:
- aa) (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE;
- ab) (V/F/I)(L/I)HRD(L/M)K(P/S/T)(Q/N/S/G)N(L/I)L(V/L/I);
- 10 ac) (I/L)(G/N)(E/R)G(T/A)YG(V/I)V(Y/C)(R/K/S)(A/G/S)
(R/L/T/I)(D/N)(K/R/E)(V/K/A/S/T/N)T(N/S/G)(E/K/Q)
(T/L/I/K)(I/V)A(L/V/I)KK;
- ad) LK(I/L)(C/A)DFGL(A/S)R;
- ae) WYRAPE(L/I)L(L/F)(C/G);
- 15 af) GCI(F/M)AE(I/L/M); e
- ag) DLL(Q/N/S/R)(K/Q/R)(L/M)(L/F)(I/T/I/C)(F/Y/L)DP
(T/E/D/R/S)(K/Q)RI;
- g) uma molécula de ácido nucleico isolada capaz de hibridizar com um ácido nucleico de (i) a (iii) acima, ou seu complemento, sendo que a
- 20 seqüência de hibridização ou o seu complemento codifica uma proteína CDK de planta que compreende vantajosamente um motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I);

pelo qual a molécula de ácido nucleico tem atividades intensificadoras de crescimento em plantas.

- 25 A presente invenção também proporciona uma molécula de ácido nucleico isolada (= seqüência de ácido nucleico) selecionada do grupo consistindo de:

- a) uma molécula de ácido nucleico isolada como representada em SEQ ID NO: 45, 47, 49, 51, 53 ou 55;

- b) uma molécula de ácido nucleico isolada codificadora da seqüência de aminoácidos como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56;
- 5 c) uma molécula de ácido nucleico isolada cuja seqüência pode ser deduzida de uma seqüência de polipeptídeos como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56 como um resultado da degeneração do código genético;
- 10 d) uma molécula de ácido nucleico isolada que codifica um polipeptídeo que tem pelo menos 80% de identidade com a seqüência de aminoácidos do polipeptídeo codificado pela molécula de ácido nucleico de (a) a (c);
- 15 e) uma molécula de ácido nucleico isolada codificadora de um homólogo, derivado ou fragmento ativo da molécula de aminoácido como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56, cujo homólogo, derivado ou fragmento é de uma origem de planta ou compreende vantajosamente um motivo um (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE;
- 20 f) uma molécula de ácido nucleico isolada codificadora de uma proteína compreendendo uma seqüência de aminoácidos selecionada do grupo consistindo de:
- i) (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE, preferivelmente PSTAIRE;
- ii) (V/F/I)(L/I)HRD(L/M)K(P/S/T)(Q/N/S/G)N(L/I)L(V/L/I); preferivelmente HRDXKXXNXL;
- 25 iii) (I/L)(G/N)(E/R)G(T/A)YG(V/I)V(Y/C)(R/K/S)(A/G/S)(R/L/T/I)(D/N)(K/R/E)(V/K/A/S/T/N)T(N/S/G)(E/K/Q)(T/L/I/K)(I/V)A(L/V/I)KK; preferivelmente GXVXXXXXXXXTXXXXAXKK;
- iv) LK(I/L)(C/A)DFGL(A/S)R, preferivelmente

LKXXDFGLXR;

v) WYRAPE(L/I)L(L/F)(C/G), preferivelmente WYRAPE;

vi) GCI(F/M)AE(I/L/M), preferivelmente GCIXAEX; e

vii) DLL(Q/N/S/R)(K/Q/R)(L/M)(L/F)(I/T/I/C)(F/Y/L)DP

5 (T/E/D/R/S)(K/Q)RI, preferivelmente

DLLXXXXXXXXDPXXRI.

g) uma molécula de ácido nucleico isolada capaz de hibridizar com um
ácido nucleico de (a) a (c) acima, ou seu complemento, sendo que a
seqüência de hibridização ou o seu complemento codifica uma
10 proteína CDK de planta que compreende vantajosamente um
motivo um (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE;

h) variantes alélicas de um ácido nucleico de acordo com qualquer um
de (a) a (d) acima, cujas variantes alélicas codificam uma CDK de
planta; e

15 i) variantes de união alternativas de um ácido nucleico de
acordo com qualquer um de (a) a (d), cujas variantes
de união alternativas codificam uma proteína CDK de
planta compreendendo vantajosamente um motivo
(P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE;

20 pelo qual a variável X significa qualquer aminoácido e
pelo qual a proteína codificada confere um aumento em
rendimento.

Com relação à seqüência de ácido nucleico como representada
por um construto de ácido nucleico que contém uma seqüência de ácido
25 nucleico aqui mencionada ou um organismo (= organismo transgênico) que
está transformado com a dita seqüência de ácido nucleico ou o dito construto
de ácido nucleico, “transgene” significa todos os construtos que têm sido
preparados por métodos de manipulação genética, preferivelmente nos quais
quer

- a) a seqüência de ácido nucleico como mostrada em tabela I A e/ou I B, aplicação N^o 1, colunas 5 e 7 ou um seu derivado, quer
- b) um elemento regulatório genético, por exemplo um promotor, que está funcionalmente ligado na seqüência de ácido nucleico como
- 5 mostrado em tabela I A e/ou I B, aplicação N^o 1, colunas 5 e 7 ou um seu derivado, ou
- c) (a) e (b)

está/estão não presente(s) em seu ambiente genético natural ou tem/têm sido modificado(s) por meio de métodos de manipulação genética, sendo possível que a modificação seja, por intermédio de exemplo, uma substituição, adição, deleção, inversão ou inserção de um ou mais nucleotídeos. “Ambiente genético natural” significa o locus cromossômico natural no organismo de origem ou a presença em uma biblioteca genômica. No caso de uma biblioteca genômica, o ambiente genético, natural, da

10 seqüência de ácido nucleico está preferivelmente pelo menos parcialmente ainda preservado. O ambiente flanqueia a seqüência de ácido nucleico pelo menos em um lado e tem um comprimento de seqüência de pelo menos 50 pb, preferivelmente pelo menos 500 pb, particularmente preferivelmente pelo menos 1.000 pb, muito particularmente preferivelmente pelo menos 5000 pb.

20 A não ser se descritos de outro modo, os termos “polinucleotídeos”, “ácido nucleico” e “molécula de ácido nucleico” como aqui usados são intercambiáveis. A não ser se descritos de outro modo, os termos “peptídeo”, “polipeptídeo” e “proteína” são intercambiáveis no presente contexto. O termo “seqüência” pode se referir aos polinucleotídeos,

25 ácido nucleicos, moléculas de ácido nucleico, peptídeos, polipeptídeos e proteínas, dependendo do contexto no qual o termo “seqüência” é usado. Os termos “gene(s)”, “polinucleotídeo”, “seqüência de ácido nucleico”, “seqüência de nucleotídeos”, ou “molécula(s) de ácido nucleico” como aqui usados referem-se a uma forma polimérica de nucleotídeos de qualquer

comprimento, quer ribonucleotídeos quer desoxirribonucleotídeos. Os termos referem-se apenas à estrutura primária da molécula.

Assim, os termos “gene(s)”, “polinucleotídeo”, “seqüência de ácido nucleico”, “seqüência de nucleotídeos”, ou “molécula(s) de ácido nucleico” como aqui usados incluem RNA e DNA de fita simples ou dupla. Também incluem tipos conhecidos de modificações, por exemplo, metilação, “bloqueios”, substituições de um ou mais nucleotídeos naturalmente ocorrentes por um análogo. Preferivelmente, a seqüência de DNA ou RNA da invenção compreende uma seqüência codificadora do polipeptídeo aqui definido.

Uma “seqüência codificadora” é uma seqüência de nucleotídeos, que é transcrita em mRNA e/ou traduzida em um polipeptídeo quando posta sob o controle das seqüências regulatórias apropriadas. Os limites da seqüência codificadora são determinados por um códon de iniciação de tradução na terminação-5’ e um códon de terminação de tradução na terminação-3’. Uma seqüência codificadora pode incluir, mas não é limitada a mRNA, cDNA, seqüências de nucleotídeos recombinantes ou DNA genômico, embora íntrons possam estar presentes igualmente sob certas circunstâncias.

Um(a) polinucleotídeo ou molécula de ácido nucleico “isolado(a)” está separado(a) de outros(as) polinucleotídeos ou moléculas de ácido nucleico, que estão presentes na fonte natural da molécula de ácido nucleico. Uma molécula de ácido nucleico isolada pode ser um fragmento cromossômico de vários kb, ou preferivelmente, uma molécula apenas compreendendo a região codificadora do gene. Conseqüentemente, uma molécula de ácido nucleico isolada da invenção pode compreender regiões cromossômicas, que são 5’ e 3’ adjacentes ou outras regiões cromossômicas adjacentes, mas preferivelmente não compreende tais seqüências que naturalmente flanqueiam a seqüência de molécula de ácido nucleico no

contexto genômico ou cromossômico no organismo do qual a molécula de ácido nucleico se origina (por exemplo seqüências que são adjacentes às regiões codificadoras de 5'- e 3'-UTRs da molécula de ácido nucleico). Em várias modalidades, a molécula de ácido nucleico isolada usada no processo de acordo com a invenção pode, por exemplo compreender menos do que aproximadamente 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb ou 0,1 kb de seqüências de nucleotídeos que naturalmente flanqueiam a molécula de ácido nucleico no DNA genômico da célula da qual a molécula de ácido nucleico se origina.

Uma molécula de ácido nucleico incluindo uma seqüência completa das moléculas de ácido nucleico usadas no processo, por exemplo o polinucleotídeo da invenção, ou uma sua parte pode ser adicionalmente isolada por reação em cadeia da polimerase, iniciadores oligonucleotídeo baseados nesta seqüência ou em suas partes sendo usadas. Por exemplo, a molécula de ácido nucleico compreendendo a seqüência completa ou sua parte pode ser isolada por reação em cadeia da polimerase usando iniciadores oligonucleotídeo que têm sido gerados baseando-se nesta seqüência completa. Por exemplo, mRNA pode se isolado de células (por exemplo por intermédio do método de extração com tiocianato de guanidínio de Chirgwin et al. (1979) *Biochemistry* 18:5294-5299) e cDNA pode ser gerado por intermédio de transcriptase reversa (por exemplo transcriptase reversa de Moloney MLV, disponível da Gibco/BRL, Bethesda, MD, ou transcriptase reversa AMV, obtenível da Seikagaku America, Inc., St.Petersburg, FL).

Moléculas de ácido nucleico que são vantajosamente para o processo de acordo com a invenção podem ser isoladas baseando-se em sua homologia em relação às moléculas de ácido nucleico aqui descritas usando as seqüências ou sua parte como sonda de hibridização e segundo técnicas de hibridização padrão sob condições de hibridização estringentes. Neste contexto, é possível usar, por exemplo, moléculas de ácido nucleico isoladas de pelo menos 15, 20, 25, 30, 35, 40, 50, 60 ou mais nucleotídeos, de

preferência de pelo menos 15, 20 ou 25 nucleotídeos em comprimento que hibridizam sob condições estridentes com as moléculas de ácido nucleico acima descritas, em particular com aquelas que incluem uma sequência de nucleotídeos da molécula de ácido nucleico usada no processo da invenção ou
5 codificadora de uma proteína usada na invenção ou da molécula de ácido nucleico da invenção. Moléculas de ácido nucleico com 30, 50, 100, 250 ou mais nucleotídeos também podem ser usadas.

As seqüências de ácido nucleico usadas no processo da invenção, que são mostradas no protocolo de seqüência em particular SEQ ID
10 NO: 45, 47, 49, 51, 53 ou 55 são vantajosamente introduzidas em um construto de ácido nucleico, preferivelmente um cassete de expressão, que torna possível a expressão de moléculas de ácido nucleico em uma planta.

Conseqüentemente, a invenção também se refere a um construto de ácido nucleico, preferivelmente a um construto de expressão,
15 compreendendo a molécula de ácido nucleico da presente invenção funcionalmente ligada em um ou mais sinais ou elementos regulatórios.

Como aqui descrito, o construto de ácido nucleico também pode compreender outros genes, que são para serem introduzidos nos organismos ou nas células. É possível e vantajoso introduzir nos, e expressar
20 nos, organismos hospedeiros genes regulatórios tais como genes para indutores, repressores ou enzimas, que, devido à sua atividade enzimática, participam na regulação de um ou mais genes de uma rota biossintética. Estes genes podem ser de origem heteróloga ou homóloga. Além disso, outros genes de biossíntese podem estar vantajosamente presentes, ou então estes
25 genes podem estar localizados em um ou mais outros construtos de ácido nucleico. Genes, que são vantajosamente empregados são genes, que influenciam o crescimento de plantas tais como fatores ou seqüências reguladores(as). Uma intensificação de elementos reguladores pode ocorrer vantajosamente em nível de transcrição pelo uso de sinais de transcrição tais

como promotores e/ou intensificadores. Em adição, contudo, uma intensificação de tradução também é possível, por exemplo pelo aumento de estabilidade de mRNA ou pela inserção de uma sequência intensificadora de tradução.

5 Em princípio, o construto de ácido nucleico pode compreender as seqüências reguladoras aqui descritas e outras seqüências relevantes para a expressão dos genes incluídos. Assim, o construto de ácido nucleico da invenção pode ser usado como cassete de expressão e portanto pode ser utilizado diretamente para introdução na planta, ou então podem ser
10 introduzidos em um vetor. Conseqüentemente em uma modalidade o construto de ácido nucleico é um cassete de expressão compreendendo um promotor de microorganismo ou um terminador de microorganismo ou ambos. Em outra modalidade o cassete de expressão inclui um promotor de planta ou um terminador de planta ou ambos.

15 Para introduzir uma molécula de ácido nucleico em um construto de ácido nucleico, e.g. como parte de um cassete de expressão, o segmento de gene codogênico é vantajosamente submetido a uma reação de amplificação ou ligação na maneira conhecida por uma pessoa experiente na técnica. É preferido seguir um procedimento similar ao protocolo para uma
20 Pfu DNA polimerase ou uma mistura de Pfu/Taq DNA polimerases. Os iniciadores são selecionados de acordo com a seqüência a ser amplificada. Os iniciadores devem ser expedientemente escolhidos em uma tal maneira que o amplificado compreenda a seqüência codogênica do códon de iniciação para o códon de terminação. Após a amplificação, o amplificado é expedientemente
25 analisado. Por exemplo, a análise pode considerar qualidade e quantidade e ser realizada após separação por eletroforese em gel. Depois, o amplificado pode ser purificado seguindo um protocolo padrão (por exemplo Qiagen). Uma alíquota do amplificado purificado está então disponível para a etapa de clonagem subsequente. O técnico experiente geralmente conhece vetores de

clonagem adequados.

Incluem, em particular, vetores que são capazes de replicação em sistemas de clonagem fáceis de manusear como sistemas baseados em levedura bacteriana ou células de inseto (e.g. expressão em baculovírus), quer
5 dizer especialmente vetores que garantem clonagem eficiente em *E. coli*, e que tornam possível a transformação estável de plantas. Vetores, que precisam ser mencionados, em particular são vários sistemas de vetor binário ou co-integrado, que são adequados para a transformação mediada por T-DNA. Tais sistemas de vetor são geralmente caracterizados pelo fato de
10 conterem pelo menos os genes *vir*, que são exigidos para a transformação mediada por *Agrobacterium*, e as seqüências de borda de T-DNA.

Em geral, sistemas de vetor preferivelmente também compreendem outras regiões cis-regulatórias tais como promotores e terminadores e/ou marcadores de seleção por meio dos quais organismos
15 adequadamente transformados podem ser identificados. Embora os genes *vir* e as seqüências de T-DNA estejam localizados no mesmo vetor no caso de sistemas de vetor co-integrado, sistemas binários são baseados em pelo menos dois vetores, um dos quais possui genes *vir*, mas não T-DNA, enquanto que um segundo tem T-DNA, mas não gene *vir*. Devido a este fato, os últimos
20 vetores mencionados são relativamente pequenos, fáceis de manipular e capazes de replicação em *E. coli* e em *Agrobacterium*. Estes vetores binários incluem os vetores das séries pBIB-HYG, pPZP, pBecks, pGreen. Aqueles que são preferivelmente usados de acordo com a invenção são Bin19, pBI101, pBinAR, pGPTV e pCAMBIA. Um resumo de vetores binários e seu uso é
25 dado por Hellens et al., Trends in Plant Science (2000) 5, 446-451. Os vetores são preferivelmente modificados em uma tal maneira, que já contêm o ácido nucleico codificador do peptídeo de trânsito e que os ácidos nucleicos da invenção, preferivelmente as seqüências de ácido nucleico codificadoras dos polipeptídeos mostradas em tabela II, aplicação N^o 1, colunas 5 e 7 podem ser

clonados 3' iniciador da sequência codificadora de peptídeo de trânsito, dando uma preproteína funcional, que é direcionada para os plastídeos e que significa que a proteína madura cumpre com sua atividade biológica.

Em um vetor de expressão recombinante, “ligação operável”
 5 significa que a molécula de ácido nucleico de interesse está ligada nos sinais regulatórios em uma tal maneira que é possível a expressão da molécula de ácido nucleico: são ligadas uma na outra em um tal modo que duas sequências cumprem com a função predita atribuída à sequência (por exemplo em um sistema de transcrição/tradução in-vitro, ou em uma célula hospedeira se o
 10 vetor for introduzido na célula hospedeira).

O termo porção como aqui definido refere-se a um pedaço de um DNA codificador de uma CDK, compreendendo pelo menos 300, 350, 400, 450 ou 500, preferivelmente 550, 600, 650 ou 700 nucleotídeos e cuja porção codifica um polipeptídeo tendo atividade de quinase dependente de
 15 ciclina, tendo um motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE e tendo um sítio ativo da seguinte sequência VLHRDLKPQNLLI, sendo que D é o resíduo catalítico predito e sendo que as seguintes modificações de dita sequência podem ocorrer: posição 1: Val → Phe; posição 6: Leu → Met; posição 9: Gln → Asn. Ademias, dita sequência CDK pode ter
 20 vantajosamente um sítio de ligação de ATP do seguinte IGEGTYGVVYRARDKVTNETIALK. Também, no sítio de ligação de ATP algumas mutações podem ocorrer. Elas são as seguintes: posição 11 Arg → Lys; posição 12 Ala → Gly, posição 13 Arg → Leu, posição 15 Lys → Arg e posição 16 Val → Leu, Ala, Ser, Thr ou Asn.. Uma porção pode ser
 25 preparada, por exemplo, pela realização de uma ou mais deleções em um ácido nucleico CDK. As porções podem ser usadas em forma isolada ou podem ser fusionadas em outras sequências codificadoras (ou não-codificadoras) com o propósito de, por exemplo, produzir uma proteína que combina várias atividades, uma delas sendo atividade de quinase dependente

de ciclina. Quando fusionado em outras seqüências codificadoras, o polipeptídeo resultante produzido da tradução pode ser maior do que aquele predito para o fragmento de CDK. Preferivelmente, a porção funcional é uma porção de um ácido nucleico CDK, mais preferivelmente uma porção da molécula de ácido nucleico conforme representada por SEQ ID NO: 45, 47, 49, 51, 53 ou 55.

Os termos “fragmento”, “fragmento de uma seqüência” ou “parte de uma seqüência”, “porção” ou “sua porção” significam uma seqüência truncada da seqüência original mencionada. A seqüência truncada (seqüência de ácido nucleico ou proteína) pode variar amplamente em comprimento; o tamanho mínimo sendo uma seqüência de tamanho suficiente para obter uma seqüência com pelo menos uma atividade e/ou função comparável da seqüência original citada ou hibridizar com a molécula de ácido nucleico da invenção ou usada no processo da invenção sob condições estridentes, enquanto que o tamanho máximo não é crítico. Em algumas aplicações, o tamanho máximo normalmente não é substancialmente maior do que aquele exigido para proporcionar a atividade e/ou função(ões) da seqüência original.

Tipicamente, a seqüência de aminoácidos truncada variará de cerca de 5 a cerca de 310 aminoácidos em comprimento. Mais tipicamente, contudo, a seqüência terá um máximo de cerca de 250 aminoácidos em comprimento, preferivelmente um máximo de cerca de 200 ou 100 aminoácidos. É normalmente desejável a seleção de seqüências de pelo menos cerca de 10, 12 ou 15 aminoácidos, a até um máximo de cerca de 20 ou 25 aminoácidos.

Outra variante de uma molécula de ácido nucleico CDK é uma molécula de ácido nucleico capaz de hibridizar sob condições de estridência reduzida, preferivelmente sob condições estridentes, com uma molécula de ácido nucleico CDK como definida aqui antes, cuja seqüência de hibridização

codifica um polipeptídeo CDK compreendendo os motivos acima mencionados. Preferivelmente, a sequência de hibridização é uma que é capaz de hibridizar com a molécula de ácido nucleico de SEQ ID NO: 45, 47, 49, 51, 53 ou 55, ou com um ácido nucleico codificador de um dos homólogos mencionados acima, ou com uma porção de qualquer uma das sequências acima citadas. Mais preferivelmente, a sequência de hibridização é capaz de hibridizar com a molécula de ácido nucleico de SEQ ID NO: 45, 47, 49, 51, 53 ou 55.

O termo “hibridização” como aqui definido é um processo no qual sequências de nucleotídeos complementares substancialmente homólogas anelam-se umas nas outras. O processo de hibridização pode ocorrer inteiramente em solução, i.e. ambos os ácidos nucleicos complementares estão em solução. O processo de hibridização também pode ocorrer com um dos ácidos nucleicos complementares imobilizado em uma matriz tal como glóbulos magnéticos, glóbulos de Sepharose ou qualquer outra resina. O processo de hibridização pode adicionalmente ocorrer com um dos ácidos nucleicos complementares imobilizado em um suporte sólido tal como uma membrana de nitrocelulose ou de náilon ou imobilizado e.g. por fotolitografia em, por exemplo, um suporte de vidro silicoso (o último também conhecido como microarranjos ou arranjos de ácido nucleico ou como pastilhas de ácido nucleico). Com o objetivo de permitir ocorrência de hibridização, as moléculas de ácido nucleico são geralmente térmica ou quimicamente desnaturadas para fundir uma fita dupla em duas fitas únicas e/ou para remover grampos-de-cabelo ou outras estruturas secundárias dos ácidos nucleicos de fita única. A estringência de hibridização é influenciada por condições tais como temperatura, concentração de sal, força iônica e composição de tampão de hibridização.

“Condições hibridização estringentes” e “condições estringentes de lavagem em hibridização” no contexto de experimentos de

hibridização de ácido nucleico hibridização tais como hibridizações de Southern e Northern são dependentes de seqüência são diferentes sob parâmetros ambientais diferentes. O técnico experiente está ciente dos vários parâmetros que podem ser alterados durante hibridização e lavagem e que quer manterão quer mudarão as condições de estringência.

A T_m é a temperatura sob força iônica e pH definidos, na qual 50% da seqüência alvo hibridiza com uma sonda perfeitamente combinada. A T_m é dependente das condições de solução e da composição de base e do comprimento da sonda. Por exemplo, seqüências mais longas hibridizam especialmente em temperaturas mais altas. A taxa de hibridização máxima é obtida de cerca de 16°C a 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 ácido nucleico favorecendo deste modo a formação de híbrido; este efeito é visível para concentrações de sódio de até 0,4M. Formamida reduz a temperatura de fusão de dúpiques de DNA-DNA e DNA-RNA em 0,6 a 0,7°C para cada percentagem de formamida, e adição de formamida 50% permite que hibridização seja realizada a de 30 a 45°C, embora a taxa de hibridização seja abaixada. Não-combinações de pares de bases reduzem a taxa de hibridização e a estabilidade térmica dos dúpiques. Em média e para sondas grandes, a T_m diminui cerca de 1°C por % de não-combinação de base. A T_m pode ser calculada usando as seguintes equações, dependendo dos tipos de híbridos:

- Híbridos de DNA-DNA (Meinkoth e Wahl, Anal. Biochem., 138: 267-284, 1984): $T_m = 81,5^\circ\text{C} + 16,6 \times \log[\text{Na}^+]^a + 0,41 \times \%[\text{G/C}^b] - 500 \times [\text{L}^c]^{-1} - 0,61 \times \% \text{formamida}$
- Híbridos de DNA-RNA ou RNA-RNA: $T_m = 79,8 + 18,5 (\log_{10}[\text{Na}^+]^a) + 0,58 (\% \text{G/C}^b) + 11,8 (\% \text{G/C}^b)^2 - 820/\text{L}^c$
- Híbridos de oligo-DNA ou oligo-RNA^d:
Para <20 nucleotídeos: $T_m = 2 (I_n)$

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

^a ou para outro cátion monovalente, mas apenas exato na faixa de 0,01-0,4 M.

^b apenas exato para %GC na faixa de 30% a 75%.

^c L = comprimento do duplex de pares de bases.

5 ^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 uréia 6M reduz a T_m em cerca de 30°C.

Especificidade de hibridização é tipicamente a função de lavagens após hibridização. Para remover genótipo residual resultante de hibridização não específica, amostras são lavadas com soluções de sal diluídas. Fatores críticos de tais lavagens incluem a força iônica e a temperatura da solução de lavagem final: quanto menor a concentração de sal e maior a temperatura ambiente de lavagem, maior será a estringência de lavagem. Condições de lavagem são tipicamente realizadas em ou abaixo de estringência de hibridização. Geralmente, condições estringentes adequadas para ensaios de hibridização de ácido nucleico ou procedimentos de detecção de amplificação de gene são como descritas acima. Condições mais ou menos estringentes também podem ser selecionadas. Geralmente, condições de estringência baixa são selecionadas para serem cerca de 50°C mais baixas do que o ponto de fusão térmico (T_m) para a sequência específica em um pH e em uma força iônica definidos. Condições de estringência média são quando a temperatura está 20°C abaixo de T_m , e condições de estringência alta são quando a temperatura está 10°C abaixo de T_m . Por exemplo, condições estringentes são aquelas que são pelo menos tão estringentes quanto, por exemplo, condições A-L; e condições de estringência reduzida são pelo menos tão estringentes quanto, por exemplo, condições M-R. Ligação não-específica pode ser controlada usando qualquer uma de numerosas técnicas conhecidas tais como, por exemplo, bloqueio da membrana com soluções contendo proteína, adições de RNA, DNA heterólogos, e SDS no tampão de hibridização, e tratamento com Rnase.

Exemplos de condições de hibridização e de lavagem são

listadas em Tabela 1:

Tabela 1:

Condições de Estringência	Híbrido de Polinucleotídeo [‡]	Comprimento do Híbrido (pb)	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 formamida 50%	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 formamida 50%	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 formamida 50%	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 formamida 50%	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 formamida 50%	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 formamida 50%	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 formamida 50%	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 formamida 50%	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 formamida 50%	60°C.; 2×SSC
R	RNA:RNA	<50	Tr*; 4 ×SSC	Tr*; 4×SSC

[‡] O “comprimento do híbrido” é o comprimento previsto para o ácido nucleico hibridizando. Quando ácidos nucleicos de sequência conhecida são hibridizados, o comprimento do híbrido pode ser determinado pelo alinhamento das seqüências e identificação das regiões conservadas aqui descritas.

[†] SSPE (1×SSPE é NaCl 0,15M, NaH₂PO₄ 10mM, e EDTA 1,25mM, pH7,4) pode substituir SSC (1×SSC é NaCl 0,15M e citrato de sódio 15mM) nos tampões de hibridização e lavagem; lavagens são realizadas por 15 minutos após completude da hibridização. As hibridizações e lavagens podem adicionalmente incluir reagente de Denhard 5 x, SDS 0,5-1,0%, DNA de esperma de salmão fragmentado, desnaturado 100 µg/ml, pirofosfato de sódio 0,5%, e formamida até 50%.

* Tb-Tr: A temperatura de hibridização para híbridos previstos em serem menores do que 50 pares de base em comprimento deve ser 5-10°C mais baixa 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 inclui a substituição de qualquer um, ou mais parceiros de híbrido de DNA ou RNA quer com um PNA, quer com um ácido nucleico modificado.

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

Após hibridização e lavagem, os duplícies podem ser detectados por auto-radiografia (quando sondas radiomarcadas foram usadas) ou por quimioluminescência, imunodeteção, por detecção fluorescente ou cromogênica, dependendo do tipo de marcação de sonda. Alternativamente, um ensaio de proteção de ribonuclease pode ser realizado para detecção de híbridos RNA:RNA.

A molécula de ácido nucleico CDK ou sua variante pode ser derivada de qualquer planta ou fonte artificial. Este ácido nucleico pode ser modificado de sua forma nativa em composição e/ou ambiente genômico através de manipulação humana deliberada. O ácido nucleico é de preferência de origem de planta, da mesma espécie de planta (por exemplo em relação a uma na qual é para ser introduzido) ou de uma espécie de planta diferente. O ácido nucleico pode ser isolado de uma espécie monocotiledônea, preferivelmente da família Poaceae, mais preferivelmente de *Oryza sativa* ou *Zea mays*. Mais preferivelmente, a CDK isolada de *Oryza sativa* é SEQ ID NO: 45 ou de *Zea mays* e é SEQ ID NO: 53. Em outra modalidade da invenção o ácido nucleico pode ser isolado de uma espécie dicotiledônea, preferivelmente da família Brassicaceae, Aceraceae, Linaceae ou Asteraceae adicionalmente preferivelmente de *Brassica napus*, *Glycine max*, *Linum usitatissimum* ou *Helianthus annuus*. Mais preferivelmente, a CDK isolada de *Brassica napus* é SEQ ID NO: 47, *Glycine max* é SEQ ID NO: 49, *Linum usitatissimum* é SEQ ID NO: 51 ou *Helianthus annuus* é SEQ ID NO: 55.

A atividade de um polipeptídeo CDK ou de um seu homólogo

e/ou expressão de um ácido nucleico codificador de uma tal CDK pode ser modulada por introdução de uma modificação genético (preferivelmente no locus de um gene CDK). O locus de um gene como aqui definido é adotado para significar uma região genômica, que inclui o gene de interesse e 10 kb a montante ou a jusante da região codificadora.

A modificação genética pode ser introduzida, por exemplo, por qualquer um (ou mais) dos seguintes métodos: TILLING, mutagênese sítio-direcionada, evolução dirigida e recombinação homóloga ou por introdução e expressão em uma planta de um ácido nucleico codificador de um polipeptídeo CDK ou de um seu homólogo, cuja(o) CDK ou homólogo compreende um motivo como mencionado acima. Após introdução da modificação genética segue uma etapa de seleção da expressão aumentada de um ácido nucleico codificador de um polipeptídeo CDK com um motivo como mencionado acima e/ou seleção de atividade aumentada de dito polipeptídeo CDK, com aumento em expressão e/ou atividade dá plantas tendo características de crescimento melhoradas.

Uma modificação genética também pode ser introduzida no locus de um gene CDK usando a técnica de TILLING (“Targeted Induced Local Lesions In Genomes”). Esta é uma tecnologia de mutagênese útil para gerar e/ou identificar, e para eventualmente isolar variantes mutadas de uma molécula de ácido nucleico codificadora de uma CDK com seqüências como aqui mencionadas capazes de exibirem atividade de quinase dependente de ciclina. TILLING também permite seleção de plantas trazendo tais variantes mutantes. TILLING combina mutagênese de densidade alta com métodos de seleção de produtividade elevada. As etapas tipicamente seguidas em TILLING são: (a) mutagênese EMS (Redei e Koncz (1992), em: C Koncz, N-H Chua, J Schell, eds, *Methods in Arabidopsis Research*. World Scientific, Singapore, pp 16-82; Feldmann et al., (1994) em: EM Meyerowitz, CR Somerville, eds, *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold

Spring Harbor, NY, pp 137-172; Lightner e Caspar (1998), em: J Martinez-Zapater, J Salinas, eds, *Methods on Molecular Biology*, Vol. 82. Humana Press, Totowa, NJ, pp 91-104); (b) preparação de DNA e agrupamento de indivíduos; (c) amplificação por PCR de uma região de interesse; (d) 5 desnaturação e anelamento para permitir formação de heterodúplex; (e) DHPLC, onde a presença de um heteroduplex em um agrupamento é detectada como um pico extra no cromatograma; (f) identificação de indivíduo mutante; e (g) seqüenciamento do produto mutante de PCR. Métodos para TILLING são bem conhecidos na técnica (McCallum Nature 10 Biotechnol. 18, 455-457, 2000, Stemple Nature Rev. Genet. 5, 145-150, 2004).

Mutagênese sítio-direcionada pode ser usada para gerar variantes de ácido nucleicos CDK ou suas porções que retêm atividade (tal como atividade de quinase dependente de ciclina). Estão 15 disponíveis vários métodos para realizar mutagênese sítio-direcionada, os mais comuns sendo métodos baseados em PCR (veja por exemplo Ausubel et al., *Current Protocols in Molecular Biology*. Wiley Eds. <http://www.4ulr.com/products/currentprotocols/index.html>).

Evolução dirigida também pode ser usada para gerar variantes 20 de ácidos nucleicos CDK. Esta consiste de iterações de rearranjo de DNA seguidas por triagem e/ou seleção apropriada para gerar variantes de ácidos nucleicos CDK ou suas porções codificadores de polipeptídeos CDK ou seus homólogos ou suas porções tendo uma atividade biológica modificada (Castle et al., (2004) *Science* 304(5674): 1151-4; US 5,811,238 e US 6,395,547).

25 TILLING, mutagênese sítio-direcionada e evolução dirigida são exemplos de tecnologias que permitem a geração de novos alelos e variantes de CDK que retêm a função de CDK e que são portanto úteis nos métodos da invenção.

Recombinação homóloga permite a introdução em um genoma

de um ácido nucleico selecionado em uma posição selecionada definida. Recombinação homóloga é uma tecnologia padrão usada rotineiramente em ciência biológica para organismos inferiores tais como levedura ou o musgo *Physcomitrella*. Métodos para realizar recombinação homóloga em plantas

5 têm sido descritos não apenas para plantas modelo (Offringa et al. (1990) EMBO J. 9, 3077-3084) mas também para plantas de colheita, por exemplo arroz (Terada et al., (2002) *Nature Biotechnol.* 20, 1030-1034; ou Iida e Terada (2004) *Curr. Opin. Biotechnol.* 15, 132-138). O ácido nucleico a ser selecionado (que pode ser uma molécula de ácido nucleico CDK ou sua

10 variante como definido aqui acima) não precisa ser selecionado para o locus de um gene CDK, mas pode ser introduzido, por exemplo, nas regiões de expressão alta. O ácido nucleico a ser selecionado pode ser um alelo melhorado usado para substituir genes endógenos ou pode ser introduzido em adição ao gene endógeno.

15 Um método preferido para introdução de uma modificação genética (que neste caso não precisa estar no locus de um gene CDK) é introdução e expressão em uma planta de um ácido nucleico codificador de um polipeptídeo CDK, ou de um seu homólogo. Um polipeptídeo CDK ou um seu homólogo como mencionado acima, e adequado para praticar a

20 presente invenção, é um tendo atividade de quinase dependente de ciclina e, em ordem crescente de preferência, tendo pelo menos 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% ou mais de identidade de sequência em relação à sequência de aminoácidos representada por 46, 48, 50, 52, 54 ou 56, e cujo polipeptídeo CDK compreende um motivo como aqui descrito. O ácido

25 nucleico a ser introduzido em uma planta pode ser uma porção ou uma sequência de hibridização como definida aqui anteriormente.

“Homólogos” de uma proteína incluem peptídeos, oligopeptídeos, polipeptídeos, proteínas e enzimas tendo substituições, deleções e/ou inserções de aminoácido em relação à proteína não modificada

em questão e tendo atividade biológica e funcional similar à da proteína não modificada da qual são derivados. Isto significa que têm um antepassado comum.

Estão incluídas no termo “homólogos” as seqüências parálogas e ortólogas, duas formas especiais de homologia, que incluem conceitos evolucionários usados para descrever relações ancestrais de genes. Preferivelmente ortólogos ou parálogos úteis na presente invenção têm as mesmas estrutura e atividade que as de uma CDK e têm a similaridade mais SEQ ID NO: 46, 48, 50, 52, 54 ou 56 em uma pesquisa BLAST recíproca.

O termo “parálogos” refere-se aos genes homólogos que resultam de uma ou mais duplicações de gene dentro do genoma de uma espécie. Parálogos de uma CDK podem ser facilmente identificados por realização de análise BLAST contra um conjunto de seqüências de mesma espécie que a seqüência em questão.

O termo “ortólogos” refere-se aos genes homólogos em organismos diferentes devido à relação antepassado destes genes. Ortólogos, por exemplo, em espécies de plantas monocotiledôneas ou dicotiledôneas podem ser facilmente encontrados pela realização de pesquisa *blast* recíproca. Esta pode ser feita por um primeiro *blast* envolvendo *blasting* da seqüência em questão (por exemplo, SEQ ID NO 45, 47, 49, 51, 53 ou 55, sendo de espécie monocotiledônea *Oryza sativa* ou *Zea mays* ou de espécie dicotiledônea *Brassica napus*, *Glycine max*, *Linum usitatissimum* ou *Helianthus annuus*) contra qualquer banco de dados de seqüências, tal como o banco de dados publicamente disponível NCBI que pode ser encontrado em: <http://www.ncbi.nlm.nih.gov>. BLASTn ou tBLASTX pode ser usado quando se começa dos nucleotídeos ou BLASTP ou TBLASTN quando se começa da proteína, com valores pré-estabelecidos padrão. Os resultados de *blast* podem ser filtrados. As seqüências de comprimento total quer de resultados filtrados quer de resultados não filtrados são então *blasted* de volta (segundo *blast*)

contra as seqüências do organismo do qual a seqüência em questão é derivada, no caso *Oryza sativa*, *Zea mays*, *Brassica napus*, *Glycine max*, *Linum usitatissimum* ou *Helianthus annuus*. Os resultados dos primeiro e segundo *blasts* são então comparados. Um parólogo é identificado se um êxito de classificação alta do segundo *blast* for de mesma espécie que aquela da qual a seqüência em questão é derivada; um ortólogo é identificado se um êxito de classificação mais alta não for da mesma espécie que aquela da qual a seqüência em questão é derivada. Um tal parólogo ou ortólogo também é considerado um homólogo de CDK, desde que este homólogo compreenda um domínio de serina/treonina quinase e compreenda um motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE. No caso de famílias grandes, ClustalW pode ser usado, seguido pela construção de uma árvore de junção de vizinhos, para ajudar a visualizar o agrupamento de genes relacionados e identificar ortólogos e parálogos.

Um homólogo pode estar na forma de uma “variante substituinte” de uma proteína, i.e. onde pelo menos um resíduo nem uma seqüência de aminoácidos tem sido removido e um resíduo diferente inserido em seu lugar. Substituições de aminoácido são tipicamente de resíduos únicos, mas podem ser agrupados dependendo das restrições funcionais aplicadas ao polipeptídeo; inserções normalmente serão da ordem de cerca de 1, 2, 3, 4 ou 5, preferivelmente 6, 7, 8, 9 ou 10 resíduos de aminoácido. Preferivelmente, substituições de aminoácido compreendem substituições de aminoácido conservativas (Tabela 2). Para produzir tais homólogos, aminoácidos da proteína podem ser substituídos por outros aminoácidos tendo propriedades similares (tais como similares hidrofobicidade, hidrofiliidade, antigenicidade, propensão para formar ou quebrar estruturas α -helicoidais ou estruturas de folha- β). Tabelas de substituições conservativas são bem conhecidas na técnica (veja por exemplo Creighton (1984) *Proteins*. W.H. Freeman and Company). A variante substituinte útil nos métodos da presente

invenção é uma variante substituinte de um polipeptídeo CDK e compreende um motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE e os outros motivos acima mencionados.

Tabela 2: Exemplos de substituições de aminoácido conservativas:

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

5 Substituições menos conservativas podem ser feitas no caso de as propriedades dos aminoácidos acima mencionados não serem tão críticas.

Um homólogo também pode estar na forma de uma “variante de inserção” de uma proteína, i.e. onde um ou mais resíduos de aminoácido são introduzidos em um sítio predeterminado na proteína. Inserções podem

10 compreender fusões amino-terminais e/ou carbóxi-terminais bem como inserções de intra-seqüência de único aminoácido ou de múltiplos aminoácidos. Geralmente, inserções dentro da seqüência de aminoácidos serão menores do que as fusões amino- ou carbóxi-terminais, da ordem de cerca de 1 a 10 resíduos. Exemplos de peptídeos ou proteínas de fusão amino-

15 ou carbóxi-terminal incluem o domínio de ligação ou o domínio de ativação de um ativador de transcrição como usado no sistema bi-híbrido de levedura, proteínas de capa de fago, etiqueta (histidina)₆, etiqueta glutathione S-transferase, proteína A, proteína ligante de maltose, di-hidro-folato-redutase, epitopo Tag-100, epitopo c-myc, epitopo FLAG®, lacZ, CMP (peptídeo

20 ligante de calmodulina), epitopo HA, epitopo de proteína C e epitopo VSV. A variante de inserção útil nos métodos da presente invenção é uma variante de inserção de um polipeptídeo CDK e compreende os motivos aqui mencionados.

Homólogos na forma de “variantes de deleção” de uma proteína são caracterizados pela remoção de um ou mais aminoácidos de uma proteína, e incluem fragmentos ativos.

5 Variantes de aminoácido de uma proteína podem ser prontamente preparadas usando técnicas de síntese de peptídeo bem conhecidas na técnica, tais como síntese de peptídeo em fase sólida e semelhante, ou por manipulações de DNA recombinante. Métodos para manipulação de uma sequência de DNA para produzir variantes de substituição, inserção ou deleção de uma proteína são bem conhecidos na
10 técnica. Por exemplo, técnicas para realizar mutações em sítios predeterminados em DNA são bem conhecidos por aqueles pessoas experientes na técnica e incluem mutagênese M13, mutagênese T7-Gen in vitro (USB, Cleveland, OH), mutagênese Sítio-Direcionada QuickChange (Stratagene, San Diego, CA), mutagênese sítio-direcionada mediada por PCR
15 ou outros protocolos de mutagênese sítio-direcionada.

O polipeptídeo CDK ou seu homólogo com um motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE, também podem ser um derivado. “Derivados” incluem peptídeos, oligopeptídeos, polipeptídeos, proteínas e enzimas que podem compreender substituições, deleções ou adições de
20 resíduos de aminoácido naturalmente e não-naturalmente ocorrentes comparados com a sequência de aminoácidos de uma forma naturalmente ocorrente da proteína, por exemplo, como apresentado em seqüências 46, 48, 50, 52, 54 ou 56. “Derivados” de uma proteína incluem peptídeos, oligopeptídeos, polipeptídeos, proteínas e enzimas que podem compreender
25 resíduos de aminoácido naturalmente ocorrentes alterados, glicosilados, acetilados ou não-naturalmente ocorrentes comparados com a sequência de aminoácidos de uma forma naturalmente ocorrente da qual é derivado, por exemplo uma molécula repórter ou outro ligante, covalente ou não-covalentemente ligado na sequência de aminoácidos, tal como uma molécula

repórter que é ligada para facilitar sua detecção, e resíduos de aminoácido não-naturalmente ocorrentes em relação à sequência de aminoácidos de uma proteína naturalmente ocorrente. O derivado útil nos métodos da presente invenção é um derivado de um polipeptídeo CDK, tendo a atividade biológica das CDKs e os motivos aqui mencionados.

As quinases do tipo CDK em plantas têm uma estrutura modular, consistindo de um lobo-N e um lobo-C compreendendo uma fenda catalítica e uma alça-T (De Bondt et al. 1993). Portanto, é imaginado que engenharia dos domínios da quinase em uma tal maneira que a atividade da proteína CDK é mantida ou modificada, pode resultar na criação de um mutante CDKA que é útil para realizar os métodos da invenção. Um tipo preferido de variante inclui aquelas geradas por deleção, empilhamento ou rearranjo de domínio (veja por exemplo He et al., Science 288, 2360-2363, 2000; ou patetes US 5.811.238 e 6.395.547), desde que a CDK resultante compreenda (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE, motivo central ativo e ligante de ATP.

O polipeptídeo CDK ou seu homólogo pode ser codificado por uma variante de união alternativa de uma molécula de ácido nucleico CDK ou gene. O termo “variante de união alternativa” como aqui usado inclui variantes de uma sequência de ácido nucleico nas quais éxons e/ou íntrons selecionados têm sido excisados, substituídos ou adicionados. Tais variantes serão aquelas que codificam polipeptídeos que compreendem mutações e nas quais a atividade biológica da proteína é mantida, o que pode ser alcançado pela retenção seletiva de segmentos funcionais da proteína. Tais variantes de união podem ser encontradas em natureza ou podem ser preparadas. Métodos de preparação de tais variantes de união são bem conhecidas na técnica. Variantes de união preferidas são variantes de união derivadas do ácido nucleico representado por SEQ ID NO 45, 47, 49, 51, 53 ou 55. Mais preferidas são as variantes de união codificadoras de um polipeptídeo

mantendo a atividade de quinase dependente de ciclina e tendo os motivos como aqui mencionados.

O homólogo também pode ser codificado por uma variante alélica de um ácido nucleico codificador de um polipeptídeo CDK ou de um seu homólogo, preferivelmente uma variante alélica do ácido nucleico representado por SEQ ID NO 45, 47, 49, 51, 53 ou 55, desde que o polipeptídeo codificado pela variante alélica tem atividade de quinase dependente de ciclina e compreende os motivos como mencionados acima. Variantes alélicas existem na natureza, e o uso destes alelos naturais está incluído nos métodos da presente invenção. Variantes alélicas incluem Polimorfismos de Nucleotídeo Único (SNPs), bem como Polimorfismos de Inserção/Deleção Pequena (INDELs). O tamanho de INDELs é normalmente menor do que 100 pb. SNPs e INDELs formam o maior conjunto de variantes de seqüência em cepas polimórficas da maioria dos organismos.

De acordo com um aspecto preferido da presente invenção, é considerada a expressão intensificada ou aumentada da molécula de ácido nucleico CDK ou de sua variante de acordo com a invenção. Métodos para obter expressão intensificada ou aumentada (sobreexpressão) de genes ou de produtos de gene estão bem documentados na técnica e incluem, por exemplo, sobreexpressão conduzida por promotores apropriados, o uso de intensificadores de transcrição e de intensificadores de tradução. Ácidos nucleicos isolados que servem como elementos promotores ou intensificadores podem ser introduzidos em uma posição apropriada (tipicamente a montante) de uma forma não-heteróloga de modo a suprarregular a expressão de um ácido nucleico CDK ou de sua variante de acordo com a invenção. Por exemplo, promotores endógenos podem ser alterados in vivo por mutação, deleção, e/ou substituição (veja, Kmiec, Pat. U.S. Nº 5,565,350; Zarling et al., PCT/US93/03868), ou promotores isolados podem ser introduzidos em uma célula de planta nas apropriadas orientação e

distância de um gene modificado de acordo com a presente invenção de maneira a controlar a expressão do gene.

Se expressão de polipeptídeo é desejada, é geralmente desejável incluir uma região de poliadenilação na extremidade- 3' de uma região codificadora de polinucleotídeo. A região de poliadenilação pode ser derivada do gene natural, de uma variedade de outros genes de planta, ou de T-DNA. A sequência de extremidade 3' a ser adicionada pode ser derivada de, por exemplo, os genes de nopalina sintase ou octopina sintase, ou alternativamente de outro gene de planta, ou menos preferivelmente de qualquer outro gene eucariótico.

Uma sequência de íntron também pode ser adicionada na região não-traduzida 5' ou na sequência codificadora da sequência codificadora parcial para aumentar a quantidade da mensagem madura que acumula no citosol. Tem sido mostrado que a inclusão de um íntron editorável na unidade de transcrição em construtos de expressão tanto de planta quanto de animal aumenta a expressão de gene em ambos os níveis de mRNA e de proteína em até 1.000 vezes (Buchman e Berg, Mol. Cell Biol. 8, 4395-4405 (1988); Callis et al., Genes Dev. 1, 1183-1200 (1987)). Tal intensificação de íntron da expressão de gene é tipicamente mais elevada quando posicionada próxima da extremidade 5' da unidade de transcrição. São conhecidos na técnica os usos dos íntrons de milho: íntrons 1, 2, e 6 de Adh1-S, íntron Bronze-1. Veja geralmente, The Maize Handbook, Chapter 116, Freeling e Walbot, Eds., Springer, N.Y. (1994).

A invenção também proporciona construtos e vetores genéticos para facilitar a introdução e/ou a expressão das seqüências de nucleotídeos úteis nos métodos de acordo com a invenção.

Portanto, é proporcionado um construto de gene compreendendo:

(i) a molécula de ácido nucleico CDK ou sua variante

funcional, cujo(a) ácido nucleico ou variante codifica uma CDK compreendendo motivo de sítio ativo e de ligação de ATP, (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE;

5 (ii) uma ou mais seqüência(s) de controle capaz(es) de conduzir expressão em uma planta de uma seqüência de ácido nucleico de (i); e opcionalmente

(iii) uma seqüência de terminação de transcrição.

Construtos úteis nos métodos de acordo com a presente invenção podem ser construídos usando tecnologia de DNA recombinante bem conhecida pelas pessoas experientes na técnica. Os construtos de gene
10 podem ser inseridos em vetores, que podem ser comercialmente disponíveis, adequados para transformação em plantas e apropriados para expressão do gene de interesse nas células transformadas.

Plantas são transformadas com um vetor compreendendo a
15 seqüência de interesse (i.e., um ácido nucleico CDK ou sua variante de acordo com a presente invenção). A seqüência de interesse é operacionalmente ligada em uma ou mais seqüências de controle (pelo menos em um promotor). Os termos “elemento regulatório”, “seqüência(s) regulatória(s)”, “seqüência de controle” e “promoter” são todos usados intercambiavelmente aqui e são para
20 serem considerados em um contexto amplo para se referirem às seqüências regulatórias de ácido nucleico capazes de efetuarem expressão das seqüências nas quais são ligadas. Incluídas pelos termos acima mencionados estão as seqüências regulatórias de transcrição derivadas de um gene genômica eucariótico clássico (incluindo o box TATA que é exigido para iniciação de
25 transcrição exata, com ou sem uma seqüência de box CCAAT) e elementos regulatórios adicionais (i.e. silenciadores, intensificadores e seqüências ativadoras a montante) que modulam expressão de gene em resposta aos estímulos desenvolvendo e/ou externos, ou em uma maneira tecido-específica. Também está incluída dentro do termo uma seqüência regulatória

de transcrição de um gene procariótico clássico, em cujo caso pode incluir uma sequência de box -35 e/ou sequências regulatórias de transcrição de box -10. O termo “elemento regulatório” também inclui um derivado ou uma molécula de fusão sintética(o), que confere, ativa ou intensifica a expressão de uma molécula de ácido nucleico em uma célula, um tecido ou um órgão. O termo “operacionalmente ligado” como aqui usado refere-se a uma ligação funcional entre a sequência de promotor e o gene de interesse, de tal modo que a sequência de promotor é capaz de iniciar a transcrição do gene de interesse.

Sequências regulatórias podem ser operacionalmente ligadas na sequência codificadora de uma proteína transgênica ou endógena e controlam sua transcrição e/ou tradução ou a estabilidade ou deterioração do mRNA codificador ou da proteína expressada. Com o objetivo de modificar e controlar a expressão de uma sequência codificadora seus elementos regulatórios tais como promotores, UTRs, sítios de união, sinais de processamento, sítios de poliadenilação, terminadores, intensificadores, indutores, repressores, sítios de modificação de pós-transcrição ou de pós-tradução podem ser modificados, adicionados ou emendados. As sequências regulatórias incluem, em particular, sequências de planta como os promotores e terminadores aqui descritos. Por exemplo, a ativação de genes de planta por integrações aleatórias de elementos intensificadores tem sido descrita por Hayashi et al., 1992 (*Science* 258:1350-1353) ou Weigel et al., 2000 (*Plant Physiol.* 122, 1003-1013) e outros aqui citados. Por exemplo, o nível de expressão da proteína endógena pode ser modulado por substituição do promotor endógeno por um promotor transgênico mais forte ou pela substituição de 3'UTR endógeno por um 3'UTR, que proporciona maior estabilidade sem emendar a região codificadora. Adicionalmente, a regulação de transcrição pode ser modulada pela introdução de um fator de transcrição artificial como descrito nos exemplos. Promotores, terminadores e UTR

alternativos são descritos abaixo.

Seqüências regulatórias são intencionadas para permitir a expressão específica dos genes e a expressão de proteína. Dependendo da planta hospedeira, isto pode significar, por exemplo, que o gene é expressado e/ou sobreexpressado após indução apenas, ou que ele é expressado e/ou sobreexpressado constitutivamente. Estas seqüências regulatórias são, por exemplo, seqüências nas quais os indutores ou repressores se ligam e que portanto regulam a expressão do ácido nucleico. Em adição às seqüências regulatórias novas, ou em vez destas seqüências, a regulação natural destas seqüências ainda pode estar presente antes dos genes estruturais reais e, se apropriado, podem ter sido geneticamente modificadas de modo que a regulação natural tenha sido desligada e a expressão de gene tenha sido aumentada. Como uma regra, citadas seqüências regulatórias estão localizadas a montante (5'), dentro, e/ou a jusante (3') em relação à seqüência codificadora de uma seqüência de ácido nucleico, que deve ser expressada. Contudo, o construto de ácido nucleico (= cassete de expressão, construto de expressão ou construto de gene) usado no processo da invenção e aqui descrito também pode ser mais simples em construção, isto é sem sinais regulatórios adicionais tendo sido inseridos antes de uma seqüência de ácido nucleico ou seus derivados, e o promotor natural junto com sua regulação não tem sido removido. Em vez disso, a seqüência regulatória natural tem sido mutada em uma tal maneira que uma regulação não mais ocorre e/ou a expressão de gene é aumentada. Estes promotores modificados também podem ser introduzidos independentemente antes do gene natural na forma de seqüências de parte (= promotor com partes de uma seqüências de ácido nucleico de acordo com a invenção) com o propósito de aumentar a atividade. Além disso, o construto de gene também pode vantajosamente compreender uma ou mais do que é conhecido como seqüências intensificadoras em ligação operável com o promotor, e estas permitem uma expressão aumentada de uma

seqüência de ácido nucleico. Também, é possível inserir seqüências vantajosas adicionais na extremidade 3' da seqüência de DNA, tais como, por exemplo, outros terminadores ou elementos regulatórios.

Seqüências regulatórias incluem sinais ou seqüências regulatórias de transcrição e tradução, e.g. seqüências localizadas a montante (5'), que concernem em particular a regulação de iniciação de transcrição ou de tradução, tais como promotores ou códons de iniciação, e seqüências localizadas a jusante (3'), que concernem em particular a regulação de terminação de transcrição ou de tradução e estabilidade de transcrito, tais como códons de terminação e sinais de poliadenilação. Seqüências regulatórias também podem estar presentes em regiões codificadoras transcritas bem como em regiões não-codificadoras transcritas, e.g. em íntrons, como por exemplo sítios de união, promotores para a regulação de expressão da molécula de ácido nucleico de acordo com a invenção em uma célula e que podem ser empregadas, em princípio, todas aquelas que são capazes de estimular a transcrição de genes nas plantas em questão.

A “seqüência codificadora” é uma seqüência de nucleotídeos, que é transcrita em mRNA e/ou traduzida em um polipeptídeo quando posta sob o controle de seqüências regulatórias apropriadas. Os limites da seqüência codificadora são determinados por um códon de iniciação de tradução na terminação-5' e um códon de terminação de tradução na terminação-3'. Uma seqüência codificadora pode incluir, mas não é limitada a mRNA, cDNA, seqüências de nucleotídeos recombinantes ou DNA genômico, enquanto que íntrons podem estar presentes também sob certas circunstâncias.

Os fatores ou as seqüências regulatórias, como descritos acima, possuem um efeito positivo sobre, a expressão dos genes introduzidos, aumentando assim a sua expressão. Portanto, uma intensificação da expressão pode vantajosamente ocorrer em nível de transcrição pelo uso de sinais de transcrição fortes tais como promotores fortes e/ou intensificadores fortes. Em

adição, intensificação de expressão em nível de tradução também é possível, por exemplo pela introdução de seqüências intensificadoras de tradução, e.g., o intensificador Ω e.g. melhorando a ligação ribossômica no transcrito, ou pelo aumento da estabilidade do mRNA, e.g. pela substituição de região codificadora de 3'UTR por uma região codificadoras de uma 3'UTR conhecida em conferir uma estabilidade alta do transcrito ou por estabilização do transcrito através da eliminação de instabilidade de transcrito, de modo que a molécula de mRNA é traduzida mais freqüentemente do que a de tipo selvagem. Por exemplo em plantas elementos ricos em AU (AREs) e elementos DST (a jusante) desestabilizaram transcritos. Estudos de mutagênese têm demonstrado que resíduos dentro de dois dos domínios conservados, as regiões ATAGAT e GTA, são necessários para função de instabilidade. Portanto remoção ou mutação de tais elemento obviamente levaria a transcritos mais estáveis, taxas de transcrição maiores e atividade de proteína mais elevada. Intensificadores de tradução também são a “seqüência overdrive“, que compreende a seqüência líder 5'-não-traduzida do vírus do mosaico do tabaco e que aumenta a razão de proteína/RNA (Gallie et al., 1987, Nucl. Acids Research 15:8693-8711).

Intensificadores são geralmente definidos como elementos ativos-cis, que podem estimular transcrição de gene independente de posição e orientação. Intensificadores diferentes têm sido identificados em plantas, que podem estimular transcrição constitutivamente, ou tecido ou estímulos específicos. Exemplos bem conhecidos de intensificadores constitutivos são o intensificador do promotor 35S (Odell et al., 1985, Nature 313:810-812) ou o intensificador ocs (Fromm et al., 1989, Plant Cell 1: 977:984). Outros exemplos são o tetrâmero de motivo de box-G que confere expressão constitutiva de nível alto em plantas dicotiledôneas e monocotiledôneas (Ishige et al., 1999, Plant Journal, 18, 443-448) ou petE, uma seqüência rica em A/T que atuam como intensificadores quantitativos de expressão de gene

em plantas transgênicas batateira e tabaco (Sandhu et al., 1998; Plant Mol Biol. 37(5):885-96). Além disso, tem sido descrita uma grande variedade de elementos ativos-cis que contribuem para padrão de expressão específica, como expressão específica em órgão ou expressão induzida em resposta a estresse biótico ou abiótico. Exemplos são elementos, que proporcionam expressão induzida por ferimento ou patógeno (Rushton, 2002, Plant Cell, 14, 749-762) ou expressão específica em célula-guarda (Plesch, 2001, Plant Journal 28, 455-464).

Vantajosamente, qualquer tipo de promotor pode ser usado para conduzir a expressão de uma sequência de ácido nucleico. O promotor pode ser um promotor induzível, i.e. tendo iniciação de transcrição induzida ou aumentada em resposta a um estímulo desenvolvvente, químico, ambiental ou físico. Adicional ou alternativamente, o promotor pode ser um promotor constitutivo, i.e. um promotor que é expressado predominantemente em pelo menos um tecido ou um órgão e predominantemente em qualquer estágio de vida da planta. Adicional ou alternativamente, o promotor pode ser um promotor preferido de tecido ou preferido de célula, i.e. um que é capaz de preferivelmente iniciar transcrição em certos tecidos, tais como as folhas, raízes, tecido de semente etc., ou até mesmo em células específicas. Promotores capazes de iniciar transcrição apenas em certos tecidos são respectivamente aqui chamados de “tecido-específicos”, e “célula-específicos”.

Promotores adequados, que são funcionais nestas plantas, são geralmente conhecidos. Podem tomar a forma de promotores constitutivos ou induzíveis. Promotores adequados podem ser capazes de expressão desenvolvvente e/ou tecido-específica em eucariotos multicelulares; assim, promotores folha-, raiz-, flor-, semente-, estoma-, tubérculo- ou fruta-específicos podem ser vantajosamente usados em plantas.

Promotores de planta diferentes utilizáveis em plantas são

promotores tais como, por exemplo, o promotor USP, o promotor LegB4-, o promotor DC3 ou o promotor de ubiquitina de salsa.

Um promotor de “planta” compreende elementos regulatórios, que medeiam a expressão de um segmento de sequência codificadora em células de planta. Conseqüentemente, um promotor de planta não precisa ser
5 de origem de planta, mas pode se originar de vírus ou microorganismos, em particularmente por exemplo de vírus que atacam células de planta.

O promotor de “planta” também pode se originar de uma célula de planta, e.g. da planta, que é transformada com uma sequência de ácido nucleico a ser expressada no processo da invenção e aqui descrita. Isto
10 também se aplica a outros sinais regulatórios de “planta”, por exemplo em terminadores de “planta”.

Para expressão em plantas, a molécula de ácido nucleico precisa, como descrito acima, estar ligada operacionalmente em ou
15 compreender um promotor adequado que expressa o gene no instante de tempo certo e em uma maneira célula- ou tecido-específica. Promotores utilizáveis são promotores constitutivos (Benfey et al., EMBO J. 8 (1989) 2195-2202), tais como aqueles que se originam de vírus de planta, tais como 35S CAMV (Franck et al., Cell 21 (1980) 285-294), 19S CaMV (veja também US 5352605 e WO 84/02913), 34S FMV (Sanger et al., Plant. Mol. Biol., 14,
20 1990: 433-443), o promotor de ubiquitina de salsa, ou promotores de planta tais como o promotor de subunidade pequena Rubisco descritos em US 4.962.028 ou um promotor plantum PRP1 [Ward et al., Plant. Mol. Biol. 22 (1993)], SSU, PGEL1, OCS [Leisner (1988) Proc Natl Acad Sci USA 85(5):
25 2553-2557], lib4, usp, mas [Comai (1990) Plant Mol Biol 15 (3):373-381], STLS1, ScBV (Schenk (1999) Plant Mol Biol 39(6):1221-1230), B33, SAD1 ou SAD2 (promotores de linho, Jain et al., Crop Science, 39 (6), 1999: 1696-1701) ou nos [Shaw et al. (1984) Nucleic Acids Res. 12(20):7831-7846]. Outros exemplos de promotores constitutivos de planta são os promotores de

V-ATPase de beterraba (WO 01/14572). Exemplos de promotores constitutivos sintéticos são o promotor Super (WO 95/14098) e promotores derivados de boxes-G (WO 94/12015). Se apropriados, promotores induzíveis por compostos químicos também podem ser adicionalmente usados, comparar
5 EP-A 388186, EP-A 335528, WO 97/06268. Expressão constitutiva, estável das proteínas de acordo com a invenção de uma planta pode ser vantajosa. Contudo, expressão induzível do polipeptídeo da invenção é vantajosa, se uma expressão tardia antes da colheita for vantajosa, porque manipulação metabólica pode levar à retardação de crescimento de planta.

10 A expressão de genes de planta também pode ser facilitada via um promotor induzível por composto químico (para uma revisão, veja Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48:89-108). Promotores quimicamente induzíveis são particularmente adequados quando for desejado expressar o gene em uma maneira tempo-específica. Exemplos de tais
15 promotores são um promotor induzível por ácido salicílico (WO 95/19443), e promotor induzível por ácido abscísico (EP 335 528), um promotor induzível por tetraciclina (Gatz et al. (1992) Plant J. 2, 397-404), um promotor induzível por ciclo-hexanol ou etanol (WO 93/21334) ou outros como aqui descrito.

20 Outros promotores adequados são aqueles que reagem às condições de estresse biótico ou abiótico, por exemplo o promotor de gene PRP1 induzido por patógeno (Ward et al., Plant. Mol. Biol. 22 (1993) 361-366), o promotor hsp80 de tomateiro induzível por calor (US 5.187.267), o promotor de alfa-amilase de batateira induzível por frio (WO 96/12814) ou o
25 promotor pinII induzível por ferimento (EP-A-0 375 091) ou outros como aqui descrito.

Promotores preferidos são em particular aqueles que conduzem expressão de gene em tecidos e órgãos, em células de semente, tais como células de endosperma e células do embrião em desenvolvimento.

Promotores adequados são o promotor de gene de napina de colza de semente oleosa (US 5.608.152), the Vicia faba USP promotor (Baeumlein et al., Mol Gen Genet, 1991, 225 (3): 459-67), o promotor de oleosina de Arabidopsis (WO 98/45461), o promotor de faseolina de Phaseolus vulgaris (US 5.504.200), o promotor Bce4 de Brassica (WO 91/13980), o promotor arc5 de feijão, o promotor DcG3 de cenoura, ou o promotor de Legumina B4 (LeB4; Baeumlein et al., 1992, Plant Journal, 2 (2): 233-9), e promotores que conduzem a expressão semente-específica em plantas monocotiledôneas tais como milho, cevada, trigo, centeio, arroz e semelhante. Promotores semente-específicos vantajosos são o promotor de proteína ligante de sacarose (WO 00/26388), o promotor de faseolina e o promotor de napina. Promotores adequados que precisam ser considerados são o promotor de gene lpt2 ou lpt1 de cevada (WO 95/15389 e WO 95/23230), e os promotores descritos em WO 99/16890 (promotores do gene de hordeína de cevada, o gene de glutelina de arroz, o gene de orizina de arroz, o gene prolamina de arroz, o gene de gliadina de trigo, o gene de glutelina de trigo, o gene de zeína de milho, o gene de glutelina de aveia, o gene de casirina de sorgo e o gene de secalina de centeio). Outros promotores adequados são Amy32b, Amy 6-6 e Aleurain [US 5.677.474], Bce4 (colza de semente oleosa) [US 5.530.149], glicinina (soja) [EP 571.741], fosfoenolpiruvato carboxilase (soja) [JP 06/62870], ADR12-2 (soja) [WO 98/08962], isocitrato liase (colza de semente oleosa) [US 5.689.040] ou α -amilase (cevada) [EP 781.849]. Outros promotores que estão disponíveis para a expressão de genes em plantas são promotores folha-específicos tais como aqueles descritos em DE-A 19644478 ou promotores luz-regulados tais como, por exemplo, o promotor petE de ervilha.

Outros promotores de planta adequados são o promotor FBPase citosólico ou o promotor ST-LSI de batateira (Stockhaus et al., EMBO J. 8, 1989, 2445), O promotor de fosfo-ribosil-pirofosfato amidotransferase de Glycine max (N^o de Acesso no GenBank de U87999) ou

o promotor nodo-específico descrito em EP-A-0.249.676.

Preferivelmente, o ácido nucleico CDK ou sua variante de acordo com a invenção é operacionalmente ligado em um promotor broto-específico. O termo “broto-específico” como aqui definido refere-se a um
5 promotor que é expressado predominantemente no broto e em qualquer estágio na vida da planta. O termo “broto” como aqui usado inclui todas as partes aéreas da planta, incluindo caules e galhos, folhas, botões de flor, órgãos reprodutivos, incluindo estruturas derivadas de broto tais como estolhos, cormos, rizomas ou tubérculos. Preferivelmente, o promotor broto-
10 específico capaz de preferencialmente expressar o ácido nucleico em todo o broto é um promotor fraco. Padrão de expressão e/ou força de promotor pode ser analisado por exemplo por copulação do promotor em um gene repórter e ensaiando a expressão do gene repórter em vários tecidos da planta. Um gene repórter adequado bem conhecido pelas pessoas experientes na técnica é beta-
15 glicuronidase. Padrão de expressão e/ou força de promotor pode ser comparado com aquele de um promotor de referência broto-específico bem caracterizado, tal como o promotor Cab27 (expressão fraca, GenBank AP004700), ou o promotor de protoclorofilida redutase putativa (expressão forte, GenBank AL606456). Referência a um “promotor fraco” indica um
20 promotor que conduz a expressão de uma sequência codificadora em um nível baixo, a saber em níveis de cerca de 1/10.000 transcritos a cerca de 1/100.000 transcritos, a cerca de 1/5.000.000 transcritos por célula. De modo inverso, um “promotor forte” conduz a expressão de uma sequência codificadora em um nível alto, ou a cerca de 1/10 transcritos a cerca de 1/100 transcritos a
25 cerca de 1/1.000 transcritos por célula. Mais preferivelmente, o promotor capaz de preferencialmente expressar o ácido nucleico em toda a planta é um promotor de metalotioneína de arroz. Deve ficar claro que a aplicabilidade da presente invenção não é restrita aos ácidos nucleicos CDK como representados no protocolo de sequência, preferivelmente como representados

em SEQ ID NO: 45, 47, 49, 51, 53 ou 55.

Opcionalmente, uma ou mais seqüências de terminador também podem ser usadas no construto introduzido em uma planta. O termo “terminador” inclui uma seqüência de controle que é uma seqüência de DNA na extremidade de unidade de transcrição que sinaliza processamento 3' (atrás do códon de terminação) e poliadenilação de um transcrito primário e terminação de transcrição. Um terminador, que pode ser usado no processo da invenção é, por exemplo, o terminador OCS1, o terminador nos3 ou o terminador 35S. Como é o caso com os promotores, seqüências de terminador diferentes devem ser usadas para cada gene. Terminadores, que são úteis em microorganismos são por exemplo o terminador fimA, terminador txn ou terminador trp. Tais terminadores também podem ser rho-dependentes ou rho-independentes. Elementos regulatórios adicionais podem incluir intensificadores de transcrição e de tradução. Aqueles pessoas experientes na técnica estão cientes das seqüências de terminador e intensificador que podem ser adequadas para uso na realização da invenção. Tais seqüências seriam conhecidas ou poderiam ser prontamente obtidas por uma pessoa experiente na técnica.

Os construtos genéticos da invenção podem adicionalmente incluir uma origem de seqüência de replicação, que é exigida para manutenção e/ou replicação em um tipo de célula específico. Um exemplo é quando um construto de gene é exigido para ser mantido em uma célula bacteriana como um elemento genético episomal (e.g. molécula de plasmídeo ou cosmídeo). Origens de replicação preferidas incluem, mas não são limitadas a, fl-ori e colE1.

Para a detecção e/ou seleção da transferência bem sucedida de uma seqüências de ácido nucleico como representada no protocolo de seqüência e usada no processo da invenção, é vantajoso usar genes marcadores (=genes repórter). Estes genes marcadores permitem a

identificação de uma transferência bem sucedida das moléculas de ácido nucleico via uma série de princípios diferentes, por exemplo via identificação visual com o auxílio de fluorescência, luminescência ou na faixa de comprimento de ondas de luz que é discernível para o olho humano, por uma

5 resistência aos herbicidas ou antibióticos, via o que são conhecidos como marcadores nutritivos (marcadores auxotrofismo) ou marcadores antinutritivos, via ensaios de enzima ou via fito-hormônios. Exemplos de tais marcadores que podem ser mencionados são GFP (= proteína fluorescente verde); o sistema de luciferina/luciferase, a β -galactosidase com seus

10 substratos coloridos, por exemplo X-Gal, as resistências a herbicida, por exemplo, imidazolinona, glifosato, fosfinotricina ou sulfonil-uréia, as resistências a antibióticos, por exemplo, a bleomicina, higromicina, estreptomicina, canamicina, tetraciclina, cloranfenicol, ampicilina, gentamicina, geneticina (G418), espectinomomicina ou blasticidina, para

15 mencionar apenas uns poucos, marcadores nutritivos tal como a utilização de manose ou xilose, ou marcadores antinutritivos tal como a resistência a 2-desóxi-glicose. Esta lista é um número pequeno de marcadores possíveis. O técnico experiente está muito familiarizado com tais marcadores. Marcadores diferentes são preferidos, dependendo do organismo e do método de seleção.

20 Portanto o construto genético pode opcionalmente compreender um gene marcador selecionável. Como aqui usado, o termo “marcador selecionável ou gene marcador selecionável” inclui qualquer gene, que confere um fenótipo a uma célula na qual ele é expressado para facilitar a identificação e/ou seleção de células que são transfectadas ou transformadas

25 com um construto de ácido nucleico da invenção. Exemplos de genes marcadores selecionáveis codificadores de proteínas que conferem resistência a antibióticos (tais como nptII que fosforila neomicina e canamicina, ou hpt, fosforilação de higromicina), a herbicidas (por exemplo bar que proporciona resistência a Basta; aroA ou gox proporcionando resistência contra glifosato),

ou genes que proporcionam uma feição metabólica (tal como manA que permite que as plantas usem manose como única fonte de carbono). Genes codificadores de proteínas marcadoras visuais resultam na formação de cor (por exemplo β -glicuronidase, GUS), luminescência (tal como luciferase) ou fluorescência (Proteína Fluorescente Verde, GFP, e seus derivados).

É sabido da integração transiente ou estável de ácidos nucleicos em células de planta que apenas uma minoria das células captura DNA estranho e, se desejado, o integra em seu genoma, dependendo do vetor de expressão usado e da técnica de expressão utilizada. Para identificar e selecionar estes integrantes, um gene codificador de um marcador selecionável (como descrito acima, por exemplo resistência a antibióticos) é normalmente introduzido nas células hospedeiras junto com o gene de interesse. Marcadores selecionáveis preferidos em plantas compreendem aqueles, que conferem resistência a um herbicida tal como glifosato ou glifosinato. Outros marcadores adequados são, por exemplo, marcadores que codificam genes envolvidos em rotas biossintéticas de, por exemplo, açúcares ou aminoácidos, tais como β -galactosidase, *ura3* ou *ilv2*. Marcadores, que codificam genes tais como luciferase, *gfp* ou outros genes de fluorescência, são igualmente adequados. Estes marcadores e os marcadores acima mencionados podem ser usados em mutantes nos quais estes genes não são funcionais porque, por exemplo, têm sido deletados por métodos convencionais. Adicionalmente, moléculas de ácido nucleico, que codificam um marcador selecionável, podem ser introduzidas em uma célula hospedeira no mesmo vetor que aqueles, que codificam os polipeptídeos da invenção ou usadas no processo ou então em um vetor separado. Células que têm sido transfectadas estavelmente com o ácido nucleico introduzido podem ser identificadas por exemplo por seleção (por exemplo células que têm integrado o marcador selecionável sobrevivem enquanto que as outras células morrem).

Visto que os genes marcadores, como uma regra

especificamente o gene para resistência aos antibióticos e aos herbicidas, não são mais exigidos ou são indesejados em célula hospedeira transgênica uma vez os ácidos nucleicos tendo sido introduzidos com sucesso, o processo de acordo com a invenção para introduzir os ácidos nucleicos vantajosamente utiliza técnicas que permitem a remoção, ou excisão, destes genes marcadores. Um tal método é o que é conhecido como cotransformação. O método de cotransformação usa dois vetores simultaneamente para a transformação, um vetor possuindo o ácido nucleico de acordo com a invenção e um segundo possuindo o(s) gene(s) marcador(es). Uma proporção grande dos transformantes recebe ou, no caso de plantas, compreende (até 40% dos transformantes e acima), ambos vetores. No caso de transformação com *Agrobactérias*, os transformantes normalmente recebem apenas uma parte do vetor, a sequência flanqueada pelo T-DNA, que normalmente representa o cassete de expressão. Os genes marcadores podem ser subsequentemente removidos da planta transformada por realização de cruzamentos. Em outro método, genes marcadores integrados em um transposon são usados para a transformação juntos com o ácido nucleico desejado (conhecido como a tecnologia *Ac/Ds technology*). Os transformantes podem ser cruzados com um recurso de transposase ou os transformantes são transformados com um construto de ácido nucleico que confere expressão de uma transposase, transientemente ou estável. Em alguns casos (aprox. 10%), o transposon pula para fora do genoma da célula hospedeira uma vez tendo ocorrido a transformação com sucesso e é perdido. Em outros casos numerosos, o transposon pula para uma localização diferente. Nestes casos, o gene marcador precisa ser eliminado por realização de cruzamentos. Em microbiologia, oram desenvolvidas técnicas que tornam possível, ou facilitam, a detecção de tais eventos. Um outro método vantajoso baseia-se no que são conhecidos como sistemas de recombinação; cuja vantagem é que eliminação por cruzamento pode ser abolida. O sistema melhor conhecido

deste tipo é o que é conhecido como o sistema Cre/lox. Cre1 é uma recombinase, que remove as seqüências localizadas entre as seqüências loxP. Se o gene marcador estiver integrado entre as seqüências loxP, ele é removido, uma vez tendo ocorrido com sucesso a transformação, pela
 5 expressão da recombinase. Outros sistemas de recombinação são os sistemas HIN/HIX, FLP/FRT e REP/STB (Tribble et al., J. Biol. Chem., 275, 2000: 22255-22267; Velmurugan et al., J. Cell Biol., 149, 2000: 553-566). Uma integração sítio-específica em um genoma de planta de uma seqüências de ácido nucleico de acordo com a invenção é possível. Naturalmente, estes
 10 métodos também podem ser aplicados aos microorganismos tais como levedura, fungos ou bactérias.

A presente invenção também inclui plantas ou células de planta obteníveis pelos métodos de acordo com a presente invenção. A presente invenção portanto proporciona plantas ou células de planta
 15 obteníveis pelo método de acordo com a presente invenção, cujas plantas ou células de planta têm introduzido nelas um ácido nucleico CDK ou sua variante, codificador de uma CDK compreendendo motivo de sítio ativo e de ligação de ATP, (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE como aqui descrito.

20 A invenção também proporciona um método para a produção de células de planta transgênicas ou plantas transgênicas tendo características de crescimento melhoradas, compreendendo introdução e expressão em uma planta de um ácido nucleico CDK ou a sua variante, codificador de uma CDK que compreende motivo de sítio ativo e de ligação de ATP,
 25 (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE como aqui descrito.

Mais especificamente, a presente invenção proporciona um método para a produção de plantas transgênicas tendo características de crescimento melhoradas, cujo método compreende:

(i) introduzir em uma planta ou célula de planta um ácido

nucleico codificador de uma CDK ou de um seu homólogo; e

(ii) cultivar uma célula de planta sob condições promotoras de crescimento e desenvolvimento de planta.

5 O ácido nucleico pode ser diretamente introduzido em uma célula de planta ou em uma própria planta (incluindo introdução em um tecido, órgão ou qualquer outra parte de uma planta). De acordo com um aspecto preferido da presente invenção, o ácido nucleico é preferivelmente introduzido em uma planta por transformação.

10 Os termos “transformação” ou “introdução” como aqui referidos incluem a transferência de um polinucleotídeo exógeno em uma célula hospedeira, independente do método usado para transferência. Tecido de planta capaz de subsequente propagação clonal, quer por organogênese quer por embriogênese, pode ser transformado com um construto genético da presente invenção e uma planta inteiro pode ser regenerada do mesmo. O
15 tecido particular escolhido variará dependendo dos sistemas de propagação clonal disponíveis, e melhor adaptados para, a espécie particular sendo transformada. Alvos de tecido exemplares incluem discos de folha, pólen, embriões, cotilédones, hipocótilos, megagametófitos, tecido de calo, tecido meristemático existente (e.g., meristema apical, botões auxiliares, e
20 meristemas de raiz), e tecido de meristema induzido) (e.g., meristema de cotilédone e meristema de hipocótilo). O polinucleotídeo pode ser transiente ou estavelmente introduzido em uma célula hospedeira e pode ser mantido não-integrado, por exemplo, como um plasmídeo. Alternativamente, pode ser integrado no genoma hospedeiro. A célula de planta transformada pode ser
25 então usada para regenerar uma planta transformada em uma maneira conhecida pelas pessoas experientes na técnica.

A transferência de genes estranhos para dentro do genoma de uma planta é chamada de transformação. Assim fazendo os métodos descritos para a transformação e a regeneração de plantas a partir de tecidos de planta

ou de células de planta são utilizados para transformação transiente ou estável. Um método de transformação vantajoso é a transformação em planta. Para este fim, é possível, por exemplo, permitir que agrobactéria atue sobre sementes de planta ou inocular um meristema de planta com Agrobactérias.

- 5 Tem se mostrado particularmente expediente de acordo com a invenção permitir que uma suspensão de agrobactérias transformadas atue sobre planta intacta ou pelo menos outros estágios iniciais de flor. A planta é subsequente-mente crescida até que as sementes da planta tratada sejam obtidas (Clough e Bent, Plant J. (1998) 16, 735-743). Para selecionar plantas transformadas, um material de planta obtido na transformação é, como uma
- 10 regra, submetido às condições seletivas de modo que plantas transformadas possam ser distinguidas das plantas não transformadas. Por exemplo, as sementes obtidas na maneira descrita acima podem ser plantadas e, após um período de crescimento inicial, submetida a uma seleção adequada por
- 15 pulverização. Uma outra possibilidade consiste em crescimento das sementes, se apropriado após esterilização, sobre placas de ágar usando um agente de seleção adequado de modo que apenas as sementes transformadas possam crescer em plantas. Outros métodos de transformação vantajosos, em particular para plantas, são conhecidos pelo técnico experiente e são descritos
- 20 aqui abaixo.

- Transformação de espécie de planta é agora uma técnica bastante rotineira. Vantajosamente, qualquer um dos vários métodos de transformação podem ser usados para introduzir o gene de interesse em um vetor antepassado adequado. Métodos de transformação incluem o uso de
- 25 lipossomos, eletroporação, compostos químicos que aumentam a captação de DNA livre, injeção de DNA diretamente na planta, bombardeio por canhão de partículas, transformação usando vírus ou pólen ou microprojecção. Métodos podem ser selecionados do método de cálcio/poli(etileno-glicol) para protoplastos (Krens et al. (1982) Nature 296, 72-74; Negrutiu et al. (1987)

Plant Mol. Biol. 8, 363-373); eletroporação para protoplastos (Shillito et al. (1985) Bio/Technol 3, 1099-1102); microinjeção em material de planta (Crossway et al. (1986) Mol. Gen. Genet. 202, 179-185); bombardeio de partícula revestida com DNA ou RNA (Klein et al. (1987) Nature 327, 70) 5 infecção com vírus (não-integrativo) e semelhante. Plantas transgênicas expressando uma CDK de acordo com a presente invenção são preferivelmente produzidas via transformação mediada por *Agrobacterium* usando qualquer um dos métodos bem conhecidos por exemplo para transformação de Brassica, feijão-soja, milho ou arroz, tais como descritos em 10 qualquer um dos seguintes: pedido de patente européia publicada EP 1198985 A1, Aldemita e Hodges (Planta 199, 612-617, 1996); Chan et al. (Plant Mol. Biol. 22, 491-506, 1993), Hiei et al. (Plant J. 6, 271-282, 1994), cujas descrições são aqui incorporadas como referências se totalmente descritas. No caso de transformação de milho, o método preferido é como descrito em 15 Ishida et al. (Nature Biotechnol. 14, 745-50, 1996) ou Frame et al. (Plant Physiol. 129, 13-22, 2002), cujas descrições são aqui incorporadas como referências se totalmente descritas. Ditos métodos são adicionalmente descritos por meio de exemplo em B. Jenes et al., Techniques for Gene Transfer, em: Transgenic Plants, Vol. 1, Engineering e Utilization, eds. S.D. 20 Kung e R. Wu, Academic Press (1993) 128-143 e em Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205-225). O ácidos nucleicos ou o construto a serem expressados é preferivelmente clonado em um vetor, que é adequado para transformação de *Agrobacterium tumefaciens*, por exemplo pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984) 8711). *Agrobactérias* 25 transformadas por um tal vetor podem ser então usadas em maneira conhecida para a transformação de plantas, em particular de plantas de colheita tais como por exemplo plantas de tabaco, por exemplo, pelo banho de folhas esmagadas ou folhas cortadas em uma solução agrobacteriana e então pelo cultivo delas em meios adequados. A transformação de plantas por meio de

Agrobacterium tumefaciens é descrita, por exemplo, por Höfgen e Willmitzer em Nucl. Acid Res. (1988) 16, 9877 ou é conhecida inter alia de F.F. White, Vectors for Gene Transfer in Higher Plants; em Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S.D. Kung e R. Wu, Academic Press, 1993, pp. 15-38.

Geralmente após transformação, células de planta ou agrupamentos de células de planta são selecionadas para a presença de um ou mais marcadores que são codificados por genes expressáveis em planta co-transferidos com gene de interesse, após o qual o material transformado é regenerado em uma planta inteira.

Como mencionado Agrobactérias transformadas com um vetor de expressão de acordo com a invenção também pode ser usado na maneira por si conhecida para a transformação de plantas tais como plantas experimentais como plantas Arabidopsis ou de colheita, tais como, por exemplo, cereais, milho, aveia, centeio, cevada, trigo, soja, arroz, algodoeiro, beterraba, canola, girassol, linho, cânhamo, batateira, tabaco, tomateiro, cenoura, pimentões, colza de semente oleosa, tapioca, mandioca, araruta, tagetes, alfafa, alface e as várias árvores, nogueiras, e espécies de parreira, em particular plantas de colheita contendo óleo tais como soja, amendoim, planta de óleo de rícino, girassol, milho, algodoeiro, linho, colza de semente oleosa, coqueiro, palma oleosa, cártamo (Carthamus tinctorius) ou amêndoas de cacau, por exemplo, pelo banho de folhas ou segmentos de folha escarificados em uma solução agrobacteriana e subsequente crescimento delas em meios adequados.

Em adição à transformação de células somáticas, que então têm que ser regeneradas em plantas intactas, também é possível transformar as células de meristemas de planta e em particular aquelas células que desenvolvem em gametas. Neste caso, os gametas transformados seguem o desenvolvimento natural da planta, originando as plantas transgênicas. Assim,

por exemplo, sementes de *Arabidopsis* são tratadas com Agrobactérias e sementes são obtidas das plantas desenvolventes das quais uma certa proporção é transformada e assim transgênica (Feldman, KA e Marks MD (1987). *Mol Gen Genet* 208:274-289; Feldmann K (1992). In: C Koncz, N-H Chua e J Shell, eds, *Methods in Arabidopsis Research*. Word Scientific, Singapore, pp. 274-289). Métodos alternativos são baseados na remoção repetida das inflorescências e incubação do sítio de excisão no centro do involúcro de flor composta com Agrobactérias transformadas, pelo qual sementes transformadas podem ser igualmente obtidas em um instante de tempo posterior (Chang (1994). *Plant J.* 5: 551-558; Katavic (1994). *Mol Gen Genet*, 245: 363-370). Contudo, um método especialmente eficaz é o método de infiltração a vácuo com suas modificações tal como o método de “banho floral”. No caso de infiltração a vácuo de *Arabidopsis*, plantas intactas sob pressão reduzida são tratadas com uma suspensão agrobacteriana (Bechthold, N (1993). *C R Acad Sci Paris Life Sci*, 316: 1194-1199), enquanto que no caso do método de “banho floral” o tecido floral desenvolve-se é incubado brevemente com uma suspensão agrobacteriana tratada com tensoativo (Clough, SJ und Bent, AF (1998). *The Plant J.* 16, 735-743). Uma certa proporção de sementes transgênicas é colhida em ambos os casos, e estas sementes podem ser distinguidas de sementes não-transgênicas pelo crescimento sob as condições seletivas descritas acima. Em adição a transformação estável de plastídeos é vantajosa porque plastídeos são maternalmente herdados na maioria das colheitas reduzindo ou eliminando o risco de fluxo de transgene através de pólen. A transformação do genoma de cloroplasto é geralmente realizada por um processo, quantidade efetiva tem sido esquematicamente exibido em Klaus et al., 2004 (*Nature Biotechnology* 22(2), 225-229). Resumidamente as seqüências a serem transformadas são clonadas juntas com um gene marcador selecionável entre seqüências flanqueadoras homólogas ao genoma de cloroplasto. Estas seqüências

flanqueadoras homólogas conduzem integração sítio-específica para dentro do plastoma. Transformação plastidal tem sido descrita para muitas espécies diferentes de planta e um resumo pode ser obtido de Bock (2001) “Transgenic plastids in basic research and plant biotechnology”. *J Mol Biol.* 2001 Sep 21; 312 (3):425-38 ou Maliga, P (2003) “Progress towards commercialization of plastid transformation technology”. *Trends Biotechnol.* 21, 20-28. Progreso biotecnológico adicional tem sido recentemente relatado em forma de transformantes de plastídeo livres de marcador, que podem ser produzidos por um gene marcador cointegrado transiente (Klaus et al., 2004, *Nature Biotechnology* 22(2), 225-229).

As células de planta geneticamente modificadas podem se regeneradas via todos os métodos com os quais um técnico experiente está familiarizado. Métodos adequados podem ser encontrados nas publicações acima mencionadas de S.D. Kung e R. Wu, Potrykus ou Höfgen e Willmitzer.

Após transferência de DNA e regeneração, as plantas putativamente transformadas podem ser avaliadas. Análise de Southern, para a presença do gene de interesse, número de cópias e/ou organização genômica. Alternativa ou genomicamente, níveis de expressão de DNA recém introduzido podem ser monitorados usando análise de Northern e/ou Western, ambas técnicas bem conhecidas pelas pessoas tendo experiência ordinária na técnica.

As plantas transformadas geradas podem ser propagadas por uma variedade de meios, tais como por propagação clonal ou técnicas de geração clássicas. Por exemplo, a planta transformada de primeira geração (ou T1) pode ser fertilizada para dar transformantes homozigóticos de segunda geração (ou T2), e as plantas T2 são posteriormente propagadas através de técnicas de geração clássicas.

Os organismos transformados gerados podem tomar uma variedade de formas. Por exemplo, podem ser quimeras de células

transformadas e células não-transformadas; transformantes clonais (e.g., todas células transformadas para conterem o cassete de expressão); enxertos de tecidos transformados e não-transformados (e.g., em plantas, um rizoma transformado enxertado em um rebento não-transformado).

5 A presente invenção claramente se estende a qualquer planta ou célula de planta produzida ou obténível por qualquer um dos métodos aqui descritos, e a todas as plantas e seus propágulos. A presente invenção se estende adicionalmente para incluir a progênie de uma célula, tecido, órgão ou planta inteira primária transformada ou transfectada que tem sido
10 produzida por quaisquer métodos acima mencionados, a única exigência sendo que a progênie exiba a(s) mesma(s) característica(s) genotípica(s) e/ou fenotípica(s) como aquelas produzidas na planta parental pelos métodos de acordo com a invenção.

 A invenção também inclui células hospedeiras contendo um
15 ácido nucleico CDK isolado de planta ou sua variante, codificador de uma CDK compreendendo as características aqui descritas. Células hospedeiras preferidas de acordo com a invenção são células de planta.

 A invenção também se estende às partes capazes de serem colhidas de uma planta de acordo com a invenção tais como mas não
20 limitadas a sementes, folhas, flores, caules, rizomas, tubérculos e bulbos. A invenção adicionalmente se refere aos produtos derivados, preferivelmente diretamente derivados, de uma parte capaz de ser colhida de uma tal planta, tal como pós ou pétalas secas, óleo, gordura e ácidos graxos, amido e proteínas.

25 A presente invenção adicionalmente inclui o uso de um gene CDK e da proteína codificada para melhorar as características de crescimento de plantas; tais características de crescimento melhoradas são como definidas aqui acima.

 A presente invenção também inclui o uso de ácidos nucleicos

CDK ou suas variantes, e ao uso de polipeptídeos CDK ou seus homólogos, cuja CDK ou cujo homólogo compreende motivo de sítio ativo e de ligação de ATP, (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE como aqui descrito, ou cujo ácido nucleico CDK ou variante codifica uma tal proteína. Um tal uso
5 refere-se ao melhoramento de características de crescimento de plantas, em particular ao melhoramento de rendimento, especialmente de rendimento de sementes. O rendimento de sementes pode incluir qualquer um ou mais dos seguintes: número total de sementes aumentado, número aumentado de sementes cheias, peso de semente aumentado, índice de colheita aumentado,
10 dentre outros.

Ácidos nucleicos CDK ou suas variantes, ou polipeptídeos CDK ou seus homólogos, podem encontrar uso em programas de geração nos quais um marcador de DNA é identificado o qual pode ser geneticamente ligado em um gene CDK ou sua variante. A CDK ou suas variantes, ou
15 CDKA ou seus homólogos, pode ser usada para definir um marcador molecular. Este marcador de DNA ou proteína pode ser então usado em programas de geração para selecionar plantas tendo características de crescimento melhoradas. O gene CDK ou sua variante pode, por exemplo, ser um ácido nucleico como representado no protocolo de seqüência
20 preferivelmente como representado em SEQ ID NO: 45, 47, 49, 51, 53 ou 55, ou um ácido nucleico codificador de qualquer um dos homólogos como aqui definido.

Variantes alélicas de uma CDK também podem encontrar uso em programas de geração auxiliados por marcador. Tais programas de
25 geração algumas vezes exigem a introdução de variação alélica por tratamento mutagênico das plantas, usando por exemplo mutagênese EMS; alternativamente, o programa pode iniciar com uma coleção de variantes alélicas de denominada origem “natural” não intencionalmente. Identificação de variantes alélicas então ocorrem por, por exemplo, PCR. Esta é seguida por

uma etapa de seleção para seleção de variantes alélicas superiores da sequência em questão e que dá características de crescimento melhoradas em uma planta. Seleção tipicamente realizada por monitoração de desempenho de crescimento de plantas contendo variantes alélicas diferentes da sequência em questão por exemplo, variantes alélicas de SEQ ID NO: 45, 47, 49, 51, 53 ou 55, ou de ácidos nucleicos codificadores de qualquer um dos homólogos acima mencionados. Desempenho de crescimento pode ser monitorado em uma estufa ou no campo. Outras etapas opcionais incluem cruzamento de plantas, nas quais a variante alélica foi identificada, com outra planta. Isto poderia ser usado, por exemplo, para preparar uma combinação de características fenotípicas interessantes.

Ácidos nucleicos CDK ou suas variantes de acordo com a invenção também podem ser usados como sondas para genética e fisicamente mapear os genes dos quais são uma parte, e como marcadores para feições ligadas àqueles genes. Tal informação pode ser útil em geração de planta com o propósito de desenvolver linhagens com fenótipos desejados. Tal uso de ácidos nucleicos CDK ou suas variantes requer apenas uma sequência de ácido nucleico de pelo menos 15 nucleotídeos em comprimento. Os ácidos nucleicos CDK ou suas variantes podem ser usados como marcadores de polimorfismo de comprimento de fragmento de restrição (RFLP). Southern blots de DNA genômico de planta digerido por enzima de restrição podem ser sondados com os ácidos nucleicos CDK ou suas variantes. Os padrões de banda resultantes podem ser então submetidos às análises genéticas usando programas de computador tal como MapMaker (Lander et al. (1987) Genomics 1, 174-181) com o objetivo 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 pais e progênie de um cruzamento definido. Segregação dos polimorfismos de DNA é anotada e usada para

calcular a posição do ácido nucleico CDK ou sua variante no mapa genético previamente obtido usando esta população (Botstein et al. (1980) *Am. J. Hum. Genet.* 32, 314-331).

5 A produção e o uso de sondas derivadas de gene de planta para uso em mapeamento genético são descritos em Bernatzky e Tanksley (*Genetics* 112, 887-898, 1986). Numerosas publicações descrevem mapeamento genético de clones de cDNA específicos usando a metodologia descrita acima ou suas variações. Por exemplo, populações de intercruzamento F2, populações de retro-cruzamento, populações
10 aleatoriamente acasaladas, linhagens isogênicas próximas, e outros conjuntos de indivíduos podem ser usados para mapeamento. Tais metodologias são bem conhecidas por aqueles pessoas experientes na técnica.

As sondas de ácido nucleico também podem ser usadas para mapeamento físico (i.e., posicionamento de seqüências sobre mapas físicos; veja Hoheisel et al. em: *Nonmammalian Genomic Analysis: A Practical Guide*, Academic press 1996, pp. 319-346, e as referências lá citadas).
15

Em outra modalidade, as sondas de ácido nucleico podem ser usadas em mapeamento de hibridização in situ com fluorescência direta (FISH) (Trask (1991) *Trends Genet.* 7, 149-154). Embora métodos correntes de FISH favoreçam o uso de clones grandes (várias a várias centenas de kb; veja Laan et al. (1995) *Genome Res.* 5, 13-20), melhoramentos em sensibilidade podem permitir execução de mapeamento de FISH usando sondas mais curtas.
20

Uma variedade de métodos de mapeamento genético e físico baseados em amplificação de ácido nucleico podem ser realizados usando os ácidos nucleicos. Exemplos incluem amplificação alelo-específica (Kazazian (1989) *J. Lab. Clin. Med.* 11, 95-96), polimorfismo de fragmentos amplificador por PCR (CAPS; Sheffield et al. (1993) *Genomics* 16, 325-332), ligação alelo-específica (Landegren et al. (1988) *Science* 241, 1077-1080),
25

reações de extensão de nucleotídeo (Sokolov (1990) Nucleic Acid Res. 18, 3671), Mapeamento de Híbridos por Radiação (Walter et al. (1997) Nat. Genet. 7, 22-28) e Happy Mapping (Dear e Cook (1989) Nucleic Acid Res. 17, 6795-6807). Para estes métodos, a seqüência e um ácido nucleico é usada para planejar e produzir pares de iniciadores para uso na reação de amplificação ou em reações de extensão de iniciador. O planejamento de tais iniciadores é bem conhecido por aqueles pessoas experientes na técnica. Em métodos empregando mapeamento genético baseado em PCR, pode ser necessário identificar as diferenças de seqüência de DNA entre os pais do cruzamento de mapeamento na região correspondendo à presente seqüência de ácido nucleico. Isto, contudo geralmente não é necessário para métodos de mapeamento.

Nesta maneira, podem ser realizadas geração, identificação, e/ou isolamento de plantas melhoradas com atividade modulada de quinase dependente de ciclina exibindo características de crescimento melhoradas.

Ácidos nucleicos CDK ou suas variantes ou polipeptídeos CDK ou seus homólogos de acordo com a presente invenção também podem encontrar uso como reguladores de crescimento. Visto que tem sido mostrado que estas moléculas são úteis no melhoramento de características de crescimento de plantas, também seriam reguladores de crescimento úteis, tais como herbicidas ou estimuladores de crescimento. A presente invenção portanto proporciona uma composição compreendendo uma CDK ou sua variante ou um polipeptídeo CDK ou seu homólogo, junto com um veículo, diluente ou excipiente adequado, para uso como um regulador de crescimento, cuja(o) CDK ou homólogo compreende o motivo de sítio ativo e de ligação de ATP, (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE como aqui descrito, ou cuja CDK ou variante codifica tal proteína.

Os métodos de acordo com a presente invenção resultam em plantas tendo características de crescimento melhoradas, como descrito aqui

anteriormente. Estas características de crescimento vantajosas também podem ser combinadas com outras feições economicamente vantajosas, tais como outras feições intensificadoras de rendimento, tolerância aos vários estresses, feições modificadoras de várias características arquitetônicas e/ou características bioquímicas e/ou fisiológicas.

Descrição das figuras

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

Fig. 1 mostra o vetor EG073qcz, que também é representado no protocolo de seqüência como SEQ ID NO: 57.

Fig. 2 mostra o vetor EG065qcz, que também é representado no protocolo de seqüência como SEQ ID NO: 58.

Fig. 3 mostra o vetor pMME0607, que também é representado no protocolo de seqüência como SEQ ID NO: 59.

Fig. 4 mostra as seqüências de vetor de EG073qcz, EG065qcz e pMME0607, que também são representadas no protocolo de seqüência como SEQ ID NO: 57 a 59.

Outras modalidades da invenção são:

O uso de uma seqüência de ácido nucleico de acordo com a invenção ou do construto de ácido nucleico de acordo com a invenção para a geração de plantas transgênicas.

Exemplos

A presente invenção agora será descrita com referência aos seguintes exemplos, que são apenas como via de ilustração.

Manipulação de DNA: a não ser que seja enunciado de outra maneira, técnicas de DNA recombinante são realizadas de acordo com protocolos padrão descritos em (Sambrook (2001) Molecular Cloning: a laboratory manual, 3rd Edition Cold Spring Harbor Laboratory Press, CSH, New York) ou em Volumes 1 e 2 de Ausubel

et al. (1994), Current Protocols in Molecular Biology, Current Protocols (<http://www.4ulr.com/products/currentprotocols/index.html>). Métodos e materiais padrão para trabalho molecular de planta são descritos em Plant Molecular Biology Labfax (1993) por R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (UK) e Blackwell Scientific Publications (UK).

Exemplo 1: Clonagem de gene

SEQ ID NO: 45 pode ser clonada nos plasmídeos pBR322 (Sutcliffe, J.G. (1979) Proc. Natl Acad. Sci. USA, 75: 3737-3741); pACYC177 (Change & Cohen (1978) J. Bacteriol. 134: 1141-1156);
10 plasmídeos da série pBS (pBSSK+, pBSSK- e outros; Stratagene, LaJolla, USA) ou cosmídeos tal como SuperCos1 (Stratagene, LaJolla, USA) ou Lorist6 (Gibson, T.J. Rosenthal, A., e Waterson, R.H. (1987) Gene 53: 283-286) para expressão em E. coli usando procedimentos bem estabelecidos, conhecidos (veja, por exemplo, Sambrook, J. et al. (1989) “Molecular
15 Cloning: A Laboratory Manual”. Cold Spring Harbor Laboratory Press ou Ausubel, F.M. et al. (1994) “Current Protocols in Molecular Biology”, John Wiley & Sons).

Exemplo 2: Seqüenciamento de DNA e análise funcional computadorizada

O DNA foi seqüenciado por procedimentos padrão, em particular pelo método de determinação de cadeia, usando seqüenciadores ABI377 (veja, por exemplo, Fleischman, R.D. et al. (1995) “Whole-genome Random Sequencing and Assembly of Haemophilus Influenzae Rd.”, Science 269; 496-512).

Exemplo 3: Transferência de DNA entre microorganismos diferentes tais como Escherichia coli e Agrobacterium tumefaciens

Vetores bifuncionais tais como pYE22m, pPAC-ResQ, pClasper, pAUR224, pAMH10, pAML10, pAMT10, pAMU10, pGMH10, pGML10, pGMT10, pGMU10, pPGAL1, pPADH1, pTADH1, pTAex3, pNGA142, pHT3101 e seus derivados que permitem a transferência de

seqüências de ácido nucleico entre microorganismos diferentes estão disponíveis para o técnico experiente. Um método fácil para isolar tais vetores bifuncionais é descrito por Soni R. e Murray J.A.H. [Nucleic Acid Research, vol. 20 no. 21, 1992: 5852]: se necessário tais vetores bifuncionais podem ser

5 construídos facilmente usando vetores padrão para *E. coli* (Sambrook, J. et al., (1989), “Molecular Cloning: A Laboratory Manual”, Cold Spring Harbor Laboratory Press ou Ausubel, F.M. et al. (1994) “Current Protocols in Molecular Biology”, John Wiley & Sons) e/ou os vetores acima mencionados, que têm uma origem de replicação para, e marcador adequado de, *Escherichia*

10 *coli* ou *Agrobacterium tumefaciens* adicionados. Tais origens de replicação são preferivelmente obtidas de plasmídeos endógenos, que têm sido isolados de espécies usadas para a produção de plantas usadas no processo da invenção. Genes, que são usados em particular como marcadores de transformação para estas espécies são genes para resistência à canamicina

15 (tais como aqueles que se origina do transposon Tn5 ou Tn-903) ou para resistência a cloranfenicol (Winnacker, E.L. (1987) “From Genes to Clones - Introduction to Gene Technology”, VCH, Weinheim) ou para outros genes de resistência a antibiótico tais como para resistência a G418, gentamicina, neomicina, higromicina ou tetraciclina.

20 Usando métodos padrão, é possível clonar um gene de interesse em um dos vetores bifuncionais acima descritos e introduzir tais vetores híbridos nas cepas de microorganismo usadas no processo da invenção.

Exemplo 4: Determinação da expressão de proteína mutante / transgênica

25 As observações da atividade de uma proteína mutada, ou transgênica, em uma célula hospedeira transformada são baseadas no fato de que a proteína é expressada em uma maneira similar e em uma quantidade similar à da proteína de tipo selvagem. Um método adequado para determinar a quantidade de transcrição do gene mutante, ou transgênico (um sinal para a

quantidade de mRNA que está disponível para a tradução do produto de gene) é realizado por um Northern blot (veja, por exemplo, Ausubel et al., (1988) *Current Protocols in Molecular Biology*, Wiley: New York), onde um iniciador que está planejado em uma maneira que se liga no gene de interesse é obtido com um marcador detectável (normalmente um marcador radioativo ou quimioluminescente) e modo que, quando o RNA total de uma cultura do organismo é extraído, separado sobre um gel, aplicado em uma matriz estável e incubado com esta sonda, a ligação e a quantidade da ligação da sonda indicam a presença e também a quantidade de mRNA para este gene. Outro método é uma PCR quantitativa. Esta informação detecta a extensão na qual o gene tem sido transcrito. RNA celular total pode ser isolado por exemplo de leveduras ou *E. coli* por uma variedade de métodos, que são conhecidos na técnica, por exemplo com o kit Ambion de acordo com as instruções do fabricante ou como descrito em Edgington et al., *Promega Notes Magazine* Number 41, 1993, p. 14.

Técnicas padrão, tal como Western blot, podem ser empregadas para determinar a presença ou quantidade relativa de proteína traduzida deste mRNA (veja, por exemplo, Ausubel et al. (1988) “*Current Protocols in Molecular Biology*”, Wiley, New York). Neste método, proteínas celulares totais são extraídas, separadas por eletroforese em gel, transferidas para uma matriz tal como nitrocelulose e incubadas com uma sonda, tal como um anticorpo, que se liga especificamente na proteína desejada. Esta sonda é normalmente proporcionada direta ou indiretamente com um marcador quimioluminescente ou colorimétrico, que pode ser detectado prontamente. A presença e a quantidade observada de marcador indicam a presença e a quantidade da proteína mutante almejada na célula. Contudo, também outros métodos também são conhecidos.

Exemplo 5: Crescimento de microorganismo geneticamente modificado: meios e condições de cultura

Microorganismos geneticamente modificados tal como *Escherichia coli* podem ser crescidos em meios de crescimento sintéticos ou naturais conhecidos pelo técnico experiente. Numerosos meios de crescimento diferentes para microorganismos tal como *Escherichia coli* são bem conhecidos e estão amplamente disponíveis.

Exemplo 6: Transformação de Agrobactérias

Plasmídeos podem ser transformados em *Agrobacterium tumefaciens* (GV3101pMP90; Koncz e Schell, 1986, Mol. Gen. Genet. 204: 383-396) usando protocolos de choque calorífico e ou eletroporação. Colônias transformadas podem ser crescidas sobre meio YEP e selecionadas por anticorpos respectivos (Rif/Gent/Km) por 2d a 28°C. Estas culturas de *Agrobacterium* foram usadas para transformação de planta.

Arabidopsis thaliana pode ser crescida e transformada de acordo com condições padrão Bechtold 1993 (Bechtold, N., Ellis, J., Pelletier, G. 1993, em “plant *Agrobacterium* mediated gene transfer by infiltration of *Arabidopsis thaliana* plants” C.R. Acad.Sci.Paris. 316:1194-1199); Bent et al. 1994 (Bent, A., Kunkel, B.N., Dahlbeck, D., Brown, K.L., Schmidt, R., Giraudat, J., Leung, J., e Staskawicz, B.J. 1994; “PPCS2 of *Arabidopsis thaliana*: A leucin-rich repeat class of plant disease resistant genes”; Science 265: 1856-1860).

Plantas *A. thaliana* transgênicas podem ser crescidas individualmente em potes contendo uma mistura 4:1 (v/v) de solo e areia de quartzo em uma câmara de crescimento York. Condições de crescimento padrão são: fotoperíodo de 16 h de luz e 8 h de escuro, 20°C, 60% de umidade relativa, e uma densidade de fluxo de fótons de 150 μ E. Para induzir germinação, sementes semeadas são mantidas a 4°C, no escuro, por 3 dias. Plantas são regadas diariamente até estarem com aproximadamente 3 semanas de idade em cujo momento estiagem é imposta pelo fechamento da água. Paralelamente, a umidade relativa foi reduzida em incrementos de 10% a cada

dois dias para 20%. As plantas podem ser ensaiadas para crescimento melhorado sob ditas condições.

Em geral é útil conduzir os ditos experimentos em três experimentos independentes sucessivos. No primeiro experimento, 10
5 linhagens T2 independentes devem ser semeadas para cada gene sendo testado. A percentagem de planta não mostrando sintomas visuais de lesão é determinada. No segundo experimento as linhagens positivas devem ser então confirmadas em um procedimento experimental idêntico. Em um terceiro experimento, pelo menos 7 linhagens positivas devem ser então confirmadas
10 em um procedimento experimental idêntico. Em um terceiro experimento, pelo menos 7 replicados das melhores linhagens mostrando crescimento melhorado devem ser então de novo confirmados.

Em um outro experimento, para linhagens maiores individuais, outras linhagens contendo o mesmo construto de gene, mas resultantes de um
15 evento de transformação diferente devem ser de novo testadas. Todos os resultados são resumidos e analisados.

Exemplo 6: Construção de vetor e transformação de arroz

Para a expressão em arroz um vetor tal como aqueles mostrados em Figuras 1 a 4 contendo o cassete de expressão SEQ ID NO: 60
20 é útil. No citado vetor SEQ ID NO: 1 como mostrado em protocolo de seqüência pode ser introduzido. Dito vetor pode ser transformado na cepa LBA4404 de Agrobacterium e subsequêntemente em plantas Oryza sativa. Plantas de arroz transformadas são permitidas crescer e são então examinadas para os parâmetros descritos em Exemplo 7.

25 Exemplo 7: Avaliação de transformantes: medições de crescimento

Aproximadamente 15 a 20 transformantes T0 independentes são comumente gerados. Os transformantes primários são transferidos de câmaras de cultura de tecido para uma estufa para crescimento e colhidos de semente T1. São retidos quatro eventos dos quais a progênie T1 segregou 3:1

para presença/ausência do transgene. Para cada um destes eventos, 10 mudas T1 contendo o transgene (hetero- e homo-zigotos), e 10 mudas T1 faltantes de transgene (nulizigotos), são selecionadas por seleção de marcador visual. As plantas T1 selecionadas são transferidas para uma estufa. Cada planta recebeu
5 um rótulo de código de barras único para claramente ligar os dados de fenotipificação à planta correspondente. As plantas T1 selecionadas são crescidas sobre solo em potes de diâmetro de 10 cm sob as condições ambientais: fotoperíodo = 11,5 h, intensidade de luz de dia = 30,000 lux ou mais, temperatura durante o dia = 28°C ou mais alta, temperatura durante a
10 noite = 22°C, umidade relativa = 60-70%. As plantas transgênicas e os nulizigotos correspondentes são crescidos lado a lado em posições aleatórias. Do estágio de semeadura até o estágio de maturidade as plantas são passadas várias vezes através de uma cabina de imagem digital. Em cada instante de tempo imagens digitais (2048x1536 pixéis, 16 milhões de cores) são tiradas
15 de cada planta de pelo menos 6 ângulos diferentes.

As panículas primárias maduras são colhidas, ensacadas, rotuladas com código de barras e então secas por três dias no forno a 37°C. As panículas são então debulhadas e todas as sementes colhidas. As cascas cheias são separadas das vazias usando um dispositivo de sopro de ar. Após
20 separação, ambos os lotes de semente são então contados usando uma máquina contadora comercialmente disponível. As cascas vazias são jogadas fora. As cascas cheias são pesadas em uma balança analítica e a área de seção transversal das sementes foi medida usando imagem digital. Este procedimento resulta no conjunto de parâmetros relacionados com semente
25 descritos abaixo.

Estes parâmetros são derivados em um modo automático das imagens digitais usando programa de computador de análise de imagem e são analisados estatisticamente. Uma ANOVA (análise de variância) de dois fatores é corrigida para o planejamento desequilibrado e é usada como

modelo estatístico para a avaliação total das características fenotípicas de planta. Um teste-F é realizado sobre todos os parâmetros medidos de todas as plantas de todos os eventos transformadas com aquele gene. O teste-F é realizado para checar um efeito do gene, também aqui chamado de um “efeito de gene global”. Se o valor do teste-F mostra que os dados são significativos, então é concluído que há um efeito de “gene”, significando que não apenas a presença ou a posição do gene causa o efeito. O limite para significância para um efeito de gene global verdadeiro é ajustado em nível de probabilidade de 5% para o teste-F.

Para checar um efeito dos genes dentro de um evento, i.e., para um efeito linhagem-específico, um teste-t é realizado dentro de cada evento usando conjuntos de dados das plantas transgênicas e das plantas nulas correspondentes. “Plantas nulas” ou “segregantes nulos” ou “nulizigotos” referem-se às plantas tratadas na mesma maneira que a planta transgênica, mas da qual o transgene tem sido segregado. Plantas nulas também podem ser descritas como as plantas transformadas negativas homozigóticas. O limite para significâncias para o teste-t é ajustado em um nível de probabilidade de 10%. Os resultados para alguns dos eventos podem estar acima ou abaixo deste limite. Isto é baseado na hipótese de que um gene poderia apenas ter um efeito em certas posições no genoma, e que a ocorrência deste efeito dependente de posição não é incomum. Este tipo de efeito de gene é também aqui chamado de um “efeito de linhagem do gene”. O valor-p é obtido por comparação do valor-t com a distribuição-t ou alternativamente, por comparação de valor-F com a distribuição-F. O valor-p então dá a probabilidade na qual a hipótese nula (i.e., não há efeito do transgene) está correta.

Os dados obtidos no primeiro experimento são confirmados em um segundo experimento com plantas T2. Três linhagens são selecionadas para análise adicional. Bateladas de sementes das plantas positivas (de hetero-

e homo-zigotos) em T1, são selecionadas por monitoração da expressão de marcador. Para cada evento escolhido, as bateladas de sementes heterozigóticas são então retidas para avaliação T2. Dentro de cada batelada de semente um número igual de plantas positivas e negativas é crescido na estufa para avaliação.

Um número total de 120 plantas transformadas é avaliado na geração T2, isto é 40 plantas por evento das quais 20 são positivas para o transgene e 20 negativas.

Devido ao fato de serem realizados dois experimentos com eventos sobrepostos, uma análise combinada foi conduzida. Esta é útil para checar a consistência dos efeitos sobre os dois experimentos, e se este for o caso, para acumular evidência de ambos os experimentos com o propósito de aumentar a confiança na conclusão. O método usado é uma abordagem de modelo misto que leva em consideração a estrutura de múltiplos níveis dos dados (i.e. experimento - evento - segregantes). Valores-p são obtidos por comparação do teste de razão de probabilidade com distribuições de qui-quadrado.

Exemplo 8: Avaliação de transformantes: medição de parâmetros relacionados com rendimento

Com a análise das sementes descrita acima, os inventores são capazes de verificar que as plantas transformadas com o construto de gene CDK codificador de uma CDK com os motivos aqui mencionados têm um número aumentado de sementes cheias, um peso total aumentado de sementes e um índice de colheita aumentado comparadas com as plantas faltantes do transgene CDK.

Resultados positivos são obtidos para plantas na geração T1 e são de novo obtidos na geração T2. Estes dados T2 são reavaliados em uma análise combinada com os resultados para a geração T1, e os valores-p obtidos mostram que os efeitos observados são significativos.

Número de sementes cheias

O número de sementes cheias é determinado pela contagem do número de cascas cheias que permanecem após a etapa de separação. Tipicamente 3 das 4 linhagens testadas estão mostrando um aumento significativo em números de sementes cheias.

Rendimento total de sementes

O rendimento total de sementes (peso total de sementes) por planta é medido pela pesagem de todas as cascas cheias colhidas de uma planta. Tipicamente 3 das 4 linhagens T1 transgênicas estão mostrando um aumento no peso total de sementes.

Índice de colheita

O índice de colheita na presente invenção é aqui definido como a razão entre o peso total de sementes e a área acima do solo (mm^2), multiplicada por um fator de 10^6 . Todas as linhagens testadas estão mostrando um índice de colheita aumentado.

Adicionalmente, há uma tendência geral para um número total de sementes aumentado.

Exemplo 9: Cultura de planta para análises bioanalíticas

Para análises bioanalíticas das plantas transgênicas, as últimas são crescidas como descrito acima.

Exemplo 10: Análise metabólica de plantas transgênicas

As modificações identificadas de acordo com a invenção são identificadas pelo seguinte procedimento:

a) Homogeneização das amostras

Homogeneização das amostras é realizada usando um moinho de bolas (Retsch). Dez a trinta grãos de arroz são transferidos para dentro de tubos plásticos (Eppendorf, Safe-Lock, 2 mL) e homogeneizados com uma bola de aço inoxidável sob esfriamento com nitrogênio líquido.

b) Liofilização

Durante o experimento, cuidado é tomado para que as amostras quer permaneçam no estado profundamente congelado (temperaturas $< -40^{\circ}\text{C}$) quer estejam livres de água por liofilização do material homogeneizado até o primeiro contato com solventes.

5 As amostras são transferidas para secador por congelamento pré-esfriado (-40°C). A temperatura inicial durante a fase de secagem principal é -35°C e a pressão é 12 Pa. Durante a fase de secagem, os parâmetros são alterados seguindo um programa de pressão e temperatura. A temperatura final após 12 horas é $+30^{\circ}\text{C}$ e a pressão final é final is 0, 10 1 a 0,4 Pa. Após a bomba de vácuo e a máquina de refrigeração terem sido desligadas, o sistema é submetido ao fluxo de ar (seco via um tubo de secagem) ou argônio.

c) Extração

15 Imediatamente após o aparelho de liofilização ter sido submetido ao fluxo, os tubos com o material de planta liofilizado são hermeticamente vedados para proteger o material da umidade do ar. Para a extração uma porção de 50 mg do material de planta homogeneizado seco é pesada em tubos de extração de fibra de vidro para dentro de cartuchos de extração de 5 ml do dispositivo ASE (Accelerated Solvent Extractor ASE 200 com Solvent Controller e AutoASE software (DIONEX)).

20 As 24 posições de amostra de um dispositivo ASE (Accelerated Solvent Extractor ASE 200 com Solvent Controller e AutoASE software (DIONEX)) são cheias com as amostras de planta, incluindo algumas amostras para teste de controle de qualidade.

25 As substâncias polares são extraídas com aproximadamente 10 ml de metanol/água (80/20, v/v) a $T = 70^{\circ}\text{C}$ e $p = 14.000 \text{ kPa}$, fase de aquecimento de 5 minutos, extração estática de 1 minuto. As substâncias mais lipofílicas são extraídas com aproximadamente 10 ml de metanol/diclorometano(40/60, v/v) a $T = 70^{\circ}\text{C}$ e $p = 14.000 \text{ ka}$, fase de

aquecimento de 5 minutos, extração estática de 1 minuto. As duas misturas de solventes são extraídas para dentro dos mesmos tubos de vidro (tubos de centrífuga, 50 ml, equipados com tampa rosqueada e septo perfurável para o ASE (DIONEX)).

5 A solução é tratada com padrões internos comercialmente disponíveis, tais como ribitol, L-glicina-2,2-d₂, L-alanina-2,3,3,3-d₄, metionina-d₃, Arginina (¹³C), Triptofano-d₅, e α-metil-glicopiranosídeo e nonadecanoato de metila, undecanoato de metila, tridecanoato de metila, pentadecanoato de metila, nonacosanoato de metila.

10 O extrato total é extraído com 8 ml de água. O resíduo sólido da amostra de planta e a luva de extração são rejeitados.

 O extrato é agitado e então centrifugado por 5 a 10 minutos pelo menos a 1.400 g com o propósito de acelerar a separação de fases. 1 ml do sobrenadante metanol/água (“fase polar”, incolor) é removido para análise por GC posterior. O restante da fase de metanol/água é rejeitado. 0,75 ml da fase orgânica (“fase lipídica”, verde escura) é removido para análise por GC posterior e 0,75 ml é removido para análise por LC. Todas as porções removidas são evaporadas até a secura usando o evaporador a vácuo em infravermelho IR Dancer (Hettich). A temperatura máxima durante o processo de evaporação não ultrapassou 40°C. Pressão no aparelho não é menor do que 1,0 kPa.

15

20

d) Processamento das fases lipídica e polar para a análise por LC/MS ou LC/MS/MS

 O extrato lipídico, que tem sido evaporado até a secura é recolhido na fase móvel. O extrato polar, que tem sido evaporado até a secura é recolhido na fase móvel.

25

e) Análise por LC-MS

 A parte LC é realizada em um sistema LCMS comercialmente disponível da Agilent Technologies, USA. Para extratos polares 10 µl são

injetados no sistema em uma vazão de fluxo de 200 µl/min. A coluna de separação (Fase Reversa C18) é mantida a 15°C durante a cromatografia. Para extratos lipídicos 5 µl são injetados no sistema em uma vazão de fluxo de 200 µl/min. A coluna de separação (Fase Reversa C18) é mantida a 30°C. HPLC é realizada com eluição em gradiente.

A análise espectrométrica de massa é conduzida em um instrumento de quadrupolo triplo Applied Biosystems API 4000 com fonte de pulverização iônica turbo. Para extratos polares o instrumento mede no modo de íon negativo no modo de varredura total de 100-1.000 amu. Para extratos lipídicos o instrumento mede no modo de íon positivo no modo de varredura total de 100-1.000 amu

f) Derivação da fase lipídica para análise por GC/MS

Para a transesterificação, uma mistura de 140 µl de clorofórmio, 37 µl de ácido clorídrico (37% em peso de HCl em água), 320 µl de metanol e 20 µl de tolueno é adicionada no extrato evaporado. O vaso é vedado hermeticamente e aquecido por 2 horas a 100°C, com agitação. A solução é subsequentemente evaporada até a secura. O resíduo é seco completamente.

A metoximação dos grupos carbonila é realizada pela reação com cloridrato de metóxi-amina (5 mg/ml em piridina, 100 µl por 1,5 horas a 60°C) em um vaso hermeticamente fechado. 20 µl de uma solução de ácidos graxos de cadeia linear, de número ímpar (solução de cada 0,3 mg/mL de ácidos graxos de 7 a 25 átomos de carbono e cada 0,6 mg/mL de ácidos graxos com 27, 29 e 31 átomos de carbono em piridina/tolueno 3/7 (v/v)) são adicionados como padrões de tempo. Finalmente, a derivação com 100 µl de N-metil-N-(trimetil-silil)-2,2,2-trifluoro-acetamida (MSTFA) é realizada por 30 minutos a 60°C, de novo no vaso hermeticamente fechado. O volume final antes da injeção no GC foi 220 µl.

g) Derivação da fase polar para a análise por GC/MS

A metoximação dos grupos carbonila é realizada pela reação

com cloridrato de metóxi-amina (5 mg/ml em piridina, 50 µl por 1,5 horas a 60°C) em um vaso hermeticamente fechado. 10 µl de uma solução de ácidos graxos de cadeia linear, de número ímpar (solução de cada 0,3 mg/mL de ácidos graxos de 7 a 25 átomos de carbono e cada 0,6 mg/mL de ácidos graxos com 27, 29 e 31 átomos de carbono em piridina/tolueno 3/7 (v/v)) são adicionados como padrões de tempo. Finalmente, a derivação com 50 µl de N-metil-N-(trimetil-silil)-2,2,2-trifluoro-acetamida (MSTFA) é realizada por 30 minutos a 60°C, de novo no vaso hermeticamente fechado. O volume final antes da injeção no GC foi 110 µl..

10 h) Análise por GC-MS

Os sistemas de GC-MS consistem de um Agilent 6890 GC acoplado em um Agilent 5973 MSD. Os autoamostradores são CompiPal ou GCPal de CTC. Para análise colunas de separação capilares comerciais normais (30 m x 0,25 mm x 0,25 µm) com fases estacionárias diferentes de poli(metil-siloxano) contendo 0 % a até 35% de grupos aromáticos, dependendo dos materiais de amostra analisados e das frações da etapa de separação de fases, são usadas (por exemplo: DB-1ms, HP-5ms, DB-XLB, DB-35ms, Agilent Technologies). Até 1 µL do volume final é injetado sem divisão e o programa de temperatura de forno é iniciado em 70°C e terminado em 340°C com taxas de aquecimento diferentes dependendo do material da amostra e da fração da etapa de separação de fases com o propósito de obter uma separação cromatográfica suficiente e número de varreduras suficiente dentro de cada pico de analito. São usadas condições padrão normais de GC-MS, por exemplo fluxo constante com 1 a 1,7 ml/min. nominal e hélio como gás de fase móvel. Ionização é feita por impacto de elétrons com 70 eV, varredura dentro de uma faixa de m/z de 15 a 600 com velocidades de varredura de 2,5 a 3 varreduras/segundo e condições de ajuste padrão.

25 i) Análise de várias amostras de planta

As amostras são medidas em série individual de 20 amostras

de planta cada (também chamadas de seqüências). Nos experimentos cada seqüência continha pelo menos 3 replicados por linhagem transgênica mais pelo menos 3 plantas da respectiva linhagem segregante nula como controles. As áreas de pico para cada analito são ajustadas para o peso seco estabelecido para a planta (área normalizada). Valores de razão são calculados para normalização adicional em relação ao controle. Nos experimentos valores de razão são calculados pela divisão da área normalizada pela média dos dados correspondentes do grupo de controle da mesma seqüência. Os valores obtidos são chamados de razão_por_controle. São comparáveis entre seqüências e indicam quanto de concentração de analito no mutante difere do grupo de controle, que são as plantas das linhagens segregantes nulas respectivas em uma dada seqüência. Controles apropriados são feitos antes de provar que os próprios vetor e procedimento de transformação não têm influência significativa sobre a composição metabólica das plantas. Portanto as mudanças descritas em comparação com o grupo de controle são indubitavelmente causadas pela mutação.

Os resultados das diferentes análises de planta podem ser vistos na seguinte tabela 3:

São analisadas as sementes de plantas de arroz contendo genes codificadores de proteínas CDK como aqui descritas.

20 Tabela 3: Resultados da análise metabólica de proteínas CDK em plantas de arroz

Metabólito	razão min.	razão máx.	Método
Metionina	1,254	1,456	GC
Prolina	0,627	0,628	GC
Asparagina	1,845	2,255	GC
Cisteína	1,526	2,029	GC
Luteína	1,793	2,593	LC
Zeaxantina	2,165	4,267	LC
Coenzima Q9	1,368	1,391	LC

Coluna 1 mostra o metabólito analisado. Colunas 2 e 3 estão mostrando a faixa de aumento do metabólico analisado como verificado entre as plantas transgênicas em comparação com as linhagens de controle. Coluna 4 indica o método analítico.

LISTAGEM DE SEQUÊNCIAS

<110> BASF Plant Science GmbH

<120> Plantas tendo características de crescimento melhoradas e método de preparar as mesmas

<130> PF58042

<140> 20060503

<141> 2007 05 24

<160> 60

<170> Biomax PatentTool de acordo com PatentIN de formato 3.1

<210> 1

<211> 885

<212> DNA

<213> Oryza sativa

<220>

<221> CDS

<222> (1)..(885)

<223> quinase cdc2

<400> 1

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

atc ttt gca gaa atg gtg aac cag aaa cca ctg ttc cct ggt gat tct	624
Ile Phe Ala Glu Met Val Asn Gln Lys Pro Leu Phe Pro Gly Asp Ser	
195 200 205	
gag att gat gaa tta ttt aag ata ttc agg gta cta gga act cca aat	672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Val Leu Gly Thr Pro Asn	
210 215 220	
gaa caa agt tgg cca gga gtt agc tca tta cct gac tac aag tct gct	720
Glu Gln Ser Trp Pro Gly Val Ser Ser Leu Pro Asp Tyr Lys Ser Ala	
225 230 235 240	
ttc ccc aag tgg cag gca cag gat ctt gca act att gtc cct act ctt	768
Phe Pro Lys Trp Gln Ala Gln Asp Leu Ala Thr Ile Val Pro Thr Leu	
245 250 255	
gac cct gct ggt ttg gac ctt ctc tct aaa atg ctt cgg tac gag cca	816
Asp Pro Ala Gly Leu Asp Leu Leu Ser Lys Met Leu Arg Tyr Glu Pro	
260 265 270	
aac aaa agg atc aca gct aga cag gct ctt gag cat gaa tac ttc aag	864
Asn Lys Arg Ile Thr Ala Arg Gln Ala Leu Glu His Glu Tyr Phe Lys	
275 280 285	
gac ctt gag atg gta caa tga	885
Asp Leu Glu Met Val Gln	
290 294	

<210> 2

<211> 294

<212> PRT

<213> *Oryza sativa*

<400> 2

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

<210> 3
 <211> 882
 <212> DNA
 <213> *Oryza sativa* (grupo-cultivar japónica)

<220>
 <221> CDS
 <222> (1)..(882)

```

<400> 3
atg gag cag tac gag gag gag aag att gga gag ggc acg tac ggg gta      48
Met Glu Gln Tyr Glu Glu Glu Lys Ile Gly Glu Gly Thr Tyr Gly Val
1      5      10      15
gtg tac aag gcg cgg gac aag gtc acc aac gag acg atc gcg ctc aag      96
Val Tyr Lys Ala Arg Asp Lys Val Thr Asn Glu Thr Ile Ala Leu Lys
20      25      30
aag atc cgg ctt gag cag gag gat gag ggt gtc ccc tcc acc gcc atc      144
Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala Ile
35      40      45
cgc gag atc tcg ctc ctc aag gag atg cat cac cgc aac atc gtc agg      192
Arg Glu Ile Ser Leu Leu Lys Glu Met His His Arg Asn Ile Val Arg
50      55      60
tta cac gat gtt atc cac agt gag aag cgc ata ggt ctt gtc ttt gag      240
Leu His Asp Val Ile His Ser Glu Lys Arg Ile Gly Leu Val Phe Glu
65      70      75      80
tat ctg gat ctg gac ctg aag aag ttc atg gac tct tgt cca gag ttt      288
Tyr Leu Asp Leu Asp Leu Lys Lys Phe Met Asp Ser Cys Pro Glu Phe
85      90      95
gca aaa aac ccc act tta att aag tca tat ctc tat cag ata ctc cgc      336
Ala Lys Asn Pro Thr Leu Ile Lys Ser Tyr Leu Tyr Gln Ile Leu Arg
100      105      110
ggc gtt gct tac tgt cat tct cat aga gtt ctt cat cga gat ttg aaa      384
Gly Val Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu Lys
115      120      125
cct cag aat tta ttg ata gat cgg cgt act aat aca ctg aag ctt gca      432
Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Thr Leu Lys Leu Ala
130      135      140
gac ttt ggt tta gcc agg gca ttt gga att cct gtc cgc aca ttt act      480
Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe Thr
145      150      155      160
cat gag gtt gta acc ttg tgg tat aga gct cca gag atc ctt ctt gga      528
His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu Gly
165      170      175
tca agg cag tat tct aca cca gtt gat atg tgg tca gtt ggt tgt atc      576
Ser Arg Gln Tyr Ser Thr Pro Val Asp Met Trp Ser Val Gly Cys Ile
180      185      190
ttt gca gaa atg gtg aac cag aaa cca ctg ttc cct ggt gat tct gag      624
Phe Ala Glu Met Val Asn Gln Lys Pro Leu Phe Pro Gly Asp Ser Glu
195      200      205
att gat gaa tta ttt aag ata ttc agg gta cta gga act cca aat gaa      672
Ile Asp Glu Leu Phe Lys Ile Phe Arg Val Leu Gly Thr Pro Asn Glu
210      215      220
caa agt tgg cca gga gtt agc tca tta cct gac tac aag tct gct ttc      720
Gln Ser Trp Pro Gly Val Ser Ser Leu Pro Asp Tyr Lys Ser Ala Phe
225      230      235      240
ccc aag tgg cag gca cag gct ctt gca act att gtc cct act ctt gac      768
Pro Lys Trp Gln Ala Gln Ala Leu Ala Thr Ile Val Pro Thr Leu Asp
245      250      255
cct gct ggt ttg gac ctt ctc tct aaa atg ctt cgg tac gag cca aac      816
Pro Ala Gly Leu Asp Leu Leu Ser Lys Met Leu Arg Tyr Glu Pro Asn
260      265      270
aaa agg atc aca gct aga cag gct ctt gag cat gag tac ttc aag gac      864
Lys Arg Ile Thr Ala Arg Gln Ala Leu Glu His Glu Tyr Phe Lys Asp
275      280      285

```

ctt gag atg gaa cgc taa
 Leu Glu Met Glu Arg
 290

882

<210> 4
 <211> 293
 <212> PRT
 <213> Oryza sativa (grupo-cultivar japonica)

<400> 4
 Met Glu Gln Tyr Glu Glu Glu Lys Ile Gly Glu Gly Thr Tyr Gly Val
 1 5 10 15
 Val Tyr Lys Ala Arg Asp Lys Val Thr Asn Glu Thr Ile Ala Leu Lys
 20 25 30
 Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala Ile
 35 40 45
 Arg Glu Ile Ser Leu Leu Lys Glu Met His His Arg Asn Ile Val Arg
 50 55 60
 Leu His Asp Val Ile His Ser Glu Lys Arg Ile Gly Leu Val Phe Glu
 65 70 75 80
 Tyr Leu Asp Leu Asp Leu Lys Lys Phe Met Asp Ser Cys Pro Glu Phe
 85 90 95
 Ala Lys Asn Pro Thr Leu Ile Lys Ser Tyr Leu Tyr Gln Ile Leu Arg
 100 105 110
 Gly Val Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu Lys
 115 120 125
 Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Thr Leu Lys Leu Ala
 130 135 140
 Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe Thr
 145 150 155 160
 His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu Gly
 165 170 175
 Ser Arg Gln Tyr Ser Thr Pro Val Asp Met Trp Ser Val Gly Cys Ile
 180 185 190
 Phe Ala Glu Met Val Asn Gln Lys Pro Leu Phe Pro Gly Asp Ser Glu
 195 200 205
 Ile Asp Glu Leu Phe Lys Ile Phe Arg Val Leu Gly Thr Pro Asn Glu
 210 215 220
 Gln Ser Trp Pro Gly Val Ser Ser Leu Pro Asp Tyr Lys Ser Ala Phe
 225 230 235 240
 Pro Lys Trp Gln Ala Gln Ala Leu Ala Thr Ile Val Pro Thr Leu Asp
 245 250 255
 Pro Ala Gly Leu Asp Leu Leu Ser Lys Met Leu Arg Tyr Glu Pro Asn
 260 265 270
 Lys Arg Ile Thr Ala Arg Gln Ala Leu Glu His Glu Tyr Phe Lys Asp
 275 280 285
 Leu Glu Met Glu Arg
 290

<210> 5
 <211> 885
 <212> DNA
 <213> cepa Allium

<220>
 <221> CDS
 <222> (1)..(885)

<400> 5
 atg gat cag tat gag aaa gtg gag aag att gga gaa gga act tat gga 48
 Met Asp Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly
 1 5 10 15
 gtt gtt tac aaa gca cgt gat cgg ctg act aat gaa acg att gct ttg 96
 Val Val Tyr Lys Ala Arg Asp Arg Leu Thr Asn Glu Thr Ile Ala Leu
 20 25 30
 aag aag att agg ttg gag cag gaa gat gag gga gtt cct agt act gcc 144

Lys	Lys	Ile	Arg	Leu	Glu	Gln	Glu	Asp	Glu	Gly	Val	Pro	Ser	Thr	Ala	
		35					40					45				
att	aga	gaa	ata	tca	ctg	ttg	aag	gaa	atg	cag	cat	gct	aac	att	gtc	192
Ile	Arg	Glu	Ile	Ser	Leu	Leu	Lys	Glu	Met	Gln	His	Ala	Asn	Ile	Val	
	50					55					60					
agg	ctg	caa	gac	gta	gtt	cat	agt	gag	aag	cga	ata	tat	ctt	gtg	ttc	240
Arg	Leu	Gln	Asp	Val	Val	His	Ser	Glu	Lys	Arg	Ile	Tyr	Leu	Val	Phe	
65					70					75					80	
gag	tat	cta	gat	ctg	gac	ctt	aag	aag	cat	atg	gat	tca	tgc	cca	gat	288
Glu	Tyr	Leu	Asp	Leu	Asp	Leu	Lys	Lys	His	Met	Asp	Ser	Cys	Pro	Asp	
				85					90					95		
ttt	gct	aaa	gat	tct	cgt	ttg	gct	aaa	aca	ttt	ctc	tat	cag	ctt	ctc	336
Phe	Ala	Lys	Asp	Ser	Arg	Leu	Ala	Lys	Thr	Phe	Leu	Tyr	Gln	Leu	Leu	
			100					105					110			
cga	gga	att	gct	tat	tgt	cac	tca	cac	cga	gtt	ctt	cat	cgt	gac	tta	384
Arg	Gly	Ile	Ala	Tyr	Cys	His	Ser	His	Arg	Val	Leu	His	Arg	Asp	Leu	
		115					120					125				
aag	cct	caa	aat	tta	ttg	atc	gac	aga	cgt	acc	aat	tca	tta	aag	ctt	432
Lys	Pro	Gln	Asn	Leu	Leu	Ile	Asp	Arg	Arg	Thr	Asn	Ser	Leu	Lys	Leu	
	130					135					140					
gct	gac	ttt	gga	ctt	gca	agg	gca	ttt	ggt	atc	cca	gtc	cga	acc	ttc	480
Ala	Asp	Phe	Gly	Leu	Ala	Arg	Ala	Phe	Gly	Ile	Pro	Val	Arg	Thr	Phe	
145					150					155					160	
aca	cac	gag	gtt	gtg	aca	ctg	tgg	tat	agg	gca	cct	gaa	atc	ctc	tta	528
Thr	His	Glu	Val	Thr	Leu	Trp	Tyr	Arg	Ala	Pro	Glu	Ile	Leu	Leu		
				165					170					175		
ggt	gct	cgt	cag	tat	tct	act	cct	gta	gac	ata	tgg	tct	gtg	gga	tgt	576
Gly	Ala	Arg	Gln	Tyr	Ser	Thr	Pro	Val	Asp	Ile	Trp	Ser	Val	Gly	Cys	
			180					185					190			
atc	ttt	gct	gaa	atg	gtg	aac	caa	cga	cct	cta	ttc	cct	ggg	gac	tct	624
Ile	Phe	Ala	Glu	Met	Val	Asn	Gln	Arg	Pro	Leu	Phe	Pro	Gly	Asp	Ser	
		195					200					205				
gag	atc	gac	gag	ctg	ttc	aaa	ata	ttt	aga	att	atg	ggt	acc	cca	aac	672
Glu	Ile	Asp	Glu	Leu	Phe	Lys	Ile	Phe	Arg	Ile	Met	Gly	Thr	Pro	Asn	
	210					215					220					
gaa	gac	aca	tgg	cca	ggt	gtt	act	tcc	ttg	ccc	gac	ttc	aag	tct	gct	720
Glu	Asp	Thr	Trp	Pro	Gly	Val	Thr	Ser	Leu	Pro	Asp	Phe	Lys	Ser	Ala	
225					230					235					240	
ttt	cca	aag	tgg	ccg	gca	aag	gac	ttg	gca	act	ata	gtt	cca	aag	ctt	768
Phe	Pro	Lys	Trp	Pro	Ala	Lys	Asp	Leu	Ala	Thr	Ile	Val	Pro	Lys	Leu	
				245					250					255		
gat	tca	gct	gga	att	gat	ctt	ctt	tat	aaa	atg	ctg	cac	ctt	gaa	ccg	816
Asp	Ser	Ala	Gly	Ile	Asp	Leu	Leu	Tyr	Lys	Met	Leu	His	Leu	Glu	Pro	
			260					265					270			
agc	aaa	aga	atc	act	gct	cgg	aag	gct	ctt	gag	cat	gaa	tac	ttc	agg	864
Ser	Lys	Arg	Ile	Thr	Ala	Arg	Lys	Ala	Leu	Glu	His	Glu	Tyr	Phe	Arg	
		275					280					285				
gat	ctt	ggg	aca	att	cca	tga										885
Asp	Leu	Gly	Thr	Ile	Pro											
	290															

<210> 6

<211> 294

<212> PRT

<213> cepa Allium

<400> 6

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

65					70					75				80
Glu	Tyr	Leu	Asp	Leu	Asp	Leu	Lys	Lys	His	Met	Asp	Ser	Cys	Pro
				85					90					95
Phe	Ala	Lys	Asp	Ser	Arg	Leu	Ala	Lys	Thr	Phe	Leu	Tyr	Gln	Leu
			100					105					110	
Arg	Gly	Ile	Ala	Tyr	Cys	His	Ser	His	Arg	Val	Leu	His	Arg	Asp
		115					120					125		Leu
Lys	Pro	Gln	Asn	Leu	Leu	Ile	Asp	Arg	Arg	Thr	Asn	Ser	Leu	Lys
	130					135					140			Leu
Ala	Asp	Phe	Gly	Leu	Ala	Arg	Ala	Phe	Gly	Ile	Pro	Val	Arg	Thr
	145				150					155				160
Thr	His	Glu	Val	Val	Thr	Leu	Trp	Tyr	Arg	Ala	Pro	Glu	Ile	Leu
			165						170					175
Gly	Ala	Arg	Gln	Tyr	Ser	Thr	Pro	Val	Asp	Ile	Trp	Ser	Val	Gly
			180					185					190	Cys
Ile	Phe	Ala	Glu	Met	Val	Asn	Gln	Arg	Pro	Leu	Phe	Pro	Gly	Asp
														Ser

	195					200				205				
Glu	Ile	Asp	Glu	Leu	Phe	Lys	Ile	Phe	Arg	Ile	Met	Gly	Thr	Pro
	210					215					220			Asn
Glu	Asp	Thr	Trp	Pro	Gly	Val	Thr	Ser	Leu	Pro	Asp	Phe	Lys	Ser
	225				230					235				Ala
Phe	Pro	Lys	Trp	Pro	Ala	Lys	Asp	Leu	Ala	Thr	Ile	Val	Pro	Lys
				245					250					255
Asp	Ser	Ala	Gly	Ile	Asp	Leu	Leu	Tyr	Lys	Met	Leu	His	Leu	Glu
			260					265					270	Pro
Ser	Lys	Arg	Ile	Thr	Ala	Arg	Lys	Ala	Leu	Glu	His	Glu	Tyr	Phe
		275					280					285		Arg
Asp	Leu	Gly	Thr	Ile	Pro									
	290													

<210> 7

<211> 885

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> CDS

<222> (1)..(885)

<400> 7

atg	gac	cag	tat	gaa	aaa	gtt	gag	aag	att	ggg	gaa	gga	aca	tat	ggt	48
Met	Asp	Gln	Tyr	Glu	Lys	Val	Glu	Lys	Ile	Gly	Glu	Gly	Thr	Tyr	Gly	
1				5					10					15		
gta	gtg	tac	aag	gct	cgt	gat	cgt	gta	act	aat	gaa	act	att	gca	ctg	96
Val	Val	Tyr	Lys	Ala	Arg	Asp	Arg	Val	Thr	Asn	Glu	Thr	Ile	Ala	Leu	
			20					25					30			
aag	aaa	ata	agg	ttg	gag	cag	gaa	gac	gag	ggg	gta	cca	agc	aca	gct	144
Lys	Lys	Ile	Arg	Leu	Glu	Gln	Glu	Asp	Glu	Gly	Val	Pro	Ser	Thr	Ala	
		35				40					45					
att	aga	gaa	ata	tct	ctc	ttg	aaa	gag	atg	caa	cat	gct	aat	att	gtg	192
Ile	Arg	Glu	Ile	Ser	Leu	Leu	Lys	Glu	Met	Gln	His	Ala	Asn	Ile	Val	
	50					55					60					
agg	ttg	cag	gat	gtg	gtg	cac	agt	gaa	aag	cga	ttg	tat	cta	gtg	ttt	240
Arg	Leu	Gln	Asp	Val	Val	His	Ser	Glu	Lys	Arg	Leu	Tyr	Leu	Val	Phe	
	65				70				75					80		
gaa	tat	ctt	gac	ttg	gac	ttg	aag	aag	cac	atg	gat	tcg	tgt	cct	gaa	288
Glu	Tyr	Leu	Asp	Leu	Asp	Leu	Lys	Lys	His	Met	Asp	Ser	Cys	Pro	Glu	
			85					90						95		
ttc	tct	aag	gat	cca	cgt	ctg	gtt	aaa	atg	ttc	ttg	tat	caa	ata	ctc	336
Phe	Ser	Lys	Asp	Pro	Arg	Leu	Val	Lys	Met	Phe	Leu	Tyr	Gln	Ile	Leu	
			100					105					110			
cgt	ggt	att	gct	tat	tgt	cat	tct	cat	aga	gtt	ctt	cat	aga	gat	ttg	384
Arg	Gly	Ile	Ala	Tyr	Cys	His	Ser	His	Arg	Val	Leu	His	Arg	Asp	Leu	
		115				120						125				
aag	cct	cag	aac	tta	cta	ata	gat	cga	cgt	aca	aat	gct	tta	aag	ctg	432

Lys	Pro	Gln	Asn	Leu	Leu	Ile	Asp	Arg	Arg	Thr	Asn	Ala	Leu	Lys	Leu		
130						135					140						
gca	gac	ttt	ggt	ttg	gct	aga	gca	ttt	ggt	att	cct	gtc	aga	act	ttc	480	
Ala	Asp	Phe	Gly	Leu	Ala	Arg	Ala	Phe	Gly	Ile	Pro	Val	Arg	Thr	Phe		
145						150					155				160		
act	cat	gag	gtg	gtg	aca	ttg	tgg	tac	agg	gca	cca	gaa	ata	ctg	ctt	528	
Thr	His	Glu	Val	Val	Thr	Leu	Trp	Tyr	Arg	Ala	Pro	Glu	Ile	Leu	Leu		
				165					170					175			
gga	tca	cgc	cat	tac	tct	act	cct	ggt	gat	gtg	tgg	tca	ggt	ggg	tgc	576	
Gly	Ser	Arg	His	Tyr	Ser	Thr	Pro	Val	Asp	Val	Trp	Ser	Val	Gly	Cys		
			180					185					190				
ata	ttt	gct	gag	atg	gtg	aat	cag	ccg	cct	ctg	ttt	cct	ggg	gac	tct	624	
Ile	Phe	Ala	Glu	Met	Val	Asn	Gln	Pro	Pro	Leu	Phe	Pro	Gly	Asp	Ser		
		195					200					205					
gag	att	gat	gaa	ctt	ttc	aag	att	ttc	aga	gta	ttg	ggg	act	cca	aat	672	
Glu	Ile	Asp	Glu	Leu	Phe	Lys	Ile	Phe	Arg	Val	Leu	Gly	Thr	Pro	Asn		
	210					215					220						
gag	gat	aca	tgg	cct	gga	gtg	act	tct	ctg	cct	gat	tac	aaa	tct	gcc	720	
Glu	Asp	Thr	Trp	Pro	Gly	Val	Thr	Ser	Leu	Pro	Asp	Tyr	Lys	Ser	Ala		
225					230					235					240		
ttc	cca	aaa	tgg	cct	cct	aag	gac	ctg	gca	att	att	gta	cca	aat	ggt	768	
Phe	Pro	Lys	Trp	Pro	Pro	Lys	Asp	Leu	Ala	Ile	Ile	Val	Pro	Asn	Val		
				245					250					255			
gat	gga	gca	ggc	ctt	gat	ctt	ctt	ggt	aaa	atg	ctc	tcc	ttg	gat	ccc	816	
Asp	Gly	Ala	Gly	Leu	Asp	Leu	Leu	Gly	Lys	Met	Leu	Ser	Leu	Asp	Pro		
			260					265					270				
agt	aag	aga	atc	acc	gcg	agg	aat	gcc	ctt	gag	cat	gag	tac	ttc	aag	864	
Ser	Lys	Arg	Ile	Thr	Ala	Arg	Asn	Ala	Leu	Glu	His	Glu	Tyr	Phe	Lys		
		275					280					285					
gat	att	ggg	tat	gtg	ccg	tga										885	
Asp	Ile	Gly	Tyr	Val	Pro												
		290															

<210> 8

<211> 294

<212> PRT

<213> Lycopersicon esculentum

<400> 8

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

Glu	Ile	Asp	Glu	Leu	Phe	Lys	Ile	Phe	Arg	Val	Leu	Gly	Thr	Pro	Asn		
						210						215					220
Glu	Asp	Thr	Trp	Pro	Gly	Val	Thr	Ser	Leu	Pro	Asp	Tyr	Lys	Ser	Ala		
						225						230					235
Phe	Pro	Lys	Trp	Pro	Lys	Asp	Leu	Ala	Ile	Ile	Val	Pro	Asn	Val			
						240						245					250
Asp	Gly	Ala	Gly	Leu	Asp	Leu	Leu	Gly	Lys	Met	Leu	Ser	Leu	Asp	Pro		
						255						260					265
Ser	Lys	Arg	Ile	Thr	Ala	Arg	Asn	Ala	Leu	Glu	His	Glu	Tyr	Phe	Lys		
						270						275					280
Asp	Ile	Gly	Tyr	Val	Pro												
						285										290	

```
<210> 9
<211> 885
<212> DNA
<213> Pisum sativum
```

```
<220>  
<221> CDS  
<222> (1)..(885)
```

<400>	9															
atg	gaa	cag	tat	gag	aag	gtt	gag	aaa	ata	gga	gaa	ggt	aca	tac	ggt	48
Met	Glu	Gln	Tyr	Glu	Lys	Val	Glu	Lys	Ile	Gly	Glu	Gly	Thr	Tyr	Gly	
1				5					10					15		
gtg	gtg	tac	aag	gct	agg	gac	cgt	gtt	acc	aat	gag	acc	att	gct	ttg	96
Val	Val	Tyr	Lys	Ala	Arg	Asp	Arg	Val	Thr	Asn	Glu	Thr	Ile	Ala	Leu	
			20					25					30			
aag	aag	att	cga	ctc	gaa	cag	gaa	gat	gag	ggg	gtt	cct	agc	act	gcc	144
Lys	Lys	Ile	Arg	Leu	Glu	Gln	Glu	Asp	Glu	Gly	Val	Pro	Ser	Thr	Ala	
		35					40					45				
ata	aga	gag	att	tct	ctt	ttg	aaa	gaa	atg	cag	cat	cgg	aac	att	gtt	192
Ile	Arg	Glu	Ile	Ser	Leu	Leu	Lys	Glu	Met	Gln	His	Arg	Asn	Ile	Val	
	50					55					60					
agg	ttg	cag	gat	gtt	gtg	cat	agt	gag	aag	cga	ttg	tat	ctt	gtt	ttt	240
Arg	Leu	Gln	Asp	Val	Val	His	Ser	Glu	Lys	Arg	Leu	Tyr	Leu	Val	Phe	
65					70					75					80	
gag	tac	ctt	gac	tta	gat	cta	aag	aag	cat	atg	gat	tca	tct	ccg	gaa	288
Glu	Tyr	Leu	Asp	Leu	Asp	Leu	Lys	Lys	His	Met	Asp	Ser	Ser	Pro	Glu	
				85					90					95		
ttt	tcc	aaa	gat	caa	cgt	caa	gta	aaa	atg	ttc	ctc	tat	caa	att	ctc	336
Phe	Ser	Lys	Asp	Gln	Arg	Gln	Val	Lys	Met	Phe	Leu	Tyr	Gln	Ile	Leu	
			100					105					110			
tgt	ggc	att	gct	tac	tgt	cat	tct	cat	aga	gtt	ctt	cac	cga	gac	ctg	384
Cys	Gly	Ile	Ala	Tyr	Cys	His	Ser	His	Arg	Val	Leu	His	Arg	Asp	Leu	
		115				120						125				
aaa	cca	caa	aat	ctg	ttg	ata	gat	cgc	agc	tct	aat	gcg	cta	aag	ctt	432
Lys	Pro	Gln	Asn	Leu	Leu	Ile	Asp	Arg	Ser	Ser	Asn	Ala	Leu	Lys	Leu	
	130					135					140					
gca	gat	ttt	ggg	ttg	gct	aga	gca	ttt	gga	att	cct	gtt	agg	aca	ttt	480
Ala	Asp	Phe	Gly	Leu	Ala	Arg	Ala	Phe	Gly	Ile	Pro	Val	Arg	Thr	Phe	
145					150					155					160	
aca	cat	gag	gtg	gtg	aca	cta	tgg	tac	aga	gct	cca	gaa	ata	ttg	ctt	528
Thr	His	Glu	Val	Val	Thr	Leu	Trp	Tyr	Arg	Ala	Pro	Glu	Ile	Leu	Leu	
				165					170					175		
ggg	tcc	cgt	cat	tat	tct	acc	cca	gtt	gat	gtt	tgg	tca	gtg	gga	tgc	576
Gly	Ser	Arg	His	Tyr	Ser	Thr	Pro	Val	Asp	Val	Trp	Ser	Val	Gly	Cys	
			180					185					190			
ata	ttt	gca	gag	atg	ata	aac	cag	cga	cca	ctt	ttc	cct	ggg	gat	tct	624
Ile	Phe	Ala	Glu	Met	Ile	Asn	Gln	Arg	Pro	Leu	Phe	Pro	Gly	Asp	Ser	
		195					200					205				
gag	att															

gaa gat aca tgg cct gga gtg act tca ttg cct gat ttt aaa tcc gcc	720
Glu Asp Thr Trp Pro Gly Val Thr Ser Leu Pro Asp Phe Lys Ser Ala	
225 230 235 240	
ttt ccc aag tgg cca tct aag gac ctg gca act ctg gtc cca agt ctt	768
Phe Pro Lys Trp Pro Ser Lys Asp Leu Ala Thr Leu Val Pro Ser Leu	
245 250 255	
gag cca tct ggt ctt gat ctg tta tct agt atg ctt cgc ttg gat ccc	816
Glu Pro Ser Gly Leu Asp Leu Leu Ser Ser Met Leu Arg Leu Asp Pro	
260 265 270	
agc aga aga att act gcc agg ggc gct ctt gag cac gaa tac ttc aaa	864
Ser Arg Arg Ile Thr Ala Arg Gly Ala Leu Glu His Glu Tyr Phe Lys	
275 280 285	
gac att aaa ttt gtc ccc tga	885
Asp Ile Lys Phe Val Pro	
290	

<210> 10

<211> 294

<212> PRT

<213> Pisum sativum

<400> 10

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

<210> 11

<211> 885

<212> DNA

<213> Camellia sinensis

<220>
 <221> CDS
 <222> (1)..(885)

<400> 11

atg gag cag tac gaa aaa gtt gag aag att ggt gaa gga acc tat ggt	48
Met Glu Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly	
1 5 10 15	
gtg gtc tat aag gct cgg gac cga gtg acc aat gaa act ata gct tta	96
Val Val Tyr Lys Ala Arg Asp Arg Val Thr Asn Glu Thr Ile Ala Leu	
20 25 30	
aag aag att cgt ttg gag cag gaa gat gag ggt gta cca agc acg gca	144
Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala	
35 40 45	
att aga gaa att tct cta ttg aaa gag atg aag cat gga aat gtt gtc	192
Ile Arg Glu Ile Ser Leu Leu Lys Glu Met Lys His Gly Asn Val Val	
50 55 60	
agg tta cag gat gta gtg cac agt gag aag cgt ttg tat ctg gtt ttt	240
Arg Leu Gln Asp Val Val His Ser Glu Lys Arg Leu Tyr Leu Val Phe	
65 70 75 80	
gag tat ctg gac ttg gat ttg aag aaa cac atg gat tca tgt ccg gag	288
Glu Tyr Leu Asp Leu Asp Leu Lys Lys His Met Asp Ser Cys Pro Glu	
85 90 95	
ttc tct aag gat cca cgt ctg ata aaa atg ttc ctg tat caa att ctt	336
Phe Ser Lys Asp Pro Arg Leu Ile Lys Met Phe Leu Tyr Gln Ile Leu	
100 105 110	
cgt ggt att gct tat tgt cac tcc cat aga gtt ctg cat cga gat ctg	384
Arg Gly Ile Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu	
115 120 125	
aaa cct cag aat ctg ctg ata gat cgc cgt aca aat gct tta aag ctt	432
Lys Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Ala Leu Lys Leu	
130 135 140	
gca gat ttt gga ttg gcc aga gca ttt ggt att ccc gtc agg aca ttt	480
Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe	
145 150 155 160	
aca cat gag gtg gtg act cta tgg tac agg gca cca gaa ata ctc ctt	528
Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu	
165 170 175	
gga tcc cgc cac tat tcg act cct gtg gat gtg tgg tca gtg ggt tgt	576
Gly Ser Arg His Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys	
180 185 190	
ata ttt gct gaa atg gtg aat cag cgg cca ttg ttt cct ggg gac tct	624
Ile Phe Ala Glu Met Val Asn Gln Arg Pro Leu Phe Pro Gly Asp Ser	
195 200 205	
gag att gat gaa cta ttc aaa atc ttc aga atc ttg ggt act cca aat	672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Ile Leu Gly Thr Pro Asn	
210 215 220	
gag gat acg tgg cct gga gtg act tct ttg gct gat ttt aaa tct gcc	720
Glu Asp Thr Trp Pro Gly Val Thr Ser Leu Ala Asp Phe Lys Ser Ala	
225 230 235 240	
ttt cca aaa tgg cct tct aag gac ctg gcc act gtg gtc cca aat ctt	768
Phe Pro Lys Trp Pro Ser Lys Asp Leu Ala Thr Val Val Pro Asn Leu	
245 250 255	
gac tca gct ggc att gat ctt ctt agt aaa atg ctc tgt ctg gat cca	816
Asp Ser Ala Gly Ile Asp Leu Leu Ser Lys Met Leu Cys Leu Asp Pro	
260 265 270	
agc aga aga att aca gct aga agt gcc ctt gaa cat gaa tac ttc aag	864
Ser Arg Arg Ile Thr Ala Arg Ser Ala Leu Glu His Glu Tyr Phe Lys	
275 280 285	
gat att ggg ttt gta cct tga	885
Asp Ile Gly Phe Val Pro	
290	

<210> 12
 <211> 294
 <212> PRT

<213> *Camellia sinensis*

<400> 12

```

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

```

<210> 13

<211> 885

<212> DNA

<213> *Pinus contorta*

<220>

<221> CDS

<222> (1)..(885)

<400> 13

```

atg gaa cag tat gag aaa gtt gag aag ata gga gaa gga aca tat ggt      48
Met Glu Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly
1      5      10      15
gtg gtt tac aag gcc cgt gat cgc ttg aca aat gag acc ata gct ctc      96
Val Val Tyr Lys Ala Arg Asp Arg Leu Thr Asn Glu Thr Ile Ala Leu
20      25      30
aag aaa att cgt ttg gag caa gaa gat gag ggt gta cca agc act gcg      144
Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala
35      40      45
att aga gaa att tct ctt ctt aaa gaa atg caa cat ggg aac atc gta      192
Ile Arg Glu Ile Ser Leu Leu Lys Glu Met Gln His Gly Asn Ile Val
50      55      60
agg ttg caa gat gtt gtc cat agt gaa aag cgg ctc tat ttg gtt ttc      240
Arg Leu Gln Asp Val Val His Ser Glu Lys Arg Leu Tyr Leu Val Phe

```

65		70		75		80	
gag tat ttg gat ttg gac ctc aag aag cat atg gat tct tgc cct gag							288
Glu Tyr Leu Asp Leu Asp Leu Lys Lys His Met Asp Ser Cys Pro Glu							
		85		90		95	
cta gca aag gat cct cgt cta atc aaa aca ttt ctg tat cag att ctg							336
Leu Ala Lys Asp Pro Arg Leu Ile Lys Thr Phe Leu Tyr Gln Ile Leu							
		100		105		110	
cgt ggc att gcc tat tgt cat tct cat cgg gtt ctt cat cgt gat ctg							384
Arg Gly Ile Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu							
		115		120		125	
aag ccg caa aat ttg ctt att gac cgc aaa acc aat gcg ttg aaa ctt							432
Lys Pro Gln Asn Leu Leu Ile Asp Arg Lys Thr Asn Ala Leu Lys Leu							
		130		135		140	
gcc gac ttt gga ctt gcc agg gca ttt gga att cca gtg agg acc ttt							480
Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe							
		145		150		155	160
act cat gag gtg gtt aca ttg tgg tat cgt gca ccc gag atc tta ctt							528
Thr His Glu Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu							
		165		170		175	
ggt tcc cgg cat tat tcg act cct gtt gat gtt tgg tct gtt gga tgt							576
Gly Ser Arg His Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys							
		180		185		190	
atc ttt gct gaa atg gtc aat cag cga cca ctt ttc cca gga gac tca							624
Ile Phe Ala Glu Met Val Asn Gln Arg Pro Leu Phe Pro Gly Asp Ser							
		195		200		205	
gag att gat gag ctc ttt aag ata ttt aga gtg ctg ggg acg cca aat							672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Val Leu Gly Thr Pro Asn							
		210		215		220	
gaa gaa aca tgg cca gga gtc acc tct ctg cct gac ttc aag tca gcc							720
Glu Glu Thr Trp Pro Gly Val Thr Ser Leu Pro Asp Phe Lys Ser Ala							
		225		230		235	240
ttc cca aag tgg cca gcc aag gat ttg gca act gtg gtt tca ggt ctt							768
Phe Pro Lys Trp Pro Ala Lys Asp Leu Ala Thr Val Val Ser Gly Leu							
		245		250		255	
gag cca gca gga att gat att ctc tcg aaa atg ctg tgc ctg gag ccc							816
Glu Pro Ala Gly Ile Asp Ile Leu Ser Lys Met Leu Cys Leu Glu Pro							
		260		265		270	
agt aga cgc atc act gct cgt agt gct ctg gag cac gag tat ttc aaa							864
Ser Arg Arg Ile Thr Ala Arg Ser Ala Leu Glu His Glu Tyr Phe Lys							
		275		280		285	
gat cta ggt ttt gta ccc tga							885
Asp Leu Gly Phe Val Pro							
		290					

<210> 14

<211> 294

<212> PRT

<213> Pinus contorta

<400> 14

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

Lys Pro Gln Asn Leu Leu Ile Asp Arg Lys Thr Asn Ala Leu Lys Leu
 130 135 140
 Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe
 145 150 155 160
 Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu
 165 170 175
 Gly Ser Arg His Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys
 180 185 190
 Ile Phe Ala Glu Met Val Asn Gln Arg Pro Leu Phe Pro Gly Asp Ser
 195 200 205
 Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Val Leu Gly Thr Pro Asn
 210 215 220
 Glu Glu Thr Trp Pro Gly Val Thr Ser Leu Pro Asp Phe Lys Ser Ala
 225 230 235 240
 Phe Pro Lys Trp Pro Ala Lys Asp Leu Ala Thr Val Val Ser Gly Leu
 245 250 255
 Glu Pro Ala Gly Ile Asp Ile Leu Ser Lys Met Leu Cys Leu Glu Pro
 260 265 270
 Ser Arg Arg Ile Thr Ala Arg Ser Ala Leu Glu His Glu Tyr Phe Lys
 275 280 285
 Asp Leu Gly Phe Val Pro
 290

<210> 15

<211> 885

<212> DNA

<213> Picea abies

<220>

<221> CDS

<222> (1) .. (885)

<400> 15

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

```

ggg tcc cga cat tat tcg act cct gtt gat gtt tgg tct gtg ggg tgt      576
Gly Ser Arg His Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys
180 185 190
atc ttt gct gaa atg gtg aat cag cga cca ctt ttc cca gga gac tca      624
Ile Phe Ala Glu Met Val Asn Gln Arg Pro Leu Phe Pro Gly Asp Ser
195 200 205
gag att gat gaa ctc ttt aag ata ttt aga gtg ctg ggg aca cca aat      672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Val Leu Gly Thr Pro Asn
210 215 220
gaa gaa aca tgg cca gga gtc acc tct ctg cca gac ttc aag tca gcc      720
Glu Glu Thr Trp Pro Gly Val Thr Ser Leu Pro Asp Phe Lys Ser Ala
225 230 235 240
ttc cca aag tgg cca gcc aag gat ttg gca act gtg gtt cca ggt ctt      768
Phe Pro Lys Trp Pro Ala Lys Asp Leu Ala Thr Val Val Pro Gly Leu
245 250 255
gag cca gca gga att gat ctt ctc tcg aaa atg ttg tgc ctg gag ccc      816
Glu Pro Ala Gly Ile Asp Leu Leu Ser Lys Met Leu Cys Leu Glu Pro
260 265 270
agt aaa cgc atc act gct cgt agt gct ctg gag cat gag tat ttc aaa      864
Ser Lys Arg Ile Thr Ala Arg Ser Ala Leu Glu His Glu Tyr Phe Lys
275 280 285
gat cta ggt ttt gta ccc tga      885
Asp Leu Gly Phe Val Pro
290

```

<210> 16

<211> 294

<212> PRT

<213> Picea abies

<400> 16

```

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

```

```

                275                280                285
Asp Leu Gly Phe Val Pro
290

<210> 17
<211> 885
<212> DNA
<213> Saccharum officinarum

<220>
<221> CDS
<222> (1)..(885)

<400> 17
atg gag cag tac gag aag gtg gag aag atc ggg gag ggc acg tac ggg      48
Met Glu Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly
1      5      10      15
gtg gtg tac aag ggg ctg gac aag gcc acc aac gag acg atc gcg ctc      96
Val Val Tyr Lys Gly Leu Asp Lys Ala Thr Asn Glu Thr Ile Ala Leu
20      25      30
aag aag atc cgc ctc gag cag gag gac gag ggc gtc ccg tcc acc gcc      144
Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala
35      40      45
atc cgc gag atc tct ctc ctc aag gag atg aac cac gac aac atc gtc      192
Ile Arg Glu Ile Ser Leu Lys Glu Met Asn His Asp Asn Ile Val
50      55      60
agg tta cat gat gtt atc cac agc gag aag cgc ata tat ctt gtc ttc      240
Arg Leu His Asp Val Ile His Ser Glu Lys Arg Ile Tyr Leu Val Phe
65      70      75      80
gag ttc ctg gat ctg gac ctg aag aag ttc atg gac tct tgc ccg gag      288
Glu Phe Leu Asp Leu Asp Leu Lys Lys Phe Met Asp Ser Cys Pro Glu
85      90      95
ttt gct aag aat ccc act ttg att aag tca tac ctc tac cag ata ctc      336
Phe Ala Lys Asn Pro Thr Leu Ile Lys Ser Tyr Leu Tyr Gln Ile Leu
100      105      110
cgt ggt gtt gct tac tgc cat tct cat aga ttt ctt cat cga gat ttg      384
Arg Gly Val Ala Tyr Cys His Ser His Arg Phe Leu His Arg Asp Leu
115      120      125
aaa cct cag aac tta ttg ata gat cgg cgt act aat aca ctg aag ctt      432
Lys Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Thr Leu Lys Leu
130      135      140
gca gac ttt ggt tta tcc agg gca ttt gga att cct gtc cgt aca ttc      480
Ala Asp Phe Gly Leu Ser Arg Ala Phe Gly Ile Pro Val Arg Thr Phe
145      150      155      160
act cat gag gta gtg aca tta tgg tac aga gct cca gaa att ctg ctt      528
Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu
165      170      175
gga gcg aaa cag tat tcc aca cca gtc gat gta tgg tct gtg ggc tgt      576
Gly Ala Lys Gln Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys
180      185      190
atc ttt gcg gaa atg gtg aac caa aag cca cta ttc cct ggt gat tct      624
Ile Phe Ala Glu Met Val Asn Gln Lys Pro Leu Phe Pro Gly Asp Ser
195      200      205
gag atc gac gaa ctg ttt aag ata ttc agg gta cta ggt aca cca aat      672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Val Leu Gly Thr Pro Asn
210      215      220
gaa caa agt tgg cca gga gtc agt tgt ttg cct gac ttc aag act gct      720
Glu Gln Ser Trp Pro Gly Val Ser Cys Leu Pro Asp Phe Lys Thr Ala
225      230      235      240
ttc ccc agg tgg cag gct cag gac ctg gca aca att gtc cca aat ctt      768
Phe Pro Arg Trp Gln Ala Gln Asp Leu Ala Thr Ile Val Pro Asn Leu
245      250      255
gaa cct gct ggt ttg gac ctt ctc tcg aaa atg ctt cga tac gag cca      816
Glu Pro Ala Gly Leu Asp Leu Leu Ser Lys Met Leu Arg Tyr Glu Pro
260      265      270
agc aaa aga atc aca gct agg caa gca ctt gag cat gag tac ttc aag      864

```

Ser Lys Arg Ile Thr Ala Arg Gln Ala Leu Glu His Glu Tyr Phe Lys
 275 280 285
 gac ctt gaa atg gta cag tga
 Asp Leu Glu Met Val Gln
 290

885

<210> 18
 <211> 294
 <212> PRT
 <213> Saccharum officinarum

<400> 18
 Met Glu Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly
 1 5 10 15
 Val Val Tyr Lys Gly Leu Asp Lys Ala Thr Asn Glu Thr Ile Ala Leu
 20 25 30
 Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala
 35 40 45
 Ile Arg Glu Ile Ser Leu Leu Lys Glu Met Asn His Asp Asn Ile Val
 50 55 60
 Arg Leu His Asp Val Ile His Ser Glu Lys Arg Ile Tyr Leu Val Phe
 65 70 75 80
 Glu Phe Leu Asp Leu Asp Leu Lys Lys Phe Met Asp Ser Cys Pro Glu
 85 90 95
 Phe Ala Lys Asn Pro Thr Leu Ile Lys Ser Tyr Leu Tyr Gln Ile Leu
 100 105 110
 Arg Gly Val Ala Tyr Cys His Ser His Arg Phe Leu His Arg Asp Leu
 115 120 125
 Lys Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Thr Leu Lys Leu
 130 135 140
 Ala Asp Phe Gly Leu Ser Arg Ala Phe Gly Ile Pro Val Arg Thr Phe
 145 150 155 160
 Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu
 165 170 175
 Gly Ala Lys Gln Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys
 180 185 190
 Ile Phe Ala Glu Met Val Asn Gln Lys Pro Leu Phe Pro Gly Asp Ser
 195 200 205
 Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Val Leu Gly Thr Pro Asn
 210 215 220
 Glu Gln Ser Trp Pro Gly Val Ser Cys Leu Pro Asp Phe Lys Thr Ala
 225 230 235 240
 Phe Pro Arg Trp Gln Ala Gln Asp Leu Ala Thr Ile Val Pro Asn Leu
 245 250 255
 Glu Pro Ala Gly Leu Asp Leu Leu Ser Lys Met Leu Arg Tyr Glu Pro
 260 265 270
 Ser Lys Arg Ile Thr Ala Arg Gln Ala Leu Glu His Glu Tyr Phe Lys
 275 280 285
 Asp Leu Glu Met Val Gln
 290

<210> 19
 <211> 885
 <212> DNA
 <213> Helianthus tuberosus

<220>
 <221> CDS
 <222> (1)..(885)

<400> 19
 atg gag cag tac gag aag gtt gag aag att ggt gaa gga aca tac ggt
 Met Glu Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly
 1 5 10 15
 gtg gtc tac aag gct cgt gac aaa gtg acc aat gaa aca att gct ttg
 Val Val Tyr Lys Ala Arg Asp Lys Val Thr Asn Glu Thr Ile Ala Leu

48

96

	20		25		30	
aaa aag atc cgt ctg gaa caa gaa gat gag ggt gtc ccg agc acg gct						144
Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala						
	35		40		45	
atc aga gag att tca ctg ttg aag gag atg caa cat gga aac att gtc						192
Ile Arg Glu Ile Ser Leu Leu Lys Glu Met Gln His Gly Asn Ile Val						
	50		55		60	
agg tta cag gat gtt gtg cac agt gac aaa cga ttg tat cta gtt ttt						240
Arg Leu Gln Asp Val Val His Ser Asp Lys Arg Leu Tyr Leu Val Phe						
	65		70		75	80
gag tat ctg gac ctt gat ttg aag aag cac atg gac tcg tgt cca gag						288
Glu Tyr Leu Asp Leu Asp Leu Lys Lys His Met Asp Ser Cys Pro Glu						
	85		90		95	
ttt tca aag gat cca cgc ctt gta aaa acg ttt ctt tat caa atc cta						336
Phe Ser Lys Asp Pro Arg Leu Val Lys Thr Phe Leu Tyr Gln Ile Leu						
	100		105		110	
cgt ggc att gct tat tgt cat tcc cat aga gtt ctc cac aga gat cta						384
Arg Gly Ile Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu						
	115		120		125	
aaa cct caa aac ttg tta att gat cgc cgt acc aat gca cta aaa ctt						432
Lys Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Ala Leu Lys Leu						
	130		135		140	
gcc gat ttt gga ttg gcc agg gcc ttt ggt atc cct gtc agg acg ttc						480
Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe						
	145		150		155	160
aca cat gag gtg gtg aca ctg tgg tat aga gcc ccc gaa att ctt ctt						528
Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu						
	165		170		175	
ggg tcc cgt cac tat tct act cct gtt gat gtg tgg tcc gtt ggt tgt						576
Gly Ser Arg His Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys						
	180		185		190	
ata ttt gct gag atg gtg aac cag cga cca ctc ttt cct gga gac tca						624
Ile Phe Ala Glu Met Val Asn Gln Arg Pro Leu Phe Pro Gly Asp Ser						
	195		200		205	
gag att gat gaa ctt ttc aag att ttc agg atc atg gga act cca aat						672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Ile Met Gly Thr Pro Asn						
	210		215		220	
gaa gag aca tgg cct ggt gtc act tct ctg cct gac ttt aag tcc gcc						720
Glu Glu Thr Trp Pro Gly Val Thr Ser Leu Pro Asp Phe Lys Ser Ala						
	225		230		235	240
ttt cca aag tgg tca tca aag gat ctg gca acc gtg gtc ccg aat ctt						768
Phe Pro Lys Trp Ser Ser Lys Asp Leu Ala Thr Val Val Pro Asn Leu						
	245		250		255	
gaa aaa gcc ggt ctc gat ctc tta tgt aaa atg ctg tgg ctt gat ccc						816
Glu Lys Ala Gly Leu Asp Leu Leu Cys Lys Met Leu Trp Leu Asp Pro						
	260		265		270	
agt aaa agg att aca gcc cgg acc gca ctg gag cat gaa tac ttc aaa						864
Ser Lys Arg Ile Thr Ala Arg Thr Ala Leu Glu His Glu Tyr Phe Lys						
	275		280		285	
gat atc gga ttc gtg cca taa						885
Asp Ile Gly Phe Val Pro						
	290					

<210> 20

<211> 294

<212> PRT

<213> Helianthus tuberosus

<400> 20

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

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

<210> 21

<211> 885

<212> DNA

<213> Daucus carota

<220>

<221> CDS

<222> (1)..(885)

<400> 21

atg gac cag tat gaa aag gtt gag aag att ggg gag ggg acg tat ggg	48
Met Asp Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly	
1	5
gta gtt tat aag gct cgt gac cgc gtt aca aat gaa act att gct tta	96
Val Val Tyr Lys Ala Arg Asp Arg Val Thr Asn Glu Thr Ile Ala Leu	
	20
aag aaa att cgt tta gag cag gaa gat gag gga gtg cca agc act gct	144
Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala	
	35
att aga gaa att tct tta ctg aag gag atg cag cat gaa aat att gtc	192
Ile Arg Glu Ile Ser Leu Leu Lys Glu Met Gln His Glu Asn Ile Val	
	50
agg cta caa gat gtt gtg cat agc gag aaa cgg ctg tat ctg gta ttt	240
Arg Leu Gln Asp Val Val His Ser Glu Lys Arg Leu Tyr Leu Val Phe	
65	70
gaa tat ctg gac ctg gat ctg aag aaa cat atg gat tca tgt cca gag	288
Glu Tyr Leu Asp Leu Asp Leu Lys Lys His Met Asp Ser Cys Pro Glu	
	85
ttt gca aag gat cca cgt ctg att aaa atg ttt ctg tat caa ata ctg	336
Phe Ala Lys Asp Pro Arg Leu Ile Lys Met Phe Leu Tyr Gln Ile Leu	
	100
cgt ggt att gct tat tgt cat tcc cat aga gtt ctg cat cgc gat ctc	384
Arg Gly Ile Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu	
	115
	120
	125

```

aaa ccc caa aac ctg ctc ata gat cgc cgt acc aat gct cta aag ctt      432
Lys Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Ala Leu Lys Leu
130                               135                               140
gca gat ttt gga ctt gcc agg gca ttt gga att cca gtc agg aca ttt      480
Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe
145                               150                               155                               160
acc cac gag gtg gtg aca ctt tgg tac agg gcc cca gag ata ctc ctt      528
Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu
165                               170                               175
gga tcc cgc cac tac tcc aca cca gtt gat gtg tgg tca gtt ggt tgt      576
Gly Ser Arg His Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys
180                               185                               190
att ttt gct gag atg gtg aac caa cag cca ttg ttt ccg ggc gat tct      624
Ile Phe Ala Glu Met Val Asn Gln Gln Pro Leu Phe Pro Gly Asp Ser
195                               200                               205
gag att gat gaa tta ttt aag att ttc aga att gtg ggt acc cca aat      672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Ile Val Gly Thr Pro Asn
210                               215                               220
gag gat acc tgg cct gga gta aca gct ctt cct gat ttt aag tct gcc      720
Glu Asp Thr Trp Pro Gly Val Thr Ala Leu Pro Asp Phe Lys Ser Ala
225                               230                               235                               240
ttc cca aaa tgg cca tct aag gaa ctg gga aac gta gtc cca aat ctc      768
Phe Pro Lys Trp Pro Ser Lys Glu Leu Gly Asn Val Val Pro Asn Leu
245                               250                               255
gat gtg gcc ggt ctc aat ctc ctc aaa aaa atg ctt tgc ttg gat ccc      816
Asp Val Ala Gly Leu Asn Leu Leu Lys Met Leu Cys Leu Asp Pro
260                               265                               270
agc aga aga att act gcc agg agt gca ctt gag cat gaa tac ttc aag      864
Ser Arg Arg Ile Thr Ala Arg Ser Ala Leu Glu His Glu Tyr Phe Lys
275                               280                               285
gat atc ggg att gtt cct taa
Asp Ile Gly Ile Val Pro
290

```

<210> 22

<211> 294

<212> PRT

<213> Daucus carota

<400> 22

```

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

```

Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Ile Val Gly Thr Pro Asn
 210 215 220
 Glu Asp Thr Trp Pro Gly Val Thr Ala Leu Pro Asp Phe Lys Ser Ala
 225 230 235 240
 Phe Pro Lys Trp Pro Ser Lys Glu Leu Gly Asn Val Val Pro Asn Leu
 245 250 255
 Asp Val Ala Gly Leu Asn Leu Leu Lys Lys Met Leu Cys Leu Asp Pro
 260 265 270
 Ser Arg Arg Ile Thr Ala Arg Ser Ala Leu Glu His Glu Tyr Phe Lys
 275 280 285
 Asp Ile Gly Ile Val Pro
 290

<210> 23

<211> 885

<212> DNA

<213> Coffea arabica

<220>

<221> CDS

<222> (1)..(885)

<400> 23

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

Glu	Asp	Thr	Trp	Pro	Gly	Val	Thr	Ser	Leu	Pro	Asp	Phe	Lys	Ser	Ala		
225					230					235					240		
ttt	cca	agg	tgg	ctt	tca	cag	gat	ctg	gca	act	gtg	ggt	ccc	aat	ctt		768
Phe	Pro	Arg	Trp	Leu	Ser	Gln	Asp	Leu	Ala	Thr	Val	Val	Pro	Asn	Leu		
				245					250					255			
gat	gct	gct	ggt	ctt	gat	ctc	ctg	cgt	aaa	atg	ctg	tgc	ctg	gat	ccc		816
Asp	Ala	Ala	Gly	Leu	Asp	Leu	Leu	Arg	Lys	Met	Leu	Cys	Leu	Asp	Pro		
			260					265					270				
agc	aag	aga	atc	aca	gct	agg	aat	gca	ctt	gag	cat	gag	tac	ttc	aaa		864
Ser	Lys	Arg	Ile	Thr	Ala	Arg	Asn	Ala	Leu	Glu	His	Glu	Tyr	Phe	Lys		
		275					280					285					
gat	atc	ggt	ttt	gtc	cct	taa											885
Asp	Ile	Gly	Phe	Val	Pro												
		290															

<210> 24

<211> 294

<212> PRT

<213> Coffea arabica

<400> 24

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

<210> 25

<211> 885

<212> DNA

<213> Helianthus annuus

<220>

<221> CDS

<222> (1)..(885)

<400> 25

atg gag cag tac gag aag gtt gag aag att ggt gaa gga aca tat ggt	48
Met Glu Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly	
1 5 10 15	
gtg gtc tac aag gct cgt gac aaa gtg acc aat gaa aca att gct ttg	96
Val Val Tyr Lys Ala Arg Asp Lys Val Thr Asn Glu Thr Ile Ala Leu	
20 25 30	
aaa aag atc cgt ttg gag caa gaa gat gag ggt gtc ccg agc acg gct	144
Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala	
35 40 45	
atc aga gag att tca ctg ttg aag gag atg caa cat gga aac att gtc	192
Ile Arg Glu Ile Ser Leu Leu Lys Glu Met Gln His Gly Asn Ile Val	
50 55 60	
agg tta cag gat gtt gtg cac agt gac aaa cga ttg tat cta gtt ttt	240
Arg Leu Gln Asp Val Val His Ser Asp Lys Arg Leu Tyr Leu Val Phe	
65 70 75 80	
gag tat ctg gac ctt gat ttg aag aag cac atg gac tcg tgt cca gag	288
Glu Tyr Leu Asp Leu Asp Leu Lys Lys His Met Asp Ser Cys Pro Glu	
85 90 95	
ttt tca aag gat cca cgc ctt gta aaa acg ttt ctt tat caa atc cta	336
Phe Ser Lys Asp Pro Arg Leu Val Lys Thr Phe Leu Tyr Gln Ile Leu	
100 105 110	
cgt ggc att gct tat tgt cat tcc cat aga gtt ctc cac aga gat cta	384
Arg Gly Ile Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu	
115 120 125	
aaa cct caa aac ttg tta att gat cgc cgt acc aat gca cta aaa ctt	432
Lys Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Ala Leu Lys Leu	
130 135 140	
gcc gat ttt gga ttg gcc agg gcc ttt ggt atc cct gtc agg aca ttc	480
Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe	
145 150 155 160	
aca cat gag gtg gtg aca ctg tgg tat aga gcc ccc gaa att ctt ctt	528
Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu	
165 170 175	
ggg tcc cgt cac tat tct act cct gtt gat gtg tgg tcc gtt ggt tgt	576
Gly Ser Arg His Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys	
180 185 190	
ata ttt gct gag atg gtg aac cag cga cca ctc ttt cct gga gac tca	624
Ile Phe Ala Glu Met Val Asn Gln Arg Pro Leu Phe Pro Gly Asp Ser	
195 200 205	
gag att gat gaa ctt ttc aag att ttc agg atc atg gga act cca aat	672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Ile Met Gly Thr Pro Asn	
210 215 220	
gaa gag aca tgg cct ggc gtc act tct ctg cct gac ttt aag tcc gcc	720
Glu Glu Thr Trp Pro Gly Val Thr Ser Leu Pro Asp Phe Lys Ser Ala	
225 230 235 240	
ttt cca aag tgg tca tca aag gat ctg gca acc gtg gtc ccg aat ctt	768
Phe Pro Lys Trp Ser Ser Lys Asp Leu Ala Thr Val Val Pro Asn Leu	
245 250 255	
gaa aaa acc ggt ctc gat ctc tta cgt aaa atg ctg tgc ttg gat ccc	816
Glu Lys Thr Gly Leu Asp Leu Leu Arg Lys Met Leu Cys Leu Asp Pro	
260 265 270	
agt aaa aga att aca gcc cgg acc gca ctg gag cat gaa tac ttc aaa	864
Ser Lys Arg Ile Thr Ala Arg Thr Ala Leu Glu His Glu Tyr Phe Lys	
275 280 285	
gat atc gga ttc gtg cca taa	885
Asp Ile Gly Phe Val Pro	
290	

<210> 26

<211> 294

<212> PRT

<213> Helianthus annuus

<400> 26

```

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

```

<210> 27

<211> 885

<212> DNA

<213> Populus tremula x Populus tremuloides

<220>

<221> CDS

<222> (1)..(885)

<400> 27

```

atg gat cag tat gag aaa gtg gaa aaa att ggt gaa gga acc tac gga      48
Met Asp Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly
1      5      10      15
gtg gtc tac aaa gct cgt gat cgt gtc acc aat gag acc att gct ttg      96
Val Val Tyr Lys Ala Arg Asp Arg Val Thr Asn Glu Thr Ile Ala Leu
20      25      30
aag aag atc cgt ttg gag cag gaa gac gag ggc gta ccc agc act gct      144
Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala
35      40      45
atc cga gaa att tct ctc ttg aaa gag atg cag cat ggt aac att gtc      192
Ile Arg Glu Ile Ser Leu Leu Lys Glu Met Gln His Gly Asn Ile Val
50      55      60
aga ctg cag gat gtg gtg cac agt gag aag cgc ctt tac ttg gtt ttt      240
Arg Leu Gln Asp Val Val His Ser Glu Lys Arg Leu Tyr Leu Val Phe
65      70      75      80

```

gag tat cta gac ttg gat ttg aag aag cac atg gat tct tct cct gaa	288
Glu Tyr Leu Asp Leu Asp Leu Lys Lys His Met Asp Ser Ser Pro Glu	
85 90 95	
ttt gct aag gat cca cgc ctt gtg aaa aca ttt ctt tat caa att ctc	336
Phe Ala Lys Asp Pro Arg Leu Val Lys Thr Phe Leu Tyr Gln Ile Leu	
100 105 110	
cgt ggc att gct tac tgc cat tct cat aga gtt ctg cat cga gat ttg	384
Arg Gly Ile Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu	
115 120 125	
aaa cct cag aat ttg ctt att gat cgc cgt acc aac gct ctg aag ctt	432
Lys Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Ala Leu Lys Leu	
130 135 140	
gca gat ttt ggt ctg gct aga gca ttt ggt ata cct gtt agg aca ttt	480
Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe	
145 150 155 160	
aca cat gag gtt gtt acc ctg tgg tat aga gcc cct gaa att ctg ctt	528
Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu	
165 170 175	
gga tct cgc cat tac tca act cca gtt gat gtg tgg tcg gtg gga tgt	576
Gly Ser Arg His Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys	
180 185 190	
ata ttt gct gag atg gtg aac cag aag cca tta ttc cca ggg gat tcc	624
Ile Phe Ala Glu Met Val Asn Gln Lys Pro Leu Phe Pro Gly Asp Ser	
195 200 205	
gag att gat gaa cta ttc aaa att ttc aga atc ttg ggt act cca aat	672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Ile Leu Gly Thr Pro Asn	
210 215 220	
gag gac acg tgg cct gga gtt act tct ttg ccc gac ttc aag agt gca	720
Glu Asp Thr Trp Pro Gly Val Thr Ser Leu Pro Asp Phe Lys Ser Ala	
225 230 235 240	
ttc cct aag tgg cct tct aag gat ttg gca act gta gtt cca acc ctt	768
Phe Pro Lys Trp Pro Ser Lys Asp Leu Ala Thr Val Val Pro Thr Leu	
245 250 255	
gaa aaa gct ggt gtg gat ctt ctc tct aaa atg ctt ttc ttg gat ccc	816
Glu Lys Ala Gly Val Asp Leu Leu Ser Lys Met Leu Phe Leu Asp Pro	
260 265 270	
act aaa aga att act gcc agg agt gct ttg gag cat gaa tac ttc aag	864
Thr Lys Arg Ile Thr Ala Arg Ser Ala Leu Glu His Glu Tyr Phe Lys	
275 280 285	
gat att ggt ttt gta cca taa	885
Asp Ile Gly Phe Val Pro	
290	

<210> 28

<211> 294

<212> PRT

<213> Populus tremula x Populus tremuloides

<400> 28

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

130	135	140
Ala Asp Phe Gly Leu	Ala Arg Ala Phe Gly Ile	Pro Val Arg Thr Phe
145	150	155
Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala	Pro Glu Ile Leu Leu	160
	165	170
Gly Ser Arg His Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys		175
	180	185
Ile Phe Ala Glu Met Val Asn Gln Lys Pro Leu Phe Pro Gly Asp Ser		190
	195	200
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Ile Leu Gly Thr Pro Asn		205
	210	215
Glu Asp Thr Trp Pro Gly Val Thr Ser Leu Pro Asp Phe Lys Ser Ala		220
225	230	235
Phe Pro Lys Trp Pro Ser Lys Asp Leu Ala Thr Val Val Pro Thr Leu		240
	245	250
Glu Lys Ala Gly Val Asp Leu Leu Ser Lys Met Leu Phe Leu Asp Pro		255
	260	265
Thr Lys Arg Ile Thr Ala Arg Ser Ala Leu Glu His Glu Tyr Phe Lys		270
	275	280
Asp Ile Gly Phe Val Pro		285
290		

<210> 29

<211> 882

<212> DNA

<213> *Nicotiana tabacum*

<220>

<221> CDS

<222> (1)..(882)

<400> 29

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

Gly	Ser	Arg	His	Tyr	Ser	Thr	Pro	Val	Asp	Val	Trp	Ser	Val	Gly	Cys	
			180					185					190			
ata	ttt	gct	gag	atg	gtg	act	cag	cgc	cct	ctg	ttt	cct	ggt	gac	tcc	624
Ile	Phe	Ala	Glu	Met	Val	Thr	Gln	Arg	Pro	Leu	Phe	Pro	Gly	Asp	Ser	
		195					200					205				
gag	att	gat	gaa	ctt	ttc	aag	att	ttc	aga	gtg	atg	ggt	act	cca	aat	672
Glu	Ile	Asp	Glu	Leu	Phe	Lys	Ile	Phe	Arg	Val	Met	Gly	Thr	Pro	Asn	
	210					215				220						
gag	gat	aca	tgg	cct	gga	gtg	act	act	ctg	cct	gat	ttt	aaa	tct	gcc	720
Glu	Asp	Thr	Trp	Pro	Gly	Val	Thr	Thr	Leu	Pro	Asp	Phe	Lys	Ser	Ala	
225					230				235						240	
ttc	cca	aaa	tgg	cct	tct	aag	gac	ctg	gca	act	att	gtc	cca	aat	ctt	768
Phe	Pro	Lys	Trp	Pro	Ser	Lys	Asp	Leu	Ala	Thr	Ile	Val	Pro	Asn	Leu	
			245						250					255		
gat	gga	gca	ggc	ctt	gat	ctt	ctt	gat	aaa	atg	ctc	cgc	ttg	gat	ccc	816
Asp	Gly	Ala	Gly	Leu	Asp	Leu	Leu	Asp	Lys	Met	Leu	Arg	Leu	Asp	Pro	
		260						265					270			
agc	aag	aga	atc	act	gcc	agg	aat	gcc	ctt	gag	cat	gag	tac	ttc	aag	864
Ser	Lys	Arg	Ile	Thr	Ala	Arg	Asn	Ala	Leu	Glu	His	Glu	Tyr	Phe	Lys	
		275					280						285			
gat	att	ggg	tat	gtt	ccg											882
Asp	Ile	Gly	Tyr	Val	Pro											
		290														

<210> 30

<211> 294

<212> PRT

<213> Nicotiana tabacum

<400> 30

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

Asp Ile Gly Tyr Val Pro
290

<210> 31
<211> 885
<212> DNA
<213> Triticum aestivum
<220>
<221> CDS
<222> (1)..(885)

```

<400> 31
atg gag cag tac gag aag gtg gag aag atc ggg gag ggc acg tac ggg      48
Met Glu Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly
1      5      10      15
gtg gtg tac aag gcc cgg gac agg acc acc aac gag acc atc gcg ctc      96
Val Val Tyr Lys Ala Arg Asp Arg Thr Thr Asn Glu Thr Ile Ala Leu
20      25      30
aag aag atc cgc ctg gag cag gag gac gag ggc gtc ccc tcc acc gcc      144
Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala
35      40      45
atc cgc gag atc tcg ctc ctc aag gag atg cag cac ggc aac atc gtc      192
Ile Arg Glu Ile Ser Leu Leu Lys Glu Met Gln His Gly Asn Ile Val
50      55      60
aag ctg cac gat gtt gtc cac agc gag aag cgc ata tgg ctc gtc ttt      240
Lys Leu His Asp Val Val His Ser Glu Lys Arg Ile Trp Leu Val Phe
65      70      75      80
gag tac ctg gat ctg gac ctg aag aag ttc atg gac tcc tgt cca gag      288
Glu Tyr Leu Asp Leu Asp Leu Lys Lys Phe Met Asp Ser Cys Pro Glu
85      90      95
ttt gcc aag agc ccc gcc ttg atc aag tca tat ctc tat cag ata ctc      336
Phe Ala Lys Ser Pro Ala Leu Ile Lys Ser Tyr Leu Tyr Gln Ile Leu
100      105      110
cgc ggc gtt gct tac tgt cat tct cat aga gtt ctt cat cga gat ttg      384
Arg Gly Val Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu
115      120      125
aaa cct cag aat tta ttg ata gac cgg cgt act aat gca ctg aag ctt      432
Lys Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Ala Leu Lys Leu
130      135      140
gca gac ttt ggt tta gca agg gca ttt gga att cct gtc cgt aca ttt      480
Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe
145      150      155      160
act cat gag gta gta aca tta tgg tac aga gct cct gaa atc ctt ctt      528
Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu
165      170      175
gga gca agg cag tat tcc aca cca gtt gac gtg tgg tca gtg ggc tgt      576
Gly Ala Arg Gln Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys
180      185      190
atc ttt gca gaa atg gtg aac cag aaa cca ctg ttc cct ggc gat tct      624
Ile Phe Ala Glu Met Val Asn Gln Lys Pro Leu Phe Pro Gly Asp Ser
195      200      205
gag att gat gag cta ttt aag ata ttc agg gta ctc ggc act cca aat      672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Val Leu Gly Thr Pro Asn
210      215      220
gaa caa act tgg cca ggc gtg agt tcc ttg cct gac tac aag tcc gcc      720
Glu Gln Thr Trp Pro Gly Val Ser Ser Leu Pro Asp Tyr Lys Ser Ala
225      230      235      240
ttc ccc agg tgg cag gca gag gac ctt gca acc gtt gtc ccc aat ctt      768
Phe Pro Arg Trp Gln Ala Glu Asp Leu Ala Thr Val Val Pro Asn Leu
245      250      255
gaa cct gtt ggc ctg gac ctt ctc tcg aaa atg ctt cgg ttc gag cca      816
Glu Pro Val Gly Leu Asp Leu Leu Ser Lys Met Leu Arg Phe Glu Pro
260      265      270
aac aag agg atc acg gct agg cag gct ctt gag cat gag tac ttc aag      864
Asn Lys Arg Ile Thr Ala Arg Gln Ala Leu Glu His Glu Tyr Phe Lys
275      280      285

```

gac atg gag atg gta cag tga
 Asp Met Glu Met Val Gln
 290

885

<210> 32
 <211> 294
 <212> PRT
 <213> Triticum aestivum

<400> 32
 Met Glu Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly
 1 5 10 15
 Val Val Tyr Lys Ala Arg Asp Arg Thr Thr Asn Glu Thr Ile Ala Leu
 20 25 30
 Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala
 35 40 45
 Ile Arg Glu Ile Ser Leu Leu Lys Glu Met Gln His Gly Asn Ile Val
 50 55 60
 Lys Leu His Asp Val Val His Ser Glu Lys Arg Ile Trp Leu Val Phe
 65 70 75 80
 Glu Tyr Leu Asp Leu Asp Leu Lys Lys Phe Met Asp Ser Cys Pro Glu
 85 90 95
 Phe Ala Lys Ser Pro Ala Leu Ile Lys Ser Tyr Leu Tyr Gln Ile Leu
 100 105 110
 Arg Gly Val Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu
 115 120 125
 Lys Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Ala Leu Lys Leu
 130 135 140
 Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe
 145 150 155 160
 Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu
 165 170 175
 Gly Ala Arg Gln Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys
 180 185 190
 Ile Phe Ala Glu Met Val Asn Gln Lys Pro Leu Phe Pro Gly Asp Ser
 195 200 205
 Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Val Leu Gly Thr Pro Asn
 210 215 220
 Glu Gln Thr Trp Pro Gly Val Ser Ser Leu Pro Asp Tyr Lys Ser Ala
 225 230 235 240
 Phe Pro Arg Trp Gln Ala Glu Asp Leu Ala Thr Val Val Pro Asn Leu
 245 250 255
 Glu Pro Val Gly Leu Asp Leu Leu Ser Lys Met Leu Arg Phe Glu Pro
 260 265 270
 Asn Lys Arg Ile Thr Ala Arg Gln Ala Leu Glu His Glu Tyr Phe Lys
 275 280 285
 Asp Met Glu Met Val Gln
 290

<210> 33
 <211> 885
 <212> DNA
 <213> Medicago sativa

<220>
 <221> CDS
 <222> (1)..(885)

<400> 33
 atg gaa cag tac gag aaa gtt gag aag ata gga gaa ggt act tac ggt 48
 Met Glu Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly
 1 5 10 15
 gtg gtt tac aag gct cgt gac cgt gct acc aat gag acg ata gct ttg 96
 Val Val Tyr Lys Ala Arg Asp Arg Ala Thr Asn Glu Thr Ile Ala Leu
 20 25 30
 aag aag att cgt ctt gag cag gaa gat gag gga gtt ccg agt acc gct 144

Lys	Lys	Ile	Arg	Leu	Glu	Gln	Glu	Asp	Glu	Gly	Val	Pro	Ser	Thr	Ala		
		35					40					45					
att	cga	gag	att	tct	ctc	ttg	aag	gaa	atg	cag	cac	agg	aac	att	gtt	192	
Ile	Arg	Glu	Ile	Ser	Leu	Leu	Lys	Glu	Met	Gln	His	Arg	Asn	Ile	Val		
	50					55					60						
agg	ttg	cag	gat	gtg	gtg	cac	agt	gag	aag	cga	ttg	tat	ctg	gtt	ttt	240	
Arg	Leu	Gln	Asp	Val	Val	His	Ser	Glu	Lys	Arg	Leu	Tyr	Leu	Val	Phe		
65				70						75					80		
gag	tac	ctt	gac	ttg	gat	cta	aag	aag	ttt	atg	gat	tca	tct	cca	gaa	288	
Glu	Tyr	Leu	Asp	Leu	Asp	Leu	Lys	Lys	Phe	Met	Asp	Ser	Ser	Pro	Glu		
				85					90					95			
ttt	gca	aaa	gat	caa	cgg	caa	ata	aag	atg	ttc	ctt	tat	caa	att	ctc	336	
Phe	Ala	Lys	Asp	Gln	Arg	Gln	Ile	Lys	Met	Phe	Leu	Tyr	Gln	Ile	Leu		
			100					105					110				
tgt	ggc	att	gct	tac	tgt	cat	tct	cat	aga	gtt	ctt	cat	aga	gac	ttg	384	
Cys	Gly	Ile	Ala	Tyr	Cys	His	Ser	His	Arg	Val	Leu	His	Arg	Asp	Leu		
		115				120						125					
aaa	cca	cag	aat	ctg	ctg	att	gat	cgc	agc	tct	aat	gcc	gta	aag	ctt	432	
Lys	Pro	Gln	Asn	Leu	Leu	Ile	Asp	Arg	Ser	Ser	Asn	Ala	Val	Lys	Leu		
	130					135					140						
gca	gat	ttt	gga	ttg	gcc	agg	gca	ttt	gga	att	cct	gtc	agg	aca	ttt	480	
Ala	Asp	Phe	Gly	Leu	Ala	Arg	Ala	Phe	Gly	Ile	Pro	Val	Arg	Thr	Phe		
145				150						155					160		
aca	cat	gag	gtg	gtg	aca	ctc	tgg	tac	aga	gct	cca	gaa	ata	ttg	ctt	528	
Thr	His	Glu	Val	Val	Thr	Leu	Trp	Tyr	Arg	Ala	Pro	Glu	Ile	Leu	Leu		
				165						170				175			
ggg	tct	cgt	cat	tat	tct	acc	ccg	gtt	gat	gtc	tgg	tca	gtg	gga	tgc	576	
Gly	Ser	Arg	His	Tyr	Ser	Thr	Pro	Val	Asp	Val	Trp	Ser	Val	Gly	Cys		
			180					185					190				
ata	ttt	gca	gag	atg	ata	aac	caa	cgg	cca	ctt	ttc	cca	ggg	gac	tct	624	
Ile	Phe	Ala	Glu	Met	Ile	Asn	Gln	Arg	Pro	Leu	Phe	Pro	Gly	Asp	Ser		
		195				200						205					
gag	att	gat	gaa	ttg	ttt	aaa	ata	ttc	aga	atc	acg	ggt	aca	ccg	aat	672	
Glu	Ile	Asp	Glu	Leu	Phe	Lys	Ile	Phe	Arg	Ile	Thr	Gly	Thr	Pro	Asn		
	210					215					220						
gag	gaa	aca	tgg	cct	gga	gtg	act	tca	ttg	cct	gat	ttt	aaa	tca	gcc	720	
Glu	Glu	Thr	Trp	Pro	Gly	Val	Thr	Ser	Leu	Pro	Asp	Phe	Lys	Ser	Ala		
225				230						235				240			
ttt	ccc	aag	tgg	cca	gct	aag	gac	ctg	gca	act	caa	gtc	cca	aat	ctg	768	
Phe	Pro	Lys	Trp	Pro	Ala	Lys	Asp	Leu	Ala	Thr	Gln	Val	Pro	Asn	Leu		
				245					250					255			
gag	cca	gct	ggt	ctt	gat	ctt	cta	tcc	agt	act	tgt	cgc	ttg	gat	ccc	816	
Glu	Pro	Ala	Gly	Leu	Asp	Leu	Leu	Ser	Ser	Thr	Cys	Arg	Leu	Asp	Pro		
			260					265					270				
acc	aga	aga	att	act	gcc	agg	gga	gct	ctt	gag	cat	gaa	tac	ttc	aaa	864	
Thr	Arg	Arg	Ile	Thr	Ala	Arg	Gly	Ala	Leu	Glu	His	Glu	Tyr	Phe	Lys		
		275				280						285					
gac	att	aag	ttt	gtc	cca	tga										885	
Asp	Ile	Lys	Phe	Val	Pro												
	290																

<210> 34

<211> 294

<212> PRT

<213> Medicago sativa

<400> 34

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

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

<210> 35

<211> 1283

<212> DNA

<213> Antirrhinum majus

<220>

<221> CDS

<222> (29)..(913)

<223> n significa a ou g ou c ou t/u, s significa g ou c

<400> 35

gtacatgggtt tcattctcatc gatctcttg atg gag cag tat gaa aag gtt gag	52
Met Glu Gln Tyr Glu Lys Val Glu	
1 5	
aag att ggg gag gga acg tat gga gtg gta tac aag gct cgt gat cgt	100
Lys Ile Gly Glu Gly Thr Tyr Gly Val Val Tyr Lys Ala Arg Asp Arg	
10 15 20	
gta aca aat gag act ata gcc ttg aag aaa atc cgt cta gag cag gaa	148
Val Thr Asn Glu Thr Ile Ala Leu Lys Lys Ile Arg Leu Glu Gln Glu	
25 30 35 40	
gat gag gga gtg ccc agc aca gct atc aga gag att tct ctc ttg aaa	196
Asp Glu Gly Val Pro Ser Thr Ala Ile Arg Glu Ile Ser Leu Leu Lys	
45 50 55	
gag atg caa cat ggg aat atc gtg agg ttg cag gat gtg gtg cac agt	244
Glu Met Gln His Gly Asn Ile Val Arg Leu Gln Asp Val Val His Ser	
60 65 70	
gag aag cgc ttg tac ctg gtg ttt gaa tat ctg gac ttg gat ttg aaa	292
Glu Lys Arg Leu Tyr Leu Val Phe Glu Tyr Leu Asp Leu Asp Leu Lys	
75 80 85	
aaa cat atg gat tca tgc cca gaa ttc tcc cag gat cct cgt ttg gtt	340
Lys His Met Asp Ser Cys Pro Glu Phe Ser Gln Asp Pro Arg Leu Val	
90 95 100	
aaa atg ttt ctg tat caa ata cta cgt ggt atc gcc tac tgt cat tct	388
Lys Met Phe Leu Tyr Gln Ile Leu Arg Gly Ile Ala Tyr Cys His Ser	
105 110 115 120	
cat cgt gtc ctt cat cga gat ttg aag cct caa aac ttg ctg ata gac	436

His	Arg	Val	Leu	His	Arg	Asp	Leu	Lys	Pro	Gln	Asn	Leu	Leu	Ile	Asp	
			125						130					135		
cgc	cgt	acc	aat	gca	tta	aag	ctt	gct	gac	ttt	gga	ttg	gcc	aga	gca	484
Arg	Arg	Thr	Asn	Ala	Leu	Lys	Leu	Ala	Asp	Phe	Gly	Leu	Ala	Arg	Ala	
			140					145					150			
ttt	ggt	att	cca	gtc	agg	act	ttt	aca	cat	gag	gtt	gtg	aca	ctg	tgg	532
Phe	Gly	Ile	Pro	Val	Arg	Thr	Phe	Thr	His	Glu	Val	Val	Thr	Leu	Trp	
		155					160					165				
tac	agg	gct	cca	gaa	ata	cta	ctt	gga	tct	cgc	cat	tat	tct	act	cca	580
Tyr	Arg	Ala	Pro	Glu	Ile	Leu	Leu	Gly	Ser	Arg	His	Tyr	Ser	Thr	Pro	
		170				175					180					
gtg	gat	gtc	tgg	tca	gtt	ggt	tgt	ata	ttt	gct	gaa	atg	gtt	aac	caa	628
Val	Asp	Val	Trp	Ser	Val	Gly	Cys	Ile	Phe	Ala	Glu	Met	Val	Asn	Gln	
		185			190					195				200		
cgg	cct	ttg	ttt	cct	ggg	gac	tct	gag	att	gat	gaa	cta	ttc	aaa	att	676
Arg	Pro	Leu	Phe	Pro	Gly	Asp	Ser	Glu	Ile	Asp	Glu	Leu	Phe	Lys	Ile	
			205					210					215			
ttt	aga	gtc	atg	ggt	acc	cca	aat	gaa	gaa	aca	tgg	cca	gga	gtg	act	724
Phe	Arg	Val	Met	Gly	Thr	Pro	Asn	Glu	Glu	Thr	Trp	Pro	Gly	Val	Thr	
		220						225				230				
tct	ttg	cct	gat	ttt	aag	tca	gca	ttt	cca	aaa	tgg	cca	gct	aag	gag	772
Ser	Leu	Pro	Asp	Phe	Lys	Ser	Ala	Phe	Pro	Lys	Trp	Pro	Ala	Lys	Glu	
		235					240					245				
ctg	gct	gct	gta	gtt	ccg	aat	ctt	gat	gca	tct	ggc	ctt	gat	ctc	ctt	820
Leu	Ala	Ala	Val	Val	Pro	Asn	Leu	Asp	Ala	Ser	Gly	Leu	Asp	Leu	Leu	
		250				255					260					
gat	aaa	atg	ctt	cgt	ttg	gac	ccc	agc	aaa	aga	att	acg	gcc	agg	aat	868
Asp	Lys	Met	Leu	Arg	Leu	Asp	Pro	Ser	Lys	Arg	Ile	Thr	Ala	Arg	Asn	
		265			270				275			280				
gct	ctt	cag	cat	gag	tac	ttc	aag	gat	att	ggt	ttt	gta	ccc	tgattggtgc	920	
Ala	Leu	Gln	His	Glu	Tyr	Phe	Lys	Asp	Ile	Gly	Phe	Val	Pro			
			285					290								
ccctcattct	ggtacgagta	tatatattg	tta	tatgacgtct	gggggttttat	tctgttccat	980									
aggaattcgt	gacagacgaa	cgttatctct	tgtttttgat	tccttgggtg	taattccatt	1040										
tatatgaag	ctgtgttgg	tgaagcaagt	taggantggc	ctctgctggt	gctttcactt	1100										
gctttaaac	ccttgtgtga	ttttgtcgat	tttttgttcc	ttttccattt	ttaatttccc	1160										
tgtaacatca	tgctgatgta	taacgtttga	gtttttgtta	tctggtttta	tatataaata	1220										
tggtgtgcct	tttagttgtt	caaaaaaaaa	aaaaaaaaaa	aaaaaacca	ggtacccgga	1280										
tcc						1283										

<210> 36

<211> 294

<212> PRT

<213> Antirrhinum majus

<400> 36

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

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

<210> 37

<211> 885

<212> DNA

<213> *Chenopodium rubrum*

<220>

<221> CDS

<222> (1)..(885)

<400> 37

atg gat cag tat gaa aaa gtt gag aag ata ggg gaa gga acc tat gga	48
Met Asp Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly	
1 5 10 15	
gtg gtt tat aag gcg cgg gac aag gtt acg aac gag act ata gct ttg	96
Val Val Tyr Lys Ala Arg Asp Lys Val Thr Asn Glu Thr Ile Ala Leu	
20 25 30	
aag aaa att cgg ttg gaa cag gag gat gag gga gtt ccg agc acg gcg	144
Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala	
35 40 45	
atc aga gaa ata tca ctt ttg aag gag atg cag cat ggc aac att gtc	192
Ile Arg Glu Ile Ser Leu Leu Lys Glu Met Gln His Gly Asn Ile Val	
50 55 60	
agg ttg cag gat gtg gtg cat agt gag aag cgc tta tat ctg gtt ttt	240
Arg Leu Gln Asp Val Val His Ser Glu Lys Arg Leu Tyr Leu Val Phe	
65 70 75 80	
gag tat ttg gac ctt gat ttg aag aaa cac atg gat tca tgc cct gat	288
Glu Tyr Leu Asp Leu Asp Leu Lys Lys His Met Asp Ser Cys Pro Asp	
85 90 95	
ttt gca aag gat cca cgt atg att aaa agg ttt ctt tat cag att ctt	336
Phe Ala Lys Asp Pro Arg Met Ile Lys Arg Phe Leu Tyr Gln Ile Leu	
100 105 110	
cgt ggt atc gct tat tgt cac tct cat agg gtc ctg cac cga gat ctg	384
Arg Gly Ile Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu	
115 120 125	

aag ccg cag aat ctg ttg ata gat cgc caa act aat gca cta aaa ctt 432
Lys Pro Gln Asn Leu Leu Ile Asp Arg Gln Thr Asn Ala Leu Lys Leu
130 135 140
gca gat ttt gga ttg gca agg gca ttt ggt att cct gtg agg act ttt 480
Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe
145 150 155 160
aca cat gag gtg gtg aca ttg tgg tac aga gct cca gaa ata ttg ctt 528
Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu
165 170 175
gga tct cga cat tac tct act cct gtg gat gtg tgg tct gtg ggt tgt 576
Gly Ser Arg His Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys
180 185 190
atc ttt gct gag atg gtg aat cag aag cca tta ttt cct gga gat tcc 624
Ile Phe Ala Glu Met Val Asn Gln Lys Pro Leu Phe Pro Gly Asp Ser
195 200 205
gag att gat gaa ctt ttc aag att ttc agg acc ttg ggt aca cca aat 672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Thr Leu Gly Thr Pro Asn
210 215 220
gag gag aca tgg cct gga gtg acc tcc ctt ccc gat ttc aaa tct tca 720
Glu Glu Thr Trp Pro Gly Val Thr Ser Leu Pro Asp Phe Lys Ser Ser
225 230 235 240
ttt cct aaa tgg atc tcc aag gat ttg tct gca gta gta cca aat ctt 768
Phe Pro Lys Trp Ile Ser Lys Asp Leu Ser Ala Val Val Pro Asn Leu
245 250 255
gat cca gct ggt att gat ctt cta aat aaa atg ctt tgc ttg gat ccg 816
Asp Pro Ala Gly Ile Asp Leu Leu Asn Lys Met Leu Cys Leu Asp Pro
260 265 270
agc aaa agg att aca gcc agg aat gct ctt gaa cat gaa tac ttc aag 864
Ser Lys Arg Ile Thr Ala Arg Asn Ala Leu Glu His Glu Tyr Phe Lys
275 280 285
gac att ggt ttt gta ccc tga 885
Asp Ile Gly Phe Val Pro
290

<210> 38

<211> 294

<212> PRT

<213> *Chenopodium rubrum*

<400> 38

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

Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Thr Leu Gly Thr Pro Asn
 210 215 220
 Glu Glu Thr Trp Pro Gly Val Thr Ser Leu Pro Asp Phe Lys Ser Ser
 225 230 235 240
 Phe Pro Lys Trp Ile Ser Lys Asp Leu Ser Ala Val Val Pro Asn Leu
 245 250 255
 Asp Pro Ala Gly Ile Asp Leu Leu Asn Lys Met Leu Cys Leu Asp Pro
 260 265 270
 Ser Lys Arg Ile Thr Ala Arg Asn Ala Leu Glu His Glu Tyr Phe Lys
 275 280 285
 Asp Ile Gly Phe Val Pro
 290

<210> 39

<211> 885

<212> DNA

<213> Zea mays

<220>

<221> CDS

<222> (1)..(885)

<400> 39

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

[illegible]

```
<210> 40
<211> 294
<212> PRT
<213> Zea mays
```

[illegible]

```
<210> 41
<211> 885
<212> DNA
<213> Vigna aconitifolia
```

<220>

<221> CDS

<222> (1)..(885)

<400> 41

atg gaa cag tac gag aag gtg gag aag ata ggg gag gga aca tac ggc	48
Met Glu Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly	
1 5 10 15	
gtc gtt tac aag gct cgc gac cgc gtc acc aat gag acc atc gct ctt	96
Val Val Tyr Lys Ala Arg Asp Arg Val Thr Asn Glu Thr Ile Ala Leu	
20 25 30	
aag aag att cgc ctc gag cag gaa gac gag ggg gtt ccc agc acc gcc	144
Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala	
35 40 45	
att cgt gag att tcg ctc ctc aaa gag atg cag cat agg aac att gtt	192
Ile Arg Glu Ile Ser Leu Leu Lys Glu Met Gln His Arg Asn Ile Val	
50 55 60	
agg ttg cag gat gta gtg cac agt gag aag cga ttg tat ctg gtt ttc	240
Arg Leu Gln Asp Val Val His Ser Glu Lys Arg Leu Tyr Leu Val Phe	
65 70 75 80	
gag tat ctg gac ttg gat cta aag aaa cac atg gat tca tct cca gag	288
Glu Tyr Leu Asp Leu Asp Leu Lys Lys His Met Asp Ser Ser Pro Glu	
85 90 95	
ttt gtg aaa gat cca cgg caa gta aaa atg ttc ctc tat caa att ctc	336
Phe Val Lys Asp Pro Arg Gln Val Lys Met Phe Leu Tyr Gln Ile Leu	
100 105 110	
tgt ggc att gct tac tgc cat tcg cac aga gtt ctt cat cga gac ttg	384
Cys Gly Ile Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu	
115 120 125	
aaa cca cag aat ttg ttg ata gac cgt cgt aca aat tcc tta aaa ctt	432
Lys Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Ser Leu Lys Leu	
130 135 140	
gca gat ttt gga ttg gct agg gca ttt ggc att cct gtc agg aca ttt	480
Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe	
145 150 155 160	
act cat gag gtg gtg aca tta tgg tac aga gct cca gaa ata ttg ctt	528
Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu	
165 170 175	
ggg tct cgt cat tat tct acc cca gtt gat gtt tgg tca gtg gga tgt	576
Gly Ser Arg His Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys	
180 185 190	
ata ttt gca gag atg gta aac cga cga cct cta ttc cca ggg gac tct	624
Ile Phe Ala Glu Met Val Asn Arg Arg Pro Leu Phe Pro Gly Asp Ser	
195 200 205	
gag att gat gaa tta ttt aaa ata ttc aga ata ttg ggt aca cct aat	672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Ile Leu Gly Thr Pro Asn	
210 215 220	
gaa gaa aca tgg cct gga gtt act gca tta ccg gat ttt aaa tca aca	720
Glu Glu Thr Trp Pro Gly Val Thr Ala Leu Pro Asp Phe Lys Ser Thr	
225 230 235 240	
ttt ccc aaa tgg cca cct aag gat tta gca act gtg gtt cca aat ctt	768
Phe Pro Lys Trp Pro Pro Lys Asp Leu Ala Thr Val Val Pro Asn Leu	
245 250 255	
gat gca gcg ggt ctt aat ctt ctt tct agt atg cta tgc ttg gat ccc	816
Asp Ala Ala Gly Leu Asn Leu Leu Ser Ser Met Leu Cys Leu Asp Pro	
260 265 270	
agc aaa aga att act gcc agg atc gct gtg gag cac gaa tac ttc aaa	864
Ser Lys Arg Ile Thr Ala Arg Ile Ala Val Glu His Glu Tyr Phe Lys	
275 280 285	
gac att aaa ttt gta ccc taa	885
Asp Ile Lys Phe Val Pro	
290	

<210> 42

<211> 294

<212> PRT

<213> Vigna aconitifolia

<400> 42

```

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

```

<210> 43

<211> 885

<212> DNA

<213> *Oryza sativa*

<220>

<221> CDS

<222> (1)..(885)

<400> 43

```

atg gag cag tac gag aag gag gag aag att ggg gag ggc acg tac ggg      48
Met Glu Gln Tyr Glu Lys Glu Glu Lys Ile Gly Glu Gly Thr Tyr Gly
1      5      10      15
gtg gtg tac agg gcg cgg gac aag gtc acc aac gag acg atc gcg ctc      96
Val Val Tyr Arg Ala Arg Asp Lys Val Thr Asn Glu Thr Ile Ala Leu
20      25      30
aag aag atc cgg ctt gag cag gag gat gag ggc gtc ccc tcc acc gca      144
Lys Lys Ile Arg Leu Glu Gln Glu Asp Glu Gly Val Pro Ser Thr Ala
35      40      45
atc cgc gag atc tcg ctc ctc aag gag atg cat cac ggc aac atc gtc      192
Ile Arg Glu Ile Ser Leu Leu Lys Glu Met His His Gly Asn Ile Val
50      55      60
agg tta cac gat gtt atc cac agt gag aag cgc ata tat ctt gtc ttt      240
Arg Leu His Asp Val Ile His Ser Glu Lys Arg Ile Tyr Leu Val Phe
65      70      75      80

```

gag tat ctg gat ctg gac cta aag aag ttc atg gac tct tgt cca gag	288
Glu Tyr Leu Asp Leu Asp Leu Lys Lys Phe Met Asp Ser Cys Pro Glu	
85 90 95	
ttt gcg aaa aac ccc act tta att aag tca tat ctc tat cag ata ctc	336
Phe Ala Lys Asn Pro Thr Leu Ile Lys Ser Tyr Leu Tyr Gln Ile Leu	
100 105 110	
cgc ggc gtt gct tac tgt cat tct cat aga gtt ctt cat cga gat ttg	384
Arg Gly Val Ala Tyr Cys His Ser His Arg Val Leu His Arg Asp Leu	
115 120 125	
aaa cct cag aat tta ttg ata gat cgg cgt act aat gca ctg aag ctt	432
Lys Pro Gln Asn Leu Leu Ile Asp Arg Arg Thr Asn Ala Leu Lys Leu	
130 135 140	
gca gac ttt ggt tta gcc agg gca ttt gga att cct gtc cgc acg ttt	480
Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe	
145 150 155 160	
act cac gag gtt gta acc ttg tgg tat aga gct cca gag atc ctt ctt	528
Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu	
165 170 175	
gga tca agg cag tat tct aca cca gtt gat atg tgg tca gtt ggt tgt	576
Gly Ser Arg Gln Tyr Ser Thr Pro Val Asp Met Trp Ser Val Gly Cys	
180 185 190	
atc ttt gca gaa atg gtg aac cag aaa cca ctg ttc cct ggt gat tct	624
Ile Phe Ala Glu Met Val Asn Gln Lys Pro Leu Phe Pro Gly Asp Ser	
195 200 205	
gag att gat gaa tta ttt aag ata ttc agg gta cta gga act cca aat	672
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Val Leu Gly Thr Pro Asn	
210 215 220	
gaa caa agt tgg cca gga gtt agc tca tta cct gac tac aag tct gct	720
Glu Gln Ser Trp Pro Gly Val Ser Ser Leu Pro Asp Tyr Lys Ser Ala	
225 230 235 240	
ttc ccc aag tgg caa gca cag gat ctt gca act att gtc cct act ctt	768
Phe Pro Lys Trp Gln Ala Gln Asp Leu Ala Thr Ile Val Pro Thr Leu	
245 250 255	
gac cct gct ggt ttg gac ctt ctc tct aaa atg ctt cgg tac gag cca	816
Asp Pro Ala Gly Leu Asp Leu Leu Ser Lys Met Leu Arg Tyr Glu Pro	
260 265 270	
aac aaa agg atc aca gct aga cag gct ctt gag cat gaa tac ttc aag	864
Asn Lys Arg Ile Thr Ala Arg Gln Ala Leu Glu His Glu Tyr Phe Lys	
275 280 285	
gac ctt gag atg gta caa tga	885
Asp Leu Glu Met Val Gln	
290 294	

<210> 44

<211> 294

<212> PRT

<213> Oryza sativa

<400> 44

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

Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Ile Pro Val Arg Thr Phe
 145 150 155 160
 Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu
 165 170 175
 Gly Ser Arg Gln Tyr Ser Thr Pro Val Asp Met Trp Ser Val Gly Cys
 180 185 190
 Ile Phe Ala Glu Met Val Asn Gln Lys Pro Leu Phe Pro Gly Asp Ser
 195 200 205
 Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Val Leu Gly Thr Pro Asn
 210 215 220
 Glu Gln Ser Trp Pro Gly Val Ser Ser Leu Pro Asp Tyr Lys Ser Ala
 225 230 235 240
 Phe Pro Lys Trp Gln Ala Gln Asp Leu Ala Thr Ile Val Pro Thr Leu
 245 250 255
 Asp Pro Ala Gly Leu Asp Leu Leu Ser Lys Met Leu Arg Tyr Glu Pro
 260 265 270
 Asn Lys Arg Ile Thr Ala Arg Gln Ala Leu Glu His Glu Tyr Phe Lys
 275 280 285
 Asp Leu Glu Met Val Gln
 290

<210> 45

<211> 1113

<212> DNA

<213> *Oryza sativa*

<220>

<221> CDS

<222> (1)..(1113)

<400> 45

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

tat gtt gtc acc cgc tgg tat aga gct cct gag ttg ctc ctt tgc tgt	624
Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro Glu Leu Leu Leu Cys Cys	
195 200 205	
gac aac tat ggc act tcc att gat gtt tgg tct gtt ggt tgc atc ttc	672
Asp Asn Tyr Gly Thr Ser Ile Asp Val Trp Ser Val Gly Cys Ile Phe	
210 215 220	
gct gag cta ctt ggt cgg aaa cct att ttc cca gga act gag tgc cta	720
Ala Glu Leu Leu Gly Arg Lys Pro Ile Phe Pro Gly Thr Glu Cys Leu	
225 230 235 240	
aat cag ctc aag ctc att gtc aat gtt ctt ggc acc atg agc gag tct	768
Asn Gln Leu Lys Leu Ile Val Asn Val Leu Gly Thr Met Ser Glu Ser	
245 250 255	
gac ttg gag ttc att gac aac cca aaa gct cgc aga tat atc aaa tcc	816
Asp Leu Glu Phe Ile Asp Asn Pro Lys Ala Arg Arg Tyr Ile Lys Ser	
260 265 270	
ctc ccc tac act ccc ggt gtg ccc ctc gcg agt atg tat ccg cat gca	864
Leu Pro Tyr Thr Pro Gly Val Pro Leu Ala Ser Met Tyr Pro His Ala	
275 280 285	
cac cct ctt gcc att gat ctt tta cag aag atg ctc ata ttt gat cct	912
His Pro Leu Ala Ile Asp Leu Leu Gln Lys Met Leu Ile Phe Asp Pro	
290 295 300	
acc aaa aga atc agt gtc acc gag gct ctt gag cac cct tac atg tcc	960
Thr Lys Arg Ile Ser Val Thr Glu Ala Leu Glu His Pro Tyr Met Ser	
305 310 315 320	
cct ctg tat gat cca agt gca aat cct ccc gcg caa gtg cct atc gat	1008
Pro Leu Tyr Asp Pro Ser Ala Asn Pro Pro Ala Gln Val Pro Ile Asp	
325 330 335	
ctg gag ata gag gag aac atc agt gca gat atg atc agg gaa atg atg	1056
Leu Asp Ile Asp Glu Asn Ile Ser Ala Asp Met Ile Arg Glu Met Met	
340 345 350	
tgg cac gag atg ctc cac tac cac cct gaa gtt gtt gca gca atg agt	1104
Trp His Glu Met Leu His Tyr His Pro Glu Val Val Ala Ala Met Ser	
355 360 365	
gcc cga tga	1113
Ala Arg	
370	
<210> 46	
<211> 370	
<212> PRT	
<213> Oryza sativa	
<400> 46	
Met Ala Ile Met Val Asp Pro Pro Asn Gly Met Gly Asn Gln Gly Lys	
1 5 10 15	
Tyr Tyr Tyr Ser Met Trp Gln Thr Leu Phe Glu Ile Asp Thr Lys Tyr	
20 25 30	
Val Pro Ile Lys Pro Ile Gly Arg Gly Ala Tyr Gly Ile Val Cys Ser	
35 40 45	
Ser Ile Asn Arg Glu Thr Asn Glu Lys Val Ala Ile Lys Lys Ile His	
50 55 60	
Asn Val Phe Asp Asn Arg Val Asp Ala Leu Arg Thr Leu Arg Glu Leu	
65 70 75 80	
Lys Leu Leu Arg His Leu Arg His Glu Asn Val Ile Ala Leu Lys Asp	
85 90 95	
Ile Met Met Pro Val His Arg Arg Ser Phe Lys Asp Val Tyr Leu Val	
100 105 110	
Tyr Glu Leu Met Asp Thr Asp Leu His Gln Ile Ile Lys Ser Pro Gln	
115 120 125	
Gly Leu Ser Asn Asp His Cys Gln Tyr Phe Leu Phe Gln Leu Leu Arg	
130 135 140	
Gly Leu Lys Tyr Leu His Ser Ala Glu Ile Leu His Arg Asp Leu Lys	
145 150 155 160	
Pro Gly Asn Leu Leu Val Asn Ala Asn Cys Asp Leu Lys Ile Cys Asp	
165 170 175	
Phe Gly Leu Ala Arg Thr Asn Ser Ser Lys Gly Gln Phe Met Thr Glu	
180 185 190	

Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro Glu Leu Leu Leu Cys Cys
 195 200 205
 Asp Asn Tyr Gly Thr Ser Ile Asp Val Trp Ser Val Gly Cys Ile Phe
 210 215 220
 Ala Glu Leu Leu Gly Arg Lys Pro Ile Phe Pro Gly Thr Glu Cys Leu
 225 230 235 240
 Asn Gln Leu Lys Leu Ile Val Asn Val Leu Gly Thr Met Ser Glu Ser
 245 250 255
 Asp Leu Glu Phe Ile Asp Asn Pro Lys Ala Arg Arg Tyr Ile Lys Ser
 260 265 270
 Leu Pro Tyr Thr Pro Gly Val Pro Leu Ala Ser Met Tyr Pro His Ala
 275 280 285
 His Pro Leu Ala Ile Asp Leu Leu Gln Lys Met Leu Ile Phe Asp Pro
 290 295 300
 Thr Lys Arg Ile Ser Val Thr Glu Ala Leu Glu His Pro Tyr Met Ser
 305 310 315 320
 Pro Leu Tyr Asp Pro Ser Ala Asn Pro Pro Ala Gln Val Pro Ile Asp
 325 330 335
 Leu Asp Ile Asp Glu Asn Ile Ser Ala Asp Met Ile Arg Glu Met Met
 340 345 350
 Trp His Glu Met Leu His Tyr His Pro Glu Val Val Ala Ala Met Ser
 355 360 365
 Ala Arg
 370

<210> 47
 <211> 1044
 <212> DNA
 <213> Brassica napus

<220>
 <221> CDS
 <222> (1)..(1044)

<400> 47
 atg tcg aaa tcg ggc gga aac caa ccc gta gat cga tac cta aga cgc 48
 Met Ser Lys Ser Gly Gly Asn Gln Pro Val Asp Arg Tyr Leu Arg Arg
 1 5 10 15
 caa gtg ctc gga gaa gga aca tac ggt gtc gtt tac aaa gcc acc gac 96
 Gln Val Leu Gly Glu Gly Thr Tyr Gly Val Val Tyr Lys Ala Thr Asp
 20 25 30
 acc aag aca ggt aag acc gta gca gtg aag aag atc agg tta gga aac 144
 Thr Lys Thr Gly Lys Thr Val Ala Val Lys Lys Ile Arg Leu Gly Asn
 35 40 45
 cag aaa gaa ggc gtc aat ttc acg gcc ctt agg gag atc aag ctc cta 192
 Gln Lys Glu Gly Val Asn Phe Thr Ala Leu Arg Glu Ile Lys Leu Leu
 50 55 60
 aaa gag ctg aac cat ccc cac atc gtc gag ctc atc gac gcg ttt cct 240
 Lys Glu Leu Asn His Pro His Ile Val Glu Leu Ile Asp Ala Phe Pro
 65 70 75 80
 cac gac ggg agc ttg cat ctg gtt ttc gag tat atg cag acg gat ttg 288
 His Asp Gly Ser Leu His Leu Val Phe Glu Tyr Met Gln Thr Asp Leu
 85 90 95
 gaa gct gtg att cgt gac cgt aac att ttt ctt tct cct ggt gat atc 336
 Glu Ala Val Ile Arg Asp Arg Asn Ile Phe Leu Ser Pro Gly Asp Ile
 100 105 110
 aag tct tat atg ttg atg act ctt aag ggt tta gct tat tgt cac aag 384
 Lys Ser Tyr Met Leu Met Thr Leu Lys Gly Leu Ala Tyr Cys His Lys
 115 120 125
 aaa tgg gtt ctt cat agg gat atg aag cct aat aat cta ctg att gga 432
 Lys Trp Val Leu His Arg Asp Met Lys Pro Asn Asn Leu Leu Ile Gly
 130 135 140
 gag aat ggt ctg ttg aag tta gct gat ttt ggt ttg gcg aga gtg ttt 480
 Glu Asn Gly Leu Leu Lys Leu Ala Asp Phe Gly Leu Ala Arg Val Phe
 145 150 155 160
 ggg agt ccg aat cgg agg ttt act cat cag gta ttt gcg acg tgg tac 528

Gly	Ser	Pro	Asn	Arg	Arg	Phe	Thr	His	Gln	Val	Phe	Ala	Thr	Trp	Tyr	
			165						170					175		
aga	gcc	ccc	gag	ttg	ctg	ttt	ggg	agt	aga	cag	tat	ggg	gca	gga	gtt	576
Arg	Ala	Pro	Glu	Leu	Leu	Phe	Gly	Ser	Arg	Gln	Tyr	Gly	Ala	Gly	Val	
			180					185					190			
gat	gtc	tg	gct	gca	ggc	tgt	atc	ttt	gct	gag	cta	tta	ctt	cgt	cga	624
Asp	Val	Trp	Ala	Ala	Gly	Cys	Ile	Phe	Ala	Glu	Leu	Leu	Leu	Arg	Arg	
		195					200					205				
ccc	ttt	cta	cct	ggg	tcc	act	gag	ata	gat	caa	ctg	gga	aag	atc	ttt	672
Pro	Phe	Leu	Pro	Gly	Ser	Thr	Glu	Ile	Asp	Gln	Leu	Gly	Lys	Ile	Phe	
	210					215					220					
caa	gct	ttt	gga	act	cca	aca	tcc	caa	tgg	tct	gac	atg	atc	tat		720
Gln	Ala	Phe	Gly	Thr	Pro	Thr	Gln	Ser	Gln	Trp	Ser	Asp	Met	Ile	Tyr	
225					230				235					240		
ctc	cca	gag	tac	atg	gag	ttc	tcc	tac	aca	cct	gct	cca	cca	cta	cgt	768
Leu	Pro	Glu	Tyr	Met	Glu	Phe	Ser	Tyr	Thr	Pro	Ala	Pro	Pro	Leu	Arg	
			245					250				255				
acc	att	ttc	cct	atg	gca	agc	gac	gac	gct	ttg	gat	ctt	cta	tcc	aaa	816
Thr	Ile	Phe	Pro	Met	Ala	Ser	Asp	Asp	Ala	Leu	Asp	Leu	Leu	Ser	Lys	
			260				265					270				
atg	ttc	atc	tac	gac	cca	cgc	caa	cgt	atc	acc	ata	cag	caa	gct	ttg	864
Met	Phe	Ile	Tyr	Asp	Pro	Arg	Gln	Arg	Ile	Thr	Ile	Gln	Gln	Ala	Leu	
	275						280					285				
gac	cac	agg	tat	ttc	tcg	tct	tct	cca	tca	cca	act	gag	cca	ggg	aag	912
Asp	His	Arg	Tyr	Phe	Ser	Ser	Ser	Pro	Ser	Pro	Thr	Glu	Pro	Gly	Lys	
	290					295					300					
ctt	cag	att	cca	gct	tcg	aaa	gga	gac	gcc	ctc	gaa	cca	aag	gca	tct	960
Leu	Gln	Ile	Pro	Ala	Ser	Lys	Gly	Asp	Ala	Leu	Glu	Pro	Lys	Ala	Ser	
305					310				315					320		
gag	caa	aac	aac	caa	cac	gga	aac	agc	ccc	gcc	gtg	cta	tca	cct	cct	1008
Glu	Gln	Asn	Asn	Gln	His	Gly	Asn	Ser	Pro	Ala	Val	Leu	Ser	Pro	Pro	
			325					330					335			
gga	aaa	atg	agg	aga	gtg	atg	ggg	cct	gag	gga	taa					1044
Gly	Lys	Met	Arg	Arg	Val	Met	Gly	Pro	Glu	Gly						
			340				345									

<210> 48

<211> 347

<212> PRT

<213> Brassica napus

<400> 48

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

Asp	Val	Trp	Ala	Ala	Gly	Cys	Ile	Phe	Ala	Glu	Leu	Leu	Leu	Arg	Arg
		195					200					205			
Pro	Phe	Leu	Pro	Gly	Ser	Thr	Glu	Ile	Asp	Gln	Leu	Gly	Lys	Ile	Phe
	210					215					220				
Gln	Ala	Phe	Gly	Thr	Pro	Thr	Gln	Ser	Gln	Trp	Ser	Asp	Met	Ile	Tyr
	225					230				235					240
Leu	Pro	Glu	Tyr	Met	Glu	Phe	Ser	Tyr	Thr	Pro	Ala	Pro	Pro	Leu	Arg
				245					250					255	
Thr	Ile	Phe	Pro	Met	Ala	Ser	Asp	Asp	Ala	Leu	Asp	Leu	Leu	Ser	Lys
			260					265						270	
Met	Phe	Ile	Tyr	Asp	Pro	Arg	Gln	Arg	Ile	Thr	Ile	Gln	Gln	Ala	Leu
		275					280					285			
Asp	His	Arg	Tyr	Phe	Ser	Ser	Ser	Pro	Ser	Pro	Thr	Glu	Pro	Gly	Lys
	290					295					300				
Leu	Gln	Ile	Pro	Ala	Ser	Lys	Gly	Asp	Ala	Leu	Glu	Pro	Lys	Ala	Ser
	305					310				315					320
Glu	Gln	Asn	Asn	Gln	His	Gly	Asn	Ser	Pro	Ala	Val	Leu	Ser	Pro	Pro
			325						330					335	
Gly	Lys	Met	Arg	Arg	Val	Met	Gly	Pro	Glu	Gly					
			340					345							

<210> 49

<211> 1353

<212> DNA

<213> Glycine max

<220>

<221> CDS

<222> (1)..(1353)

<400> 49

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

	180		185		190	
ctc ttg tct ttt aat cat ccc tcc att gtg aat gtt aaa gaa gta gtt						624
Leu Leu Ser Phe Asn His Pro Ser Ile Val Asn Val Lys Glu Val Val						
	195		200		205	
gtt gat gat ttt gat ggt act ttt atg gta atg gag cac atg gag tat						672
Val Asp Asp Phe Asp Gly Thr Phe Met Val Met Glu His Met Glu Tyr						
	210		215		220	
gac ctg aag ggg ctg atg gag gtt aag aag cat ccc ttt agc atg agt						720
Asp Leu Lys Gly Leu Met Glu Val Lys Lys His Pro Phe Ser Met Ser						
	225		230		235	240
gaa att aaa tcc ttg gta cgg caa ctt ctt gaa ggt gtc aag tat ctc						768
Glu Ile Lys Ser Leu Val Arg Gln Leu Glu Gly Val Lys Tyr Leu						
	245		250		255	
cat gat aat tgg gtt atc cat agg gac ttg aaa tca tca aac atc ctg						816
His Asp Asn Trp Val Ile His Arg Asp Leu Lys Ser Ser Asn Ile Leu						
	260		265		270	
ctg aac cat gat ggg gag ctt aaa ata tgc gac ttt ggg ttg tca agg						864
Leu Asn His Asp Gly Glu Leu Lys Ile Cys Asp Phe Gly Leu Ser Arg						
	275		280		285	
cag tat gga agc ccg ctg aag cca tat act cct gtt gtg gtt aca cta						912
Gln Tyr Gly Ser Pro Leu Lys Pro Tyr Thr Pro Val Val Val Thr Leu						
	290		295		300	
tgg tac aga gca cct gaa ctt ttg ctg ggt gct aaa gaa tac tca act						960
Trp Tyr Arg Ala Pro Glu Leu Leu Leu Gly Ala Lys Glu Tyr Ser Thr						
	305		310		315	320
gca att gac atg tgg tca gtg ggt tgc ata atg gct gaa tta att gcc						1008
Ala Ile Asp Met Trp Ser Val Gly Cys Ile Met Ala Glu Leu Ile Ala						
	325		330		335	
aag gaa cct ctg ttc agg ggt aaa agt gaa ctt gaa caa ctt gat aag						1056
Lys Glu Pro Leu Phe Arg Gly Lys Ser Glu Leu Glu Gln Leu Asp Lys						
	340		345		350	
att ttt cga acc ctg ggt acg cct gat gag aaa att tgg cca gga tta						1104
Ile Phe Arg Thr Leu Gly Thr Pro Asp Glu Lys Ile Trp Pro Gly Leu						
	355		360		365	
tcc aaa tta cct ggg gct aaa gca aat ttt gtt aag caa ctg ttt aat						1152
Ser Lys Leu Pro Gly Ala Lys Ala Asn Phe Val Lys Gln Leu Phe Asn						
	370		375		380	
aca cta agg aag aag ttt cca gct gcg tcg ttt att ggt ttg cca gtt						1200
Thr Leu Arg Lys Lys Phe Pro Ala Ala Ser Phe Ile Gly Leu Pro Val						
	385		390		395	400
tta tct gag ctt gga ttt gac ttg ttg cag cag ctg cta act tat gac						1248
Leu Ser Glu Leu Gly Phe Asp Leu Leu Gln Gln Leu Leu Thr Tyr Asp						
	405		410		415	
cct gaa aag agg ata aca gca gaa gat gct ctc ctt cat gat tgg ttc						1296
Pro Glu Lys Arg Ile Thr Ala Glu Asp Ala Leu Leu His Asp Trp Phe						
	420		425		430	
cat gaa gcc cct ctt ccc aaa tct gat ttc aag cca att ttt cct tct						1344
His Glu Ala Pro Leu Pro Lys Ser Asp Phe Lys Pro Ile Phe Pro Ser						
	435		440		445	
tgg cag tga						1353
Trp Gln						
	450					

<210> 50

<211> 450

<212> PRT

<213> Glycine max

<400> 50

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

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

<210> 51
 <211> 885
 <212> DNA
 <213> Linum usitatissimum

<220>
 <221> CDS
 <222> (1)..(885)

<400> 51
 atg gag cag tac gag aaa gtg gag aag att ggg gag gga act tac gga
 Met Glu Gln Tyr Glu Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly
 1 5 10 15
 gtg gta tat aag gcc cgt gat cga aac acc aat cag act att gct ttg

48

96

Val	Val	Tyr	Lys	Ala	Arg	Asp	Arg	Asn	Thr	Asn	Gln	Thr	Ile	Ala	Leu	
			20					25					30			
aag	aag	att	cgc	ctt	gaa	cag	gaa	gat	gag	ggc	gtg	cca	agc	acc	gcc	144
Lys	Lys	Ile	Arg	Leu	Glu	Gln	Glu	Asp	Glu	Gly	Val	Pro	Ser	Thr	Ala	
		35					40					45				
att	aga	gaa	att	tcg	ctt	ctg	aaa	gaa	atg	cag	cat	ggg	aac	att	gtg	192
Ile	Arg	Glu	Ile	Ser	Leu	Leu	Lys	Glu	Met	Gln	His	Gly	Asn	Ile	Val	
	50					55					60					
agg	ttg	cag	gat	gtg	gtg	cac	agt	gag	aaa	cgg	cta	tat	ttg	gtt	ttt	240
Arg	Leu	Gln	Asp	Val	Val	His	Ser	Glu	Lys	Arg	Leu	Tyr	Leu	Val	Phe	
65				70						75					80	
gag	tac	ctg	gac	tta	gat	ttg	aaa	aag	cac	atg	gac	tca	agt	ccg	gag	288
Glu	Tyr	Leu	Asp	Leu	Asp	Leu	Lys	Lys	His	Met	Asp	Ser	Ser	Pro	Glu	
			85						90					95		
ttt	gct	aag	gat	ccg	cgt	ata	gtc	aaa	act	ttc	cta	tac	caa	atc	ctc	336
Phe	Ala	Lys	Asp	Pro	Arg	Ile	Val	Lys	Thr	Phe	Leu	Tyr	Gln	Ile	Leu	
			100					105					110			
cgt	ggc	att	tct	tat	tgc	cac	tct	cac	agg	gtt	ctt	cat	cga	gat	cta	384
Arg	Gly	Ile	Ser	Tyr	Cys	His	Ser	His	Arg	Val	Leu	His	Arg	Asp	Leu	
		115					120					125				
aag	cca	cag	aat	ctg	ctt	att	gat	cgt	cgt	acc	aat	gca	ctc	aag	ctt	432
Lys	Pro	Gln	Asn	Leu	Leu	Ile	Asp	Arg	Arg	Thr	Asn	Ala	Leu	Lys	Leu	
	130					135					140					
gca	gat	ttt	ggg	ctg	gcc	aga	gct	ttt	ggg	ata	cct	gtc	agg	aca	ttc	480
Ala	Asp	Phe	Gly	Leu	Ala	Arg	Ala	Phe	Gly	Ile	Pro	Val	Arg	Thr	Phe	
145				150					155					160		
acc	cat	gag	gtg	gtg	act	ctc	tgg	tat	aga	gcc	cca	gag	att	ctt	ctc	528
Thr	His	Glu	Val	Val	Thr	Leu	Trp	Tyr	Arg	Ala	Pro	Glu	Ile	Leu	Leu	
			165					170						175		
gga	tct	cac	cac	tac	tct	act	cca	gtt	gat	gtg	tgg	tct	ata	gga	tgt	576
Gly	Ser	His	His	Tyr	Ser	Thr	Pro	Val	Asp	Val	Trp	Ser	Ile	Gly	Cys	
		180						185				190				
att	ttc	gcc	gaa	atg	gtg	aac	cag	cgc	ccc	tta	ttc	ccc	ggg	gac	tct	624
Ile	Phe	Ala	Glu	Met	Val	Asn	Gln	Arg	Pro	Leu	Phe	Pro	Gly	Asp	Ser	
		195				200					205					
gag	att	gat	gag	cta	ttt	aag	att	ttc	aga	atc	atg	ggg	acg	ccc	aat	672
Glu	Ile	Asp	Glu	Leu	Phe	Lys	Ile	Phe	Arg	Ile	Met	Gly	Thr	Pro	Asn	
	210					215					220					
gaa	gac	acc	tgg	gcc	gga	gta	aca	tct	tta	ccc	gat	ttc	aag	tcc	acc	720
Glu	Asp	Thr	Trp	Ala	Gly	Val	Thr	Ser	Leu	Pro	Asp	Phe	Lys	Ser	Thr	
225				230						235				240		
ttc	cca	aag	tgg	cct	tca	aag	gac	ctg	gca	acc	atg	gtt	caa	act	ctc	768
Phe	Pro	Lys	Trp	Pro	Ser	Lys	Asp	Leu	Ala	Thr	Met	Val	Gln	Thr	Leu	
			245					250						255		
gag	tca	acc	ggg	gtt	gac	ctc	tta	cga	aaa	atg	ctg	tgc	ttg	gat	cca	816
Glu	Ser	Thr	Gly	Val	Asp	Leu	Leu	Arg	Lys	Met	Leu	Cys	Leu	Asp	Pro	
		260					265					270				
agc	aaa	aga	att	aca	gcc	agg	act	gcc	cta	gag	cat	gaa	tac	ttc	aag	864
Ser	Lys	Arg	Ile	Thr	Ala	Arg	Thr	Ala	Leu	Glu	His	Glu	Tyr	Phe	Lys	
		275					280					285				
gac	att	gga	ttt	gtg	cca	taa										885
Asp	Ile	Gly	Phe	Val	Pro											
	290				294											

<210> 52

<211> 294

<212> PRT

<213> Linum usitatissimum

<400> 52

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

```

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

```

<210> 53

<211> 1152

<212> DNA

<213> Zea mays

<220>

<221> CDS

<222> (1)..(1152)

<400> 53

```

atg tca gat act gac tca gac aat gaa ata cac aga cct gaa acc cct      48
Met Ser Asp Thr Asp Ser Asp Asn Glu Ile His Arg Pro Glu Thr Pro
1          5          10          15
gaa aga gag aag gca cca cac agg tgt atc aat atg ctt caa gat tgc      96
Glu Arg Glu Lys Ala Pro His Arg Cys Ile Asn Met Leu Gln Asp Cys
      20          25          30
aga agt gtt gat gag ttt gag agg ctc aac aaa att aac gag ggc aca      144
Arg Ser Val Asp Glu Phe Glu Arg Leu Asn Lys Ile Asn Glu Gly Thr
      35          40          45
tat ggt gtt gta tac aga gca agg gat aag aag aca agt gag att gtt      192
Tyr Gly Val Val Tyr Arg Ala Arg Asp Lys Lys Thr Ser Glu Ile Val
      50          55          60
gca ttg aag aag gta aag atg gag aag gag cga gaa ggt ttc cca tta      240
Ala Leu Lys Lys Val Lys Met Glu Lys Glu Arg Glu Gly Phe Pro Leu
      65          70          75          80
acc tct ctt agg gaa ata aat atc ctc ttg tct ttc cac aat cct tca      288
Thr Ser Leu Arg Glu Ile Asn Ile Leu Leu Ser Phe His Asn Pro Ser
      85          90          95
att gtg gat gtt aag gaa gta gta gtt gga agt agt cta gat agt att      336
Ile Val Asp Val Lys Glu Val Val Val Gly Ser Ser Leu Asp Ser Ile
      100        105        110
ttt atg gtt atg gag tac atg gaa cat gat ctt aaa ggt gtc atg gag      384
Phe Met Val Met Glu Tyr Met Glu His Asp Leu Lys Gly Val Met Glu

```

115	120	125	
acc atg aaa caa cca tat acc caa agc gag gta aag tgt ttg atg ctt			432
Thr Met Lys Gln Pro Tyr Thr Gln Ser Glu Val Lys Cys Leu Met Leu			
130	135	140	
cag ctg cta gaa ggt gta aaa tat cta cat gac aat tgg gtg ctt cat			480
Gln Leu Leu Glu Gly Val Lys Tyr Leu His Asp Asn Trp Val Leu His			
145	150	155	160
agg gat ctg aag acc tca aat ctt ttg ttg aat aac cgt ggt gag tta			528
Arg Asp Leu Lys Thr Ser Asn Leu Leu Leu Asn Asn Arg Gly Glu Leu			
165	170	175	
aaa ata tgt gat ttt gga ctg tct cgt caa tat gga agc cca tta aaa			576
Lys Ile Cys Asp Phe Gly Leu Ser Arg Gln Tyr Gly Ser Pro Leu Lys			
180	185	190	
cct tat act caa ctg gtt gtg act ttg tgg tac agg gcc cca gaa ttg			624
Pro Tyr Thr Gln Leu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Leu			
195	200	205	
ttg tta gga aca aag gag tat tct act gct att gat atg tgg tct gtg			672
Leu Leu Gly Thr Lys Glu Tyr Ser Thr Ala Ile Asp Met Trp Ser Val			
210	215	220	
ggc tgt ata atg gct gag ctt ctt gcc aaa gaa ccg tta ttc aat gga			720
Gly Cys Ile Met Ala Glu Leu Leu Ala Lys Glu Pro Leu Phe Asn Gly			
225	230	235	240
aaa aca gag ttt gaa cag cta gat aag att ttt aga aca ctc ggc aca			768
Lys Thr Glu Phe Glu Gln Leu Asp Lys Ile Phe Arg Thr Leu Gly Thr			
245	250	255	
cct aat gag aag att tgg cct ggc tat gct aaa tta cca ggt gtg aaa			816
Pro Asn Glu Lys Ile Trp Pro Gly Tyr Ala Lys Leu Pro Gly Val Lys			
260	265	270	
gta aat ttt gtt aaa caa ccg tat aat aga cta agg gat aag ttc cca			864
Val Asn Phe Val Lys Gln Pro Tyr Asn Arg Leu Arg Asp Lys Phe Pro			
275	280	285	
gcc gct tct ttt tct ggg cga cca atc ctg tct gaa gct ggt ttt gat			912
Ala Ala Ser Phe Ser Gly Arg Pro Ile Leu Ser Glu Ala Gly Phe Asp			
290	295	300	
cta ttg aac aga ctg ctg act tat gac cct gat aag cgc ata tca gca			960
Leu Leu Asn Arg Leu Leu Thr Tyr Asp Pro Asp Lys Arg Ile Ser Ala			
305	310	315	320
gac gat gct ctt aag cac aaa tgg ttc tct gaa gtt cct ctg cct aaa			1008
Asp Asp Ala Leu Lys His Lys Trp Phe Ser Glu Val Pro Leu Pro Lys			
325	330	335	
tca aag gac ttc atg cca aca ttc cct gct ctt aat gaa ctt gac agg			1056
Ser Lys Asp Phe Met Pro Thr Phe Pro Ala Leu Asn Glu Leu Asp Arg			
340	345	350	
cgt tcc aga agg tat ctg aag agt cct gat cct ctg gaa gag caa cgt			1104
Arg Ser Arg Arg Tyr Leu Lys Ser Pro Asp Pro Leu Glu Glu Gln Arg			
355	360	365	
ttg aaa gaa ctg caa ggg aac ata ggc aac cat ggg ctt ttt ggg tga			1152
Leu Lys Glu Leu Gln Gly Asn Ile Gly Asn His Gly Leu Phe Gly			
370	375	380	

<210> 54

<211> 383

<212> PRT

<213> Zea mays

<400> 54

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

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

<210> 55

<211> 885

<212> DNA

<213> Helianthus annuus

<220>

<221> CDS

<222> (1)..(885)

<400> 55

atg	gac	cag	tat	gag	aag	gtt	gag	aag	att	ggt	gaa	gga	aca	tat	ggt	48
Met	Asp	Gln	Tyr	Glu	Lys	Val	Glu	Lys	Ile	Gly	Glu	Gly	Thr	Tyr	Gly	
1				5					10					15		
gtg	gtc	tac	aag	gct	cgt	gac	aaa	gtt	acc	aat	gaa	aca	atc	gct	ttg	96
Val	Val	Tyr	Lys	Ala	Arg	Asp	Lys	Val	Thr	Asn	Glu	Thr	Ile	Ala	Leu	
			20					25					30			
aag	aag	att	cgt	ctg	gag	caa	gaa	gac	gag	ggt	gtc	cca	agc	aca	gct	144
Lys	Lys	Ile	Arg	Leu	Glu	Gln	Glu	Asp	Glu	Gly	Val	Pro	Ser	Thr	Ala	
		35					40				45					
att	aga	gaa	atc	tcg	tta	ttg	aaa	gag	atg	caa	cat	gga	aat	att	gtc	192
Ile	Arg	Glu	Ile	Ser	Leu	Leu	Lys	Glu	Met	Gln	His	Gly	Asn	Ile	Val	
	50					55				60						
aag	tta	cag	gat	gtt	gtg	cac	agt	gac	aaa	cga	ttg	tat	ctg	gtt	ttc	240
Lys	Leu	Gln	Asp	Val	Val	His	Ser	Asp	Lys	Arg	Leu	Tyr	Leu	Val	Phe	
65				70					75					80		
gag	tat	ctg	gac	ctg	gat	ttg	aag	aag	cat	atg	gac	tca	tgt	ccc	gag	288
Glu	Tyr	Leu	Asp	Leu	Asp	Leu	Lys	Lys	His	Met	Asp	Ser	Cys	Pro	Glu	

				85					90				95				
ttt	tca	aag	gat	cca	cac	att	gta	aaa	atg	ttt	cta	tat	cag	atc	cta		336
Phe	Ser	Lys	Asp	Pro	His	Ile	Val	Lys	Met	Phe	Leu	Tyr	Gln	Ile	Leu		
			100						105				110				
cgt	ggc	att	gct	tat	tgc	cat	tcc	cat	aga	gtt	ctc	cat	aga	gac	tta		384
Arg	Gly	Ile	Ala	Tyr	Cys	His	Ser	His	Arg	Val	Leu	His	Arg	Asp	Leu		
		115					120					125					
aaa	cct	caa	aac	ctt	ttg	att	gat	cgc	cgt	acc	aat	gca	cta	aag	ctc		432
Lys	Pro	Gln	Asn	Leu	Leu	Ile	Asp	Arg	Arg	Thr	Asn	Ala	Leu	Lys	Leu		
	130					135				140							
gca	gat	ttt	gga	ttg	gcc	agg	gcc	ttt	ggc	att	cct	gtc	agg	aca	ttc		480
Ala	Asp	Phe	Gly	Leu	Ala	Arg	Ala	Phe	Gly	Ile	Pro	Val	Arg	Thr	Phe		
	145				150				155						160		
aca	cac	gag	gtg	gtt	aca	ctg	tgg	tat	aga	gca	cca	gaa	att	ctt	ctt		528
Thr	His	Glu	Val	Val	Thr	Leu	Trp	Tyr	Arg	Ala	Pro	Glu	Ile	Leu	Leu		
			165					170					175				
gga	tcg	cgt	cat	tat	tct	act	cct	gtt	gat	gtg	tgg	tcg	gtc	ggg	tgt		576
Gly	Ser	Arg	His	Tyr	Ser	Thr	Pro	Val	Asp	Val	Trp	Ser	Val	Gly	Cys		
		180					185					190					
ata	ttt	gct	gag	atg	gtg	aac	caa	cgg	cca	ctc	ttc	cct	ggg	gaa	tct		624
Ile	Phe	Ala	Glu	Met	Val	Asn	Gln	Arg	Pro	Leu	Phe	Pro	Gly	Glu	Ser		
		195				200			205								
gag	att	gat	gag	ctg	ttc	aaa	att	ttc	aga	atc	atg	gga	act	cca	aat		672
Glu	Ile	Asp	Glu	Leu	Phe	Lys	Ile	Phe	Arg	Ile	Met	Gly	Thr	Pro	Asn		
	210				215				220								
gaa	gat	aca	tgg	cat	ggg	gtg	aca	tca	ctg	cct	gac	ttt	aag	tcc	gcc		720
Glu	Asp	Thr	Trp	His	Gly	Val	Thr	Ser	Leu	Pro	Asp	Phe	Lys	Ser	Ala		
	225				230				235						240		
ttt	cca	aaa	tgg	tca	tca	aag	gac	ctt	gcg	act	atc	gtc	ccg	aat	ctt		768
Phe	Pro	Lys	Trp	Ser	Ser	Lys	Asp	Leu	Ala	Thr	Ile	Val	Pro	Asn	Leu		
			245				250						255				
gaa	aag	ccc	ggg	ctt	gat	ctc	cta	cgt	aaa	atg	ctg	tgc	ttg	gat	ccc		816
Glu	Lys	Pro	Gly	Leu	Asp	Leu	Leu	Arg	Lys	Met	Leu	Cys	Leu	Asp	Pro		
		260					265					270					
agt	aga	aga	att	aca	gcc	cgg	gcc	gca	ctg	gaa	cac	gaa	tac	ttt	aaa		864
Ser	Arg	Arg	Ile	Thr	Ala	Arg	Ala	Ala	Leu	Glu	His	Glu	Tyr	Phe	Lys		
		275				280						285					
gat	atc	ggg	ttt	gtg	cca	taa											885
Asp	Ile	Gly	Phe	Val	Pro												
	290				294												

<210> 56

<211> 294

<212> PRT

<213> Helianthus annuus

<400> 56

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

145		150		155		160
Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu						
		165		170		175
Gly Ser Arg His Tyr Ser Thr Pro Val Asp Val Trp Ser Val Gly Cys						
		180		185		190
Ile Phe Ala Glu Met Val Asn Gln Arg Pro Leu Phe Pro Gly Glu Ser						
		195		200		205
Glu Ile Asp Glu Leu Phe Lys Ile Phe Arg Ile Met Gly Thr Pro Asn						
		210		215		220
Glu Asp Thr Trp His Gly Val Thr Ser Leu Pro Asp Phe Lys Ser Ala						
		225		230		235
Phe Pro Lys Trp Ser Ser Lys Asp Leu Ala Thr Ile Val Pro Asn Leu						
		245		250		255
Glu Lys Pro Gly Leu Asp Leu Leu Arg Lys Met Leu Cys Leu Asp Pro						
		260		265		270
Ser Arg Arg Ile Thr Ala Arg Ala Ala Leu Glu His Glu Tyr Phe Lys						
		275		280		285
Asp Ile Gly Phe Val Pro						
		290				

<210> 57
 <211> 11449
 <212> DNA
 <213> Desconhecido

<220>
 <221>
 <222> (1)..(11449)
 <223> vetor

<400> 57							
gtttaccgcg	caatatatcc	tgtcaaacac	tgatagtttg	tggaattoga	gctcgggtacc		60
cgggggtcct	ctagagtcga	cctgcaggca	tgcaagcttt	gcagtgccgc	gtgaccgggt		120
cgtgcccttc	tctagagata	atgagcattg	catgtctaag	ttataaaaaa	ttaccacata		180
ttttttttgt	cacacttggt	tgaagtgcag	tttatctatc	tttatacata	tattttaaact		240
ttactctacg	aataatataa	tctatagtac	tacaataata	tcagtgtttt	agagaatcat		300
ataaatgaac	agttagacat	ggtctaaagg	acaattgagt	attttgacaa	caggactcta		360
cagttttatc	tttttagtgt	gcatgtgttc	tccttttttt	ttgcaaatag	cttcacctat		420
ataatacttc	atccatttta	ttagtacatc	catttagggg	ttagggttaa	tgggttttat		480
agactaattt	ttttagtaca	tctattttat	tctatttttag	cctctaaatt	aagaaaacta		540
aaactctatt	ttagtttttt	tatttaatat	tttagatata	aaatagaata	aaataaagtg		600
actaaaaatt	aaacaaatac	cctttaagaa	attaaaaaaa	ctaaggaaac	atttttcttg		660
tttcgagtag	ataatgccag	cctgttaaag	gccgtcgacg	agtctaacgg	acaccaacca		720
gcgaaccagc	agcgtcgcg	cgggccaagc	gaagcagacg	gcacggcacc	tctgtcgctg		780
cctctggacc	cctctcgaga	gttccgctcc	accgttgagc	ttgctccgct	gtcggcatcc		840
agaaattg	tgccggagcg	gcagacgtga	gccggcacgg	caggcgccct	cctcctcttc		900
tcacggcacc	ggcagctacg	ggggattcct	ttcccaccgc	tccttcgctt	tccttctctc		960
gcccgcgcgt	ataaatagac	acccctccca	cacctctctt	ccccaacctc	gtgttggtcg		1020
gagcgcacac	acacacaacc	agatctcccc	caaattccacc	cgtcggcacc	tcgcttcaa		1080
ggtacgccgc	tcgtcctccc	cccccccccc	cctctctacc	ttctctagat	cggcgttccg		1140
gtccatgggt	agggcccggt	agttctactt	ctgttcatgt	ttgtgttaga	tccgtgtttg		1200
tgtttagatcc	gtgctgctag	cgttcgtaga	cggatgcgac	ctgtacgtca	gacacgttct		1260
gattgctaac	ttgccagtg	ttctcttttg	ggaatcctgg	gatggctcta	gccgttccgc		1320
agacgggac	gatttcatga	ttttttttgt	ttcgttgcat	agggtttgg	ttgccctttt		1380
cctttatttc	aatatatgcc	gtgcacttgt	ttgtcgggtc	atcttttcat	gctttttttt		1440
gtcttggttg	tgatgatgtg	gtctggttgg	gcggtcgttc	tagatcggag	tagaattctg		1500
tttcaaacta	cctggtggat	ttattaattt	tgatctgtga	tgtgtgtgcc	atacatattc		1560
atagttacga	attgaagatg	atggatggaa	atatcgatct	aggataggta	tacatgttga		1620
tgccgggtttt	actgatgcac	atacagagat	gcttttttgt	cgcttggttg	tgatgatgtg		1680
gtgtggttgg	gcggtcgttc	attcgttcta	gatcggagta	gaatactgtt	tcaaactacc		1740
tggtgtattt	attaattttg	gaactgtatg	tgtgtgtcat	acatcttcat	agttacgagt		1800
ttaatgatga	tggaatatatc	gatctaggat	aggatatacat	ggtgatgtgg	gttttactga		1860
tgcatataca	tgatggcata	tgacgacatc	attcatatgc	tctaaccttg	agtacctatc		1920
tattataata	aacaagtatg	ttttataatt	atttcgatct	tgatataact	ggatgatggc		1980
atatgcagca	gctatatgtg	gattttttta	gccctgcctt	catacgctat	ttatttgctt		2040

ggtactgttt	cttttgtcga	tgctcaccct	gttgttttgg	gttactttctg	caggggtaccc	2100
ccgggggatcc	actagttcta	gaaaccatgg	ccaccgcgcg	cgccgcgtct	accgcgctca	2160
ctggcgccac	taccgctgcg	cccaaggcga	ggcgccgggc	gcacctcctg	gccaccgcgc	2220
gcgcccctgc	cgcgcccac	aggtgctcag	cgcgctcacc	cgccatgccg	atggctcccc	2280
cggccacccc	gtcccgccg	tggggcccca	ccgatccccg	caagggcgc	gacatcctcg	2340
tcgagtcacct	cgagcgctgc	ggcgctccgc	acgtcttcgc	ctaccccggc	ggcgcgcca	2400
tggagatcca	ccaggcactc	acccgctccc	ccgtcatcgc	caaccacctc	ttccgccacg	2460
agcaagggga	ggccttttgcg	gcctccggct	acgcgcgctc	ctcgggcgcg	gtcggcgctc	2520
gcacgcgccac	ctccggcccc	ggcgccacca	accttgtctc	cgcgctcgcc	gacgcgctgc	2580
tcgattccgt	ccccatggtc	gccatcacgg	gacagggtgc	gcgacgcatg	attggcaccg	2640
acgccttcca	ggagacgccc	atcgctcgagg	tcacccgctc	catcaccaag	cacaactacc	2700
tggctctcga	cgtcgacgac	atcccccgcg	tcgtgcagga	ggctttcttc	ctcgctcctc	2760
ctggtcgacc	ggggccggtg	cttgtcgaca	tcccacaagg	catccagcag	cagatggcgg	2820
tgcctgtctg	ggacaagccc	atgagctctgc	ctgggtacat	tgcgcgcctt	cccaagcccc	2880
ctgcgactga	gttgcttgag	caggtgctgc	gtcttgttgg	tgaatcccgg	cgccctgttc	2940
tttatgttgg	cgggtggctgc	gcagcatctg	gtgaggagtt	gcgacgcttt	gtggagctga	3000
ctggaatccc	ggtcacaact	actcttatgg	gcctcggcaa	cttccccagc	gacgaccac	3060
tgtctctgcg	catgctaggt	atgcatggca	cggtgtatgc	aaattatgca	gtggataaag	3120
ccgatctggt	gctgacactt	ggtgtgcggt	ttgatgatcg	tgtgacaggg	aagattgagg	3180
cttttgcga	cagggtctaa	attgtgcacg	ttgatattga	tccggctgag	attggcaaga	3240
acaagcagcc	acatgtgtcc	atctgtgcag	atgttaagct	tgccttgcag	ggcatgaatg	3300
ctcttcttga	aggaagcaca	tcaaagaaga	gctttgactt	tggctcatgg	aacgatgagt	3360
tggatcagca	gaagaggga	ttcccccttg	ggtataaaac	atctaataag	gagatccagc	3420
cacaatatgc	tattcaggtt	cttgatgagc	tgacgaaaag	cgaggccatc	atcggcacag	3480
gtgttgggca	gcaccagatg	tgggcggcac	agtactacac	ttacaagcgg	ccaaggcagt	3540
ggttgtcttc	agctggtctt	gggctatgg	gatttggttt	gccggctgct	gctggtgctt	3600
ctgtggccaa	cccagggtgtt	actgttgttg	acatcgatgg	agatggtagc	tttctcatga	3660
acgttcagga	gctagctatg	atccgaattg	agaacctccc	ggtgaaggtc	tttgtgtctaa	3720
acaaccagca	cctggggatg	gtggtgcagt	gggaggacag	gttctataag	gccaacagag	3780
cgcacacata	cttgggaaac	ccagagaatg	aaagtgcgat	atatccagat	ttcgtgacga	3840
tcgccaagg	gttcaacatt	ccagcggctc	gtgtgacaaa	gaagaacgaa	gtccgcgcag	3900
cgataaagaa	gatgctcgag	actccagggc	cgtacctctt	ggatataatc	gtcccacacc	3960
aggagcatgt	gttgccatag	atccctaagt	gtggggcttt	caaggatatg	atcctggatg	4020
gtgatggcag	gactgtgtac	tgatctaaaa	tccagcaagc	aactgatcta	aaatccagca	4080
agcaccgcct	ccctgctagt	acaagggtga	tatgttttta	tctgtgtgat	gttctcctgt	4140
attctatctt	ttttttagtg	ccgtcagcta	tctgttatgg	taatcctatg	tagcttccga	4200
ccttgtaatt	gtgtagtctg	ttgttttctt	tctggcatgt	gtcataagag	atcatttaag	4260
tgctttttgc	tacatataaa	taagataata	agcactgcta	tgcagtgggt	ctgaattggc	4320
ttctgttgcc	aaatttaagt	gtccaaactg	tcttgccttt	tgttttgcgt	atttttttcc	4380
tttttttagt	attattatat	tggtaatttc	aactcaacat	atgatgtatg	gaataatgct	4440
agggtgcaa	tttcaaacta	ttttacaaac	cagaatggca	ttttcgtggg	ttgaggggag	4500
tgaaaaaaaa	tgaggcattt	gactgaatta	gttacctgat	ccattttcgt	ggtttggatc	4560
attggaatta	aattccattc	taataatagt	aattttggca	tatatcaatt	aagttaattc	4620
ggttttatgc	aaaatatatt	tgtatactat	tattatcaag	atgtcggaga	tatttatatg	4680
ctacattttt	actatacagg	agtgcagatg	agagtgtcat	gtaagttaca	cagtagaagc	4740
aaattctatt	aatgcataaa	atcattttcca	tcatccaccc	tatgaatttg	agatagacct	4800
atatctaaac	tttgaaaagt	ggttgaatat	caaattccaa	attaaataag	ttattttatt	4860
gagtgaattc	taatttctct	aaaacgaagg	gatctaaacg	ccctctaaag	ctaatttgga	4920
aactcaaaact	ttcttagcat	tggaggggat	tgagaaaaaa	tattaattca	ttttcatctc	4980
aatcattcaa	tctccaaaga	gatttgagtt	ccttattagt	ctgttccatg	catcaaatcg	5040
gctcaatgtg	tcattatttg	ccatgacgat	tgacgagttg	ttctggggcc	tagcgctttc	5100
cacgccgatg	tgctggggcc	tggtcctgga	gaagacagct	tgatatttaa	agctatcaat	5160
tgtttcaatt	gattcccact	tcatttttct	aaatgtagaa	aacgggtgacg	tataagaaaa	5220
agaatgaatt	aggactttta	ttccgtacac	taatctagag	cgccccctta	aggcgctgcg	5280
atcgcgtaa	cagcttgctg	aggaggcctc	ggaccgttaa	ttaacacgtg	ggcgcgccac	5340
tagtcaattc	agtacattaa	aaacgtccgc	aatgtgttat	taagttgtct	aagcgtcaat	5400
ttgtttacac	cacaatatat	cctgccacca	gccagccaac	agctccccga	ccggcagctc	5460
ggcacaataa	caccactcga	tacaggcagc	ccatcagtc	gggacggcgt	cagcgggaga	5520
gccgttgtaa	ggcggcagac	tttgctcatg	ttaccgatgc	tattcggaag	aacggcaact	5580
aagctgccgg	gtttgaaaca	cggatgatct	cgcgaggagg	agcatgttga	ttgtaacgat	5640
gacagagcgt	tgctgcctgt	gatcaaatat	catctccctc	gcagagatcc	gaattatcag	5700
ccttcttatt	catttctcgc	ttaaccgtga	caggctgtcg	atcttgagaa	ctatgccgac	5760
ataataggaa	atcgctggat	aaagccgctg	aggaagctga	gtggcgctat	ttcttttagaa	5820
gtgaacgttg	acgatcgctg	accgtacccc	gatgaattaa	ttcggacgta	cgttctgaac	5880
acagctggat	acttacttgg	gcgattgtca	tacatgacat	caacaatgta	cccgtttgtg	5940
taaccgtctc	ttggagggttc	gtatgacact	agtgggtccc	ctcagcttgc	gactagatgt	6000

tgaggcctaa	cattttatta	gagagcaggc	tagttgctta	gatacatgat	cttcaggccg	6060
ttatctgtca	gggcaagcga	aaattggcca	tttatgacga	ccaatgcccc	gcagaagctc	6120
ccatctttgc	cgccatagac	gccgcgcccc	ccttttgggg	tgtagaacat	ccttttgcca	6180
gatgtgga	agaagttcgt	tgtcccattg	ttggcaatga	cgtagtagcc	ggcgaaagtg	6240
cgagacccat	ttgcgctata	tataagccta	cgatttccgt	tgcgactatt	gtcgttaattg	6300
gatgaactat	tatcgtagtt	gctctcagag	ttgtcgtaat	ttgatggact	attgtcgtaa	6360
ttgcttatgg	agttgtcgt	gttgcttgg	gaaatgtcgt	agttggatgg	ggagtagtca	6420
tagggaagac	gagcttcac	cactaaaaca	attggcaggt	cagcaagtgc	ctgccccgat	6480
gccatcgcaa	gtacgaggct	tagaaccacc	ttcaacagat	cgcgcatagt	cttccccagc	6540
tctctaacgc	ttgagttaag	ccgcgcgcgc	aagcggcgct	ggcttgaacg	aattgttaga	6600
cattatttgc	cgactacctt	ggtgatctcg	ccttttcacgt	agtgaacaaa	ttcttccaac	6660
tgatctggcg	gcgaggccaa	gcgatcttct	tgtccaagat	aagcctgcct	agcttcaagt	6720
atgacgggct	gatactgggc	cggcaggcgc	tccattggcc	agtcggcgag	gacatccttc	6780
ggcgcgattt	tgcgggttac	tgcgctgtac	caaatgcggg	acaacgtaag	cactacattt	6840
cgctcatcgc	cagcccagtc	ggcgcgcgag	ttccatagcg	ttaaggtttc	atttagcgcc	6900
tcaaatagat	cctgttcagg	aaccggatca	aagagttcct	ccgcgcgctg	acctaccaag	6960
gcaacgctat	gttctcttgc	ttttgtcagc	aagatagcca	gatcaatgtc	gatcgtggct	7020
ggctcgaaga	tacctgcaag	aatgtcattg	cgctgccatt	ctccaaattg	cagttcgcgc	7080
ttagctggat	aacgccacgg	aatgatgtcg	tcgtgcacaa	caatgggtgac	ttctacagcg	7140
cggagatact	cgctctctcc	aggggaagcc	gaagtttcca	aaaggctcgt	gatcaaagct	7200
cgccgcgttg	tttcatcaag	ccttacggtc	accgtaacca	gcaaatacat	atcactgtgt	7260
ggcttcaggc	cgccatccac	tgcggagccg	tacaaatgta	cgccagcaaa	cgtcggttcg	7320
agatggcgct	cgatgacgcc	aactacctct	gatagttgag	tcgatacttc	ggcgatcacc	7380
gcttccctca	tgatgtttaa	ctcctgaatt	aagccgcgcc	gcgaagcggg	gtcggcttga	7440
atgaattggt	aggcgtcatc	ctgtgtctcc	gagaaccagt	accagtacat	cgctgtttcg	7500
ttcgagactt	gaggtctagt	tttatacgtg	aacaggctcaa	tgccgcgcgag	agtaaaagcca	7560
cattttgcgt	acaaattgca	ggcaggtaga	ttgttcgttt	gtgtctctaa	tcgtatgcca	7620
aggagctgtc	tgcttagtgc	ccactttttc	gcaaattcga	tgagactgtg	cgcgactcct	7680
ttgcctcgg	gcgtgtgcga	cacaacaatg	tgttcgatag	aggctagatc	gttccatgtt	7740
gagttgagtt	caatcttccc	gacaagctct	tggtcgatga	atgcgccata	gcaagcagag	7800
tcttcatcag	agtcatcatc	cgagatgtaa	tccttccggg	aggggctcac	acttctggta	7860
gatagttcaa	agccttggtc	ggataggtgc	acatcgaaac	cttcacgaac	aatgaaatgg	7920
ttctcagcat	ccaatgtttc	cgccacctgc	tcagggtatca	ccgaaatcct	catatgacgc	7980
ctaacgcctg	gcacagcgga	tcgcaaacct	ggcgcggctt	ttggcacaaa	aggcgtgaca	8040
ggttttcgaa	tccgttgctg	ccacttggtt	acccttttgc	cagattttgg	aactataatt	8100
tatgttagag	gcgaagtctt	gggtaaaaac	tgccctaaaa	ttgctgggga	tttcaggaaa	8160
gtaaacatca	ccttccggct	cgatgtctat	tgtagatata	tgtagtgat	ctacttgatc	8220
gggggatctg	ctgcctcgcg	cgtttcgggt	atgacggtga	aaacctctga	cacatgcagc	8280
tcccgagac	ggtcacagct	tgtctgtaa	cggatgccgg	gagcagacaa	gcccgtcagg	8340
gcgcgtcagc	gggtgttggc	gggtgtcggg	gcgcagccat	gacccagtc	cgtagcgata	8400
gcggagtgt	tactggctta	actatgcggc	atcagagcag	attgtactga	gagtgcacca	8460
tatgcggtgt	gaaataccgc	acagatgcgt	aaggagaaaa	taccgcatca	ggcgctcttc	8520
cgcttcctcg	ctcactgact	cgctgcgctc	ggtcggttcg	ctgcggcgag	cggtatcagc	8580
tactcaaaag	gcggtaatat	ggttatccac	agaatcaggg	gataacgcag	gaaagaacat	8640
gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaaa	gccgcgttgc	tggcggtttt	8700
ccatagcgct	cgccccctcg	acgagcatca	caaaaaatga	cgctcaagtc	agaggtggcg	8760
aaacccgaca	ggactataaa	gataccaggc	gtttccccct	ggaagctccc	tcgtgcgctc	8820
tcctgttccg	accctgccgc	ttaccggata	cctgtccgcc	tttctccctt	cggaagcggt	8880
ggcgctttct	catagctcac	gctgtaggta	tctcagttcg	gtgtagggtc	ttcgctccaa	8940
gctgggctgt	gtgcacgaac	cccccgttca	gcccagaccg	tgccgcttat	ccggttaacta	9000
tcgtcttgag	tccaacccgg	taagacacga	cttatcgcca	ctggcagcag	ccactggtta	9060
caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	ggtggcctaa	9120
ctacgggctac	actagaagga	cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	9180
cggaaaaaga	gttggtagct	cttgatccgg	caaaacaaacc	accgctggta	gcggtgggtt	9240
ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	9300
cttttctacg	gggtctgacg	ctcagtgga	cgaaaactca	cgtaaaggga	ttttggctcat	9360
gagattatca	aaaaggatct	tcacctagat	cctttttaaat	taaaaatgaa	gttttaaatc	9420
aatctaaagt	atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	tcagtgaggc	9480
acctactcta	gcgactctgt	tatttcgttc	atccatagtt	gctgactcc	ccgctgtgta	9540
gataactacg	atacggggag	gcttaccatc	tggtcccgag	gctgcaatga	taccgcgaga	9600
cccacgctca	ccggctccag	atttatcagc	aataaaccag	ccagccggaa	gggcccagcg	9660
cagaagtgg	cctgcaactt	tatccgcctc	catccagctc	attaattgtt	gccgggaagc	9720
tagagtaagt	agttcgccag	ttaatagttt	gcgcaacggt	gttgccattg	ctgcaggggg	9780
gggggggggg	ggggacttcc	attgttcatt	ccacggacaa	aaacagagaa	aggaaacgac	9840
agaggccaaa	aagcctcgct	ttcagcacct	gtcgtttcct	ttcttttcag	agggtatttt	9900
aaataaaaaa	attaagttat	gacgaagaag	aacggaaaac	ccttaaaccg	gaaaattttc	9960

ataaatagcg	aaaacccgcg	aggtcgccgc	cccgtaacct	gtcggatcac	cggaaaggac	10020
ccgtaaagtg	ataatgatta	tcatctacat	atcacaacgt	gcgtggaggc	catcaaacca	10080
cgtcaaataa	tcaattatga	cgcaggatc	gtattaattg	atctgcatca	acttaacgta	10140
aaaacaactt	cagacaatac	aatcagcga	cactgaatac	ggggcaacct	catgtccccc	10200
ccccccccc	ccctgcaggc	atcgtggtgt	cacgctcgtc	gtttggtatg	gcttcattca	10260
gctccggttc	ccaacgatca	aggcgagtta	catgatcccc	catgttggtc	aaaaaagcgg	10320
ttagctcctt	cggtcctccg	atcgttggtc	gaagtaagtt	ggccgcagtg	ttatcactca	10380
tggttatggc	agcactgcat	aattctctta	ctgtcatgcc	atccgtaaga	tgcttttctg	10440
tgactggtga	gtactcaacc	aagtcattct	gagaatagtg	tatgcggcga	ccgagttgct	10500
cttgcccggc	gtcaacacgg	gataataccg	cgccacatag	cagaacttta	aaagtgtc	10560
tcattgga	acgttcttcg	gggcgaaaac	tctcaaggat	cttaccgctg	ttgagatcca	10620
gttcgatgta	accactcgt	gcacccaact	gatcttcagc	atcttttact	ttcaccagcg	10680
tttctgggtg	agcaaaaaa	ggaaggcaaa	atgccgcaaa	aaagggaata	aggcgacac	10740
ggaaatgttg	aatactcata	ctcttctttt	ttcaatatta	ttgaagcatt	tatcagggtt	10800
attgtctcat	gagcggatac	atatttgaat	gtatttagaa	aaataaaca	ataggggttc	10860
cgcgcacatt	tccccgaaaa	gtgccacctg	acgtctaaga	aaccattatt	atcatgacat	10920
taacctataa	aataggcgt	atcacgaggc	cctttcgtct	tcaagaattg	gtcgacgac	10980
ttgtcgcgtt	cggatatttt	cgtggagttc	ccgccacaga	cccggattga	aggcgagatc	11040
cagcaactcg	cgccagatca	tcctgtgacg	gaactttggc	gcgtgatgac	tggccaggac	11100
gtcggccgaa	agagcgacaa	gcagatcacg	cttttcgaca	gcgtcgatt	tgcatcgag	11160
gatttttcgg	cgctgcgcta	cgtccgcgac	cgcgttgagg	gatcaagcca	cagcagccca	11220
ctcgaccttc	tagccgaccc	agacgagcca	agggatcttt	ttggaatgct	gctccgctgt	11280
caggctttcc	gacgtttggg	tggttgaaca	gaagtcatta	tcgtacggaa	tgccaagcac	11340
tcccagaggg	aaccctgtgg	ttggcatgca	catacaaatg	gacgaacgga	taaacctttt	11400
cacgcccttt	taaatatccg	ttattcta	aaacgctctt	ttctcttag		11449

<210> 58

<211> 4510

<212> DNA

<213> Desconhecido

<220>

<221>

<222> (1)..(4510)

<223> vetor

<400> 58

acacaggaaa	cagctatgac	catgattacg	ccaagctatc	gtttaaacct	taaggcgatc	60
gcgctgaggg	ggaccgcacg	tggaattagc	ttggcgcgcc	aattcccgat	ctagtaacat	120
agatgacacc	gcgcgcgata	atttatccta	gtttgcgcgc	tatatattgt	tttctatcgc	180
gtattaaatg	tataattgcg	ggactcta	cataaaaacc	catctcataa	ataacgtcat	240
gcattacatg	ttaattatta	catgcttaac	gtaattcaac	agaaattata	tgataatcat	300
cgcaagaccg	gcaacaggat	tcaatcttaa	gaaactttat	tgccaaatgt	ttgaacgac	360
ggggaaaattc	gagctccacc	gcggtggcgg	ccgctctaga	actagtggat	ccccgggct	420
gcaggaattc	gatatcaagc	ttatcgatac	cgctcgacctc	gagggggggc	ccggtacctt	480
ggtgaactac	cgatgatcgt	aagagcttac	agacggtgcc	ttatataggc	agagcgctcg	540
aaggggggtg	gtgtcacaca	cgcactgcga	tccttgctta	caagctaagc	aaaggcatcg	600
tggcagacaa	ggaataaagt	ggcacagggtg	ccaaaagaaa	gtggacagca	cacttgtccg	660
aaaagcacac	aataatacaa	ctcagtggag	catccactga	atgggcccag	tactctccgc	720
tgcggtcatc	cctgacgtca	atgcctcctc	aatcgtcatt	gcttacgtct	tgagatctta	780
gcacatcttc	tgcggtact	ggcgacatag	cccttggtgt	ggcctggtat	tcctcgaact	840

cctgctcttc tatcttcagc ttggtgactt ccttcataag ttcctcctcc attgtttgca	900
cgccaccggt cttcagaact tcaagattct tgagaaattc atcaattctg tcttgaatgt	960
gttcttcaat gagatctgcc cagtaccagc aatggcattt gttcattgcg cacttgaaga	1020
atttccttcc tgggttggca gatgttctgg acactaactg aattgcaggc ttcctgcatg	1080
cacagagtag aactggttcc tctgttagca taaaagcttg ctctgttctt tcgaacatgt	1140
ccttcttgaa tgtggggaat ggatccagaa ttttgtttcc ccattctctg agttgctcag	1200
tgtatggatg atctggataa ggaattactt cccttatggc ttgtgtaagc aggatcatct	1260
cttcggttgg ttcattctga gctaatttgg ctttgagcct ggacaagatg tcagctaaac	1320
cattgctctt cccttttatg tgttcaatga ctatctctgg tcctgcacca gtgatgtagt	1380
ccatgaacct gatccatctg atctcagaag gcttgtgttc agcactcttg ttgtagaacc	1440
tttcgattgc actactgtca gttctgactg tgatctctct tttgtccaag tagaacaatc	1500
tcactctttc taagccattc ataaccccat agatttctgc atcacagggt ccttttggct	1560
tatcaaattt tccactggca tacctacaga tttgctctgt atttcttggg tctgccttgt	1620
ttttcttcca ctgcatact gctcccatc cagttgcaca tgcatctgtt tcaatgataa	1680
tgtatgcatac ttctgggtgga atagtgagat ttggaagcgt tctcaccatt gtcttgatcc	1740
tattgatcag cttccaatct tctgaattga gccttcgctc acctttctct gaggtctttg	1800
gatataatgg gccaaagaagc ttgcccatac ctttgatgtg gtttctggca tagttcagtg	1860
ttgctagcca ggattaatta aaggcctgtt aacagcgctg ggcccgataa ttcactggcc	1920
gtcgttttac aacgtcgtga ctgggaaaac cctggcggtta cccaacttaa tcgccttgca	1980
gcacatcccc ctttcgccag ctggcgtaat agcgaagagg ccgcaccga tcgcccttcc	2040
caacagttgc gcagcctgaa tggcgaatgg cgctgatgc ggtattttct ccttacgcac	2100
ctgtgcggta tttcacaccg catatggtgc actctcagta caatctgctc tgatgccgca	2160
tagttaagcc agccccgaca cccgccaaca cccgctgacg cgccctgacg ggcttgtctg	2220
ctcccgcat ccgcttacag acaagctgtg accgtctccg ggagctgcat gtgtcagagg	2280
ttttcaccgt catcaccgaa acgcgcgaga cgaaagggcc tcgtgatacg cctattttta	2340
taggttaatg tcatgataat aatggtttct tagacgtcag gtggcacttt tcggggaaat	2400
gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg	2460
agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa	2520
catttccgtg tcgcccttat tccctttttt gcggcatttt gccttctgt ttttgctcac	2580
ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac	2640
atcgaactgg atctcaacag cggtaaagac cttgagagtt ttcgccccga agaacgtttt	2700
ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg tattgacgcc	2760
gggcaagagc aactcggctg ccgcatacac tattctcaga atgacttggt tgagtactca	2820

ccagtcacag	aaaagcatct	tacggatggc	atgacagtaa	gagaattatg	cagtgtctgcc	2880
ataaccatga	gtgataacac	tgcggccaac	ttactttctga	caacgatcgg	aggaccgaag	2940
gagctaaccg	cttttttgca	caacatgggg	gatcatgtaa	ctcgccttga	tcgttgggaa	3000
ccggagctga	atgaagccat	accaaacgac	gagcgtgaca	ccacgatgcc	tgtagcaatg	3060
gcaacaacgt	tgcgcaaact	attaactggc	gaactactta	ctctagcttc	ccggcaacaa	3120
ttaatagact	ggatggaggc	ggataaagtt	gcaggaccac	ttctgcgctc	ggcccttcg	3180
gctggctggt	ttattgctga	taaatctgga	gccggtgagc	gtgggtctcg	cggtatcatt	3240
gcagcactgg	ggccagatgg	taagccctcc	cgtatcgtag	ttatctacac	gacggggagt	3300
caggcaacta	tggatgaacg	aaatagacag	atcgctgaga	taggtgcctc	actgattaag	3360
catttgtaac	tgtcagacca	agtttactca	tatatacttt	agattgattt	aaaacttcat	3420
ttttaattta	aaaggatcta	ggtgaagatc	cttttttgata	atctcatgac	caaaatccct	3480
taacgtgagt	tttcgttcca	ctgagcgtca	gaccccgtag	aaaagatcaa	aggatcttct	3540
tgagatcctt	tttttctgcg	cgtaatctgc	tgcttgcaaa	caaaaaaacc	accgctacca	3600
gcggtggttt	gtttgccgga	tcaagagcta	ccaactcttt	ttccgaaggt	aactggcttc	3660
agcagagcgc	agataccaaa	tactgttctt	ctagtgtagc	cgtagttagg	ccaccacttc	3720
aagaactctg	tagcaccgcc	tacatacctc	gctctgctaa	tcctgttacc	agtggctgct	3780
gccagtggcg	ataagtcgtg	tcttaccggg	ttggactcaa	gacgatagtt	accggataag	3840
gcgcagcggc	cgggctgaac	ggggggttcg	tgcacacagc	ccagcttgga	gcgaacgacc	3900
tacaccgaac	tgagatacct	acagcgtgag	ctatgagaaa	gcgccacgct	tcccgaaggg	3960
agaaaggcgg	acaggatatcc	ggtaagcggc	agggtcggaa	caggagagcg	cacgagggag	4020
cttccagggg	gaaacgcctg	gtatctttat	agtcctgtcg	ggtttcgcca	cctctgactt	4080
gagcgtcgat	ttttgtgatg	ctcgtcaggg	gggaggagcc	tatggaaaaa	cgccagcaac	4140
gcggcctttt	tacggttcct	ggccttttgc	tggccttttg	ctcacatggt	ctttcctgcg	4200
ttatccccctg	attctgtgga	taaccgtatt	accgcctttg	agtgagctga	taccgctcgc	4260
cgcagccgaa	cgaccgagcg	cagcgagtca	gtgagcgagg	aagcgggaaga	gcgccaata	4320
cgcaaaccgc	ctctccccgc	gcgttggccg	attcattaat	gcagctggca	cgacaggttt	4380
cccgactgga	aagcgggcag	tgagcgcaac	gcaattaatg	tgagttagct	cactcattag	4440
gcaccccagg	ctttacactt	tatgcttccg	gctcgtatgt	tgtgtggaat	tgtgagcgga	4500
taacaatttc						4510

<210> 59

<211> 13170

<212> DNA

<213> Desconhecido

<220>

<221>

<222> (1)..(13170)

<223> vetor

<400> 59

gtttacccgc caatatatcc tgtcaaacac tgatagtttg tggaattcga gctcgggtacc	60
cggggatcct ctagagtcga cctgcaggca tgcaagcttc cggctgcagt gcagcgtgac	120
ccggtcgtgc ccctctctag agataatgag cattgcatgt ctaagttata aaaaattacc	180
acatatTTTT tttgtcacac ttgtttgaag tgcagtttat ctatctttat acatatTTTT	240
aaactttact ctacgaataa tataatctat agtactacaa taatatcagt gtttttagaga	300
atcatataaa tgaacagtta gacatggctt aaaggacaat tgagtatttt gacaacagga	360
ctctacagtt ttatctTTTT agtgtgcatg tgttctcctt tttttttgca aatagcttca	420
cctatataat acttcatcca ttttattagt acatccattt agggtttagg gttaatgggt	480
tttatagact aattttttta gtacatctat ttatttctat ttttagcctct aaattaagaa	540
aactaaaact ctatttttagt ttttttattt aatagtttag atataaaata gaataaaata	600
aagtgactaa aaattaaaca aatacccttt aagaaattaa aaaaactaag gaaacatttt	660
tcttgtttcg agtagataat gccagcctgt taaacgccgt cgacgagtct aacggacacc	720
aaccagcgaa ccagcagcgt cgcgtcgggc caagcgaagc agacggcacg gcatctctgt	780
cgctgcctct ggacccctct cgagagttcc gctccaccgt tggacttgct ccgctgtcgg	840
catccagaaa ttgctgggcg gagcggcaga cgtgagccgg cacggcaggc ggctctctcc	900
tcctctcacg gcaccggcag ctacggggga ttcttttccc accgctcctt cgctttccct	960
tcctcgcccg ccgtaataaa tagacacccc ctccacaccc tctttcccca acctcgtgtt	1020
gttcggagcg cacacacaca caaccagatc tcccccaaat ccaccgctcg gcaacctcgc	1080
ttcaaggtac gccgctcgtc ctcccccccc cccccctct ctacctctc tagatcggcg	1140
ttccggtcca tggttagggc ccggtagttc tacttctgtt catgtttgtg ttagatccgt	1200
gtttgtgtta gatccgtgct gctagcgttc gtacacggat gcgacctgta cgtcagacac	1260
gttctgattg ctaacttgcc agtgtttctc tttggggaat cctgggatgg ctctagccgt	1320
tccgcagacg ggatcgattt catgattttt tttgtttcgt tgcatagggt ttggtttgcc	1380
cttttccttt atttcaatat atgccgtgca cttgtttgtc gggtcactct ttcatgcttt	1440
tttttgcctt ggttgtgatg atgtggtctg gttgggcggc cgttctagat cggagtagaa	1500
ttctgtttca aactacctgg tggatttatt aattttggat ctgtatgtgt gtgccataca	1560
tattcatagt tacgaattga agatgatgga tggaaatata gatctaggat aggtatacat	1620
gttgatgcgg gttttactga tgcataaca gagatgcttt ttgttcgctt ggttgtgatg	1680
atgtggtgtg gttgggcggc cgttcattcg ttctagatcg gagtagaata ctgtttcaaa	1740
ctacctggtg tatttattaa ttttggaact gtatgtgtgt gtcatacatc ttcatagtta	1800
cgagtttaag atggatggaa atatcgatct aggataggta tacatgttga tgtgggtttt	1860

actgatgcat atacatgatg gcatatgcag catctattca tatgctctaa ccttgagtac	1920
ctatctatta taataaacia gtatgtttta taattatttc gatcttgata tacttgatg	1980
atggcatatg cagcagctat atgtggattt ttttagccct gccttcatac gctatttatt	2040
tgcttggtac tgtttctttt gtcgatgtc accctgttgt ttggtgttac ttctgcaggg	2100
tacccccggg gatccactag ttctagaaac catggccacc gccgccgccg cgtctaccgc	2160
gctcactggc gccactaccg ctgcgcccac gccgaggcgc cgggcgcacc tcctggccac	2220
ccgccgcgc ctcgccgcgc ccatcagggtg ctcagcggcg tcacccgccca tgccgatggc	2280
tcccccgcc accccgctcc ggccgtgggg cccaccgat ccccgcaagg gcgcgcacat	2340
cctcgtcgag tccctcgagc gctgcggcgt ccgcgacgtc ttgcctacc ccggcggcgc	2400
gtccatggag atccaccagg cactcaccgc cccccgtc atcgccaacc acctcttcgc	2460
ccacgagcaa ggggaggcct ttgcggcctc cggctacgcg cgctcctcgg gccgcgtcgg	2520
cgtctgcac gccacctccg gccccggcgc caccaacctt gtctccgcgc tcgccgacgc	2580
gctgctcgat tccgtcccca tggtcgccat cacgggacag gtgccgcgac gcatgattgg	2640
caccgacgcc ttccaggaga cgcctatcgt cgaggtcacc cgctccatca ccaagcacia	2700
ctacctggtc ctcgacgtcg acgacatccc ccgcgtcgtg caggaggctt tcttctcgc	2760
ctcctctggt cgaccggggc cgggtgcttg cgacatcccc aaggacatcc agcagcagat	2820
ggcgggtgct gtctgggaca agcccatgag tctgcctggg tacattgcgc gccttcccaa	2880
gccccctgc actgagttgc ttgagcagggt gctgcgtctt gttggtgaat cccggcgccc	2940
tgttctttat gttggcgggtg gctgcgcagc atctggtgag gagttgcgac gctttgtgga	3000
gctgactgga atccccgtca caactactct tatgggcctc ggcaacttcc ccagcgacga	3060
cccactgtct ctgcgcatgc taggtatgca tggcacgggtg tatgcaaatt atgcagtgga	3120
taaggccgat ctgttgcttg cacttggtgt gcggtttgat gatcgtgtga cagggaagat	3180
tgaggctttt gcaagcaggg ctaagattgt gcacgttgat attgatccgg ctgagattgg	3240
caagaacaag cagccacatg tgtccatctg tgcagatgtt aagcttgctt tgcagggcat	3300
gaatgctctt cttgaaggaa gcacatcaaa gaagagcttt gactttggct catggaacga	3360
tgagttggat cagcagaaga gggaattccc ccttggttat aaaacatcta atgaggagat	3420
ccagccacia tatgctattc aggttcttga tgagctgacg aaaggcgagg ccatcatcgg	3480
cacagggtgtt gggcagcacc agatgtgggc ggcacagtac tacacttaca agcggccaag	3540
gcagtgggtt tcttcagctg gtcttggggc tatgggattt ggtttgccgg ctgctgctgg	3600
tgcttctgtg gccaaaccag gtgttactgt tggtgacatc gatggagatg gtagctttct	3660
catgaacgtt caggagctag ctatgatccg aattgagaac cccccggtga aggtctttgt	3720
gctaaacaac cagcacctgg ggatggtggt gcagtgggag gacaggttct ataaggccaa	3780
cagagcgcac acatacttgg gaaaccaga gaatgaaagt gagatatatc cagatttctg	3840

gacgatcgcc aaaggggttca acattccagc ggtccgtgtg acaaagaaga acgaagtccg	3900
cgcagcgata aagaagatgc tcgagactcc agggccgtac ctcttgata taatcgtccc	3960
acaccaggag catgtgttgc ctatgatccc taatgggtggg gctttcaagg atatgatcct	4020
ggatgggtgat ggcaggactg tgtactgatc taaaatccag caagcaactg atctaaaatc	4080
cagcaagcac cgcctccctg ctagtacaag ggtgatatgt ttttatctgt gtgatgttct	4140
cctgtattct atcttttttt gtaggccgtc agctatctgt tatggtaatc ctatgtagct	4200
tccgaccttg taattgtgta gtctgttgtt ttccttctgg catgtgtcat aagagatcat	4260
ttaagtgcct tttgctacat ataaataaga taataagcac tgctatgcag tggttctgaa	4320
ttggcttctg ttgccaaatt taagtgtcca actggtcctt gcttttgttt tcgctatttt	4380
tttccttttt tagttattat tatattggta atttcaactc aacatatgat gtatggaata	4440
atgctagggc tgcaatttca aactatttta caaaccagaa tggcattttc gtggtttgag	4500
gggagtga aaatgatgagg catttgactg aattagttac ctgatccatt ttcgtgggtt	4560
ggatcattgg aattaaattc cattctaata atagtaattt tggcatatat caattaagtt	4620
aattcgggtt tatgcaaaat atatttgtat actattatta tcaagatgtc ggagatat	4680
atatgctaca tttttactat acaggagtga gatgaagagt gtcatgtaag ttacacagta	4740
gaaacaaatt ctattaatgc ataaaatcat ttccatcatc caccctatga atttgagata	4800
gacctatatc taaactttga aaagtgggtg aatatcaaat tccaaattaa ataagttatt	4860
ttattgagtg aattctaatt tctctaaaac gaagggatct aaacgccctc taaagcta	4920
ttggaaactc aaactttctt agcattggag gggattgaga aaaaatatta attcattttc	4980
atctcaatca ttcaatctcc aaagagattt gagttcctta ttagtctgtt ccatgcatca	5040
aatcggctca atgtgtcatt atttgccatg acgattgacg agttgttctg gggcctagcg	5100
ctttccacgc cgatgtgctg gggcctggtc ctggagaaga cagcttgata tttaaagcta	5160
tcaattgttt caattgattc ccacttcatt tttctaaatg tagaaaacgg tgacgtataa	5220
gaaaaagaat gaattaggac ttttattccg tactactaatc tagagcggcc ccttaaggcg	5280
ctgcgatcgc gttaacagct tgctgaggag gcctcggacc gttaattaat cctggctagc	5340
aactactgaac tatgccagaa accacatcaa agatatgggc aagcttcttg gccattata	5400
tccaaagacc tcagagaaaag gtgagcgaag gctcaattca gaagattgga agctgatcaa	5460
taggatcaag acaatgggtga gaacgcttcc aaatctcact attccaccag aagatgcata	5520
cattatcatt gaaacagatg catgtgcaac tggatgggga gcagtatgca agtggagaa	5580
aaacaaggca gacccaagaa atacagagca aatctgtagg tatgccagtg gaaaatttga	5640
taagccaaaa ggaacctgtg atgcagaaat ctatgggggtt atgaatggct tagaaaagat	5700
gagattgttc tacttggaca aaagagagat cacagtcaga actgacagta gtgcaatcga	5760
aaggttctac aacaagagtg ctgaacacaa gccttctgag atcagatgga tcaggttcat	5820

ggactacatc	actggtgcag	gaccagagat	agtcattgaa	cacataaaaag	ggaagagcaa	5880
tggtttagct	gacatcttgt	ccagggtcaa	agccaaaatta	gctcagaatg	aaccaacgga	5940
agagatgatc	ctgcttacac	aagccataag	ggaagtaatt	ccttatccag	atcatccata	6000
cactgagcaa	ctcagagaat	ggggaacaa	aattctggat	ccattcccca	cattcaagaa	6060
ggacatgttc	gaaagaacag	agcaagcttt	tatgctaaca	gaggaaccag	ttctactctg	6120
tgcatgcagg	aagcctgcaa	ttcagttagt	gtccagaaca	tctgccaacc	caggaaggaa	6180
attcttcaag	tgcgcaatga	acaaatgcc	ttgctggtac	tgggcagatc	tcattgaaga	6240
acacattcaa	gacagaattg	atgaatttct	caagaatctt	gaagttctga	agaccggtgg	6300
cgtgcaaaca	atggaggagg	aacttatgaa	ggaagtcacc	aagctgaaga	tagaagagca	6360
ggagttcgag	gaataccagg	ccacaccaag	ggctatgtcg	ccagtagccg	cagaagatgt	6420
gctagatctc	caagacgtaa	gcaatgacga	ttgaggaggc	attgacgtca	gggatgaccg	6480
cagcggagag	tactgggccc	attcagtgga	tgctccactg	agttgtatta	ttgtgtgctt	6540
ttcggacaag	tgtgctgtcc	actttctttt	ggcacctgtg	ccactttatt	ccttgtctgc	6600
cacgatgcct	ttgcttagct	tgtaagcaag	gatcgcagtg	cgtgtgtgac	accaccccc	6660
ttccgacgct	ctgcctatat	aaggcacctg	ctgtaagctc	ttacgatcat	cggtagttca	6720
ccaaggtacg	cccgggtcgc	tcctacgcgt	caatgatccg	cggacgccga	gcccgagctc	6780
gaatttcccc	gatcgttcaa	acatttgcca	ataaagtttc	ttaagattga	atcctgttgc	6840
cggctcttgcg	atgattatca	tataatttct	gttgaattac	gttaagcatg	taataattaa	6900
catgtaatgc	atgacgttat	ttatgagatg	ggtttttatg	attagagtcc	cgcaattata	6960
catttaatac	gcgatagaaa	acaaaatata	gcgcgcgcaa	taggataaat	tatcgcgcg	7020
ggtgtcatct	atgttactag	atcggaatt	ggcgcgccac	tagtcaattc	agtacattaa	7080
aaacgtccgc	aatgtgttat	taagttgtct	aagcgtcaat	ttgtttacac	cacaatatat	7140
cctgccacca	gccagccaac	agctccccga	ccggcagctc	ggcacaaaat	caccactcga	7200
tacaggcagc	ccatcagtcc	gggacggcgt	cagcgggaga	gccgttgtaa	ggcggcagac	7260
tttgctcatg	ttaccgatgc	tattcggaag	aacggcaact	aagctgccgg	gtttgaaaca	7320
cggatgatct	cgcggagggt	agcatgttga	ttgtaacgat	gacagagcgt	tgctgcctgt	7380
gatcaaatat	catctccctc	gcagagatcc	gaattatcag	ccttcttatt	catttctcgc	7440
ttaaccgtga	caggctgtcg	atcttgagaa	ctatgccgac	ataataggaa	atcgctggat	7500
aaagccgctg	aggaagctga	gtggcgctat	ttctttagaa	gtgaacgttg	acgatcgtcg	7560
accgtacccc	gatgaattaa	ttcggacgta	cgttctgaac	acagctggat	acttacttgg	7620
gcgattgtca	tacatgacat	caacaatgta	cccgtttgtg	taaccgtctc	ttggaggttc	7680
gtatgacact	agtggttccc	ctcagcttgc	gactagatgt	tgaggcctaa	cattttatta	7740
gagagcaggc	tagttgctta	gatacatgat	cttcaggccg	ttatctgtca	gggcaagcga	7800

aaattggcca	tttatgacga	ccaatgcccc	gcagaagctc	ccatctttgc	cgccatagac	7860
gccgcgcccc	ccttttgggg	tgtagaacat	ccttttgcca	gatgtggaaa	agaagttcgt	7920
tgtcccatg	ttggcaatga	cgtagtagcc	ggcgaaagt	cgagacccat	ttgcgctata	7980
tataagccta	cgatttcctg	tgcgactatt	gtcgttaatt	gatgaactat	tatcgtagtt	8040
gctctcagag	ttgtcgtaat	ttgatggact	attgtcgtaa	ttgcttatgg	agttgtcgta	8100
gttgcttggg	gaaatgtcgt	agttggatgg	ggagtagtca	tagggaagac	gagcttcac	8160
cactaaaaca	attggcaggt	cagcaagtgc	ctgccccgat	gccatcgcaa	gtacgaggct	8220
tagaaccacc	ttcaacagat	cgcgcatagt	cttccccagc	tctctaacgc	ttgagttaag	8280
ccgcgcccgc	aagcggcgtc	ggcttgaacg	aattgttaga	cattatttgc	cgactacctt	8340
ggtgatctcg	cctttcacgt	agtgaacaaa	ttcttccaac	tgatctgcgc	gcgaggccaa	8400
gcgatcttct	tgtccaagat	aagcctgcct	agcttcaagt	atgacgggct	gatactgggc	8460
cggcaggcgc	tccattgccc	agtcggcagc	gacatccttc	ggcgcgattt	tgccggttac	8520
tgcgctgtac	caaatgcggg	acaacgtaag	cactacattt	cgctcatcgc	cagcccagtc	8580
gggcggcgag	ttccatagcg	ttaaggtttc	atttagcgcc	tcaaatagat	cctgttcagg	8640
aaccggatca	aagagttcct	ccgccgctgg	acctaccaag	gcaacgctat	gttctcttgc	8700
ttttgtcagc	aagatagcca	gatcaatgtc	gatcgtggct	ggctcgaaga	tacctgcaag	8760
aatgtcattg	cgctgccatt	ctccaaattg	cagttcgcg	ttagctggat	aacgccacgg	8820
aatgatgtcg	tcgtgcacaa	caatggtgac	ttctacagcg	cggagaatct	cgctctctcc	8880
aggggaagcc	gaagtttcca	aaaggtcgtt	gatcaaagct	cgccgcgttg	tttcatcaag	8940
ccttacggtc	accgtaacca	gcaaatacat	atcactgtgt	ggcttcaggc	cgccatccac	9000
tgcgagccg	tacaaatgta	cggccagcaa	cgctcggttc	agatggcgct	cgatgacgcc	9060
aactacctct	gatagttgag	tcgatacttc	ggcgatcacc	gcttccctca	tgatgtttaa	9120
ctcctgaatt	aagccgcgcc	gcgaagcgg	gtcggcttga	atgaattggt	aggcgtcatc	9180
ctgtgctccc	gagaaccagt	accagtacat	cgctgtttcg	ttcgagactt	gaggtctagt	9240
tttatacgtg	aacaggtcaa	tgccgccgag	agtaaagcca	cattttgcgt	acaaattgca	9300
ggcaggtaca	ttgttcgttt	gtgtctctaa	tcgtatgcca	aggagctgtc	tgcttagtgc	9360
ccactttttc	gcaaattcga	tgagactgtg	cgcgactcct	ttgcctcggg	gcgtgtgcga	9420
cacaacaatg	tggtcgatag	aggctagatc	gttccatggt	gagttgagtt	caatcttccc	9480
gacaagctct	tggtcgatga	atgcgccata	gcaagcagag	tcttcatcag	agtcatcatc	9540
cgagatgtaa	tccttccggg	aggggctcac	acttctggta	gatagttcaa	agccttggtc	9600
ggataggtgc	acatcgaaca	cttcacgaac	aatgaaatgg	ttctcagcat	ccaatgtttc	9660
cgccacctgc	tcagggatca	ccgaaatctt	catatgacgc	ctaacgcctg	gcacagcgga	9720
tcgcaaacct	ggcgcgggct	ttggcacaaa	aggcgtgaca	ggtttgcgaa	tccgttgctg	9780

ccacttgтта acccttttgc cagatttggт аactataatt tatgttagag gcgaagtctt	9840
gggtaaaaac tggcctaaaa ttgctgggga tttcaggaaa gtaaacaтca ccttcaggct	9900
cgatgtctat tgtagatata tgtagtgtat ctacttgatc gggggatctg ctgcctcgcg	9960
cgtttcggтg atgacggтga aaacctctga cacatgcagc тcccgгagac ggtcacagct	10020
tgtctgтаag cggatgccгg gagcagacaa gcccgtcagg gcgcgtcagc gggтgttgгc	10080
gggtgtcggg gcgcagccat gaccagтca cgtagcgata gcggagtгta tactggctta	10140
actatgcggc atcagagcag attgtactga gagtgcacca tatgcggтgt gaaataccgc	10200
acagatgcgt aaggagaaaa тaccgcatca ggcgtctctc cgcttcctcg ctactgact	10260
cgctgcgctc ggtcgttcgг ctgcggcgag cggtatcagc тactcaaaг gcggтаatac	10320
ggttatccac agaatcaggг gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa	10380
aggccaggaa ccgtaaaaaг gccgcgttgс tggcgttttt ccataggctc cgccccctг	10440
acgagcatca caaaaatcga cgctcaagтc agaggтggcg aaacccgaca ggactataaa	10500
gataccaggc gtttccccct ggaagctccc тcgtgcgctc тcctgttcг accctgcgc	10560
ttaccggata cctgtccгc tttctccctt cgгgaagcgт ggcgctttct catagctcac	10620
gctgtagгta тctcagttcg gtgtagгtcг тtcgctccaa gctgggctgt gtgcacgaac	10680
ccccgттca gcccgaccгc тgcgcттat ccgгtaacta тcgtcttgag тccaaccгg	10740
таagacacga cttatcgcca ctggcagcag ccactггтаa caggattagc agagcgaggт	10800
atgtagгcgг tgctacagag тtcttgaaгt gтггcctaa ctacggctac actagaagga	10860
cagtatttgг tatctgcгct ctgctgaagc cagttacctt cgгaaaaга gttggtagct	10920
cttgatccгg caaacaаacc accgctггta gcgгtgгttt тtttgттгc aagcagcaga	10980
ttacgcгcag aaaaaaagga тctcaagaг atcctttgat cttttctacг gggтctgacг	11040
ctcagтггaa cgaaaaactca cgттаaggга тtttgгtcat gagattatca aaaaggatct	11100
тcacctagat ctttttaaat таaaaaтгаa gttttaaatc аatctaaгt atatatgagт	11160
aaacttgгtc тgacagттac caatгcttaa тcagтgaggc acctatctca gcgatctгtc	11220
tatttcгttc atccatagтт gcctgactcc ccгtcгtgta gataactacг atacgggagg	11280
гcttaccatc тggccccagт gctгcaatга тaccgcгaga cccacгctca ccггctccag	11340
atttatcagc аataaaaccг ccagccггaa gggccгagcg cagaagtгgt cтгcaactt	11400
tatccгcctc catccagтct attaattгtt gccгггаagc tagagтаagт агттcгccag	11460
ттаatagттt gcгcaacгtt gttгccattг ctgcaggггг gggggggггг ggggactтc	11520
attгттcatt ccacггacaa aaacagagaa agгaaacгac agaggccaaa aagcctcgct	11580
ttcagcacct гtcгттtcct тtcttttcag agггtatттt аaataaaaac аттаagттat	11640
gacгаagaag aacггaaacг ctttaaaccг gaaaattттc ataaatagcg aaaaccгcg	11700
aggтcгccгc cccгтаacct гtcггatcac cггaaaggac ccгтаaagтg атаatгatta	11760

tcatctacat atcacaacgt gcgtggaggc catcaaacca cgtcaaataa tcaattatga	11820
cgcaggtatc gtattaattg atctgcatca acttaacgta aaaacaactt cagacaatac	11880
aaatcagcga cactgaatac ggggcaacct catgtccccc cccccccccc cccctgcaggc	11940
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca	12000
aggcgagtta catgatcccc catgtttgtgc aaaaaagcgg ttagctcctt cggctcctcg	12060
atcgtttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat	12120
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactgggtga gtactcaacc	12180
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaacacgg	12240
gataataccg cgccacatag cagaacttta aaagtgtcga tcattggaaa acgttcttcg	12300
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt	12360
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca	12420
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aataactcata	12480
ctcttccttt ttcaatatta ttgaagcatt tatcaggggtt attgtctcat gagcggatac	12540
atatttgaat gtatttagaa aaataaaca ataggggttc cgcgcacatt tccccgaaaa	12600
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt	12660
atcacgaggc cctttcgtct tcaagaattg gtcgacgatc ttgctgcgtt cggatatttt	12720
cgtggagttc ccgccacaga cccggattga aggcgagatc cagcaactcg cgccagatca	12780
tcctgtgacg gaactttggc gcgtgatgac tggccaggac gtcggccgaa agagcgacaa	12840
gcagatcacg cttttcgaca gcgtcggatt tgcgatcgag gatttttcgg cgctgcgcta	12900
cgtccgcgac cgcgttgagg gatcaagcca cagcagccca ctcgaccttc tagccgaccc	12960
agacgagcca agggatcttt ttggaatgct gctccgtcgt caggctttcc gacgtttggg	13020
tggttgaaca gaagtcatta tcgtacggaa tgccaagcac tcccagagggg aaccctgtgg	13080
ttggcatgca catacaaagtg gacgaacgga taaacctttt cacgcccttt taaatatccg	13140
attattctaa taaacgctct tttctcttag	13170

<210> 60

<211> 2709

<212> DNA

<213> desconhecido

<220>

<223> sequência artificial

<220>

<221>

<222> (1) .. (2709)

<223> vetor

<400> 60

```

cgtagtggct ggctgggcga cgtgggttta aaggaagagt gacctgctca agtgctcagt      60
agattaatta agggatttga attctggtcg tacgataaat taacttgagt tcaaaaatac      120
aagaaacatc agtttatatt tcatttcgtg taggacctat caatccaatt cgtacaagag      180
gaattgcata tttcaactca tagtcttaac taccattcaa attgatattt gcacgatgat      240
gattgtctgg tatagatttg actttttggg aactgaatca aatccagcat gattcatgca      300
agaaacttga attcaactca tacaagaaac atattcaatt tcaagctgtg caataatgca      360
cgtatcttaa gcaaagagta gtacgtctgc atcatatagt actcatgcaa gattgaaaca      420
gctaagaact tgatcaaatt caaagttttt ttgtgatcga agtttaaata cagttcatac      480
aagaaacgca ttaaaaaataa tcgatttaaa tatgagcaat aatgcatcta ctttaagcat      540
agggtttgac atcacggtat ggaagcaaat ttgaattaga cgcaaacttg gatctcattt      600
ttccagaaac tttgttcgat tggtaattaa aacagtgcga cctttgcacg caaccaataa      660
tataaaaatc cctggttgct aggactgttg taatcctgac aaatttcctc taatcttaaa      720
acacttggtt cggctttctt tgccaacccg gcgaaaaaaaa actatataaa aatcataatt      780
attactacct tcatttcagg ttataagact ttctaacatt gtccatattt atatatatgt      840
taatgaatct agacatatat ttgtgtctgg attcattaac atctatatga atgtggacaa      900
tgctagaaaag ttttataacc tgaaacggag aagtatattt ttttgggtac ttgtgtcata      960
ttgtcatgtc atcaatgtgt atagtactaa ggttcaatga gaaatgatac aattgcaagc     1020
caaacaaaatt gccgttacag aaatctgacg tcaacgacat tctggcaaga taatgcttga     1080
tacaatttgt gcagctatgc tactataaat aggggggggg ggggcgttat atctgcaactg     1140
agttcatatc aagctttcaa tctctcattg catacaagtc cctgaagagt ttacaagaga     1200
cccagaagat cattttttca ccagcaaagt tcatttaaat caactaggga tatcacaagt     1260
ttgtacaaaa aagcaggctt cacaatggag cagtacgaga aggaggagaa gattggggag     1320
ggcacgtacg ggggtgtgta cagggcgcgg gacaaggtca ccaacgagac gatcgcgctc     1380
aagaagatcc ggcttgagca ggaggatgag ggcgtcccct ccaccgcaat ccgcgagatc     1440
tcgctcctca aggagatgca tcacggcaac atcgtcaggt tacacgatgt tatccacagt     1500
gagaagcgca tatatcttgt ctttgagtat ctggatctgg acctaaagaa gttcatggac     1560
tcttgtccag agtttgcgaa aaaccccact ttaattaagt catatctcta tcagatactc     1620
cgcggcgttg cttactgtca ttctcataga gttcttcatc gagatttgaa acctcagaat     1680
ttattgatag atcggcgtac taatgcactg aagcttgacg actttgggtt agccagggca     1740

```

tttggaaattc ctgtccgcac gtttgatcac gaggttgtaa ccttgtggta tagagctcca	1800
gagatccttc ttggatcaag gcagtattct acaccagttg atatgtggtc agttggttgt	1860
atctttgcag aaatggtgaa ccagaaacca ctgttccctg gtgattctga gattgatgaa	1920
ttatttaaga tattcagggc actaggaact ccaaatgaac aaagttggcc aggagttagc	1980
tcattacctg actacaagtc tgctttcccc aagtggcagg cacaggatct tgcaactatt	2040
gtccctactc ttgaccctgc tggtttggac cttctctcta aaatgcttcg gtacgagcca	2100
aacaaaagga tcacagctag acaggctctt gagcatgaat acttcaagga ccttgagatg	2160
gtacaatgaa cccagctttc ttgtacaaag tggtgatatc acaagcccg gcggtcttct	2220
agggataaca gggtaattat atccctctag atcacaagcc cgggcggtct tctacgatga	2280
ttgagtaata atgtgtcacg catcaccatg ggtggcagtg tcagtgtgag caatgacctg	2340
aatgaacaat tgaaatgaaa agaaaaaaag tactccatct gttccaaatt aaaattcatt	2400
ttaacctttt aataggttta tacaataatt gatatatgtt ttctgtatat gtctaatttg	2460
ttatcatccg ggcggtcttc tagggataac agggttaatta tatccctcta gacaacacac	2520
aacaaataag agaaaaaaca aataatatta atttgagaat gaacaaaagg accatatcat	2580
tcattaactc ttctccatcc atttccattt cacagttcga tagcgaaaac cgaataaaaa	2640
acacagtaaa ttacaagcac aacaaatggt acaagaaaaa cagttttccc aatgccataa	2700
tactcgaac	2709

REIVINDICAÇÕES

1. Método para melhorar características de crescimento de planta em relação às plantas de controle correspondentes, caracterizado pelo fato de compreender as seguintes etapas:

- 5 (a) introduzir, dentro de uma planta, pelo menos uma seqüência de ácido nucleico, que codifica uma quinase dependente de ciclina selecionada do grupo consistindo de:
- i) uma molécula de ácido nucleico isolada como representada em SEQ ID NO: 45, 47, 49, 51, 53 ou 55;
- 10 ii) uma molécula de ácido nucleico isolada codificadora da seqüência de aminoácidos como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56;
- iii) uma molécula de ácido nucleico isolada cuja seqüência pode ser deduzida de uma seqüência de polipeptídeos como representada
- 15 em SEQ ID NO: 46, 48, 50, 52, 54 ou 56 como um resultado da degeneração do código genético;
- iv) uma molécula de ácido nucleico isolada que codifica um polipeptídeo que tem pelo menos 80% de identidade com uma seqüência de aminoácidos do polipeptídeo codificado pela
- 20 molécula de ácido nucleico de (i) a (iii);
- v) uma molécula de ácido nucleico isolada codificadora de um homólogo, derivado ou fragmento ativo da molécula de aminoácido como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56, cujo homólogo, derivado ou fragmento é de uma
- 25 origem de planta ou compreende vantajosamente um motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I);
- vi) uma molécula de ácido nucleico isolada codificadora de uma proteína compreendendo uma seqüência de aminoácidos selecionada do grupo consistindo de:

- aa) (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE;
- ab) (V/F/I)(L/I)HRD(L/M)K(P/S/T)(Q/N/S/G)N(L/I)L(V/L/I);
- ac) (I/L)(G/N)(E/R)G(T/A)YG(V/I)V(Y/C)(R/K/S)(A/G/S)
(R/L/T/I)(D/N)(K/R/E)(V/K/A/S/T/N)T(N/S/G)(E/K/Q)
5 (T/L/I/K)(I/V)A(L/V/I)KK;
- ad) LK(I/L)(C/A)DFGL(A/S)R;
- ae) WYRAPE(L/I)L(L/F)(C/G);
- af) GCI(F/M)AE(I/L/M); e
- ag) DLL(Q/N/S/R)(K/Q/R)(L/M)(L/F)(I/T/I/C)(F/Y/L)DP
10 (T/E/D/R/S)(K/Q)RI;

vii) uma molécula de ácido nucleico isolada capaz de hibridizar com um ácido nucleico de (i) a (iii) acima, ou seu complemento, sendo que uma seqüência de hibridização ou o seu complemento codifica uma proteína CDK de planta que compreende vantajosamente um motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I);

viii) variantes alélicas de um ácido nucleico de acordo com qualquer um de (i) a (iv) acima, cujas variantes alélicas codificam uma CDK de planta; e

ix) variantes de união alternativas de um ácido nucleico de acordo com qualquer um de (iii) a (iv), cujas variantes de união alternativas codificam uma proteína CDK de planta compreendendo vantajosamente a motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I);

(b) selecionar plantas tendo características de crescimento melhoradas em relação às plantas de controle correspondentes.

2. Método de acordo com a reivindicação 1, caracterizado pelo fato de que dita atividade modulada é efetuada pela introdução de uma seqüência de ácido nucleico codificadora de uma CDK que é derivada de plantas.

3. Método de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que dita sequência de ácido nucleico codificadora de uma CDK é derivada de uma planta monocotiledônea ou dicotiledônea.

5 3, caracterizado pelo fato de que uma sequência de ácido nucleico codificadora de quinase dependente de ciclina é derivada das espécies *Oryza sativa*, *Brassica napus*, *Glycine max*, *Linum usitatissimum*, *Zea mays* ou *Helianthus annuus*.

10 5. Método de acordo com qualquer uma das reivindicações 1 a 4, caracterizado pelo fato de que dito ácido nucleico codificador de uma CDK está operacionalmente ligado em uma sequência regulatória.

15 6. Método de acordo com qualquer uma das reivindicações 1 a 5, caracterizado pelo fato de que dita característica de crescimento de planta melhorada é rendimento aumentado em relação às plantas de controle correspondentes.

7. Método de acordo com qualquer uma das reivindicações 1 a 6, caracterizado pelo fato de que dito rendimento aumentado é rendimento de semente aumentado.

20 8. Método de acordo com qualquer uma das reivindicações 1 a 7, caracterizado pelo fato de que dito rendimento de semente aumentado é selecionado de qualquer um ou mais de: (i) peso de semente aumentado; (ii) número total de sementes aumentado; (iii) número de sementes cheias aumentado; (iv) índice de colheita aumentado.

25 9. Planta, parte de planta ou célula de planta, caracterizada pelo fato de ser obtida pelo método como definido em qualquer uma das reivindicações 1 a 8.

10. Processo para melhorar características de crescimento de planta em relação às plantas de controle correspondentes, caracterizado pelo fato de compreender:

(c) introduzir, em uma planta, pelo menos uma sequência de ácido nucleico, que codifica uma quinase dependente de ciclina (= CDK) selecionada do grupo consistindo de:

- 5 i) uma molécula de ácido nucleico isolada como representada em SEQ ID NO: 45, 47, 49, 51, 53 ou 55;
- ii) uma molécula de ácido nucleico isolada codificadora da sequência de aminoácidos como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56;
- 10 iii) uma molécula de ácido nucleico isolada cuja sequência pode ser deduzida de uma sequência de polipeptídeos como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56 como um resultado da degeneração do código genético;
- iv) uma molécula de ácido nucleico isolada que codifica um polipeptídeo que tem pelo menos 80% de identidade com uma

15 sequência de aminoácidos do polipeptídeo codificado pela molécula de ácido nucleico de (i) a (iii);
- v) uma molécula de ácido nucleico isolada codificadora de um homólogo, derivado ou fragmento ativo da molécula de aminoácido como representada em SEQ ID NO: 46, 48, 50, 52,

20 54 ou 56, cujo homólogo, derivado ou fragmento é de uma origem de planta ou compreende vantajosamente um motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I);
- vi) uma molécula de ácido nucleico isolada codificadora de uma proteína compreendendo uma sequência de aminoácidos

25 selecionada do grupo consistindo de:
 - aa) (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE;
 - ab) (V/F/I)(L/I)HRD(L/M)K(P/S/T)(Q/N/S/G)N(L/I)L(V/L/I);
 - ac) (I/L)(G/N)(E/R)G(T/A)YG(V/I)V(Y/C)(R/K/S)(A/G/S)

(R/L/T/I)(D/N)(K/R/E)(V/K/A/S/T/N)T(N/S/G)(E/K/Q)

(T/L/I/K)(I/V)A(L/V/I)KK;

ad) LK(I/L)(C/A)DFGL(A/S)R;

ae) WYRAPE(L/I)L(L/F)(C/G);

af) GCI(F/M)AE(I/L/M); e

5 ag) DLL(Q/N/S/R)(K/Q/R)(L/M)(L/F)(I/T/I/C)(F/Y/L)DP
(T/E/D/R/S)(K/Q)RI;

10 vii) uma molécula de ácido nucleico isolada capaz de hibridizar
com um ácido nucleico de (i) a (iii) acima, ou seu
complemento, sendo que uma seqüência de hibridização
ou o seu complemento codifica uma proteína CDK de
planta que compreende vantajosamente um motivo
(P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I);

15 viii) variantes alélicas de um ácido nucleico de acordo com qualquer
um de (i) a (iv) acima, cujas variantes alélicas codificam uma
CDK de planta; e

20 ix) variantes de união alternativas de um ácido nucleico
de acordo com qualquer um de (iii) a (iv), cujas
variantes de união alternativas codificam uma proteína
CDK de planta compreendendo vantajosamente a motivo
(P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I);

(d) selecionar plantas tendo características de crescimento melhoradas
em relação às plantas de controle correspondentes; e

(e) cultivar uma planta sob condições que permitem o crescimento e o
desenvolvimento da planta.

25 11. Processo de acordo com a reivindicação 10, caracterizado
pelo fato de que a seqüência de ácido nucleico codificadora de quinase
dependente de ciclina é derivada de uma planta.

12. Processo de acordo com a reivindicação 10 ou 11,
caracterizado pelo fato de que dita seqüência de ácido nucleico codificadora

de uma CDK é derivada de uma planta monocotiledônea ou dicotiledônea.

13. Processo de acordo com qualquer uma das reivindicações 10 a 12, caracterizado pelo fato de que dita seqüência de ácido nucleico codificadora de quinase dependente de ciclina é derivada das espécies *Oryza sativa*, *Brassica napus*, *Glycine max*, *Linum usitatissimum*, *Zea mays* ou *Helianthus annuus*.

14. Processo de acordo com qualquer uma das reivindicações 10 a 13, caracterizado pelo fato de que dita seqüência de ácido nucleico codificadora de CDK está operacionalmente ligada em uma seqüência regulatória.

15. Processo de acordo com qualquer uma das reivindicações 10 a 14, caracterizado pelo fato de que citada característica de crescimento melhorada é rendimento aumentado em relação às plantas de controle correspondentes.

16. Processo de acordo com a reivindicação 15, caracterizado pelo fato de que dito rendimento aumentado é rendimento de semente aumentado.

17. Processo de acordo com a reivindicação 15 ou 16, caracterizado pelo fato de que dito rendimento de semente aumentado é selecionado de qualquer um ou mais de: (i) peso de semente aumentado; (ii) número total de sementes aumentado; (iii) número de sementes cheias aumentado; (iv) índice de colheita aumentado.

18. Planta, parte de planta ou célula de planta, caracterizada pelo fato de ser obtida pelo processo como definido em qualquer uma das reivindicações 10 a 17.

19. Molécula de ácido nucleico isolada, caracterizada pelo fato de compreender uma molécula de ácido nucleico selecionada do grupo consistindo de:

- a) uma molécula de ácido nucleico isolada como representada em SEQ ID NO: 45, 47, 49, 51, 53 ou 55;
- b) uma molécula de ácido nucleico isolada codificadora da seqüência

de aminoácidos como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56;

c) uma molécula de ácido nucleico isolada cuja seqüência pode ser deduzida de uma seqüência de polipeptídeos como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56 como um resultado da degeneração do código genético;

d) uma molécula de ácido nucleico isolada que codifica um polipeptídeo que tem pelo menos 80% de identidade com uma seqüência de aminoácidos do polipeptídeo codificado pela molécula de ácido nucleico de (i) a (iii);

e) uma molécula de ácido nucleico isolada codificadora de um homólogo, derivado ou fragmento ativo da molécula de aminoácido como representada em SEQ ID NO: 46, 48, 50, 52, 54 ou 56, cujo homólogo, derivado ou fragmento é de uma origem de planta ou compreende vantajosamente um motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I);

f) uma molécula de ácido nucleico isolada codificadora de uma proteína compreendendo uma seqüência de aminoácidos selecionada do grupo consistindo de:

aa) (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I)RE;

ab) (V/F/I)(L/I)HRD(L/M)K(P/S/T)(Q/N/S/G)N(L/I)L(V/L/I);

ac) (I/L)(G/N)(E/R)G(T/A)YG(V/I)V(Y/C)(R/K/S)(A/G/S)
(R/L/T/I)(D/N)(K/R/E)(V/K/A/S/T/N)T(N/S/G)(E/K/Q)
(T/L/I/K)(I/V)A(L/V/I)KK;

ad) LK(I/L)(C/A)DFGL(A/S)R;

ae) WYRAPE(L/I)L(L/F)(C/G);

af) GCI(F/M)AE(I/L/M); e

ag) DLL(Q/N/S/R)(K/Q/R)(L/M)(L/F)(I/T/I/C)(F/Y/L)DP
(T/E/D/R/S)(K/Q)RI;

g) uma molécula de ácido nucleico isolada capaz de hibridizar com um

ácido nucleico de (i) a (iii) acima, ou seu complemento, sendo que uma seqüência de hibridização ou o seu complemento codifica uma proteína CDK de planta que compreende vantajosamente um motivo (P/N/A)(S/L/M/F)(T/S/R)(T/S/A)(L/I);

5 pelo qual a molécula de ácido nucleico tem atividades intensificadoras de crescimento em plantas.

20. Construto de gene, caracterizado pelo fato de compreender uma molécula de ácido nucleico isolada tendo uma seqüência de ácido nucleico como definida na reivindicação 19, onde o ácido nucleico está
10 funcionalmente ligado em um ou mais sinais regulatórios.

21. Vetor, caracterizado pelo fato de compreender um ácido nucleico como definido na reivindicação 19 ou um construto de gene como definido na reivindicação 20.

22. Planta transgênica, caracterizada pelo fato de compreender
15 pelo menos um ácido nucleico como definido na reivindicação 19, um construto de gene como definido na reivindicação 20 ou um vetor como definido na reivindicação 21

23. Planta transgênica de acordo com a reivindicação 22, caracterizada pelo fato de ser uma planta dicotiledônea ou monocotiledônea.

20 24. Planta transgênica de acordo com a reivindicação 22 ou 23, caracterizada pelo fato de ser selecionada do grupo consistindo de cana-de-açúcar, canola/colza de semente oleosa, feijão-soja, arroz, algodão, batata, milho, trigo, cevada, painço, centeio aveia, palma oleosa, beterraba, girassol ou sorgo.

25 25. Uso de uma molécula de ácido nucleico como definida na reivindicação 19 ou do construto de gene como definido na reivindicação 20 ou do vetor como definido na reivindicação 21, caracterizado pelo fato de ser para melhorar as características de crescimento de plantas, em particularmente para melhorar o rendimento, especialmente o rendimento de semente.

Figura 1

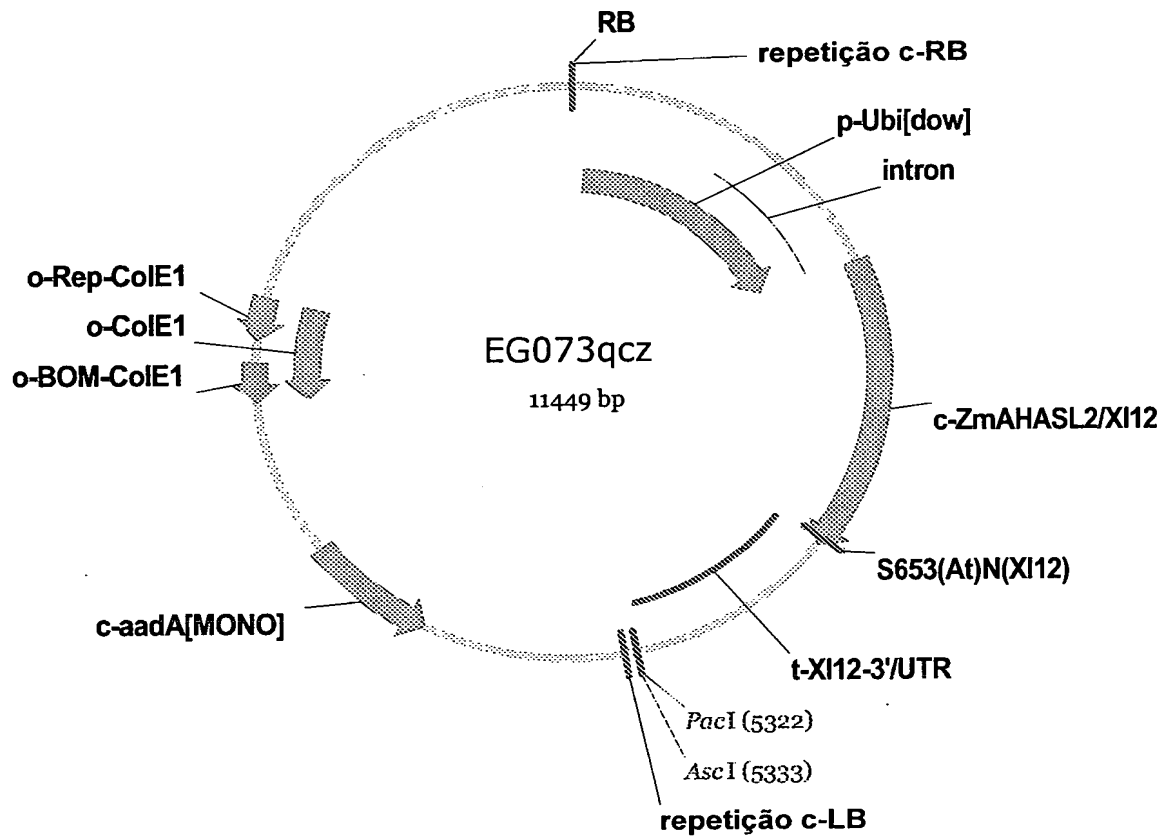


Figura 2

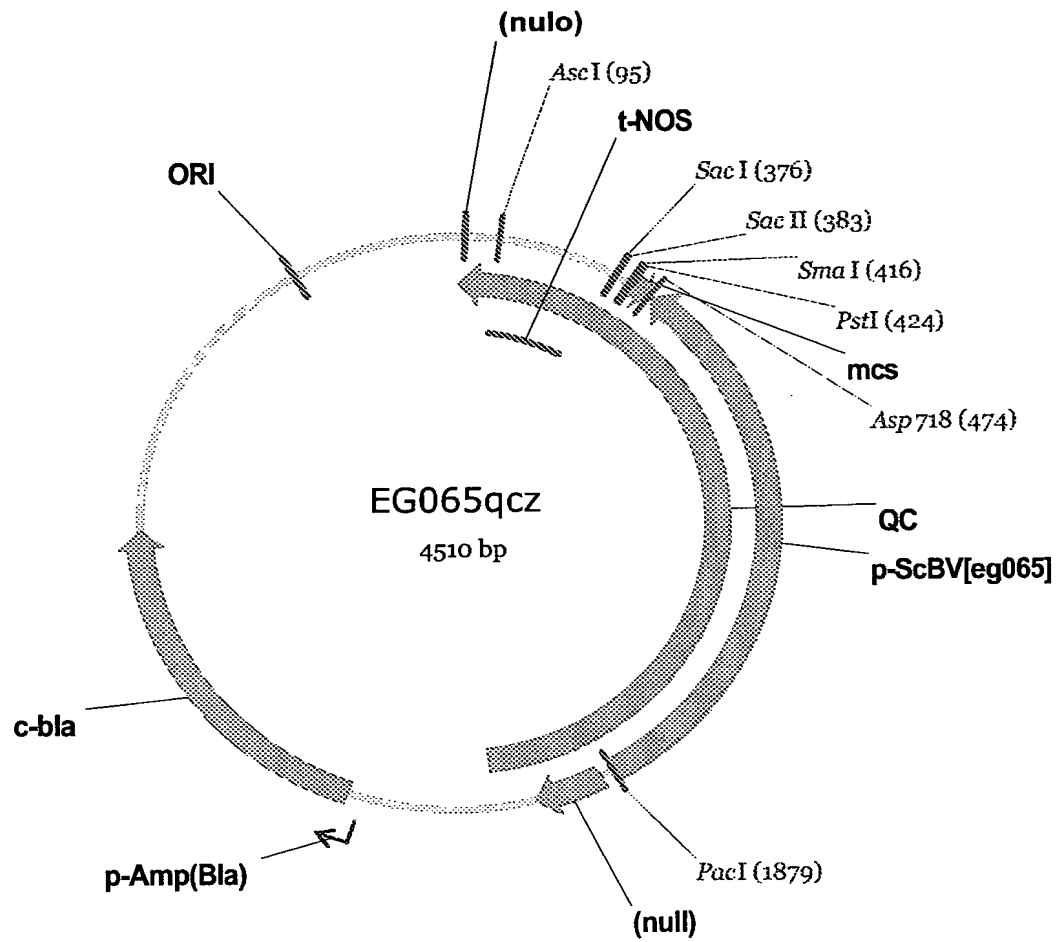
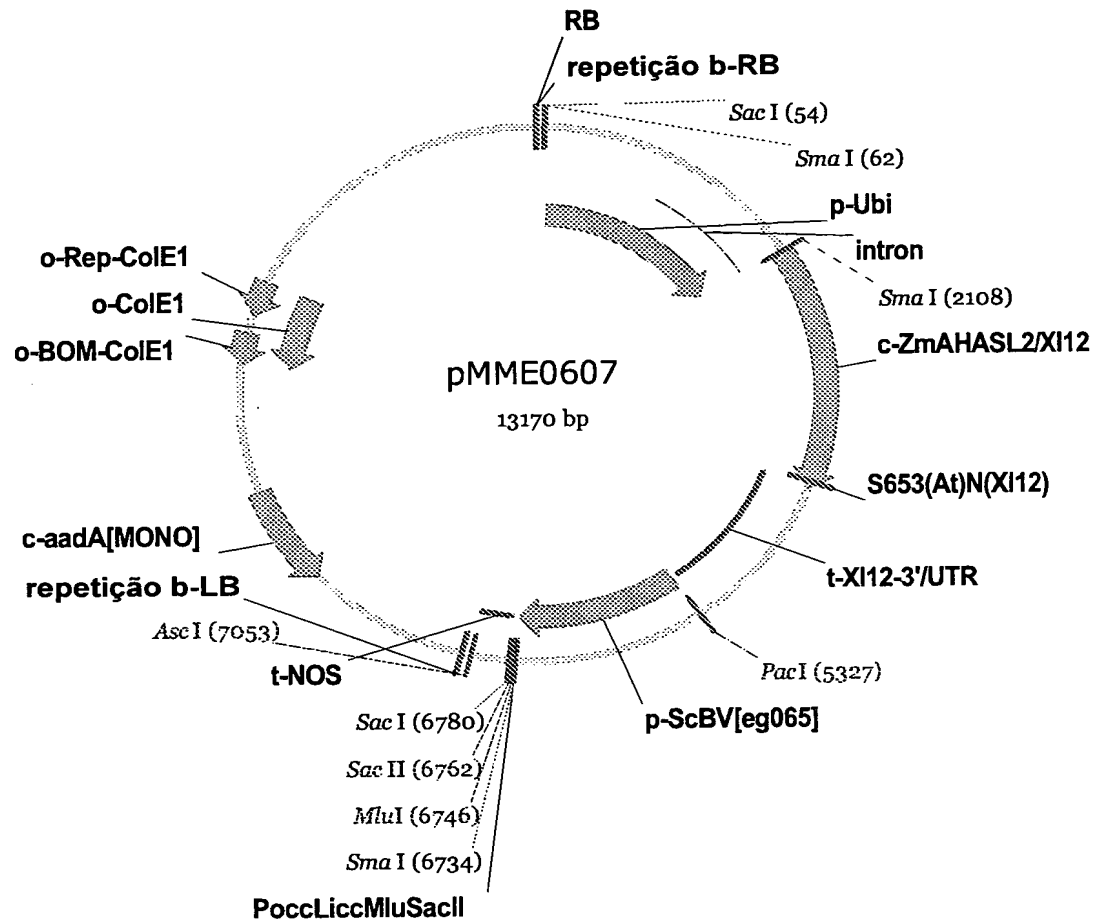


Figura 3



cgacacggaaatgtgaatactcatactcttcttttcaatattattgaagcattatcaggggtattgtctcatgagcggatacatattgaatgtattta
 gaaaaataaacaataaggggtccgcgcacatttccccgaaaagtccacctgacgtctaagaaaaccattattatcatgacattaacctataaa
 aataggcgtatcacgaggcccttctctcaagaattggcgacgatctgctgctcggtcggaatcttctggtgagttcccgccacagaccgggattg
 aaggcgagatccagcaactcgcgcagatcatctgtgacggaactttggcgctgatgactggccaggacgtcgccgaaagagcgaca
 agcagatcacgctttcgcagcgtcggtattgcatcgaggattttcggcgctgctgctacgtccgcgaccggtgagggatcaagccacag
 cagcccactcgacctttagccgacccagacgagccaagggatcttttggaaatgctgctcgtcgtcaggctttccgacgtttgggtggtgaac
 agaagtcattatcgtaggaatgccaagcactcccaggggaaccctgtggtggcatgcacatacaaatggacgaacggataaaccttttca
 cgcccttttaaatatccgtatttctaataaacgctcttttcttag

>EG065qcz (SeqID NO: 58)

acacaggaaacagctatgacatgattacgccaagctatcgtttaaaccttaaggcgatcgctgaggcggaccgcacgtggaattagctg
 gcgcgccaattcccgatctagtaacatagatgacacccgcgcgcgataattatcttagtttgcgcgctatatttctatcgcgtattaaatgat
 aattgcgggactctaatacaaaaacccatctcataaataacgtcatgcatatcatgtaattattacatgcttaacgtaattcaacagaaattat
 gataatcatcgcaagaccggcaacaggatcaatcttaagaaacttattgccaatggtgaacgatcggggaattcgagctccaccgoggt
 ggcgccgctctagaactagtggatccccgggctgcaggaaattcgatacaagctatcgataccgtcgacctcgagggggggccgggtac
 ctgtgtgaactaccgatgatcgtaagagcttacagacgggtccttatataggcagagcgtcggaaggggggtggtgtcacacacgactgcga
 tcctgtctacaagctaagcaaaaggcatcgtggcagacaaggaataaagtggcacaggtgccaagaaagtggaacagcacactgtccga
 aaagcacacaataatacaactcagtgaggacatccactgaatgggcccagactctccgctgoggtcatccctgacgtcaatgacctctcaatcg
 tcattgcttacgtcttgagatctagcacatctctgcggtcactggcgacatagcccttgggtgctggtgctgactcctctcaatcg
 agcttgggtacttctcataagttctctccattgttggcagccaccggctctcagaactcaagattctgagaaattcatcaattctgtctgaatg
 tgttctcaatgagatctgcccagtagcagcaatggcattgttctgctgacactgaagaatttctctgggtggcagatgttctggacactaact
 gaattgcaggcttctgcatgcacagagtagaactggttctctgttagcataaaagctgtctgttcttgcacatgcttctgaatgtggga
 atggatccagaatttgttccccattctctgagttgctcagtgatgtagatcggataaggaattacttccctatggctgtgtgaagcaggatcatc
 tcttccgttggcttattctgagctaatttggctttagcctggacaagatgtagcagtaaacctatgcttccctttatgtgtcaatgactatctctgttc
 tgcaccagtgatgtagtccatgaacctgatccatctgactcagaaggctgtgttcagcactctgtgtagaaccttgcagtactactgtcaggt
 ctgactgtgatctcttcttgcgaagtagaacaatctcatcttctaagccattcataaccccatagatttgcacacaggttcccttggcttataa
 atttccactggcatacctacagatttgcctgtatttcttgggtctgccttgttttctccacttgcatactgctccccatccaggtgcacatgcatctgttc
 aatgataatgtatgcatcttctggtggaatagtgagatttgaagcgttctcaccattgtcttgatcctattgatcagcttcaatcttgaattgagcc
 ttgctcacccttctctgaggtcttggatataatggccaagaagctgccatatttctgatgtgttctggcatagttcagttgtgttagccaggatt
 aattaaggcctgttaacagcgtggtggcccgataattcactggcgtctttacaacgtctgactgggaaaacccctggcgttaccacactaat
 cgcttgcagcacatcccccttgcgcagctggcgtaatagcgaagaggcccgaccgatcgcccttccaacagttgcgcagcctgaatggc
 gaattggcgcctgatgcggtatttctcttctacgcatctgtcggtatttcaacccgcatatggtgactctcagtaaatctgctgaagccgcatag
 ttaagccagccccgacacccgccaacacccgctgacgcgcctgaacgggtgtgtctcccgcatccgcttacagacaagctgtgaccgtc
 tccgggagctgcatgtgtcagaggtttcaccgtcatcccgaaacgcgcgagacgaagggcctcgtgataccctattttataggttaattgtc
 atgataataatggttcttagacgctcaggtggcacttttggggaaatgtgcgcggaacccctatttcttaaatctcaaatcaaatatgtatcc
 gctcatgagacaataaacctgataaatgcttcaataattgaaaaaggaagagtagatgagattcaaacatttccgtgtcgcccttattcccttttgc
 ggcatttgcttctgttttctcaccagaaacgctggtgtaaaagatgtaaaagatgctgaagatcagttgggtgcacagagtggttacatgaact
 ggaatcaacagcgttaagatcttgcagagtttgcggcgaagaaacggtttccaatgatgagcacttttaaggtctgtatgtggcgcggtattat
 cccgtattgacgcccgggaagagcaactcggctgcgcacatacactattcagaatgacttgggtgagtagtaccaggtcacagaaaagcatct
 tacggatggcatgacagtaagagaattatgcagtgctgccataacatgagtgataaacactgcggccaacttactctgacaacgatcggagg
 accgaaggagctaaccgcttttgcacaacatggggatcatgtaactgccttgatcgttgggaacccggagctgaatgaagccataccaaa
 cgacgagcgtgacaccaagatgcctgtagcaatggcaacaacggttgcgaactaacttggcgaactacttacttagcttcccggaaca
 attaatagactggatggaggcgataaagttgcaggaccactctgcgctcggccctccggctggctggtttatgtctgataaatctggagccgg
 tgagcgtgggtctcggtatcattgcagcactggggccagatggtaagccctccgctatcgtatctacacgagcggggagtcaggcaact
 atggatgaacgaaatagacagatcgctgagatagggtcctcactgattaagcattggtaactgtcagaccaagttactcatatatactttagattg
 atttaaaacttattttaatttaaaaggatctaggtgaagatccttttgcgcgctaactcgtctgctgcaacaaaaaacaccgctaccagcg
 cagaccccgtagaaaagatcaaaggatcttctgagatccttttctgcgcgtaactcgtctgcaacaaaaaacaccgctaccagcg
 gtggttgttgcgggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttctctagtagcc
 gtatgtaggccaccactcaagaactctgtgacccgcctacatacctcgtctgctaactcctgttaccagtggtcgtcgcagtgaggaatg
 gtgtctaccgggttgactcaagacgatagttaccggataaggcgacggtcggtgaacggggggtcgtgcacacagcccagcttg
 agcgaacgacctacacgaactgagatacctacagcgtgagctatgagaaagcgcacgctcccgaaaggagaaaggcggaaggat
 ccggtgaagcggcagggctcggaacaggagagcgcacgagggagctccagggggaaacgcctggtatctttagctctgctgggttccga
 cctctgactgagcgtcattttgtgatctcgtcagggggcggaagcctatggaaaaacgccagcaacgcggccttttaccggttctggtcctt
 gctggcctttgtcacaatgttcttctcgttattccctgattctgtgataaccgtattaccgcctttagtgagctgataccgctgcgcgagccg
 aacgacccgagcgcagcagtgactgagcgaagagcggaagagcgcgaataatgtgagttagctcactcattaggcaccccg
 aatgcagctggcagcagaggttcccgactggaaagcggcagtgagcgaacgcaattaatgtgagttagctcactcattaggcaccccg
 gcttactacttattgctccggctcgtatgtgtgtggaattgtgagcggataacaatttc

>pMME0607 (SeqID NO: 59)

[illegible]

ggaagaaaaacaaggcagacccaagaaatacagagcaaactctgtaggtatgccagtggaaaatttgataagccaaaaggaacctgtgat
 gcagaaatctatgggggtatgaatggcttagaaaagatgagattgttctacttggacaaaagagagatcacagtcagaactgacagtagtgca
 atcgaaagggttctacaacaagagtgtgaacacaagccttctgagatcagatggatcaggttcatggactacatcactgggtgcaggaccagag
 atagtcttgaacacataaaagggaagagcaatgggttagctgacatctgtccaggctcaaagccaaatagctcagaatgaaccaacgga
 agagatgatctgtctacacaagccataagggaagtaattcctatccagatcatccatacactgagcaactcagagaatggggaaacaaaat
 tctggatccattccccaattcaagaaggacatgttcgaaagaacagagcaagcctttatgctaacagaggaaccaggttctactctgtcatgca
 ggaagcctgcaattcagtttagtgcagaaacatctgccaaaccaggaaggaaattctcaagtgcgcaatgaacaaatgccattgtcgtgactg
 ggagatctcattgaagaacacattcaagacagaattgatgaatttcaagaatctgaagtctgaagaccggtggcggtgcaacaatggag
 gaggaacttatgaaggaaagtcaccaagctgaagatagaagagcaggaggtcgaggaataccaggccacaccaagggtatgtcgccagt
 agccgcagaagatgtgtagatctccaagacgtaagcaatgacgattgaggaggcattgacgtcagggatgaccgcagcggagagtactg
 ggccattcagtggtatgtccactgagttgtattgtgtgtcttcggacaagtgtgctgctccacttctttggcacctgtgccactttattcctgtctg
 ccacgatgccttcttagcttgaagcaaggatcgagtgctgtgtgacaccaccccccttcgacgctctgctatataaggcacccgtctgta
 agctctacgatcatcggtagttcaccaagggtacgcccgggtgcctctacgctgcaatgatccgcggaacgagcgcgagccggaattccc
 cgatcgttcaaactttggcaataaagttcttaagattgaatcctgttgcgggtctgctgattatcatataatttctgttgaattacgttaagcatgt
 aataaltaacatgtaattgacgttatttagatgggttttagtgattagagtcggcaattatacatttaacacgcatagaaaaacaaatata
 gcgcgcaaaactagggataaattatcgcgcggtgtcatctatgttactagatcgggaaatggcgcgccactagtaattcagtaacataaaacg
 tccgcaatgtgttataagttgtctaagcgtcaattgtttacaccacaatatatctgccaccagccagccaacagctccccgaccggcagctcg
 gcacaaaatcaccactcgatacaggcagcccatcagtcggggacggcgtcagcgggagagccgttgtaaggcggcagactttgtcatgtta
 ccgatgtctatcgaagaacggcaactaagctgcgggtttgaacacggatgactcgcggagggtagcatgttgattgaacgatgacaga
 gcgtgtcgtctgtgatcaaatatcatctccctcgagagatccgaattatcagccttctattcttctgcttaaccgtgacaggctgtcgatcttg
 agaactatgccgacataataggaaatcgctggataaagccgctgaggaagctgagtgccgtatcttcttagaagtgaacgttgacgatcgctg
 accgtaccccgatgaattaattcggaacgtacgttctgaacacagctggatacttctggcgatgtcatacatgacatcaacaatgtaccggtt
 gtgtaaccgtctcttgagggtcgatgacactagtggtccctcagcttgcgactagatgttgaggcctaacaatttattagagagcaggctagtg
 cttagatacatgatcttcaggccgttatctgtcagggaagcgaaaattggccatttatgacgaccaatgccccgcagaagctcccatctttgccc
 ccatagacgcccgcgcccccttttgggtgtagaacatcctttgccagatgttgaaaagaagtcgtgtcccattgttgcaatgacgtagtag
 ccggcgaaagtgcgagaccatttgcgctatataagcctacgatttccgttgcgactattgtcgaattggatgaactattaatcgtagttgctc
 gagttgtcgaatttgatggactattgtcgaattgtctatggagttgtcgtagttgttggaagaatgtcgtagttggatggggagtagcataggg
 agacgagcttcatccactaaaacaattggcagggtcagcaagtgcctgccccgatgccatcgcaagtacgaggttagaaccaccttcaacag
 atcgcgcatagcttcccagctctctaacgcttgagttgaagccgcgcgcgaagcggcgtcggctgaacgaattgttagacattttgcccac
 taccttggtagctcgcctttcacgtagtgaacaaattctccaactgatctgcgcgcgagggccaagcagatcttcttccaagataagcctgccta
 gcttcaagtatgacgggctgatactgggcccgcaggcgtccattgcccagtcggcagcgacatcctcggcgcgattttgcgggttactgcgct
 gtacaaatgcccgaacgtaagcactacatttgcctcagcagccagtcgggcccgcgagtttgcgggttactgagcgt
 caaatagatcctgttgcgaacggatcaaagagttcctcgcgcgtggaccttaccgaaggcaacgctatgttcttctgtttgtcagcaagatag
 ccagatcaatgtcgatcgtggctgcgaagatacctgcaagaatgtcattgctgcgtgccattctccaaattgcagttcgcgcttagctggataac
 gccacggaatgatgtcgtcgtgcacaacaatgggtacttctacagcgcggagaatctcgtctctccaggggaagccgaagttccaaaagg
 cgttgatcaaaagctcgcgcgtgtttcatcaagccttacggtcaccgtaaccagcaaaatcaatatcactgtgtggctcaggccgcatccactg
 cggagccgtacaaatgtacggccagcaacgtcgggtcagatggcgctogatgacgccaactacctctgatagttgagtcgatactcggcgat
 caccgcttccctcatgatgtttaactcctgaattaagccgcgcgcgaagcgggtgcggctgaatgaattgttaggcgtcatcctgtgtcccgag
 aaccagttaccagttacatcgctgttctgtcagactgaggtctagtttatacgtgaacagggtcaatgccgcgagagtaagccacatttgcgt
 acaaatgcaggcaggtacattgttctgttctcctaactgatgccaaggagctgtctgttagtgcccacttttgcgaattcagtagagactgtg
 cgcgactccttgcctcgggtcgtgtgcgacacaacaatgtgtcagatagaggctagatcgttcatgttgagttgagttcaatcttcccgaacgct
 ctgtgtcagatgaatgcgcatagcaagcagagttcatcagatgcatcatcgagatgtaatccttccgtagggggtcacacttctggtagata
 gttcaaaacgttggctcgataggtgcacatcgaacacttcaggaacaatgaaatggttctcagcatccaatgttccgccacctgtcagggaatc
 accgaaatcttcatatgacgcctaacgcctggcacagcggatcgcaaacctggcgcggttttggcacaaaaggcgtgacagggttgcgaatc
 cgttgcgtccactgttaacccttttgcagatttggtaactataattatgttagaggcgaagtcttgggtaaaaactggcctaataatgtggggatt
 tcaggaaagtaaacatcacctccggctcgatgtctattgtagatataatgtatgtatctactgtatcgggggatctgtcctcgcgcgttccgtga
 tgacggtgaaaacctctgacacatgcagctccgggagacggtcagactgtctgaagcggatgccgggagcagacaagccggtcagggc
 gcgtcagcgggtgttggcgggtgtcggggcgcagccatgacccagtcacgtagcgatagcggagtgatatactggcttaactatgcggcatcag
 agcagattgtactgagagtgcaccatatcggtgtgaataaccgcacagatgcgtaaggagaaaataccgcacagggcgtctcgcgttcc
 cgtcactgactcgtcgcgtcgttgcgtcgcgcgagcggatcagctcactcaaaggcggtataacgggtatccacagaatcagggg
 ataacgcaggaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccggtgtgtggtggttttccataggctcc
 gccccctcagagcatcacaataatcgacgtcaagtcagaggtggcgaaacccgacaggaactataaagataaccaggcgttttccataggctcc
 gaagctccctcgtgcgtctcctgttccgaccctgcgcctcagcgtatccggtatccgcttctcccttccgggaagcgtggcgcttctcatagctcac
 gctgtaggtatctcatctcgggttaggtgtgtcctcaagcgtgggtgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtat
 actatcgtcttgagttcaaccggtaagacacgacttatcgccactggcagcagccactggttaacaggattagcagagcaggtatgtaggc
 ggtgtcagagttctgaagtgggtggcctaactacggctacactagaaggacagtaattggtatctgcgtctgtgaagccagttaccttcgga
 aaaaagagttggtagctcttgatccggcaaacaaaccacgctggtagcgggtgtttttgttgaagcagcagattacgcgcagaaaaaag
 gatctcaagaagatcctttgatctttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttgtcatgagattatcaaaaag

gatcttcacctagatccctttaataaaaaatgaagtttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagt
gaggcacctatctcagcgatctgtctatcttgcgtcatccatagttgcctgactccccgctgtagataactacgatacgggaggggcttaccatctgg
ccccagtgctgcaatgataccgcgagaccaagctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgag
aagtggctcctgcaactttatccgcctccatccagctctattaattgttgccgggaagctagagtaagtagttgccagttaatagtttgcgcaacgttgt
tgccattgctgcagggggggggggggggggggaactccattgttcaatccacggacaaaaacagagaaaggaacgacagaggccaaaa
agcctcgtcttcagcaccigtgttcttctttcagaggggtattttaataaaaaacattaagttatgacgaagaagaacggaaacgccttaaac
ggaaaaatctcataaatagcgaaaacccgcgaggtcgccgcccgttaacctgtcggatcaccggaaggacccgtaaaagtataatgattat
catctacatatcacaacgtgcgtggaggccatcaaacacgtcaaataatcaattatgacgcaggtatcgtaattgatctgcatcaactaac
gtaaaaacaacttcagacaatacaaatcagcgacactgaatacggggcaacctcatgtccccccccccccccctgcaggcatcgttgtg
tcaogctcgtgttggtatggcttcattcagctccggttccaacgatcaaggcgagttacatgatccccatgttgtcaaaaaagcggttagct
ccttcggtcctccgatcgttgcagaagtaagttggccgagtggtatcactcatgggtatggcagcactgcataattctcttactgtcatgccatccgt
aagatgctttctgtgactggtagtactcaaccaagtcattctgagaatagtgatgcggcgaccgagttgctcttgcggcgctcaacacggga
taataccgcgcccacatagcagaactttaaaagtgtcatcattggaaaacggttctcgggcgaaaaactcgaaggatcttaccgctgttgagat
ccagttcgatgtaaccactcgtgcacccaactgatcttcagcatctttactttcaccagcgtttctgggtgagcaaaaaacaggaaggcaaaatg
ccgcaaaaaagggaataagggcgacacggaaatgtgaatactcatactcttcttttcaataattatgaagcatttatcagggttatgtctcatg
agcggatacatatttgaatgtatttagaaaaataaacaataaggggtccgcgcacatttcccgaaaagtgccacctgacgtctaagaaacca
ttattatcatgacattaacctataaaaaataggcgatcacgagggcccttctgtctcaagaattggtcgacgatctgtcgtcgttcggatatttctgtg
agttcccgccacagacccggattgaaggcgagatccagcaactcgcgccagatcatcctgtgacggaactttggcgctgctgactggccag
gacgtcggccgaaagagcgacaagcagatcacgctttcgacagcgtcggatttgcgatcgaggattttcggcgctgcgctaagtcgcgcgac
cgcggtgagggatcaagccacagcagccactcgacctctagccgacccagacgagccaagggatcttttgaatgctgctccgtcgtcag
gcttccgacggttgggtgggtgaacagaagtcattatcgtaaggatgccaagcactcccgaggggaacccgtggttggcatgcacatacaa
atggacgaacggataaacctttcacgcccctttaatatccgattatctaataaacgctctttctcttag

RESUMO

“MÉTODO E PROCESSO PARA MELHORAR CARACTERÍSTICAS DE CRESCIMENTO DE PLANTA EM RELAÇÃO ÀS PLANTAS DE CONTROLE CORRESPONDENTES, PLANTA, PARTE DE PLANTA OU CÉLULA DE PLANTA, MOLÉCULA DE ÁCIDO NUCLEICO ISOLADA, CONSTRUTO DE GENE, VETOR, PLANTA TRANSGÊNICA, E, USO DE 5 UMA MOLÉCULA DE ÁCIDO NUCLEICO OU DO CONSTRUTO DE GENE OU DO VETOR”

A presente invenção refere-se em geral ao campo de biologia 10 molecular e concerne a um método para melhorar as características de crescimento de planta. Mais especificamente, a presente invenção concerne a um método para melhorar as características de crescimento de planta pela modulação da expressão em uma planta de um ácido nucleico de planta codificador de uma quinase dependente de ciclina (CDK) e/ou pela 15 modulação da atividade em uma planta de uma proteína CDK de planta, cuja proteína CDK compreende motivos diferentes ou cujo ácido nucleico CDK codifica tal proteína. A presente invenção também concerne às plantas tendo expressão modulada de um ácido nucleico CDK de planta e/ou atividade modulada de uma proteína CDK de planta, cuja proteína CDK compreende 20 diferentes motivos de sequência ou cujo ácido nucleico codifica tal proteína e cujas plantas têm características de crescimento melhoradas em relação às plantas de tipo selvagem correspondentes. A invenção adicionalmente se refere às seqüências de ácido nucleico específicas codificadoras das proteínas anteriormente mencionadas tendo atividade melhorada de crescimento de 25 planta, aos construtos de ácido nucleico, aos vetores e às plantas contendo ditas seqüências de ácido nucleico.