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

(21) **PI0618378-6 A2**

BRPI0618378A2

(22) Data de Depósito: 08/11/2006
(43) Data da Publicação: 30/08/2011
(RPI 2121)

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

(54) Título: **MÉTODOS PARA AUMENTAR O RENDIMENTO DE PLANTA E/OU RESISTÊNCIA AO ESTRESSE EM RELAÇÃO ÀS PLANTAS DE CONTROLE E PARA PRODUZIR UMA PLANTA TRANSGÊNICA TENDO RENDIMENTO AUMENTADO E/OU RESISTÊNCIA AO ESTRESSE AUMENTADA, CONSTRUÇÃO, USOS DE UMA CONSTRUÇÃO E DE UM ÁCIDO NUCLÉICO DEL1 OU VARIANTE DO MESMO, OU DE UM POLIPEPTÍDEO DEL1 OU HOMÓLOGO DO MESMO, POLIPEPTÍDEO DEL1 ISOLADO, E, SEQÜÊNCIA DE ÁCIDO NUCLEICO ISOLADA**

(30) Prioridade Unionista: 08/11/2005 EP 05110458.6, 08/11/2005 EP 05110470.1, 08/11/2005 US 60/734262, 08/11/2005 US 60/734281

(73) Titular(es): Cropdesign N.V., Universiteit Gent, Vib Vzw

(72) Inventor(es): Christophe Reuzeau, Dirk Inze, Jan de Meutter, Kobe Vlieghe, Lieven de Veylder, Vladimir Mironov

(74) Procurador(es): MOMSEN, LEONARDOS & CIA.

(86) Pedido Internacional: PCT EP2006068248 de 08/11/2006

(87) Publicação Internacional: WO 2007/054522de 18/05/2007

(57) Resumo: MÉTODOS PARA AUMENTAR O RENDIMENTO DE PLANTA E/OU RESISTÊNCIA AO ESTRESSE EM RELAÇÃO ÀS PLANTAS DE CONTROLE E PARA PRODUZIR UMA PLANTA TRANSGÊNICA TENDO RENDIMENTO AUMENTADO E/OU RESISTÊNCIA AO ESTRESSE AUMENTADA, CONSTRUÇÃO, USOS DE UMA CONSTRUÇÃO E DE UM ÁCIDO NUCLEICO DEL1 OU VARIANTE DO MESMO, OU DE UM POLIPEPTÍDEO DEL1 OU HOMÓLOGO DO MESMO, POLIPEPTÍDEO DEL1 ISOLADO, E, SEQUÊNCIA DE ÁCIDO NUCLEICO ISOLADA. A presente invenção diz respeito a um método para aumentar o rendimento de planta e/ou resistência ao estresse aumentada pela diminuição do nível de atividade de um polipeptídeo DEL 1. Um tal método compreende introduzir em uma planta um ácido nucleico que compreende um ácido nucleico DEL 1 variante. Um outro método compreende infra regular a expressão de um gene DEL 1. A invenção também diz respeito a plantas transgênicas tendo introduzido nelas um ácido nucleico DELi variante destas, plantas estas que têm rendimento aumentado e/ou resistência ao estresse aumentada em relação às plantas de controle. A presente invenção também diz respeito a uma novo proteína DELi e sua seqüência codificadora e construções úteis nos métodos da invenção.

"MÉTODOS PARA AUMENTAR O RENDIMENTO DE PLANTA E/OU RESISTÊNCIA AO ESTRESSE EM RELAÇÃO ÀS PLANTAS DE CONTROLE E PARA PRODUZIR UMA PLANTA TRANSGÊNICA TENDO RENDIMENTO AUMENTADO E/OU RESISTÊNCIA AO ESTRESSE AUMENTADA, CONSTRUÇÃO, USOS DE UMA CONSTRUÇÃO E DE UM ÁCIDO NUCLEICO DEL1 OU VARIANTE DO MESMO, OU DE UM POLIPEPTÍDEO DEL1 OU HOMÓLOGO DO MESMO, POLIPEPTÍDEO DEL1 ISOLADO, E, SEQÜÊNCIA DE ÁCIDO NUCLEICO ISOLADA"

A presente invenção no geral diz respeito ao campo da biologia molecular e diz respeito a um método para melhorar as características de crescimento vegetal em relação às plantas de controle. Mais especificamente, a presente invenção diz respeito a um método para aumentar o rendimento vegetal e/ou aumentar a resistência ao estresse que compreende diminuir o nível de atividade de um polipeptídeo DEL1. A presente invenção também diz respeito às plantas tendo atividade diminuída de um polipeptídeo DEL1, plantas estas que têm rendimento aumentado e/ou resistência ao estresse aumentada em relação às plantas de controle. Preferivelmente a atividade diminuída de um polipeptídeo DEL1 é obtida pela modulação da expressão de um ácido nucleico que codifica um fragmento de um polipeptídeo DEL1, ou pela regulação negativa da expressão de um gene DEL1 ou uma variante deste. A invenção também fornece um novo polipeptídeo DEL1, ácidos nucleicos que codificam um tal polipeptídeo e construções úteis nos métodos da invenção.

A população mundial sempre crescente e a diminuição da oferta de terra arável disponível para a agricultura estimula a pesquisa voltada para a melhora da eficiência da agricultura. Meios convencionais para as

melhoras de safra e da horticultura utiliza técnicas de cruzamento seletivo para identificar plantas tendo as características desejáveis. entretanto, tais técnicas de cruzamento seletivo têm diversas desvantagens, a saber que estas técnicas são tipicamente intensivas em trabalho e resultam em plantas que freqüentemente contêm componentes genéticos heterogêneos que nem sempre podem resultar no traço desejável sendo passado a diante a partir das plantas precursoras. Avanços ma biologia molecular têm permitido a humanidade modificar o germoplasma de animais e plantas. A engenharia genética de plantas acarreta necessariamente a isolação e manipulação de material genético (tipicamente na forma de DNA ou RNA) e a introdução subseqüente deste material genético em uma planta. Tal tecnologia tem a capacidade para liberar safras ou plantas tendo vários traços econômicos, agronômicos ou horticulturais melhorados. Um traço de interesse econômico particular é o rendimento. O rendimento é normalmente definido como o produto mensurável de valor econômico de uma safra. Isto pode ser definido em termos de quantidade e/ou qualidade. O rendimento é diretamente dependente de diversos fatores, por exemplo, o número e tamanho dos órgãos, arquitetura da planta (por exemplo, o número de ramos), produção de semente e mais. O desenvolvimento de raiz, captação de nutriente e tolerância ao estresse também podem ser fatores importantes na determinação do rendimento. A otimização de um dos fatores supracitados pode portanto contribuir para aumentar rendimento de safra.

O rendimento de semente é um traço particularmente importante, visto que as sementes de muitas plantas são importantes para a nutrição humana e animal. As safras tais como, milho, arroz, trigo, canola e soja são responsáveis por mais da metade da ingestão calórica humana total, seja através do consumo direto das próprias sementes ou através do consumo de produtos de carne fortalecidos com sementes processadas. Elas são também uma fonte de açúcares, óleos e muitos tipos de

metabólitos usados em processos industriais. As sementes contêm um embrião (a fonte de novos brotos e raízes) e um endosperma (a fonte de nutrientes para o crescimento do embrião durante a germinação e durante o crescimento inicial das mudas). O desenvolvimento de uma semente envolve muitos genes e requer a transferência de metabólitos das raízes, folhas e hastes para dentro da semente em crescimento. O endosperma, em particular, assimila os precursores metabólicos de carboidratos, óleos e proteínas e os sintetiza nas macromoléculas de armazenagem para encher o grão.

As condições de estresse abiótico, tais como deficiência ou excesso de energia solar, água e nutrientes, salinidade, temperatura alta ou baixa e poluição (por exemplo, metais pesados), podem ter um impacto maior sobre o crescimento vegetal e pode significantemente reduzir o rendimento, por exemplo, de cultivares de safra. Embora a luz seja essencial para as plantas visto que a mesma fornece energia para o metabolismo, a luz UV é nociva para os organismos vivos. Em particular, a radiação UV-B (280-320 nm) afeta a estabilidade de proteína e ácidos nucleicos resultando em fotossíntese reduzida, redução de biomassa, síntese de proteína diminuída e função de cloroplasto prejudicada, assim como dano ao DNA. Devido à redução da camada de ozona protetora na estratosfera, problemas a cerca dos efeitos da radiação UV-B sobre as plantas tem aumentado. A resposta fisiológica a estes estresses surge de mudanças na expressão do gene celular.

A capacidade para aumentar o rendimento vegetal e a resistência ao estresse teriam muitas aplicações em áreas tais como a agricultura, incluindo na produção de plantas ornamentais, arboricultura, horticultura e silvicultura. O aumento de rendimento também pode encontrar uso na produção de algas para o uso em biorreatores (para a produção biotecnológica de substâncias tais como produtos farmacêuticos, anticorpos ou vacinas, ou para a bioconversão de resíduos orgânicos) e outras de tais áreas.

A transição no ciclo celular de plantas e animais é controlada pelo caminho RB/E2F/DP. Os fatores de transcrição E2F formam um componente crítico deste mecanismo regulatório (Stals & Inze, Trends Plant Science 6, 359-364, 2001). A família de fatores de transcrição E2F na Arabidopsis thaliana consiste de seis membros, que podem ser agrupados em duas subfamílias (Mariconti et al., J. Biol. Chem. 277, 9911-9919, 2002): por um lado E2Fa, E2Fb e E2Fc, por outro lado E2Fd, E2Fe e E2Ff. Os membros E2F da primeira subfamília têm um domínio de ligação de DNA e requer um parceiro DP influenciador para formar um ativador transcricional funcional (E2Fa e E2Fb) ou repressor (E2Fc). Os membros da segunda subfamília possuem 2 domínios de ligação de DNA equivalentes à E2F e compreendem características estruturais que os discriminam dos três membros do primeiro subgrupo. AtE2Fd e AtE2Fe têm um sinal de localização nuclear (NLS) na parte de terminal C da proteína, que é suficiente para a ligação de DNA sem a necessidade de um parceiro de DP, mas AtE2Ff carece de um NLS. Os dados experimentais sugeriram que os membros de E2F da segunda subfamília não tem nenhum domínio de ativação e foram postulados interferir com a atividade das outras proteínas AtE2F (Mariconti et al., 2002).

Na Arabidopsis, AtE2Fe (também conhecido como DEL1 ou E2L3) é expressado no começo da transição G_1/S e na transição S/G_2. A proteína é capaz de ligar sítios de E2F sem um parceiro DP. Um estudo por Vlieghe et al. (Current Biology 15, 59-63, 2005) demonstrou que AtE2Fe pode estar envolvido no controle do endociclo. Foi mostrado que em um mutante DEL1 (del1-1, com uma inserção de T-DNA entre os dois domínios de ligação de DNA) os níveis de ploidia foram aumentados, ao passo que as plantas que super expressam DEL1 tiveram endorreduplicação reduzida. As mudanças nos níveis de ploidia correlacionaram-se com os níveis de expressão alteradas para vários genes alvo de E2F envolvido na replicação de DNA. Os transcritos DEL1 foram detectados em células que se dividem mas

foram ausentes em células no endociclo, o que sugeriu que DEL1 pode estar envolvido na repressão do endociclo em células mitoticamente ativas. Barow M. e Meister A. (2003) Plant, Cell and Environment 26, 571-584 sugerem que as espécies de planta endopoliplóides têm vantagens quando crescem por exemplo em latitudes mais altas. O controle do endociclo é um processo complexo nas plantas e muitos genes estão envolvidos.

Na presente invenção foi surpreendentemente verificado que a diminuição do nível de atividade de um polipeptídeo DEL1 dá plantas tendo rendimento aumentado e/ou resistência ao estresse aumentada em relação às plantas de controle. Preferivelmente o nível diminuído de atividade é obtido pela expressão moduladora em uma planta de um ácido nucleico que codifica um fragmento de um polipeptídeo DEL1 ou pela regulação negativa da expressão de um gene DEL1 ou uma variante deste. Preferivelmente, o polipeptídeo DEL1 é de origem vegetal.

De acordo com uma forma de realização da presente invenção, é fornecido um método para aumentar o rendimento vegetal e/ou aumentar a resistência ao estresse, que compreende diminuir o nível de atividade de um polipeptídeo DEL1, preferivelmente pela modulação da expressão em uma planta de um ácido nucleico que codifica um fragmento de um polipeptídeo DEL1, ou pela regulação negativa da expressão de um gene DEL1 ou uma variante deste.

Qualquer referência daqui em diante a uma "proteína útil nos métodos da invenção" é considerado significar um polipeptídeo DEL1 como aqui definido. Qualquer referência daqui em diante a um "ácido nucleico útil nos métodos da invenção" é considerado significar um ácido nucleico capaz de codificar um tal polipeptídeo DEL1. Os termos "polipeptídeo" e "proteína" são usados aqui intercambiavelmente e referem-se aos aminoácidos em uma forma polimérica de qualquer comprimento. Os termos "polinucleotídeo(s)", "seqüência(s) de ácido nucleico", "seqüência(s) de nucleotídeo" "gene(s)" são

aqui usados intercambiavelmente e referem-se aos nucleotídeos, ribonucleotídeos ou desoxirribonucleotídeos ou uma combinação de ambos, em uma forma polimérica de qualquer comprimento.

A escolha de plantas de controle adequadas é uma parte da rotina de um cenário experimental e pode incluir plantas do tipo selvagem correspondentes ou plantas correspondentes sem o gene de interesse. A planta de controle é tipicamente da mesma espécie vegetal ou ainda da mesma variedade como a planta a ser avaliada. A planta de controle também pode ser um nulizigoto da planta a ser avaliada. Uma "planta de controle" como aqui usada refere-se não apenas a plantas inteiras, mas também a partes de plantas, incluindo sementes e partes de semente.

Vantajosamente, o desempenho dos métodos de acordo com a presente invenção resulta em plantas tendo rendimento aumentado, particularmente rendimento de semente aumentado e/ou biomassa aumentada, em relação às plantas de controle.

O termo "rendimento" no geral significa um produto mensurável de valor econômico, necessariamente relacionado com uma safra especificada, a uma área e a um período de tempo. A parte individual de plantas diretamente contribuem para o rendimento com base no seu número, tamanho e/ou peso, ao passo que o rendimento real é o rendimento por acre para uma safra e ano, que é determinada dividindo-se a produção total (inclui tanto a produção colhida quanto a estimada) pelos acres plantados.

Os termos "aumento", "melhorar" ou "melhora" são intercambiáveis e devem significar no sentido da aplicação pelo menos uns 5 %, 6 %, 7 %, 8 %, 9 % ou 10 %, preferivelmente pelo menos 15 % ou 20 %, mais preferivelmente 25 %, 30 %, 35 % ou 40 % mais de rendimento e/ou crescimento em comparação com a planta do tipo selvagem como aqui definida.

O termo "rendimento aumentado" como aqui definido é

considerado significar um aumento em qualquer um ou mais dos seguintes, cada um em relação às plantas de controle: (i) biomassa aumentada (peso) de uma ou mais partes de uma planta, particularmente as partes acima do solo (colhíveis), biomassa de raiz aumentada ou biomassa aumentada de qualquer outra parte colhível; (ii) rendimento de semente total aumentado, que inclui um aumento na biomassa de semente (peso de semente) e que pode ser um aumento no peso da semente por planta ou em uma base de semente individual; (iii) número aumentado de flores por planta; (iv) número aumentado de sementes (cheias); (v) taxa de enchimento de semente aumentada (que é expressada como a razão entre o número de sementes cheias dividido pelo número total de sementes); (vi) tamanho de semente aumentado, que também pode influenciar a composição de sementes; (vii) volume de semente aumentado, que também pode influenciar a composição de sementes (incluindo teor e composição de óleo, proteína e carboidrato totais); (viii) área da semente individual aumentada; (ix) comprimento de semente individual e/ou perímetro da semente aumentados; (x) índice de colheita aumentado, que é expressado como uma razão do rendimento de partes colhíveis, tais como sementes, em relação à biomassa total; e (xi) 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 tamanho de semente e/ou peso de semente aumentado. Um TKW aumentado também pode resultar de um aumento no tamanho do embrião e/ou tamanho do endosperma. Um aumento no rendimento também pode resultar na arquitetura modificada, ou pode ocorrer como um resultado da arquitetura modificada.

Tomando-se o milho como um exemplo, um aumento de rendimento pode ser manifestado como um ou mais dos seguintes: aumento no número de plantas estabelecidas por hectare ou acre, um aumento no número de espigas por planta, um aumento no número de fileiras, número de

grãos por fileira, peso do grão, peso de mil grãos, comprimento/diâmetro da espiga, aumento na taxa de enchimento de semente (que é o número de sementes cheias dividido pelo número total de sementes e multiplicado por 100), entre outras. Tomando-se o arroz como um exemplo, um aumento de rendimento pode manifestar-se por si só como um aumento em um ou mais dos seguintes: número de plantas per hectare ou acre, número de panículos por planta, número de espiguetas por panículo, número de flores (flósculos) por panículo (que é expressado como uma razão do número de sementes cheias em relação ao número de panículos primários), aumento na taxa de enchimento de semente (que é o número de sementes cheias dividido pelo número total de sementes e multiplicado por 100), aumento no peso de mil grãos, entre outros.

De acordo com uma característica preferida, o desempenho dos métodos da invenção resulta em plantas tendo rendimento aumentado, particularmente biomassa aumentada e/ou rendimento de semente. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento vegetal, método este que compreende diminuir o nível de atividade de um polipeptídeo DEL1, preferivelmente pela modulação da expressão em uma planta de um ácido nucleico que codifica um fragmento de um polipeptídeo DEL1 ou pela regulação negativa da expressão de um gene DEL1 ou uma variante deste.

Visto que as plantas transgênicas de acordo com a presente invenção têm rendimento aumentado, é provável que estas plantas exibam uma taxa de crescimento aumentada (durante pelo menos parte do seu ciclo de vida), em relação à taxa de crescimento de plantas de controle em um estágio correspondente no seu ciclo de vida. A taxa de crescimento aumentada pode ser específica para uma ou mais partes de uma planta (incluindo sementes), ou pode ser substancialmente por toda 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 considerado significar o tempo necessário para crescer de uma semente madura seca até o estágio onde a planta produziu sementes maduras secas, similares ao material de partida. Este ciclo de vida pode ser influenciado por fatores tais como vigor inicial, taxa de crescimento, índice de verdor, tempo de florescimento e velocidade de maturação da semente. O aumento na taxa de crescimento pode ocorrer em um ou mais estágios no ciclo de vida de uma planta ou durante substancialmente o ciclo de vida da planta inteiro. A taxa de crescimento aumentada durante os estágios iniciais no ciclo de vida de uma planta pode refletir vigor realçado. O aumento na taxa de crescimento pode alterar o ciclo de colheita de uma planta permitindo que as plantas sejam semeadas mais tarde e/ou colhidas mais cedo do que de outro modo seria possível (um efeito similar pode ser obtido com tempo de florescimento mais cedo). Se a taxa de crescimento é suficientemente aumentada, ela pode permitir a semeadura adicional de sementes da mesma espécie vegetal (por exemplo semeadura e colheita de plantas de arroz seguido pela semeadura e colheita de mais plantas de arroz todas dentro de um período de crescimento convencional). Similarmente, se a taxa de crescimento é suficientemente aumentada, ela pode permitir a semeadura adicional de sementes de espécies vegetais diferentes (por exemplo a semeadura e colheita de plantas de arroz seguida, por exemplo, pela semeadura e colheita opcional de soja, batata ou qualquer outra planta adequada). Os tempos adicionais de colheita do mesmo rizoma no caso de algumas plantas de safra também podem ser possíveis. Alterar o ciclo de colheita de uma planta pode levar a um aumento na produção de biomassa anual por acre (devido a um aumento no número de vezes (diga-se em um ano) que qualquer planta particular pode ser cultivada e colhida). Um aumento na taxa de crescimento também pode permitir o cultivo de plantas transgênicas em uma área geográfica mais ampla do que as suas contrapartes do tipo selvagem, visto que as limitações territoriais para o cultivo de uma

safra são freqüentemente determinadas pelas condições ambientais adversas no tempo de plantio (estação inicial) ou no tempo de colheita (estação final). 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 a partir das curvas de crescimento, tais parâmetros podem ser: T-Mid (o tempo levado para as plantas atingirem 50 % do seu tamanho máximo) e T-90 (time levado para as plantas atingirem 90 % do seu tamanho máximo), entre outros.

O desempenho dos métodos da invenção dá plantas tendo uma taxa de crescimento aumentada. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas, método este que compreende diminuir o nível de atividade de um polipeptídeo DEL1, preferivelmente pela modulação da expressão em uma planta de um ácido nucleico que codifica um fragmento de um polipeptídeo DEL1 ou pela regulação negativa da expressão de um gene DEL1 ou uma variante deste.

Vantajosamente, o desempenho dos métodos de acordo com a presente invenção resulta em plantas tendo resistência ao estresse aumentada, em relação às plantas de controle.

As plantas transgênicas obtidas pelo método têm resistência ao estresse aumentada; isto é, são capazes de crescer normalmente sob condições ambientais em que as plantas de controle mostram crescimento, metabolismo, viabilidade, produtividade e/ou esterilidade masculina ou feminina reduzidos. Um aumento no rendimento e/ou taxa de crescimento ocorre se a planta está sob condições que não de estresse ou se a planta é exposta a vários estresses comparada com as plantas de controle. O termo tolerância como aqui usado abrange a proteção contra o estresse variando de uma demora para uma inibição substancialmente completa de alteração no metabolismo celular, crescimento celular reduzido e/ou morte celular causada pelas condições de estresse. As plantas tipicamente respondem à exposição ao estresse pelo

crescimento mais lento. Em condições de estresse grave, a planta pode mesmo interromper o crescimento completamente. O estresse brando por outro lado é aqui definido como sendo qualquer estresse ao qual uma planta é exposta que não resulta na planta cessando de crescer completamente sem a capacidade para retomar o crescimento. O estresse brando no sentido da invenção leva a uma redução no crescimento das plantas estressadas de menos do que 40 %, 35 % ou 30 %, preferivelmente menos do que 25 %, 20 % ou 15 %, mais preferivelmente menos do que 14 %, 13 %, 12 %, 11 % ou 10 % ou menos em comparação com a planta de controle sob condições não estressantes. Devido aos avanços nas práticas agrícolas (irrigação, fertilização, tratamentos com pesticidas) os estresses graves não são freqüentemente encontradas em plantas de safra cultivadas. Como uma conseqüência, o crescimento comprometido induzido pelo estresse brando é freqüentemente uma característica indesejável para a agricultura. Os estresses brandos são os estresses típicos aos quais uma planta pode ser exposta. Estes estresses podem ser os estresses bióticos e/ou abióticos (ambientais) de todo dia aos quais uma planta é exposta. Os estresses abióticos ou ambientais típicos incluem estresses das temperaturas causadas pelas temperaturas quentes ou frias/congelantes atípicas; estresse do sal; estresse da água (seca ou água em excesso), estresse pela radiação de UV. A luz do sol não apenas contêm radiação dos comprimentos de onda certos para a fotossíntese: radiação de comprimentos de onda mais curtos, tais como radiação de ultravioleta B (UV-B, 280320 nm) também está presente. A UV-B é nociva para os organismos vivos visto que os componentes celulares tais como proteínas e ácido nucleicos absorvem esta radiação rica em energia. Muitos estudos têm mostrado os efeitos nocivos do UV-B tais como fotossíntese reduzida, redução da biomassa, síntese de proteína diminuída e função de cloroplasto prejudicada, assim como dano ao DNA. Além disso, a expressão de muitos genes é mudada pela irradiação de UV-B: por exemplo, a transcrição de genes de defesa é aumentada, ao passo que os níveis de mRNA

para os genes cloroplásticos declinam. Os produtos químicos também podem causar estresses abióticos. Os estresses bióticos são tipicamente aqueles estresses causados pelos patógenos, tais como bactérias, vírus, fungos e insetos. O termo condições "não estressantes" como aqui usado são aquelas condições ambientais que não vão significantemente além das condições de estresse climático normal e outros estresses abióticos que as plantas podem encontrar e que permitem crescimento ótimo das plantas. As pessoas habilitadas na técnica estão cientes das condições de solo normais e condições climáticas para uma dada localização geográfica.

De acordo com uma outra característica preferida, o desempenho dos métodos da invenção resulta em plantas tendo resistência ao estresse aumentada. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a resistência ao estresse de uma planta, método este que compreende diminuir o nível de atividade de um polipeptídeo DEL1, preferivelmente pela modulação da expressão em uma planta de um ácido nucleico que codifica um fragmento de um polipeptídeo DEL1 ou pela regulação negativa da expressão de um gene DEL1 ou uma variante deste.

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

O termo "planta" como aqui usado abrange plantas inteiras, ancestrais e progênie das plantas e parte de plantas, incluindo sementes, brotos, hastes, folhas, raízes (incluindo tubérculos), flores e tecidos e órgãos, em que cada um dos anteriormente mencionados compreendem o gene/ácido nucleico de interesse. O termo "planta" também abrange células de planta, culturas de suspensão, tecido de calo, embriões, regiões meristemáticas, gametófitos, esporófitos, pólen e microesporos, mais uma vez em que cada um dos anteriormente mencionados compreendem o gene/ácido nucleico de interesse.

As plantas que são particularmente úteis nos métodos da

invenção incluem todas as plantas que pertencem à superfamília Viridiplantae, em particular plantas monocotiledônea e dicotiledônea incluindo legumes de forragem ou feno, plantas ornamentais, safras de alimento, árvores e arbustos selecionados da lista que compreende Acer spp., Actinidia spp., Abelmoschus spp., Agropyron spp., Allium spp., Amaranthus spp., Ananas comosus, Annona spp., Apium graveolens, Arachis spp, Artocarpus spp., Asparagus officinalis, Avena spp. (por exemplo, Avena sativa, Avena fatua, Avena byzantina, Avena fatua var. sativa, Avena hybrida), Averrhoa carambola, Benincasa hispida, Bertholletia excelsea, Beta vulgaris, Brassica spp. (por exemplo, Brassica napus, Brassica rapa ssp. [canola, semente oleaginosa de colza, nabo]), Cadaba farinosa, Camellia sinensis, Canna indica, Capsicum spp., Carex elata, Carica papaya, Carissa macrocarpa, Carya spp., Carthamus tinctorius, Castanea spp., Cichorium endivia, Cinnamomum spp., Citrullus lanatus, Citrus spp., Cocos spp., Coffea spp., Colocasia esculenta, Cola spp., Coriandrum sativum, Corylus spp., Crataegus spp., Crocus sativus, Cucurbita spp., Cucumis spp., Cynara spp., Daucus carota, Desmodium spp., Dimocarpus longan, Dioscorea spp., Diospyros spp., Echinochloa spp., Elaeis (por exemplo, Elaeis guineensis, Elaeis oleifera), Eleusine coracana, Eriobotrya japonica, Eugenia uniflora, Fagopyrum spp., Fagus spp., Ficus carica, Fortuneila spp., Fragaria spp., Ginkgo biloba, Glycine spp. (por exemplo, Glycine max, Sofa hispida ou Sofa max), Gossypium hirsutum, Helianthus spp. (por exemplo, Helianthus annuus), Hemerocallis fulva, Hibiscus spp., Hordeum spp. (por exemplo, Hordeum vulgare), batatas Ipomoea, Juglans spp., Lactuca sativa, Lathyrus spp., Lens culinaris, Linum usitatissimum, Litchi chinensis, Lotus spp., Luffa acutangula, Lupinus spp., Luzula sylvatica, Lycopersicon spp. (por exemplo, Lycopersicon esculentum, Lycopersicon lycopersicum, Lycopersicon pyriforme), Macrotyloma spp., Malus spp., Malpighia emarginata, Mammea americana, Mangifera indica, Manihot spp., Manilkara zapota, Medicago

sativa, Melilotus spp., Mentha spp., Momordica spp., Morus nigra, Musa spp., Nicotiana spp., Olea spp., Opuntia spp., Ornithopus spp., Oryza spp. (por exemplo, Oryza sativa, Oryza latifolia), 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., Ricinus communis, Rubus spp., Saccharum spp., Sambucus spp., Secale cereale, Sesamum spp., Sinapis sp., Solanum spp. (por exemplo, Solanum tuberosum, Solanum integrifolium ou Solanum lycopersicum), Sorgo bicolor, Spinacia spp., Syzygium spp., Tagetes spp., Tamarindus indica, Theobroma cacao, Trifolium spp., Triticosecale rimpaui, Triticum spp. (por exemplo, Triticum aestivum, Triticum durum, Triticum turgidum, Triticum hybernum, Triticum macha, Triticum sativum ou Triticum vulgare), Tropaeolum minus, Tropaeolum majus, Vaccinium spp., Vicia spp., Vigna spp., Viola odorata, Vitis spp., Zea mays, Zizania palustris, Ziziphus spp., entre outras.

De acordo com uma forma de realização preferida da presente invenção, a planta é uma planta de safra. Os exemplos de plantas de safra incluem soja, girassol, canola, alfafa, semente de colza, algodão, tomate, batata e tabaco. Mais preferivelmente, a planta é uma planta monocotiledônea. Os exemplos de plantas monocotiledôneas incluem cana de açúcar. Mais preferivelmente a planta é um cereal. Os exemplos de cereais incluem arroz, milho, trigo, cevada, painço, centeio, sorgo e aveia.

O termo "polipeptídeo DEL1" como aqui definido refere-se uma proteína equivalente a E2F com dois domínios de ligação de DNA do tipo hélice de asa (domínio E2F TDP, Pfam PF02319, Interpro IPR003316), mas carecendo de um domínio de dimerização; ver por exemplo a Figura 1. Preferivelmente, o polipeptídeo DEL1 compreende a seqüência de assinatura

1 (SEQ ID NO: 9) (Y/T/P) (S/D) RK (Q/D) KSL (G/W) (L/T) LC (T/E/Q/S) (N/R/K) F (L/V) (A/S/T/R) (L/I/R) Y (N/G/P/D)

Além disso, preferivelmente, a seqüência de assinatura 1 é

(Y/T/P) (S/D) RK (Q/D)KSL (G/W) (L/T) LC (T/E/Q/S) (N/R/K)F (L/V) (A/S/T) (L/I/R) Y (N/G/P/D)

Mais preferivelmente, a seqüência de assinatura 1 é

(Y/T) (S/D) RK (Q/D)KSL (G/W) (L/T) LC (T/E/S) (N/R)F (L/V) (A/S/T) (L/I) Y (N/G/D)

O mais preferivelmente, a seqüência de assinatura 1 é YSRKQKSLGLLCTNFLALYN. As seqüências de assinatura alternativas para caracterizar um polipeptídeo DEL1 incluem:

seqüência de assinatura 2 (SEQ ID NO: 10):

(G/S) LD (D/E) (A/V) A (S/A/T/V/R/K) (K/R/S) LGVE;

seqüência de assinatura 3 (SEQ ID NO: 11):

R (R/K) (E/D)KSL (G/A/R) (I/L) (L/M) (T/S) (Q/K) (K/N)F (I/V) (K/Q/M) LF (I/V/L/T) (C/A/V /M/N/T) (S/E/T/M);

seqüência de assinatura 4 (SEQ ID NO: 12):

(I/V/L) (S/T) L (D/E) (D/T/V/E)AA (K/R) (L/I/C/R) (L/I) (L/M/I) (G/E);

e seqüência de assinatura 5 (SEQ ID NO: 13):

(T/A)K (V/I) RRLYDIAN (V/I) L (S/C/T) S (M/L) (N/R/A/H/Q) (L/F) I (E/R/K/D)K (T/V/I) (H/Q /T) (T/V/Q/H) (L/G/A/T/V/P) (D/E) (S/T/E) R (K/G) (P/R) (A/K) (F/P) (K/L/A/R) (W/F) (L/K).

Preferivelmente, a seqüência de assinatura 2 é (G/S) LDDAA (S/A/T/R/K) (K/R) LGVE, o mais preferivelmente a seqüência de assinatura 2 é GLDDAASKLGVE.

Além disso, a seqüência de assinatura 3 é preferivelmente

R (R/K) (E/D)KSL (G/A/R) (I/L) (L/M) (T/S) Q (K/N)F (I/V) (K/Q/M) LF (I/V/L/T) (C/A/V/M/N /T) (S/E/M),

mais preferivelmente, a seqüência de assinatura 3 é

R (R/K)EKSL (G/A) LL (T/S) Q (K/N)F (I/V) (K/Q) LF (I/V/L/T) (C/A/M/N/T) (S/M)

o mais preferivelmente a seqüência de assinatura 3 é RREKSLGLLTQNFIKLFICS. A Seqüência de assinatura 4 é preferivelmente

(I/V) (S/T) L (D/E) (D/T/V/E) AA (K/R) (L/I) (L/I) (L/M/I) (G/E)

Mais preferivelmente, a seqüência de assinatura 4 é

(I/V) SL (D/E) (D/T/E) AA (K/R) (L/I) (L/I) (L/M) G, o mais preferivelmente a seqüência de assinatura 4 é ISLDDAAKLLLG.

A seqüência de assinatura 5 é preferivelmente

(T/A)K (V/I) RRLYDIAN (V/I) L (S/C/T) S (M/L) (N/R/A) (L/F) I (E/R/K/D)K (T/V/I) (H/Q) (T/V/Q) (L/G/A/T) (D/E) (S/T/E) R (K/G) (P/R) (A/K) (F/P) (K/L/A/R) (W/F) (L/K)

Mais preferivelmente a seqüência de assinatura 5 é

(T/A)K (V/I) RRLYDIAN (V/I) L (S/C/T) S (M/L) (N/R) (L/F) I (E/R/D)K (T/V/I) (H/Q) (T/V/Q) (L/G/A/T) (D/E) (S/T) RKPAF (K/L/R)WL, o mais preferivelmente a seqüência de assinatura 5 é TKVRRLYDIANVLSSMNLIEKTHTLDSRKPAFKWL.

O mais preferivelmente, a proteína DEL1 é como representada pela SEQ ID NO: 8.

As proteínas da invenção são identificáveis pela presença do domínio(s) E2F TDP (mostrado na Figura 1). O termo "domínio" refere-se a um conjunto de aminoácidos conservados nas posições específicas junto com um alinhamento de seqüências de proteínas evolucionariamente relacionado. Enquanto os aminoácidos nas outras posições podem variar entre homólogos, aminoácidos que são altamente conservadas nas posições específicas indicam aminoácidos que são essenciais na estrutura, estabilidade ou atividade de uma proteína. Identificado pelo seu alto grau de conservação em seqüências

alinhadas de uma família de homólogos de proteína, estes podem ser usados como identificadores para determinar se qualquer polipeptídeo em questão pertence a uma família de polipeptídeo previamente identificado (neste caso, as proteínas úteis nos métodos da invenção e ácidos nucleicos que codificam as mesmas como aqui definido).

Os termos "motivo" ou "seqüência de consenso" ou "seqüência de assinatura" referem-se uma região conservada curta na seqüência de proteínas evolucionariamente relacionadas. Os motivos são freqüentemente partes altamente conservadas de domínios, mas também podem incluir apenas parte do domínio, ou ser localizado fora do domínio conservado (se todos os aminoácidos do motivo caem fora de um domínio definido). Uma pessoa habilitada na técnica saberá que nas seqüências de assinatura, nas posições onde 3 ou mais aminoácidos são possíveis, a seqüência não é muito conservada e que os desemparelhamentos podem ocorrer nestas posições.

O termo "polipeptídeo DEL1" também abrange homólogos da SEQ ID NO: 8. Pelo alinhamento de outras seqüências de proteína com a SEQ ID NO: 8, as seqüências de assinatura correspondentes detalhadas acima podem ser facilmente identificadas. Deste modo, os polipeptídeos DEL1 ou seus homólogos podem ser facilmente identificados, usando técnicas de rotina bem conhecidas na técnica, tal como pelo alinhamento de seqüência. As bases de dados especializadas também existem para a identificação de domínios, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acid Res 30, 242-244, InterPro (Mulder et al., (2003) Nucl. Acids Res. 31, 315-318, Prosite (Bucher e Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D.,

Eds., pp53-61, AAAIPress, Menlo Park; Hulo et al., Nucl. Acids. Res. 32: D134-D137, (2004), ou Pfam (Bateman et al., Nucleic Acid Research 30(1): 276-280 (2002). Um conjunto de ferramentas para a análise in silico de seqüências de proteína está disponível no servidor de proteômicos ExPASY (hospedado pelo Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acid Res. 31: 3784-3788 (2003)).

Os domínios também podem ser identificados usando técnicas de rotina, tais como pelo alinhamento de seqüência. Os métodos para o alinhamento de seqüências para comparação são bem conhecidos na técnica, tais métodos incluem GAP, BESTFIT, BLAST, FASTA e TFASTA. GAP usa o algoritmo de Needleman e Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar o alinhamento global (isto é, transpor as seqüências completas) de duas seqüências que maximizem o número de emparelhamentos e minimize o número de intervalos. O algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-10) calcula a identidade de seqüência percentual e realiza uma análise estatística da similaridade entre as duas seqüências. O software para realizar a análise BLAST é publicamente disponível através do National Centre for Biotechnology Information (NCBI). Os homólogos podem ser facilmente identificados usando, por exemplo, o algoritmo de alinhamento de seqüência múltiplo ClustalW (versão 1.83), com os parâmetros de alinhamento aos pares padrão e um método de contagem em porcentagem. As porcentagens globais de similaridade e identidade também podem ser determinadas usando um dos métodos disponíveis no pacote de software MatGAT (Campanella et al., BMC Bioinformatics. 10 de julho de 2003; 4: 29. MatGAT: a application that generates similarity/identity matrices using protein or DNA sequences). A edição manual menor pode ser realizada para otimizar o alinhamento entre motivos conservadas, como estaria evidente a uma pessoa habilitada na técnica. Além disso, ao invés de usar seqüências de

tamanho natural para a identificação de homólogos, os domínios específicos (tais como o domínio E2F_TDP, ou um dos motivos definidos acima) também podem ser usados. Os valores de identidade de seqüência, que são indicados abaixo no Exemplo 3 como uma porcentagem foram determinados na seqüência de ácido nucleico ou aminoácido inteira, e/ou nos domínios ou motivo(s) conservado(s) selecionado(s), usando os programas mencionados acima usando os parâmetros padrão.

Os exemplos de polipeptídeos DEL1 ou homólogos são fornecidos no Exemplo 1.

Um ensaio pode ser realizado para determinar a atividade de DEL1. Os exemplos de ensaios para determinar a atividade de DEL1 são fornecidos no Exemplo 6.

A seqüência dada na SEQ ID NO: 21 representa uma proteína DEL1 até agora desconhecida. É portanto fornecido um polipeptídeo DEL1 isolado selecionado do grupo que consiste de:

(a) um polipeptídeo como dado na SEQ ID NO 21,

(b) um polipeptídeo com uma seqüência de aminoácido que tem pelo menos 60 % de identidade de seqüência, preferivelmente 70 % de identidade de seqüência, mais preferivelmente 80 % ou 90 % de identidade de seqüência, o mais preferivelmente 95 %, 96 %, 97 %, 98 % ou 99 % de identidade de seqüência com a seqüência de aminoácido como dado na SEQ ID NO 21,

(c) um homólogo, um derivado, e/ou fragmento funcional de uma proteína como definida em (a) ou (b).

Deve ser entendido que as seqüências que caem sob a definição de "polipeptídeo DEL1 ou homólogo deste" não devem ser limitados às seqüências representadas pela SEQ ID NO: 8, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29 ou SEQ ID NO: 31, mas que

qualquer polipeptídeo que compreende dois domínios de ligação de DNA do tipo hélice com asa (domínio E2F_TDP) como descrito por Pfam record PF02319 ou Interpro record IPR003316 mas carecendo de um domínio de dimerização e/ou compreendendo uma ou mais das seqüências de assinatura detalhadas na SEQ ID NO: 9, 10, 11, 12 e 13 podem ser adequados para o uso nos métodos da invenção.

Também úteis nos métodos da invenção são homólogos de qualquer uma das seqüências de aminoácido dadas na Tabela A do Exemplo 1. "Homólogos" de uma proteína abrange peptídeos, oligopeptídeos, polipeptídeos, proteínas e enzimas tendo substituições, deleções e/ou inserções de aminoácido em relação à proteína não modificada em questão e tendo atividades biológica e funcional similares como a proteína não modificada da qual eles são derivados.

Uma deleção refere-se à remoção de um ou mais aminoácidos de uma proteína.

Uma inserção refere-se a um ou mais resíduos de aminoácido que são introduzidos em um sítio pré determinado em uma proteína. As inserções podem compreender fusões de terminal N e/ou terminal C assim como inserções intra-seqüência de aminoácidos únicos ou múltiplos. No geral, as inserções dentro da seqüência de aminoácido serão menores do que as fusões de terminal N ou C, da ordem de cerca de 1 a 10 resíduos. Os exemplos de proteínas ou peptídeos de fusão de terminal N ou C incluem o domínio de ligação ou domínio de ativação de um ativador transcricional como usado no sistema de híbrido duplo de levedura, as proteínas de revestimento de fago, rótulo de (histidina)-6, rótulo de glutationa S-transferase, proteína A, proteína de ligação de maltose, diidrofoliato redutase, epítopo Tag'100, epítopo c-myc, epítopo FLAG®, lacZ, CMP (peptídeo de ligação de calmodulina), epítopo HA, epítopo de proteína C e epítopo VSV.

Uma substituição refere-se à substituição de aminoácidos da

proteína com outros aminoácidos tendo propriedades similares. As substituições de aminoácido são tipicamente de resíduos únicos, mas podem ser agrupados dependendo das restrições funcionais colocadas no polipeptídeo; as inserções usualmente serão da ordem de cerca de 1 a 10 resíduos de aminoácido. Preferivelmente, as substituições de aminoácido compreendem substituições de aminoácido conservativas. Para produzir tais homólogos, os aminoácidos da proteína podem ser substituídos por outros aminoácidos tendo propriedades similares (tais como hidrofobicidade similar, hidrofilicidade, antigenicidade, propensão para formar ou romper estruturas α-helicoidais ou estruturas de folha β). As tabelas de substituição conservativas são bem conhecidas na técnica (ver por exemplo Creighton (1984) Proteins. W. H. Freeman and Company e Tabela 1 abaixo).

Tabela 1: Exemplos de substituições de aminoácido conservadas

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

As substituições, deleções e/ou inserções de aminoácido podem ser facilmente fabricadas usando técnicas sintéticas de peptídeo bem conhecidas na técnica, tais como a síntese de peptídeo de fase sólida e outros ou pelas manipulações de DNA recombinante. Os métodos para a manipulação de seqüências de DNA para produzir variantes de substituição, inserção ou deleção de uma proteína são bem conhecidos na técnica. Por exemplo, técnicas para fabricar mutações de substituição nos sítios pré determinados no DNA são bem conhecidos por aqueles habilitados na técnica e incluem a mutagênese M13, mutagênese in vitro T7-Gen (USB, Cleveland, OH), Mutagênese Direcionada ao Sítio QuickChange (Stratagene, San Diego,

CA), mutagênese mediada pela PCR direcionada ao sítio ou outros protocolos de mutagênese direcionada ao sítio.

Também úteis nos métodos da invenção são derivados de qualquer um dos polipeptídeos dados na Tabela A do Exemplo 1 ou ortólogos ou parálogos de qualquer uma das SEQ ID Nos anteriormente mencionadas. "Derivados" incluem peptídeos, oligopeptídeos, polipeptídeos que, comparados com a seqüência de aminoácido da forma que ocorre naturalmente da proteína, tais como aqueles apresentados na SEQ ID NO: 8, podem compreender substituições de aminoácidos com resíduos de aminoácido que não ocorrem naturalmente, ou adições de resíduos de aminoácido que não ocorrem naturalmente. Os derivados dos polipeptídeos dados na Tabela A do Exemplo 1 são outros exemplos que podem ser adequados para o uso nos métodos da invenção. Os derivados úteis nos métodos da presente invenção preferivelmente têm atividade biológica e funcional similares como a proteína não modificada a partir da qual eles são derivados.

Os "Derivados" de uma proteína também abrangem peptídeos, oligopeptídeos, polipeptídeos que compreendem resíduos de aminoácido alterados que ocorrem naturalmente (glicosilado, acilado, prenilado, fosforilado, miristoilado, sulfatado, etc) ou alterados que não ocorrem naturalmente comparados com a seqüência de aminoácido de uma forma que ocorre naturalmente do polipeptídeo. Um derivado também pode compreender um ou mais substituintes que não aminoácido ou adições comparadas com a seqüência de aminoácido a partir da qual o mesmo é derivado, por exemplo uma molécula repórter ou outro ligando, covalentemente ou não covalentemente ligado com a seqüência de aminoácido, tal como uma molécula repórter que é ligada para facilitar a sua detecção e resíduos de aminoácido que não ocorrem naturalmente em relação à seqüência de aminoácido de uma proteína que ocorre naturalmente.

É considerado que a diminuição da atividade de DEL1 também pode ser obtida pela modificação do(s) domínio(s) de ligação de DNA em uma proteína DEL1 pela deleção, adição ou substituição de aminoácidos como descrito acima. Tais proteínas DEL1 modificadas também podem ser úteis nos métodos da presente invenção. Também derivados de proteínas DEL1, em que o domínio(s) de ligação de DNA é ou são modificado(s) são considerados ser úteis nos métodos da presente invenção.

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

O termo "parálogo" diz respeito a duplicações de gene dentro do genoma de uma espécie que leva aos genes parálogos. Parálogos de DEL1 podem ser facilmente identificados pela realização de uma análise BLAST contra um conjunto de seqüências da mesma espécie como a seqüência dúvida.

Ortólogos e parálogos podem ser facilmente verificados pela realização de uma chamada pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST envolvendo submeter a BLAST uma seqüência dúvida (por exemplo usando qualquer uma das seqüências listadas na Tabela A do Exemplo 1) contra qualquer base de dados de seqüência, tal como a base de dados NCBI publicamente disponível. BLASTN ou TBLASTX (usando valores padrão) são no geral usados quando se parte de uma seqüência de nucleotídeo e BLASTP ou TBLASTN (usando valores padrão) quando partindo de uma seqüência de proteína. Os resultados de BLAST podem ser opcionalmente filtrados. As seqüências de tamanho natural dos resultados filtrados ou não filtrados são depois retro BLASTed (segundo BLAST) contra seqüências do organismo a partir dos quais a seqüência dúvida é derivada (onde a seqüência dúvida é a SEQ ID NO: 7 ou SEQ ID NO: 8, o segundo BLAST portanto seria contra seqüências de Arabidopsis thaliana). Os

resultados do primeiro e segundo BLASTs são depois comparados. Um parálogo é identificado se um acerto de alta classificação do primeiro blast é da mesma espécie como da qual a seqüência dúvida é derivada, um retro BLAST depois idealmente resulta na seqüência dúvida como o acerto mais alto; um ortólogo é identificado se um acerto de alta classificação no primeiro BLAST não é da mesma espécie como da qual a seqüência dúvida é derivada e preferivelmente resulta no retro BLAST na seqüência dúvida estando entre os acertos mais altos.

Os acertos de alta classificação são aqueles tendo um valor E baixo. Quanto mais baixo o valor E, mais significante a contagem (ou em outras palavras mais baixa chance de que o acerto fosse verificado por acaso). A computação do valor E é bem conhecida na técnica. além dos valores E, as comparações também são contadas pela identidade percentual. A identidade percentual refere-se ao número de nucleotídeos idêntico (ou aminoácidos) entre as duas seqüências de ácido nucleico (ou polipeptídeo) comparadas em um comprimento particular. No caso de famílias grandes, ClustalW pode ser usado, seguido por uma árvore de unificadora vizinha, para ajudar a visualizar a formação de grupos de genes relacionados e para identificar ortólogos e parálogos.

A Tabela A do Exemplo 1 dá exemplos de ortólogos e parálogos da proteína DEL1 representados pela SEQ ID NO 8. Outros ortólogos e parálogos podem ser facilmente identificados usando o procedimento BLAST descrito acima.

Um polipeptídeo DEL1 é codificado por um ácido nucleico/ gene DEL1. Portanto o termo "ácido nucleico/gene DEL1" como aqui definido é qualquer ácido nucleico/gene que codifique um polipeptídeo DEL1 como definido acima.

Os exemplos de ácidos nucleicos DEL1 incluem mas não são limitados àqueles representados na Tabela A do Exemplo 1.

A seqüência representada pela SEQ ID NO: 20 foi até agora desconhecida. É portanto fornecida uma seqüência de ácido nucleico isolada que compreende:

(i) uma seqüência de ácido nucleico representada pela SEQ ID NO: 20, ou o seu filamento complementar;

(ii) uma seqüência de ácido nucleico que codifica uma seqüência de aminoácido como definida acima em (a) a (c);

(iii) uma seqüência de ácido nucleico capaz de hibridizar (preferivelmente sob condições severas) com uma seqüência de ácido nucleico de (i) ou (ii) acima, seqüência de hibridização esta que preferivelmente codifica uma proteína tendo atividade que aumenta rendimento e/ou resistência ao estresse;

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

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

(vi) uma seqüência de ácido nucleico que tem em ordem crescente de preferência 60 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % ou 99 % de identidade de seqüência com a SEQ ID NO: 20 ou o seu complemento; e

(vii) uma porção de uma seqüência de ácido nucleico de acordo com qualquer um de (i) a (vi) acima, porção esta que preferivelmente codifica uma proteína tendo atividade que aumenta o rendimento e/ou a resistência ao estresse.

Os ácidos nucleicos que codificam proteínas DEL1 úteis nos métodos da invenção não precisam ser ácidos nucleicos de tamanho natural, visto que o desempenho dos métodos da invenção não contam com o uso de seqüências de ácido nucleico de tamanho natural. Os exemplos de ácidos nucleicos DEL1 adequados para o uso na realização dos métodos da invenção

incluem as seqüências de ácido nucleico dadas na Tabela A do Exemplo 1, mas não são limitados a estas seqüências. As variantes de ácido nucleico também podem ser úteis na prática dos métodos da invenção. Os exemplos de tais variantes de ácido nucleico incluem porções de ácidos nucleicos que codificam uma proteína útil nos métodos da invenção, os ácido nucleicos que hibridizam com os ácidos nucleicos que codificam uma proteína útil nos métodos da invenção, variantes de união de ácidos nucleicos que codificam uma proteína útil nos métodos da invenção, variantes alélicas de ácidos nucleicos que codificam uma proteína útil nos métodos da invenção e variantes de ácidos nucleicos que codificam uma proteína útil nos métodos da invenção que são obtidos pelo embaralhamento de gene. Os termos porção, seqüência hibridizadora, variante de união, variante alélica e embaralhamento de gene serão agora descritos.

De acordo com a presente invenção, é fornecido um método para melhorar características de crescimento de plantas, que compreende introduzir e expressar em uma planta uma porção de qualquer uma das seqüências de ácido nucleico dadas na Tabela A do Exemplo 1, ou uma porção de um ácido nucleico que codifica um ortólogo, parálogo ou homólogo de qualquer uma das seqüências de aminoácido dadas na Tabela A do Exemplo 1.

O termo porção como aqui definido refere-se um pedaço de um DNA que codifica um polipeptídeo DEL1, pedaço este que codifica um polipeptídeo que carece de pelo menos parte de um dos domínios de ligação de DNA, mas compreende pelo menos a parte localizada no terminal C dos domínios de ligação de DNA e opcionalmente também uma ou mais das seqüências de assinatura definidas acima. Uma porção pode ser preparada, por exemplo, fazendo-se uma ou mais deleções a um ácido nucleico DEL1. As porções podem ser usadas na forma isolada ou elas podem ser fundidas a outras seqüências codificadoras (ou não codificadoras) de modo a, por

exemplo, produzir uma proteína que combina diversas atividades. Quando fundida a outras seqüências codificadoras, o polipeptídeo resultante produzido na tradução pode ser maior do que aquele prognosticado para o fragmento DEL1. Preferivelmente, a porção é uma porção de um ácido nucleico como representado por qualquer uma das seqüências listadas na Tabela A do Exemplo 1. Mais preferivelmente, a porção de um ácido nucleico codifica um polipeptídeo DEL1 que compreende pelo menos a parte localizada no terminal C dos domínios de ligação de DNA e opcionalmente também uma ou mais das seqüências de assinatura definidas acima, mas que carece de pelo menos parte de um domínio de ligação de DNA ou que carece de um ou ambos os domínios de ligação de DNA, tal deleção resultando na perda de atividade de ligação de DNA. O mais preferivelmente a porção de um ácido nucleico é como representado pela SEQ ID NO: 1 ou codifica um polipeptídeo representado pela SEQ ID NO: 2.

O termo "fragmento de um polipeptídeo DEL1" como usado nesta invenção (a seguir chamado de DEL1f) refere-se a uma proteína DEL1 como definida acima em que pelo menos parte de um domínio de ligação de DNA é deletado ou que carece de um ou de ambos os domínios de ligação de DNA, mas que ainda compreende a parte localizada no terminal C dos domínios de ligação de DNA e que opcionalmente também compreende uma ou mais das seqüências de assinatura como definido acima, tal deleção resultando na perda da atividade de ligação de DNA. Preferivelmente, DEL1 f carece de parte do primeiro domínio de ligação de DNA, o mais preferivelmente, DEL1 f é como representado pela SEQ ID NO: 2.

Uma outra variante de um ácido nucleico/gene de DEL1 é um ácido nucleico capaz de hibridizar sob condições de severidade reduzida, preferivelmente sob condições severas, com um ácido nucleico/gene de DEL1 como definido mais acima, seqüência de hibridização esta que codifica um polipeptídeo que compreende pelo menos a seqüência localizada no terminal

C dos domínios de ligação de DNA e preferivelmente também uma ou mais das seqüências de assinatura definidas acima. A seqüência hibridizadora tem tipicamente pelo menos 600 nucleotídeos consecutivos no comprimento, preferivelmente pelo menos 800 nucleotídeos consecutivos no comprimento, mais preferivelmente pelo menos 1000 nucleotídeos consecutivos no comprimento e o mais preferivelmente pelo menos 1200 nucleotídeos consecutivos no comprimento, os nucleotídeos consecutivos sendo de qualquer uma das seqüências de ácido nucleico dadas na Tabela A do Exemplo 1. Preferivelmente, a seqüência hibridizadora é uma que é capaz de hibridizar a um ácido nucleico que codifica proteínas como representado na Tabela A do Exemplo 1, ou a uma porção de qualquer uma das seqüências anteriormente mencionadas. O mais preferivelmente a seqüência hibridizadora é capaz de hibridizar a SEQ ID NO: 1 ou SEQ ID NO: 7.

O termo "hibridização" como aqui definido é um processo em que seqüências de nucleotídeo complementares substancialmente homólogas recozem entre si. O processo de hibridização pode ocorrer totalmente em solução, isto é, ambos os ácidos nucleicos complementares estão em solução. O processo de hibridização também pode ocorrer com um dos ácidos nucleicos complementares imobilizados a uma matriz tais como pérolas magnéticas, pérolas de Sefarose ou qualquer outra resina. O processo de hibridização além disso pode ocorrer com um dos ácidos nucleicos complementares imobilizados a um suporte sólido tal como uma membrana de nitrocelulose ou náilon ou imobilizado por exemplo, pela fotolitografia, por exemplo, a um suporte de vidro silicoso (o último conhecido como séries ou microsséries de ácido nucleico ou como lascas de ácido nucleico). De modo a permitir que a hibridização ocorra, as moléculas de ácido nucleico são no geral térmica ou quimicamente desnaturada para fundir um filamento duplo em dois filamentos simples e/ou para remover grampos de cabelo ou outras estruturas secundárias de ácidos nucleicos de filamento único.

O termo "severidade" refere-se às condições sob as quais uma hibridização ocorre. A severidade de hibridização é influenciada por condições tais como temperatura, concentração salina, força iônica e composição do tampão de hibridização. No geral, as condições de baixa severidade são selecionados para serem cerca de 30° C mais baixo do que o ponto de fusão térmica I para a seqüência específica em uma força iônica e pH definidos. As condições de severidade média são quando a temperatura está 20° C abaixo da T_m e condições de severidade alta são quando a temperatura está 10° C abaixo da T_m. As condições de hibridização de severidade alta são tipicamente usadas para isolar seqüências hibridizadoras que têm similaridade de seqüência alta com a seqüência de ácido nucleico alvo. Entretanto, os ácidos nucleicos podem desviar na seqüência e ainda codificar um polipeptídeo substancialmente idêntico, devido à degenerescência do código genético. Portanto as condições de hibridização de severidade média algumas vezes podem ser necessário para identificar tais moléculas de ácido nucleico.

A T_m é a temperatura sob força iônica e pH definidos, em que 50 % da seqüência alvo hibridiza a uma sonda perfeitamente emparelhada. 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 especificamente em temperaturas mais altas. A taxa de hibridização máxima é obtida de cerca de 16° C até 32° C abaixo da T_m. A presença de cátions monovalentes na solução de hibridização reduz a repulsão eletrostática entre os dois filamentos de ácido nucleico promovendo deste modo a formação de híbrido; este efeito é visível para as concentrações de sódio de até 0,4 M (para concentrações mais altas, este efeito pode ser ignorado). A formamida reduz a temperatura de fusão de duplexes de DNA-DNA e DNA-RNA com 0,6 a 0,7° C para cada percentual de formamida e a adição de 50 % de formamida permite que a hibridização seja realizada de 30 a 45° C, embora a taxa de

hibridização seja diminuída. Os desemparelhamentos de par de base reduz a taxa de hibridização e a estabilidade térmica dos duplexes. Em média e para sondas maiores, a T_m diminui cerca de 1° C por % de desemparelhamento de base. A T_m pode ser calculada usando as seguintes equações, dependendo dos tipos de híbridos:

(i) híbridos de DNA-DNA (Meinkoth e Wahl, Anal. Biochem., 138: 267-284, 1984): $T_m = 81,5° C + 16,6 \times \log_{10}[Na+]^a + 0,41 \times \%[G/C^b] - 500 \times [L^c]^{-1} - 0,61 \times \%$ de formamida

(ii) híbridos de DNA-RNA ou RNA-RNA:

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

(iii) híbridos de oligo-DNA ou oligo-RNA[d]:

Para <20 nucleotídeos: Tm = 2 (ln)

Para 20 a 35 nucleotídeos: Tm = 22 + 1,46 (ln)

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

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

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

[d] Oligo, oligonucleotídeo; l_n, comprimento efetivo do iniciador = 2x (nº de G/C) + (nº de NT).

A ligação não específica pode ser controlada usando qualquer uma de várias técnicas conhecidas tais como, por exemplo, bloquear a membrana com soluções contendo proteína, adições de RNA, DNA e SDS heterólogos ao tampão de hibridização e tratamento com Rnase. Para as sondas não homólogas, uma série de hibridizações pode ser realizada variando-se um de (i) diminuir progressivamente a temperatura de recozimento (por exemplo de 68° C a 42° C) ou (ii) diminuir progressivamente a concentração de formamida (por exemplo de 50 % a 0 %). O técnico habilitado está ciente de vários parâmetros que podem ser alterados

durante a hibridização e que manterá ou trocará as condições de severidade.

Além das condições de hibridização, a especificidade de hibridização tipicamente também depende da função de lavagens pós-hibridização. Para remover o fundo que resulta da hibridização não específica, as amostras são lavadas com soluções salinas diluídas. Os fatores críticos de tais lavagens incluem a força iônica e a temperatura da solução de lavagem final: quanto mais baixa a concentração salina e quanto mais alta a temperatura de lavagem, mais alta a severidade da lavagem. As condições de lavagem são tipicamente realizadas na ou abaixo da severidade de hibridização. Uma hibridização positiva dá um sinal que é pelo menos duas vezes daquele do fundo. No geral, as condições severas adequadas para os ensaios de hibridização de ácido nucleico ou procedimentos de detecção da amplificação de gene são como apresentados acima. As condições mais ou menos severas também podem ser selecionadas. O técnico habilitado está ciente de vários parâmetros que podem ser alterados durante as lavagens e que manterá ou mudará as condições de severidade.

Por exemplo, as condições de hibridização de severidade alta típicas para híbridos de DNA mais longos do que 50 nucleotídeos abrangem a hibridização a 65° C em 1 x SSC ou a 42° C em 1 x SSC e 50 % de formamida, seguido pela lavagem a 65° C em 0,3 x SSC. Os exemplos de condições de hibridização de severidade média para híbridos de DNA mais longos do que 50 nucleotídeos abrangem a hibridização a 50° C em 4x SSC ou a 40° C em 6x SSC e 50 % de formamida, seguido pela lavagem a 50° C em 2x SSC. O comprimento do híbrido é o comprimento prognosticado para o ácido nucleico de hibridização. Quando ácidos nucleicos de seqüência conhecida são hibridizados, o comprimento do híbrido pode ser determinado pelo alinhamento das seqüências e identificação das regiões conservadas aqui descritas. 1 xSSC é 0,15 M de NaCl e 15 mM de citrato de sódio; as hibridizações e lavagens podem adicionalmente incluir 5 x o reagente de

Denhardt, 0,5 a 1,0 % de SDS, 100 μg/ml de DNA de esperma de salmão desnaturado, fragmentado, 0,5 % de pirofosfato de sódio.

Para os propósitos de definir o nível de severidade, referência pode ser feita a Sambrook et al. (2001) Molecular Cloning: a laboratory manual, 3ª Edição Cold Spring Harbor Laboratory Press, CSH, Nova Iorque ou a Current Protocols in Molecular Biology, John Wiley & Sons, N. Y. (1989 e atualizações anuais).

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante de união que codifica um polipeptídeo DEL1 como aqui acima definido. O termo "variante de união" como aqui usado abrange variantes de uma seqüência de ácido nucleico em que introns e/ou exons selecionados foram excisados, substituídos ou adicionados, ou em que introns foram encurtados ou alongados. Tais variantes serão aquelas em que a atividade biológica da proteína é substancialmente retida, que pode ser obtida retendo-se seletivamente segmentos funcionais da proteína. Tais variantes de união podem ser encontrados na natureza ou podem ser feitos pelo homem. Os métodos para fabricar tais variantes de união são conhecidos na técnica (ver por exemplo Foissac e Schiex, BMC Bioinformatics. 2005; 6: 25). As variantes de união preferidas são variantes de união de um ácido nucleico que codifica um polipeptídeo DEL1 como definido acima. Mais preferidas são as variantes de união de ácidos nucleicos que codificam ortólogos ou parálogos do polipeptídeo representado pela SEQ ID NO: 8, tal como os polipeptídeos listados na Tabela A do Exemplo 1. A mais preferida é a variante de união representada pela SEQ ID NO: 7.

Uma outra variante de ácido nucleico útil na realização dos métodos da invenção é uma variante alélica de um ácido nucleico que codifica um polipeptídeo DEL1 como aqui acima definido. Os alelos ou variantes alélicas são formas alternativas de um dado gene, localizadas na mesma posição cromossômica. As variantes alélicas existem na natureza e abrangido

dentro dos métodos da presente invenção está o uso destes alelos naturais. As variantes alélicas abrangem Polimorfismos de Nucleotídeo Único (SNPs), assim como Polimorfismos de Inserção/Deleção Pequenos (INDELs). O tamanho dos INDELs é usualmente menor do que 100 pares de base. SNPs e INDELs formam o maior conjunto de variantes de seqüência em que as cepas polimórficas ocorrem naturalmente da maioria dos organismos. As variantes alélicas preferidas são variantes alélicas de ácidos nucleicos que codificam ortólogos e parálogos do polipeptídeo da SEQ ID NO: 8, tais como os polipeptídeos listados na Tabela A do Exemplo 1. A mais preferida é a variante alélica representada pela SEQ ID NO: 7.

Uma outra variante de ácido nucleico útil nos métodos da invenção é uma variante de ácido nucleico obtida pelo embaralhamento de gene. O embaralhamento de gene ou evolução dirigida também podem ser usados para gerar variantes de ácidos nucleicos que codificam proteínas DEL1 ou DEL1f como definido acima. Isto consiste de iterações de embaralhamento de DNA seguido pela triagem e/ou seleção apropriadas para gerar variantes de DEL1s de ácido nucleico ou porções destas que codificam DEL1s de polipeptídeo ou porções destes tendo uma atividade biológica modificada (Castle et al., (2004) Science 304(5674): 1151-4; patentes US 5.811.238 e 6.395.547). Preferivelmente, o ácido nucleico variante obtido pelo embaralhamento de gene codifica um polipeptídeo DEL1 ou DEL1f que compreende qualquer um ou mais dos motivos ou domínios como aqui definidos.

Além disso, variantes de ácido nucleico, incluindo os ácidos nucleicos que codificam um polipeptídeo DEL1f, também podem ser obtidas pela mutagênese direcionada ao sítio. Diversos métodos são disponíveis para se obter a mutagênese direcionada ao sítio, o mais comum sendo os métodos com base na PCR (Current Protocols in Molecular Biology. Wiley Eds.).

O ácido nucleico DEL1 ou variante deste podem ser derivados

de qualquer fonte natural ou artificial. o ácido nucleico/gene ou variante destes podem ser isolados de uma fonte microbiana, tal como fonte de levedura ou fungos, ou de uma planta, alga ou animal (incluindo ser humano). Este ácido nucleico pode ser modificado a partir da forma nativa na composição e/ou ambiente genômico através da manipulação humana deliberada. o ácido nucleico é preferivelmente de origem vegetal, seja da mesma espécie vegetal (por exemplo daquela em que o mesmo deva ser introduzido) ou seja de uma espécie vegetal diferente. O ácido nucleico pode ser isolado de uma espécie dicotiledônea, preferivelmente da família Brassicaceae, mais preferivelmente de Arabidopsis thaliana. Mais preferivelmente, o ácido nucleico DEL1 é isolado de Arabidopsis thaliana e é representado pela SEQ ID NO: 7 e a seqüência de aminoácido DEL1 é como representada pela SEQ ID NO: 8, ao passo que DEL1f é representado pela SEQ ID NO: 1 e a seqüência de aminoácido DEL1f é como representada pela SEQ ID NO: 2.

Qualquer referência aqui a uma proteína DEL1 ou DEL1f é portanto considerada significar uma proteína DEL1 ou DEL1f como definida acima. Qualquer ácido nucleico que codifica uma tal proteína DEL1 ou DEL1f é adequado para o uso na realização dos métodos da invenção.

A expressão de um ácido nucleico que codifica um polipeptídeo DEL1 ou um homólogo ou fragmento deste pode ser modulada pela introdução de uma modificação genética (preferivelmente no local de um gene DEL1). O local de um gene como aqui definido é considerado 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: ativação de T-DNA, TILLING, mutagênese direcionada ao sítio, mutagênese de transposon, evolução dirigida e recombinação homóloga ou pela introdução e expressão

em uma planta de um ácido nucleico que codifica um polipeptídeo DEL1 ou um fragmento deste. A seguir da introdução da modificação genética, segue uma etapa de selecionar quanto a expressão modificada de um ácido nucleico que codifica um polipeptídeo DEL1 ou um fragmento deste, modificação esta que na expressão dá plantas tendo rendimento aumentado e/ou resistência ao estresse aumentada.

A ativação da rotulação de T-DNA (Hayashi et al. Science (1992) 1350-1353) envolve a inserção de T-DNA, usualmente contendo um promotor (também pode ser um realçador de tradução ou um intron), na região genômica do gene de interesse ou 10 kb a montante ou a jusante da região codificadora de um gene em uma configuração tal que o promotor direcione a expressão do gene alvejado. Tipicamente, a regulagem da expressão do gene alvejado pelo seu promotor natural é rompido e o gene cai sob o controle do promotor recém introduzido. O promotor é tipicamente embutido em um T-DNA. Este T-DNA é aleatoriamente inserido no genoma vegetal, por exemplo, através da infecção por Agrobacterium e leva à super expressão de genes próximos ao T-DNA inserido. As plantas transgênicas resultantes mostram fenótipos dominantes devido à super expressão de genes próximo ao promotor introduzido. O promotor a ser introduzido pode ser qualquer promotor capaz de direcionar a expressão de um gene no organismo desejado, neste caso um planta. Por exemplo, os promotores constitutivos, preferidos de tecido, preferidos do tipo de célula e promotores indutíveis são todos adequados para o uso na ativação de T-DNA.

Uma modificação genética também pode ser introduzida no local de um gene DEL1 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 mutagenizadas de um ácido nucleico DEL1 capaz de exibir atividade DEL1 ou DEL1f. TILLING também permite a seleção de plantas que carregam tais variantes

mutantes. Estas variantes mutantes podem exibir ainda atividade de DEL1 ou DEL1f mais alta do que aquela exibida pelo gene na sua forma natural. TILLING combina mutagênese de alta densidade com métodos de triagem de alto rendimento. As etapas tipicamente seguidas em TILLING são: (a) mutagênese EMS (Redei GP e Koncz C (1992) Em Methods in Arabidopsis Research, Koncz C, Chua NH, Schell J, eds. Singapura, World Scientific Publishing Co, pp. 16-82; Feldmann et al., (1994) Em Meyerowitz EM, Somerville CR, eds, Arabidopsis. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 137-172; Lightner J e Caspar T (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 reunião de indivíduos; (c) amplificação pela PCR de uma região de interesse; (d) desnaturação e recozimento para permitir a formação de heteroduplexes; (e) DHPLC, onde a presença de um heteroduplex em uma reunião é detectada como um pico extra no cromatograma; (f) identificação do indivíduo mutante; e (g) seqüenciamento do produto de PCR mutante. Os métodos para TILLING são bem conhecidos na técnica (McCallum et al., (2000) Nat Biotechnol 18: 455-457; revisado por Stemple (2004) Nat Rev Genet 5(2): 145-50).

A mutagênese de transposon é uma técnica de mutagênese com base na inserção de transposons em genes, que freqüentemente resulta no silenciamento de gene. A técnica foi usada para diversas espécies vegetais, incluindo arroz (Greco et al., Plant Physiol, 125, 1175-1177, 2001), milho (McCarty et al., Plant J. 44, 52-61, 2005) e Arabidopsis (Parinov e Sundaresan, Curr. Opin. Biotechnol. 11, 157-161, 2000).

A ativação de T-DNA, TILLING, mutagênese direcionada ao sítio, mutagênese de transposon e evolução dirigida são exemplos de tecnologias que permitem a geração de novos alelos e variantes DEL1, e/ou a modulação da sua expressão.

A recombinação homóloga permite a introdução em um

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

Um método preferido para introduzir uma modificação genética (que neste caso não precisa estar no local de um gene DEL1) é introduzir e expressar em uma planta um ácido nucleico que codifica um polipeptídeo DEL1 ou um fragmento deste, como definido acima. O ácido nucleico a ser introduzido em uma planta pode ser um ácido nucleico de tamanho natural ou pode ser uma porção ou uma seqüência hibridizadora como mais acima definido.

De acordo com um aspecto preferido da presente invenção, a expressão aumentada do ácido nucleico DEL1 ou variante deste é considerada. Os métodos para aumentar a expressão de genes ou produtos de gene são bem documentados na técnica e incluem, por exemplo, a super expressão direcionada pelos promotores apropriados, o uso de realçadores de transcrição ou realçadores de tradução. Se por exemplo a expressão aumentada de um gene endógeno é desejada, ácidos nucleicos isolados que servem como elementos promotores ou realçadores podem ser introduzidos em uma posição apropriada (tipicamente a montante) de uma forma não heteróloga de um polinucleotídeo de modo a super regular a expressão de um ácido nucleico DEL1 ou variante deste. Por exemplo, promotores endógenos podem ser alterados in vivo pela mutação, deleção e/ou substituição (ver, Kmiec, Pat. U.S. Nº 5.565.350; Zarling et al., PCT/US93/03868), ou promotores isolados podem ser introduzidos em uma célula vegetal na

orientação e na distância apropriadas de um gene da presente invenção de modo a controlar a expressão do gene.

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

Uma seqüência de intron também pode ser adicionada à região não traduzida ou a seqüência codificadora da seqüência codificadora parcial para aumentar a quantidade da mensagem madura que acumula no citossol. A inclusão de um intron encaixável na unidade de transcrição em construções de expressão tanto vegetal quanto animal mostrou aumentar a expressão de gene tanto ao nível de mRNA quanto de proteína até 1000 vezes, Buchman e Berg, Mol. Cell Biol. 8: 4395-4405 (1988); Callis et al., Genes Dev. 1: 1183-1200 (1987). Tal realce de intron da expressão de gene é tipicamente o máximo quando colocado próximo à extremidade 5' da unidade de transcrição. O uso dos introns do milho Adh1-S intron 1, 2 e 6, o intron Bronze-1 são conhecidos na técnica. Ver no geral, The Maize Handbook, Capítulo 116, Freeling e Walbot, Eds., Springer, N.Y. (1994).

De acordo com um outro aspecto preferido da presente invenção, a expressão diminuída (uma redução ou eliminação substancial) de uma seqüência de ácido nucleico DEL1 é considerada.

Referência aqui a "redução ou eliminação substancial" é considerada significar uma diminuição nos níveis de expressão de gene e/ou polipeptídeo endógenos e/ou atividade de polipeptídeo em relação às plantas de controle. A redução ou eliminação substancial é em ordem crescente de

preferência de pelo menos 10 %, 20 %, 30 %, 40 % ou 50 %, 60 %, 70 %, 80 %, 85 %, 90 %, ou 95 %, 96 %, 97 %, 98 %, 99 % ou mais reduzida comparada com aquela das plantas de controle.

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

Para a redução ou eliminação substancial da expressão de um gene endógeno em uma planta, um comprimento suficiente de nucleotídeos substancialmente contíguos de uma seqüência de ácido nucleico é requerida. De modo a realizar o silenciamento de gene, isto pode ser tão pouco quanto 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10 ou menos nucleotídeos, alternativamente isto pode ser tanto quanto o gene inteiro (incluindo as UTRs 5' e/ou 3', em parte ou no todo). O trecho de nucleotídeos substancialmente contíguos pode ser derivado da SEQ ID NO: 1 ou SEQ ID NO: 7, ou de qualquer uma das seqüências de ácido nucleico dadas na Tabela A do Exemplo 1, ou de qualquer ácido nucleico capaz de codificar um ortólogo, parálogo ou homólogo de qualquer uma das seqüências de aminoácido dadas na Tabela A. Uma seqüência de ácido nucleico que codifica um polipeptídeo (funcional) não é uma exigência para os vários métodos aqui debatidos para a redução ou eliminação substancial da expressão de um gene endógeno.

Esta redução ou eliminação substancial da expressão podem ser obtidas usando ferramentas e técnicas de rotina. Um método preferido para a redução ou eliminação substancial da expressão de gene endógeno é pela introdução e expressão em uma planta de uma construção genética na

qual o ácido nucleico (neste caso um trecho de nucleotídeos substancialmente contíguos derivados da SEQ ID NO: 1 ou SEQ ID NO: 7, ou de qualquer uma das seqüências de ácido nucleico dadas na Tabela A do Exemplo 1, ou de qualquer ácido nucleico capaz de codificar um ortólogo, parálogo ou homólogo de qualquer uma das seqüências de aminoácido dadas na Tabela A) é clonado como uma repetição invertida (em parte ou completamente), separado por um espaçador (DNA não codificador).

Em um tal método preferido, a expressão do gene endógeno é reduzida ou substancialmente eliminada através do silenciamento mediado pelo RNA usando uma repetição invertida de um ácido nucleico ou uma parte deste (neste caso um trecho de nucleotídeos substancialmente contíguos derivados da SEQ ID NO: 1 ou SEQ ID NO: 7, ou de qualquer uma das seqüências de ácido nucleico dadas na Tabela A do Exemplo 1, ou de qualquer ácido nucleico capaz de codificar um ortólogo, parálogo ou homólogo de qualquer uma das seqüências de aminoácido dadas na Tabela A), preferivelmente capaz de formar uma estrutura de grampo de cabelo. A repetição invertida é clonada em um vetor de expressão que compreende as seqüências de controle. Uma seqüência de ácido nucleico de DNA não codificador (um espaçador, por exemplo um fragmento da região de ligação de matriz (MAR), um intron, um poliligador, etc.) está localizado entre os dois ácidos nucleicos invertidos para formar a repetição invertida. Depois da transcrição da repetição invertida, um RNA quimérico com uma estrutura auto-complementar é formado (parcial ou completo). Esta estrutura de RNA de filamento duplo é aludida como o RNA de grampo de cabelo (hpRNA). O hpRNA é processado pela planta em siRNAs que são incorporados em um complexo silenciador induzido por RNA (RISC).

O RISC cliva ainda os transcritos de mRNA, reduzindo substancialmente deste modo o número de transcritos de mRNA a serem traduzidos em polipeptídeos. Para outros detalhes gerais ver por exemplo,

Grierson et al. (1998) WO 98/53083; Waterhouse et al. (1999) WO 99/53050).

O desempenho dos métodos da invenção não contam com a introdução e expressão em uma planta de uma construção genética em que o ácido nucleico é clonado como uma repetição invertida, mas qualquer um ou mais de diversos métodos de "silenciamento de gene" bem conhecidos podem ser usados para se obter os mesmos efeitos.

Um tal método para a redução da expressão de gene endógeno é o silenciamento mediado pelo RNA da expressão de gene (regulação negativa). O silenciamento neste caso é deflagrado em uma planta por uma seqüência de RNA de filamento duplo (dsRNA) que é substancialmente similar ao gene endógeno alvo. Este dsRNA é ainda processado pela planta em cerca de 20 a cerca de 26 nucleotídeos chamados de RNAs interferentes curtos (siRNAs). Os siRNAs são incorporados em um complexo silenciador induzido por RNA (RISC) que cliva o transcrito de mRNA do gene alvo endógeno, reduzindo substancialmente deste modo o número de transcritos de mRNA a serem traduzidos em um polipeptídeo. Preferivelmente, a seqüência de RNA de filamento duplo corresponde a um gene alvo.

Um outro exemplo de um método de silenciamento de RNA envolve a introdução das seqüências de ácido nucleico ou partes destas (neste caso um trecho de nucleotídeos substancialmente contíguos derivados da SEQ ID NO: 1 ou SEQ ID NO: 7, ou de qualquer uma das seqüências de ácido nucleico dadas na Tabela A do Exemplo 1, ou de qualquer ácido nucleico capaz de codificar um ortólogo, parálogo ou homólogo de qualquer uma das seqüências de aminoácido dadas na Tabela A) em uma orientação de sentido em uma planta. "Orientação de sentido" refere-se a uma seqüência de DNA que é homóloga a um transcrito de mRNA desta. Seria portanto introduzida em uma planta pelo menos uma cópia da seqüência de ácido nucleico. A seqüência de ácido nucleico adicional reduzirá a expressão do gene endógeno,

dando origem a um fenômeno conhecido como co-supressão. A redução da expressão de gene será mais pronunciada se diversas cópias adicionais de uma seqüência de ácido nucleico são introduzidas na planta, visto que existe uma correlação positiva entre os níveis de transcrito altos e a deflagração da co-supressão.

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

As seqüências de ácido nucleico de anti-sentido podem ser planejadas de acordo com as regras de emparelhamento de base de Watson e Crick. A seqüência de ácido nucleico de anti-sentido pode ser complementar à seqüência de ácido nucleico inteira (neste caso um trecho de nucleotídeos substancialmente contíguos derivados da SEQ ID NO: 1 ou SEQ ID NO: 7, ou de qualquer uma das seqüências de ácido nucleico dadas na Tabela A do Exemplo 1, ou de qualquer ácido nucleico capaz de codificar um ortólogo, parálogo ou homólogo de qualquer uma das seqüências de aminoácido dadas na Tabela A), mas também podem ser um oligonucleotídeo que seja de anti-

sentido apenas a uma parte da seqüência de ácido nucleico (incluindo as UTRs 5' e 3' do mRNA). Por exemplo, a seqüência de oligonucleotídeo de anti-sentido pode ser complementar à região que circunda o sítio de início de tradução de um transcrito de mRNA que codifica um polipeptídeo. O comprimento de uma seqüência de oligonucleotídeo de anti-sentido adequada é conhecida na técnica e pode começar de cerca de 50, 45, 40, 35, 30, 25, 20, 15 ou 10 nucleotídeos no comprimento ou menos. Uma seqüência de ácido nucleico de anti-sentido de acordo com a invenção pode ser construída usando a síntese química e reações de ligação enzimática usando métodos conhecidos na técnica. Por exemplo, uma seqüência de ácido nucleico de anti-sentido (por exemplo, uma seqüência de oligonucleotídeo de anti-sentido) pode ser quimicamente sintetizado usando nucleotídeos que ocorrem naturalmente ou nucleotídeos variadamente modificados planejados para aumentar a estabilidade biológica das moléculas ou para aumentar a estabilidade física do duplex formado entre as seqüências de ácido nucleico de anti-sentido e sentido, por exemplo, derivados de fosforotioato e nucleotídeos substituídos por acridina podem ser usados. Os exemplos de nucleotídeos modificados que podem ser usados para gerar as seqüências de ácido nucleico de anti-sentido são bem conhecidos na técnica. As modificações de nucleotídeo conhecidas incluem metilação, ciclização e 'encapuzamentos' e substituição de um ou mais dos nucleotídeos que ocorrem naturalmente com um análogo tal como inosina. Outras modificações de nucleotídeos são bem conhecidas na técnica.

A seqüência de ácido nucleico de anti-sentido pode ser biologicamente produzido usando um vetor de expressão no qual uma seqüência de ácido nucleico foi subclonado em uma orientação de anti-sentido (isto é, RNA transcrito a partir do ácido nucleico inserido será de uma orientação de anti-sentido a um ácido nucleico alvo de interesse). Preferivelmente, a produção de seqüências de ácido nucleico de anti-sentido em plantas ocorre por meio de uma construção de ácido nucleico

estavelmente integrada compreendendo um promotor, um oligonucleotídeo de anti-sentido operavelmente ligado e um terminador.

As moléculas de ácido nucleico usadas para silenciar nos métodos da invenção (sejam introduzidas em uma planta ou geradas in situ) hibridizam com ou se ligam a transcritos de mRNA e/ou DNA genômico que codifica um polipeptídeo para inibir deste modo a expressão da proteína, por exemplo, pela inibição da transcrição e/ou tradução. A hibridização pode ser pela complementaridade de nucleotídeo convencional para formar um duplex estável, ou, por exemplo, no caso de uma seqüência de ácido nucleico de anti-sentido que se liga aos duplexes de DNA, através de interações específicas na ranhura maior da hélice dupla. As seqüências de ácido nucleico de anti-sentido podem ser introduzidas em uma planta pela transformação ou injeção direta em um sítio de tecido específico. Alternativamente, as seqüências de ácido nucleico de anti-sentido podem ser modificadas para alvejar as células selecionadas e depois sistemicamente administradas. Por exemplo, para a administração sistêmica, as seqüências de ácido nucleico de anti-sentido podem ser modificadas tal que elas se liguem especificamente aos receptores ou antígenos expressados em uma superfície de célula selecionada, por exemplo, pela ligação da seqüência de ácido nucleico de anti-sentido aos peptídeos ou anticorpos que se ligam aos receptores ou antígenos de superfície celular. As seqüências de ácido nucleico de anti-sentido também podem ser liberadas às células usando os vetores aqui descritos.

De acordo com um outro aspecto, a seqüência de ácido nucleico de anti-sentido é uma seqüência de ácido nucleico α-anomérica. Uma seqüência de ácido nucleico α-anomérica forma híbridos de filamento duplo específicos com o RNA complementar em que, ao contrário das unidades b usuais, os filamentos correm paralelos entre si (Gaultier et al. (1987) Nucl Ac Res 15: 6625-6641). A seqüência de ácido nucleico de anti-sentido também pode compreender um 2'-o-metilribonucleotídeo (Inoue et al.

(1987) Nucl Ac Res 15, 6131-6148) ou um análogo de RNA-DNA quimérico (Inoue et al. (1987) FEBS Lett. 215, 327-330).

A redução ou eliminação substancial da expressão de gene endógeno também pode ser realizada usando ribozimas. As ribozimas são moléculas de RNA catalíticas com atividade de ribonuclease que são capazes de clivar uma seqüência de ácido nucleico de filamento único, tal como um mRNA, ao qual as mesmas têm uma região complementar. Assim, as ribozimas (por exemplo, ribozimas cabeça de martelo (descritas em Haselhoff e Gerlach (1988) Nature 334, 585-591) podem ser usadas para clivar cataliticamente transcritos de mRNA que codificam um polipeptídeo, reduzindo substancialmente deste modo o número de transcritos de mRNA a serem traduzidos em um polipeptídeo. Uma ribozima tendo especificidade quanto a uma seqüência de ácido nucleico pode ser planejada (ver por exemplo: Cech et al. Patente U.S. Nº 4.987.071; e Cech et al. Patente U.S. Nº 5.116.742). Alternativamente, os transcritos de mRNA que correspondem a uma seqüência de ácido nucleico podem ser usados para selecionar um RNA catalítico tendo uma atividade de ribonuclease específica de uma reunião de moléculas de RNA (Bartel e Szostak (1993) Science 261, 1411-1418). O uso de ribozimas para o silenciamento de gene em plantas é conhecido na técnica (por exemplo, Atkins et al. (1994) WO 94/00012; Lenne et al. (1995) WO 95/03404; Lutziger et al. (2000) WO 00/00619; Prinsen et al. (1997) WO 97/13865 e Scott et al. (1997) WO 97/38116).

O silenciamento de gene também podem ser obtido pela mutagênese de inserção (por exemplo, inserção de T-DNA ou inserção de transposon) ou pelas estratégias como descritas, entre outras, por Angell e Baulcombe ((1999) Plant J 20(3): 357-62), (Amplicon VIGS WO 98/36083), ou Baulcombe (WO 99/15682). Portanto, de acordo com um outro aspecto preferido da invenção, é fornecido um método para aumentar o rendimento e/ou aumentar a resistência ao estresse de uma planta pela regulação negativa

da expressão de um gene endógeno DEL1 na mutagênese de inserção.

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

Um outro método para o silenciamento de gene é pelo alvejamento de seqüências de ácido nucleico complementares à região reguladora do gene (por exemplo, o promotor e/ou realçadores) para formar estruturas helicoidais triplas que impedem a transcrição do gene em células alvo. Ver Helene, C., Anticancer Drug Res. 6, 569-84, 1991; Helene et al., Ann. N. Y. Acad. Sci. 660, 27-36 1992; e Maher, L. J. Bioassays 14, 807-15, 1992.

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

Alternativamente, um programa de triagem pode ser adaptado até identificar em uma população de planta variantes naturais de um gene, variantes estas que codificam polipeptídeos com atividade reduzida. Tais variantes naturais também podem ser usadas por exemplo, para realizar a recombinação homóloga.

MicroRNAs artificiais e/ou naturais (miRNAs) podem ser usados para silenciar a expressão de gene e/ou a tradução de mRNA. Os miRNAs endógenos são RNAs de filamento único pequenos tipicamente de 19 a 24 nucleotídeos de comprimento. Estes funcionam primariamente para regular a expressão de gene e/ou tradução de mRNA. A maioria dos microRNAs (miRNAs) vegetais têm complementaridade perfeita ou quase perfeita com as suas seqüências alvos. Entretanto, existem alvos naturais com até cinco desemparelhamentos. Estes são processados a partir de RNAs não codificadores mais logos com estruturas dobradas para trás características pelas Rnases específicas de filamento duplo da família Dicer. No processamento, estes são incorporados no complexo silenciador induzido por RNA (RISC) pela ligação ao seu componente principal, uma proteína Argonaute. MiRNAs servem como os componentes de especificidade de RISC, visto que eles emparelham em base para alvejar ácidos nucleicos, principalmente mRNAs, no citoplasma. Os eventos regulatórios subsequentes incluem a clivagem do mRNA alvo e a destruição e/ou inibição traducional. Os efeitos da super expressão de miRNA são assim freqüentemente refletidos nos níveis de mRNA diminuídos de genes alvos.

Os microRNAs artificiais (amiRNAs), que têm tipicamente 21 nucleotídeos no comprimento, podem ser geneticamente engendrados especificamente para regular negativamente a expressão de gene de genes únicos ou múltiplos de interesse. Os determinantes da seleção alvo de microRNA vegetal são bem conhecidos na técnica. Os parâmetros empíricos para o reconhecimento alvejado foram definidos e podem ser usados para ajudar no planejamento de amiRNAs específicos, Schwab R, 2005. As ferramentas convenientes para o planejamento e geração de amiRNAs e seus precursores também estão disponíveis para o público, Schwab et al., 2006.

Para o desempenho ótimo, as técnicas de silenciamento de gene usadas para reduzir a expressão em uma planta de um gene endógeno

requer o uso de seqüências de ácido nucleico de plantas monocotiledôneas para a transformação de plantas monocotiledôneas e de plantas dicotiledôneas para a transformação de plantas dicotiledôneas. Preferivelmente, uma seqüência de ácido nucleico de qualquer espécie vegetal dada é introduzida nesta mesma espécie. Por exemplo, uma seqüência de ácido nucleico de arroz é transformada em uma planta de arroz. Entretanto, não é uma exigência absoluta que a seqüência de ácido nucleico a ser introduzida origine-se da mesma espécie vegetal como a planta em que a mesma será introduzida. É suficiente que exista homologia substancial entre o gene alvo endógeno e o ácido nucleico a ser introduzido.

São descritos acima exemplos de vários métodos para a redução ou eliminação substancial de expressão em uma planta de um gene endógeno. Uma pessoa habilitada na técnica facilmente seria capaz de adaptar os métodos anteriormente mencionados para o silenciamento de modo a obter redução da expressão de um gene endógeno em uma planta inteira ou em partes desta através do uso de um promotor apropriado, por exemplo.

As construções genéticas intencionadas no silenciamento da expressão de gene podem compreender a seqüência de nucleotídeo DEL1, por exemplo como representado pela SEQ ID NO: 7 (ou uma ou mais porções desta) em uma orientação de sentido e/ou anti-sentido em relação à seqüência promotora. As cópias de sentido ou anti-sentido de pelo menos parte do gene endógeno na forma de repetição direta ou invertida podem ser utilizadas nos métodos de acordo com a invenção. As características de crescimento de plantas e em particular o aumento de rendimento e/ou aumento na resistência ao estresse, também podem ser melhoradas pela introdução em uma planta de pelo menos parte de uma versão de anti-sentido de uma seqüência de nucleotídeo DEL1, por exemplo representado pela SEQ ID NO: 7. Deve estar claro que parte do ácido nucleico (uma porção) poderia alcançar o resultado desejado. Os genes de anti-sentido homólogos são preferidos aos genes de

anti-sentido heterólogos, os genes homólogos sendo genes vegetais da mesma espécie vegetal em que a construção silenciadora é introduzida e genes heterólogos sendo os genes de espécie vegetal relacionada ou não relacionada.

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

Portanto, é fornecido uma construção de gene que compreende:

(i) um ácido nucleico que codifica um polipeptídeo DEL1f, ou variante deste como aqui acima definido;

(ii) uma ou mais seqüências de controle capazes de direcionar a expressão da seqüência de ácido nucleico de (i); e opcionalmente

(iii) uma seqüência de terminação de transcrição.

As construções úteis nos métodos de acordo com a presente invenção podem ser construídas usando a tecnologia de DNA recombinante bem conhecidas pelas pessoas habilitadas na técnica. As construções de ácido nucleico podem ser inseridas em vetores, que podem ser comercialmente disponíveis, adequados para transformar em plantas e adequado para a expressão do ácido nucleico de interesse nas células transformadas.

As plantas são transformadas com um vetor que compreende a seqüência de interesse (isto é, um ácido nucleico que codifica o polipeptídeo DEL1f ou um fragmento equivalente de uma proteína DEL1, ou uma seqüência de ácido nucleico planejada para a redução ou eliminação substancial da expressão de um gene endógeno DEL1). O técnico habilitado está bem ciente dos elementos genéticos que devem estar presentes no vetor de modo a transformar, selecionar e propagar com êxito as células hospedeiras contendo a seqüência de interesse. A seqüência de interesse é operavelmente ligada a uma ou mais seqüências de controle (pelo menos a um promotor). Os termos "elemento regulador", "seqüência de controle" e

"promotor" são todos usados intercambiavelmente aqui e devem ser considerados em um contexto amplo para se referir às seqüências de ácido nucleico regulatórias capazes de efetuar a expressão das seqüências às quais eles estão ligados. O termo "promotor" tipicamente refere-se a uma seqüência de ácido nucleico de controle localizada a montante do início transcricional de um gene e que está envolvida no reconhecimento e ligação da RNA polimerase e outras proteínas, direcionando deste modo a transcrição de um ácido nucleico operavelmente ligado. Abrangidos pelos termos anteriormente mencionados são seqüências reguladoras transcricionais derivados de um gene genômico eucariótico clássico (incluindo o TATA box que é requerido para o início de transcrição exata, com ou sem uma seqüência de CCAAT box) e elementos reguladores adicionais (isto é, seqüências a ativadoras, realçadores e silenciadores a montante) que alteram a expressão de gene em resposta aos estímulos desenvolvimentais e/ou externos, ou em uma maneira específica de tecido. Também incluído dentro do termo está uma seqüência reguladora transcricional de um gene procariótico clássico, caso este em que o mesmo pode incluir uma seqüência -35 box e/ou seqüências reguladoras transcricionais -10 box. O termo "elemento regulador" também abrange uma molécula de fusão sintética ou derivado que confere, ativa ou realça a expressão de uma molécula de ácido nucleico em uma célula, tecido ou órgão. O termo "operavelmente ligado" como aqui usado refere-se a uma ligação funcional entre a seqüência promotora e o gene de interesse, tal que a seqüência promotora é capaz de iniciar a transcrição do gene de interesse.

Vantajosamente, qualquer tipo de promotor pode ser usado para direcionar a expressão da seqüência de ácido nucleico. O termo "promotor" refere-se a uma seqüência de ácido nucleico de controle localizada a montante do início transcricional de um gene e que está envolvida no reconhecimento e ligação da RNA polimerase e outras proteínas, direcionando deste modo a transcrição de um ácido nucleico

operavelmente ligado. Um promotor "vegetal" compreende elementos reguladores, que medeiam a expressão de um segmento da seqüência codificadora em células de planta. Conseqüentemente, um promotor vegetal não precisa ser de origem vegetal, mas pode originar-se de vírus ou microorganismos, por exemplo de vírus que atacam células vegetais. O "promotor vegetal" também pode originar-se de uma célula vegetal, por exemplo, da planta que é transformada com a seqüência de ácido nucleico a ser expressada no processo da invenção e aqui descrito. Isto também se aplica a outros sinais reguladores "vegetais", tais como terminadores vegetais. Os promotores a montante das seqüências de nucleotídeo úteis nos métodos da presente invenção podem ser modificados por uma ou mais substituições, inserções e/ou deleções de nucleotídeo sem interferir com a funcionalidade ou atividade de seus promotores, a matriz de leitura aberta (ORF) ou a região reguladora 3' tal como terminadores ou outras regiões reguladoras 3' que estão localizadas fora da ORF. É além disso possível que a atividade dos promotores seja aumentada pela modificação da sua seqüência, ou que os mesmos sejam substituídos completamente por promotores mais ativos, mesmo promotores de organismos heterólogos. Para a expressão em plantas, a molécula de ácido nucleico, como descrito acima, deve ser ligada operavelmente a ou compreender um promotor adequado que expresse o gene no ponto certo no tempo e com o padrão de expressão espacial requerido.

O promotor pode ser um promotor constitutivo, que refere-se a um promotor que é transcricionalmente ativo durante a maioria, mas não necessariamente todas, das fases do seu crescimento e desenvolvimento e sob a maioria das condições ambientais, em pelo menos uma célula, tecido ou órgão. Alternativamente, o promotor pode ser um promotor indutível, isto é, tendo início de transcrição induzido ou aumentado em resposta a um produto químico (para uma revisão ver Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48: 89-108), estímulos ambientais ou físicos. Um outro exemplo

de um promotor indutível é um promotor indutível por estresse, isto é, um promotor ativado quando uma planta é exposta às várias condições de estresse, ou um promotor induzido por patógeno.

Adicional ou alternativamente, o promotor pode ser um promotor específico de órgão ou específico de tecido, isto é, um que é capaz de preferencialmente iniciar a transcrição em certos órgãos ou tecidos, tais como as folhas, raízes, tecido de semente etc; ou o promotor pode ser um promotor ubíquo, que é ativo em substancialmente todos os tecidos ou células de um organismo, ou o promotor pode ser desenvolvimentalmente regulado, sendo deste modo ativo durante certos estágios desenvolvimentais ou em partes da planta que passa pelas mudanças desenvolvimentais. Os promotores capazes de iniciar a transcrição em certos órgãos ou tecidos são aludidos aqui apenas como "específicos de órgão" ou "específicos de tecido" respectivamente, Similarmente, os promotores capazes de iniciar a transcrição em certas células são aqui aludidas apenas como "específicos de célula".

Preferivelmente, o ácido nucleico DEL1f (isto é, o ácido nucleico que codifica um polipeptídeo DEL1f) ou variante deste é operavelmente ligado a um promotor específico de semente. Um promotor específico de semente é transcricionalmente ativo de modo predominante em tecido de semente, mas não de modo necessariamente exclusivo em tecido de semente (em casos de expressão avariada). O promotor específico de semente pode ser ativo durante o desenvolvimento da semente e/ou durante a germinação. Os promotores específicos de semente são bem conhecidos na técnica. Preferivelmente, o promotor específico de semente é o promotor representado pela SEQ ID NO: 5 ou um promotor funcionalmente equivalente. Deve estar claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico DEL1f representado pela SEQ ID NO: 1, nem é a aplicabilidade da invenção restrita à expressão de um ácido nucleico DEL1f quando direcionado por um promotor específico de semente. Os exemplos de

outros promotores específicos de semente que também podem ser usados para direcionar a expressão de um ácido nucleico DEL1 são mostrados na Tabela 2 abaixo. Outros exemplos de promotores específicos de semente são dados em Qing Qu e Takaiwa (Plant Biotechnol. J. 2, 113-125, 2004), divulgação esta que é aqui incorporada por referência como se totalmente apresentada.

Tabela 2: Exemplos de promotores específicos de semente

Fonte de Gene	Referência
Genes específicos de semente	Simon et al., Plant Mol. Biol. 5: 191, 1985; Scofield et al., J. Biol. Chem. 262: 12202, 1987.; Baszczynski et al., Plant Mol. Biol. 14: 633, 1990.
Albumina de Castanha do Pará	Pearson et al., Plant Mol. Biol. 18: 235-245, 1992.
Legumina	Ellis et al., Plant Mol. Biol. 10: 203-214, 1988.
Glutelina (arroz)	Takaiwa et al., Mol. Gen. Genet. 208: 15-22, 1986; Takaiwa et al., FEBS Letts. 221: 43-47, 1987.
Zeína	Matzke et al Plant Mol Biol, 14(3): 323-32 1990
NapA	Stalberg et al, Planta 199: 515-519, 1996.
TrigoLMWeHMW Glutenina-1	Mol Gen Genet 216: 81-90, 1989; NAR 17: 461-2, 1989
trigo SPA	Albani et al, Plant Cell, 9: 171-184, 1997
α, β, γ-gliadinas do trigo	EMBO J. 3: 1409-15, 1984
Promotor ltr1 da cevada	Diaz et al. (1995) Mol Gen Genet 248(5): 592-8
B1, C, D, hordeína da cevada	Theor Appl Gen 98: 1253-62, 1999; Plant J 4: 343-55, 1993; Mol Gen Genet 250: 750-60, 1996
DOF de cevada	Mena et al, A Plant Journal, 116(1): 53-62, 1998
blz2	EP99106056.7
Promotor sintético	Vicente-Carbajosa et al., Plant J. 13: 629-640, 1998.
Prolamina NRP33 do Arroz	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
α-globulina Glb-1 do arroz	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
OSH1 do arroz	Sato et al, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996
α-globulina REB/OHP-1 do arroz	Nakase et al. Plant Mol. Biol. 33: 513-522, 1997
ADP-glicose do Arroz	Trans Res 6: 157-68, 1997
Pirofosforilase	
Família de ESR gene do milho	Plant J 12: 235-46, 1997
α-cafirina do sorgo	DeRose et al., Plant Mol. Biol 32: 1029-35, 1996
KNOX	Postma-Haarsma et al, Plant Mol. Biol. 39: 257-71, 1999

Oleosina do Arroz	Wu et al, J. Biochem. 123: 386, 1998
Oleosina do Girassol	Cummins et al., Plant Mol. Biol. 19: 873-876, 1992
PRO0117,putativoarroz40S Proteína ribossômica	WO 2004/070039
PRO0136, alanina amino-transferase do arroz	não publicada
PRO0147, inibidor de tripsina ITR1 (cevada)	não publicada
PRO0151, arroz WSI18	WO 2004/070039
PRO0175, arroz RAB21	WO 2004/070039
PRO005	WO 2004/070039
PRO0095	WO 2004/070039
α-amilase (Amy32b)	Lanahan et al, Plant Cell 4: 203-211, 1992; Skriver et al, Proc Natl Acad Sci USA 88: 7266-7270, 1991
Gene equivalente à catepsina p	Cejudo et al, Plant Mol Biol 20: 849-856, 1992
Cevada Ltp2	Kalla et al., Plant J. 6: 849-60, 1994
Chi26	Leah et al., Plant J. 4: 579-89, 1994
Milho B-Peru	Selinger et al., Genetics 149;1125-38,1998

Em uma outra forma de realização, o ácido nucleico DEL1 ou variante deste, é operavelmente ligado a um promotor constitutivo como definido acima. Preferivelmente, o promotor constitutivo tem um padrão de expressão comparável como o promotor CaMV 35S, mais preferivelmente, o promotor constitutivo tem o mesmo padrão de expressão como o promotor CaMV 35S, o mais preferivelmente, o promotor constitutivo é o promotor CaMV 35S. Deve estar claro que a aplicabilidade da presente invenção não é restrita ao ácido nucleico DEL1f representado pela SEQ ID NO: 1, nem é a aplicabilidade da invenção restrita à expressão de um ácido nucleico DEL1' quando direcionado por um promotor constitutivo. Os exemplos de outros promotores constitutivos que também podem ser usados para direcionar a expressão de um ácido nucleico DEL1 são mostrados na Tabela 3 abaixo.

Tabela 3: Exemplos de promotores constitutivos

Fonte de Gene	Referência
Actina	McElroy et al, Plant Cell, 2: 163-171, 1990
CAMV 35S	Odell et al, Nature, 313: 810-812, 1985
CaMV 19S	Nilsson et al., Physiol. Plant. 100: 456-462, 1997
GOS2	de Pater et al, Plant J Nov; 2(6): 837-44, 1992, WO 2004/065596
Ubiquitina	Christensen et al, Plant Mol. Biol. 18: 675-689, 1992
Ciclofilina do Arroz	Buchholz et al, Plant Mol Biol. 25(5): 837-43, 1994
Histona H3 do milho	Lepetit et al, Mol. Gen. Genet. 231: 276-285, 1992
Histona H3 da Alfafa	Wu et al. Plant Mol. Biol. 11: 641-649, 1988
Actina 2	An et al, Plant J. 10(1); 107-121, 1996

34S FMV	Sanger et al., Plant. Mol. Biol., 14, 1990: 433-443
Subunidade pequena de Rubisco	US 4.962.028
OCS	Leisner (1988) Proc Natl Acad Sci USA 85(5): 2553
SAD1	Jain et al., Crop Science, 39 (6), 1999: 1696
SAD2	Jain et al., Crop Science, 39 (6), 1999: 1696
Nos	Shaw et al. (1984) Nucleic Acid Res. 12(20): 7831-7846
V-ATPase	WO 01/14572
Super promotor	WO 95/14098
Proteínas G-box	WO 94/12015

Para a identificação de promotores funcionalmente equivalentes, a força de promotor e/ou padrão de expressão de um promotor candidato podem ser analisados por exemplo ligando-se operavelmente o promotor a um gene repórter e ensaiando o nível e o padrão de expressão do gene repórter em vários tecidos da planta. Os genes repórteres bem conhecidos adequados incluem por exemplo beta-glicuronidase ou beta galactosidase. A atividade de promotor é ensaiada medindo-se a atividade enzimática da beta-glicuronidase ou beta-galactosidase. A força de promotor e/ou padrão de expressão pode ser depois comparada com aquela de um promotor de referência (tal como aquele usado nos métodos da presente invenção). Alternativamente, a força de promotor pode ser ensaiada quantificando-se os níveis de mRNA ou comparando-se os níveis de mRNA do ácido nucleico usado nos métodos da presente invenção, com os níveis de mRNA de genes de manutenção interna tais como 18S rRNA, usando métodos conhecidos na técnica, tais como Northern blotting com análise densitométrica de autorradiogramas, PCR em tempo real quantitativa ou RT-PCR (Heid et al., 1996 Genome Methods 6: 986-994). No geral por "promotor fraco" é intencionado um promotor que direcione a expressão de uma seqüência codificadora em um nível baixo. Por "nível baixo" é intencionado em níveis de cerca de 1/10.000 transcritos a cerca de 1/100.000 transcritos, a cerca de 1/500,0000 transcritos por célula.

Ao contrário, um "promotor forte" direciona a expressão de uma seqüência codificadora em nível alto, ou a cerca de 1/10 transcritos a cerca de 1/100 transcritos a cerca de 1/1.000 transcritos por célula.

Opcionalmente, uma ou mais seqüências terminadoras também podem ser usadas na construção introduzida em uma planta. O termo "terminador" abrange uma seqüência de controle que é uma seqüência de DNA no final de uma unidade transcricional que sinaliza o processamento e a poliadenilação em 3' de um transcrito primário e a terminação de transcrição. O terminador pode ser derivado do gene natural, a partir de uma variedade de outros genes vegetais, ou de T-DNA. O terminador a ser adicionado pode ser derivado, por exemplo, dos genes da nopalina sintase ou octopina sintase, ou alternativamente de um outro gene vegetal, ou menos preferivelmente de qualquer outro gene eucariótico. Os elementos reguladores adicionais podem incluir realçadores transcricionais assim como traducionais. Aqueles habilitados na técnica estarão cientes de seqüências terminadoras e realçadoras que podem ser adequadas para o uso na realização da invenção. Tais seqüências seriam conhecidas ou podem ser facilmente obtidas por uma pessoa habilitada na técnica.

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

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

Para a detecção da transferência bem sucedida das seqüências de ácido nucleico como usada nos métodos da invenção e/ou seleção de

plantas transgênicas que compreendem estes ácidos nucleicos, é vantajoso usar genes marcadores (ou genes repórteres). Portanto, a construção genética pode opcionalmente compreender um gene marcador selecionável. Como aqui usado, os termos "marcador selecionável", "gene marcador selecionável" ou "gene repórter" incluem qualquer gene que confere um fenótipo em uma célula em que a mesma é expressada para facilitar a identificação e/ou seleção de células que são transfectadas ou transformadas com uma construção de ácido nucleico da invenção. Estes genes marcadores permitem a identificação de uma transferência bem sucedida das moléculas de ácido nucleico por intermédio de uma série de princípios diferentes. Os marcadores adequados podem ser selecionados de marcadores que conferem resistência aos antibióticos ou herbicidas, que introduzem um novo traço metabólico ou que permite a seleção visual. Os exemplos de genes marcadores selecionáveis incluem genes que conferem resistência aos antibióticos (tal como nptl1 que fosforila neomicina e canamicina, ou hpt, que fosforila higromicina, ou genes que conferem resistência, por exemplo, à bleomicina, estreptomicina, tetraciclina, cloranfenicol, ampicilina, gentamicina, geneticina (G418), espectinomicina ou blasticidina), aos herbicidas (por exemplo bar que fornece resistência a Basta; aroA ou gox que fornecem resistência contra glifosato, ou os genes que conferem resistência, por exemplo, à imidazolinona, fosfinotricina ou sulfoniluréia), ou genes que fornecem um traço metabólico (tal como manA que permite que as plantas usem a manose como a única fonte de carbono ou a xilose isomerase para a utilização de xilose, ou marcadores antinutritivos tais como a resistência à 2-desoxiglicose). A expressão de genes marcadores visuais resulta na formação de cor (por exemplo p-glicuronidase, GUS ou β-galactosidase com seus substratos coloridos, por exemplo X-Gal), luminescência (tal como o sistema de luciferina/luciferase) ou fluorescência (Proteína Fluorescente Verde, GFP e seus derivados). Esta lista representa apenas um número pequeno de

marcadores possíveis. O técnico habilitado está familiarizado com tais marcadores. Marcadores diferentes são preferidos, dependendo do organismo e do método de seleção.

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

Visto que os genes marcadores, particularmente os genes para a resistência aos antibióticos e herbicidas, não são mais requeridos ou são indesejados na célula hospedeira transgênica uma vez que os ácidos nucleicos tenham sido introduzidos com êxito, 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 de tal método é aquele que é conhecido como co-transformação. O método de co-transformação utiliza dois vetores simultaneamente para a transformação, um vetor carregando o ácido nucleico de acordo com a invenção e um segundo

carregando o(s) gene(s) marcador(es). Uma proporção grande de transformantes recebe ou, no caso de plantas, compreende (até 40 % ou mais dos transformantes), ambos os vetores. No caso de transformação com Agrobacteria, os transformantes usualmente recebe apenas uma parte do vetor, isto é, a seqüência flanqueada pelo T-DNA, que usualmente representa o cassete de expressão. Os genes marcadores podem ser subseqüentemente removidos da planta transformada pela realização de cruzamentos. Em um outro método, os genes marcadores integrados em um transposon são usados para a transformação junto com o ácido nucleico desejado (conhecido como a tecnologia Ac/Ds). Os transformantes podem ser cruzados com uma fonte de transposase ou os transformantes são transformados com uma construção de ácido nucleico que conferem expressão de uma transposase, transitoriamente ou estável. Em alguns casos (aprox. 10 %), o transposon pula para fora do genoma da célula hospedeira uma vez que a transformação tenha ocorrido com êxito e é perdido. Em um outro número de casos, o transposon pula para uma localização diferente. Nestes casos o gene marcador deve ser eliminado pela realização de cruzamentos. Em microbiologia, técnicas foram desenvolvidas que tornam possível, ou facilita, a detecção de tais eventos. Um outro método vantajoso conta com o que é conhecido como sistemas de recombinação; cuja vantagem é que a eliminação pelo cruzamento pode ser dispensada. 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 de loxP. Se o gene marcador é integrado entre as seqüências loxP, é removido uma vez que a transformação tenha ocorrido com êxito, pela expressão da recombinase. Outros sistemas de recombinação são o sistema HIN/HIX, FLP/FRT e REP/STB (Tribble et al., J. Biol. Chem., 275, 2000: 22255-22267; Velmurugan et al., J. Cell Biol., 149, 2000: 553-566). Uma integração específica de sítio no genoma vegetal das seqüências de ácido nucleico de acordo com a invenção é possível. Naturalmente, estes

métodos também podem ser aplicados aos microorganismos tais como levedura, fungos ou bactérias.

A presente invenção também abrange plantas ou partes destas (incluindo sementes) obtenível pelos métodos de acordo com a presente invenção. A presente invenção portanto fornece plantas obteníveis pelo método de acordo com a presente invenção, plantas estas ou partes de planta que têm introduzido nesse sentido um ácido nucleico DELlf ou variante deste.

A invenção também fornece um método para a produção de plantas transgênicas tendo rendimento aumentado e/ou resistência ao estresse aumentada, que compreende a introdução e expressão em uma planta de um ácido nucleico DEL1f ou uma variante deste.

Para os propósitos da invenção, "transgênica", "transgene" ou "recombinante" significa com respeito, por exemplo, a uma seqüência de ácido nucleico, um cassete de expressão, construção de gene ou um vetor que compreende a seqüência de ácido nucleico ou um organismo transformado com as seqüências de ácido nucleico, cassetes de expressão ou vetores de acordo com a invenção, todas estas construções realizadas pelos métodos recombinantes em que

a) as seqüências de ácido nucleico que codificam proteínas úteis nos métodos da invenção, ou

b) seqüência(s) de controle genético que é(são) operavelmente ligada(s) com a seqüência de ácido nucleico de acordo com a invenção, por exemplo um promotor, ou

c) a) e b)

não são localizados em seu ambiente genético natural ou foi modificado pelos métodos recombinantes, sendo possível para a modificação tomar a forma de, por exemplo, uma substituição, adição, deleção, inversão ou inserção de um ou mais resíduos de nucleotídeo. O ambiente genético natural é entendido como significando o local genômico ou cromossômico

natural na planta original ou a presença em um biblioteca genômica. No caso de uma biblioteca genômica, o ambiente genético natural da seqüência de ácido nucleico é preferivelmente retido, pelo menos em parte. O ambiente flanqueia a seqüência de ácido nucleico pelo menos em um lado e tem comprimento de seqüência de pelo menos 50 pares de base, preferivelmente pelo menos 500 pares de base, de modo especialmente preferível pelo menos 1000 pares de base, o mais preferivelmente pelo menos 5000 pares de base. Um cassete de expressão que ocorrem naturalmente - por exemplo a combinação que ocorre naturalmente do promotor natural das seqüências de ácido nucleico com a seqüência de ácido nucleico correspondente que codifica um polipeptídeo útil nos métodos da presente invenção, como definido acima - torna-se um cassete de expressão transgênico quando este cassete de expressão é modificado pelos métodos não naturais, sintéticos ("artificiais") tais como, por exemplo, tratamento mutagênico. Os métodos adequados são descritos, por exemplo, na US 5.565.350 ou WO 00/15815.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas tendo rendimento aumentado e/ou resistência ao estresse aumentada, método este que compreende:

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

(ii) cultivar a célula vegetal sob condições que promovam o crescimento e desenvolvimento de vegetais.

O ácido nucleico pode ser introduzido diretamente em uma célula vegetal ou na própria planta (incluindo a introdução em um tecido, órgão ou qualquer outra parte de uma planta). De acordo com uma característica preferida da presente invenção, o ácido nucleico é preferivelmente introduzido em uma planta pela transformação.

A transferência de genes estranhos no genoma de uma planta é chamada de transformação. A transformação de espécie vegetal é agora uma

técnica completamente de rotina. Vantajosamente, qualquer um de diversos métodos de transformação pode ser usado para introduzir o gene de interesse em uma célula ancestral adequada. Os métodos descritos para a transformação e regeneração de plantas a partir de tecidos vegetais ou células vegetais podem ser utilizados para a transformação transitória ou estável. Os métodos de transformação incluem o uso de lipossomas, eletroporação, produtos químicos que aumentam a captação de DNA livre, injeção do DNA diretamente na planta, bombardeamento de pistola de partícula, transformação usando vírus ou pólen e microprojeção. Os métodos podem ser selecionados do método do método do cálcio/polietileno glicol para protoplastos (Krens, F. A. et al., (1982) Nature 296, 7274; Negrutiu I et al. (1987) Plant Mol Biol 8: 363-373); eletroporação de protoplastos (Shillito R. D. et al. (1985) Bio/Technol 3, 1099-1102); microinjeção em material vegetal (Crossway A et al., (1986) Mol. Gen Genet 202: 179-185); bombardeamento de partícula revestida com DNA ou RNA (Klein TM et al., (1987) Nature 327: 70) infecção com vírus (não integrativa) e outros. As plantas transgênicas, incluindo plantas transgênicas de safra, são preferivelmente produzidas por intermédio da transformação mediada por Agrobacterium. Um método de transformação vantajoso é a transformação em planta. Para esta finalidade, é possível, por exemplo, permitir que as agrobactérias atuem sobre as sementes da planta ou inoculem o meristema da planta com agrobactérias. Mostrou-se particularmente conveniente de acordo com a invenção permitir que uma suspensão de agrobactérias transformadas atuem sobre a planta intacta ou pelo menos na flor primordial. A planta é subseqüentemente cultivada até que as sementes da planta tratada sejam obtidas (Clough e Bent, Plant J. (1998) 16, 735-743). Os métodos para a transformação mediada por Agrobacterium de arroz incluem métodos bem conhecidos para a transformação de arroz, tais como aqueles descritos em qualquer uma das seguintes: Pedido de patente européia EP 1198985 A1, Aldemita e Hodges (Planta 199: 612-617, 1996);

Chan et al. (Plant Mol Biol 22 (3): 491-506, 1993), Hiei et al. (Plant J 6 (2): 271-282, 1994), divulgações estas que são aqui incorporadas por referência como se completamente apresentadas. No caso de transformação de milho, o método preferido é como descrito em Ishida et al. (Nat. Biotechnol 14(6): 745-50, 1996) ou Frame et al. (Plant Physiol 129(1): 13-22, 2002), divulgações estas que são aqui incorporadas por referência como se completamente apresentadas. Os ditos métodos são ainda descritos por via de exemplo em B. Jenes et al., Techniques for Gene Transfer, em: Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S. D. Kung e R. Wu, Academic Press (1993) 128-143 e em Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205-225). Os ácidos nucleicos ou a construção a ser expressada são preferivelmente clonados em um vetor, que é adequado para transformar Agrobacterium tumefaciens, por exemplo pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984) 8711). As agrobactérias transformadas por um tal vetor podem ser depois usadas na maneira conhecida para a transformação de plantas, tais como plantas usadas como um modelo, como Arabidopsis (Arabidopsis thaliana está dentro do escopo da presente invenção não considerada como um planta de safra), ou plantas de safra tais como, por via de exemplo, as plantas de tabaco, por exemplo pela imersão das folhas machucadas ou as folhas picadas em uma solução agrobacteriana e depois cultivando-as em meio adequado. A transformação de plantas por meio de Agrobacterium tumefaciens é descrita, por exemplo, por Hofgen e Willmitzer em Nucl. Acid Res. (1988) 16, 9877 ou é conhecido 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.

Além da transformação de células somáticas, que depois devem ser regeneradas em plantas intactas, também é possível transformar as células de meristemas vegetais e em particular aquelas células que

desenvolvem em gametas. Neste caso, os gametas transformados seguem o desenvolvimento de planta natural, dando origem às plantas transgênicas. Assim, por exemplo, as sementes de Arabidopsis são tratadas com agrobactérias e sementes são obtidas a partir das plantas em desenvolvimento 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, Singapura, pp. 274-289]. Os métodos alternativos são com base na remoção repetida das inflorescências e incubação do sítio de excisão no centro da roseta com as agrobactérias transformadas, por meio do qual as sementes transformadas podem ser igualmente obtidas em um ponto tardio no tempo (Chang (1994). Plant J. 5: 551-558; Katavic (1994). Mol Gen Genet, 245: 363-370). Entretanto, um método especialmente eficaz é o método de infiltração a vácuo com as suas modificações tais como o método da "imersão floral". No caso de infiltração a vácuo de Arabidopsis, as plantas intactas sob pressão reduzida são tratadas com uma suspensão agrobacteriana [Bechthold, N (1993). C R Acad Sci Paris Life Sci, 316: 1194-1199], enquanto que no caso do método da "imersão floral" o tecido floral em desenvolvimento é ligeiramente incubado com uma suspensão agrobacteriana tratada com tensoativo [Clough, SJ e Bent, AF (1998). The Plant J. 16, 735-743]. Uma certa proporção de sementes transgênicas são colhidas em ambos os casos e estas sementes podem ser distinguidas das sementes não transgênicas pelo cultivo sob as condições seletivas descritas acima. Além disso, a transformação estável de plastídeos é de vantagens porque os plastídeos são herdados maternalmente na maioria das safras reduzindo ou eliminando o risco de fluxo de transgene através do pólen. A transformação do genoma de cloroplasto é no geral obtida por um processo que foi esquematicamente demonstrado em Klaus et al., 2004 [Nature Biotechnology 22 (2), 225-229]. Em resumo, as seqüências a serem transformadas são clonadas juntas com um

gene marcador selecionável entre seqüências homólogas flanqueadoras ao genoma de cloroplasto. Estas seqüências homólogas flanqueadoras direcionam a integração específica de sítio no plastoma. A transformação plastídica foi descrita para muitas espécies vegetais diferentes e uma vista geral é dada em Bock (2001) Transgenic plastids in basic research and plant biotechnology. J Mol Biol. 21 de setembro de 2001; 312 (3): 425-38 ou Maliga, P (2003) Progress towards commercialization of plastid transformation technology. Trends Biotechnol. 21, 20-28. Outros progressos biotecnológicos foram recentemente relatados na forma de transformantes plastídicos isentos de marcador, que podem ser produzidos por um gene marcador co-integrado transitório (Klaus et al., 2004, Nature Biotechnology 22(2), 225-229).

As células vegetais geneticamente modificadas podem ser regeneradas por intermédio de todos os métodos com que o técnico habilitado está familiarizado. Os métodos adequados podem ser encontrados nas publicações supracitadas por S. D. Kung e R. Wu, Potrykus ou Hofgen e Willmitzer.

No geral depois da transformação, as células vegetais ou agrupamentos de célula são selecionados quanto a presença de um ou mais marcadores que são codificados pelos genes expressíveis em planta co-transferidos com o gene de interesse, a seguir do que o material transformado é regenerado em uma planta inteira. Para selecionar as plantas transformadas, o material vegetal obtido na transformação é, como uma regra, submetido às condições seletivas de modo que as plantas transformadas podem ser distinguidas de plantas não transformadas. Por exemplo, as sementes obtidas na maneira descrita acima pode ser plantada e, depois de um período de crescimento inicial, submetida a uma seleção adequada por pulverização. Uma outra possibilidade consiste em cultivar as sementes, se apropriado depois da esterilização, em placas de ágar usando um agente de seleção

adequado de modo que apenas as sementes transformadas possam crescer em plantas. Alternativamente, as plantas transformadas são triadas quanto à presença de um marcador selecionável tal como aqueles descritos acima.

A seguir da transferência e regeneração de DNA, as plantas putativamente transformadas também podem ser avaliadas, por exemplo, usando a análise de Southern, quanto a presença do gene de interesse, número de cópia e/ou organização genômica. Alternativa ou adicionalmente, os níveis de expressão do DNA recém introduzido pode ser monitorado usando as análises de Northern e/ou Western, ambas as técnicas sendo bem conhecidas pelas pessoas tendo habilidade comum na técnica.

As plantas transformadas geradas pode ser propagadas por uma variedade de meios, tais como pela propagação clonal ou técnicas de cruzamento clássicas. Por exemplo, uma planta transformada de primeira geração (ou T1) pode ser auto-polinizada e transformantes de segunda geração (ou T2) homozigotos selecionados e as plantas T2 podem ser depois ainda propagadas através das técnicas de cruzamento clássicas.

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

A presente invenção claramente estende-se a qualquer célula vegetal ou planta produzidas por qualquer um dos métodos aqui descritos e a todas as partes de plantas e seus propágulos. A presente invenção estende-se ainda a abranger a progênie de uma célula primária transformada ou transfectada, tecido, órgão ou planta inteira que foram produzidos por qualquer um dos métodos anteriormente mencionados, a única exigência sendo que a progênie exibe a(s) mesma(s) característica(s) genotípica(s) e/ou

fenotípica(s) como aquelas produzidas pelo precursor nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucleico que codifica um proteína DEL1 isolado como aqui acima definida. As células hospedeiras preferidas de acordo com a invenção são células vegetais.

As plantas hospedeiras para os ácidos nucleicos ou o vetor usado no método de acordo com a invenção, o cassete de expressão ou a construção ou vetor são, em princípio, vantajosamente todas as plantas, que são capazes de sintetizar os polipeptídeos usados no método da invenção.

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

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

amido ou proteínas.

A presente invenção também abrange o uso de mutantes vegetais com expressão de um gene DEL1 diminuída para melhorar as características de crescimento de plantas, em particular no aumento do rendimento e/ou aumento da resistência ao estresse.

A presente invenção além disso abrange o uso de ácidos nucleicos DEL1 ou variantes destes e o uso de polipeptídeos DEL1 ou homólogos destes.

Um tal uso diz respeito a melhorar as características de crescimento de plantas, em particular no aumento de rendimento e/ou aumento da resistência ao estresse. O aumento de rendimento compreende pelo menos um de biomassa aumentada, peso total aumentado de sementes e número aumentado de sementes cheias.

Os ácidos nucleicos DEL1 ou variantes destes, ou polipeptídeos DEL1 ou homólogos destes podem encontrar uso em programas de cruzamento em que um marcador de DNA é identificado que pode ser geneticamente ligado a um gene DEL1 ou variante deste. Os ácidos nucleicos/genes DEL1 ou variantes destes, ou polipeptídeos DEL1 ou homólogos destes podem ser usados para definir um marcador molecular. Este marcador de DNA ou proteína pode ser depois usado em programas de cruzamento para selecionar plantas tendo rendimento aumentado e/ou resistência ao estresse aumentada. O gene DEL1 ou variante deste, por exemplo, pode ser um ácido nucleico como representado por qualquer uma das seqüências listadas na Tabela A do Exemplo 1.

As variantes alélicas de um ácido nucleico/gene de DEL1 também podem encontrar uso em programas de cruzamento auxiliados por marcador. Tais programas de cruzamento algumas vezes requerem a introdução de variação alélica pelo tratamento mutagênico das plantas, usando por exemplo a mutagênese de EMS; alternativamente, o programa pode

começar com uma coleção de variantes alélicas da assim chamada origem "natural" causada de modo não intencional. A identificação de variantes alélicas depois ocorre, por exemplo, pela PCR. Isto é seguido por uma etapa para a seleção de variante alélicas superiores da seqüência em questão e que dá rendimento aumentado e/ou resistência ao estresse aumentada. A seleção é tipicamente realizada pelo monitoramento do desempenho de crescimento de plantas contendo variantes alélicas diferentes da seqüência em questão, por exemplo, variantes alélicas diferentes de qualquer uma das seqüências listadas na Tabela A do Exemplo 1. O desempenho de crescimento pode ser monitorado em uma estufa ou no campo. Outras etapas opcionais incluem cruzamento de plantas, em que a variante alélica superior foi identificada, com uma outra planta. Isto poderia ser usado, por exemplo, para fabricar uma combinação de características fenotípicas interessantes.

Um ácido nucleico DEL1 ou variante deste também podem ser usados como sondas para mapear genética e fisicamente os genes dos quais eles são uma parte e como marcadores para traços ligados a estes genes. Tal informação pode ser útil no cruzamento de planta de modo a desenvolver linhas com fenótipos desejados. Tal uso de ácidos nucleicos DEL1 ou variantes destes requer apenas uma seqüência de ácido nucleico de pelo menos 15 nucleotídeos no comprimento. os ácidos nucleicos DEL1 ou variantes destes podem ser usados como marcadores de do polimorfismo de comprimento de fragmento de restrição (RFLP). Southern blots (Sambrook J, Fritsch EF e Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico vegetal digerido com restrição podem ser sondados com os ácidos nucleicos DEL1 ou variantes destes. Os padrões de faixa resultantes podem ser depois submetidos às análises genéticas usando programas de computador tais como MapMaker (Lander et al. (1987) Genomics 1: 174-181) de modo a construir um mapa genético. Além disso, os ácidos nucleicos podem ser usados para sondar Southern blots contendo DNAs genômicos

tratados com endonuclease de restrição de um conjunto de indivíduos que representam precursor e progênie de um cruzamento genético definido. A segregação dos polimorfismos de DNA é mencionado e usado para calcular a posição do ácido nucleico DEL1 ou variante deste no mapa genético anteriormente obtido usando esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32: 314-331).

A produção e uso de sondas derivadas de gene vegetal para o uso no mapeamento genético são descritos em Bernatzky e Tanksley (1986) Plant Mol. Biol. Repórter 4: 37-41. Numerosas publicações descrevem o mapeamento genético de clones de cDNA específicos usando a metodologia esboçada acima ou variações destes. Por exemplo, populações de intercruzamento F2, populações de retrocruzamento, populações aleatoriamente acasaladas, linhagens quase isogênicas e outros conjuntos de indivíduos podem ser usados para o mapeamento. Tais metodologias são bem conhecidas por aqueles habilitados na técnica.

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

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

Uma variedade de métodos de com base na amplificação de ácido nucleico para o mapeamento genético e físico pode ser realizada usando os ácidos nucleicos. Os exemplos incluem a amplificação específica de alelo

(Kazazian (1989) J. Lab. Clin. Med 11: 95-96), polimorfismo de fragmentos amplificados pela PCR (CAPS; Sheffield et al. (1993) Genomics 16: 325-332), ligação específica de alelo (Landegren et al. (1988) Science 241: 1077-1080), reações de extensão de nucleotídeo (Sokolov (1990) Nucleic Acid Res. 18: 3671), Mapeamento de Híbrido por Radiação (Walter et al. (1997) Nat. Genet. 7: 22-28) e Mapeamento Happy (Dear e Cook (1989) Nucleic Acid Res. 17: 6795-6807). Para estes métodos, a seqüência de um ácido nucleico é usada para planejar e produzir pares de iniciador para o uso na reação de amplificação ou em reações de extensão de iniciador. O planejamento de tais iniciadores é bem conhecido por aqueles habilitados na técnica. Em métodos que utilizam o mapeamento genético com base na PCR, pode ser necessário identificar diferenças de seqüência de DNA entre os precursores do cruzamento de mapeamento na região que corresponde à seqüência de ácido nucleico iminente. Isto, entretanto, no geral não é necessário para os métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas tendo rendimento aumentado e/ou resistência ao estresse aumentada, como descrito mais acima. Estas características de crescimento vantajosas também podem ser combinadas com outros traços economicamente vantajosos, tais como outros traços que realcem o rendimento, tolerância adicional aos vários estresses, traços que modificam várias características arquiteturais e/ou características bioquímicas e/ou fisiológicas.

Descrição das figuras

A presente invenção será agora descrita com referência às seguintes figuras em que:

A Fig. 1 mostra a estrutura de domínio típica de polipeptídeos DEL1. As proteínas DEL1 compreendem 2 domínios E2F TDP (Pfam PF02319, Interpro IPR003316), que representam um domínio de ligação de DNA de hélice com asa. Os domínios são indicados em negrito.

A Fig. 2 mostra um alinhamento múltiplo de várias proteínas DEL1. Os asteriscos representam aminoácidos idênticos nas várias seqüências, os cólons representam substituições altamente conservadas e os pontos representam substituições menos conservadas.

A Fig. 3 mostra um vetor binário p037, para a expressão em Oryza sativa de uma seqüência codificadora DEL1f de Arabidopsis thaliana sob o controle de um promotor específico de semente (referência interna P R00218).

A Fig. 4 detalha exemplos de seqüências úteis na realização dos métodos de acordo com a presente invenção.

A Fig. 5 representa as medições de biomassa por planta em resposta às doses de UV-B aumentadas. Plantas de 18 dias de idade foram irradiadas com as doses indicadas de UV-B. A biomassa de plantas foram medidas 8 dias depois dos tratamentos com UV-B. Os valores indicam a média ± SE. DEL1.2 e DEL1.4 correspondem às linhagens que super expressam DEL1 #2 e #4, respectivamente. DEL1 KO corresponde à linhagem silenciada del1-1.

A Fig. 6 mostra as medições de fotoprodutos do dímero de ciclobutilpirimidina (CPD). O nível total de CPDs foi medido para 5ª folha de plantas de 18 dias de idade diretamente depois de 5 horas de tratamento UV-B (A), ou depois de uma exposição adicional à luz branca, permitindo fotorreparar o DNA (B). os níveis de CPD são plotados relativamente ao nível medido nas plantas Col-0.

Exemplos

A presente invenção será agora descrita com referência aos seguintes exemplos, que são apenas por via de ilustração. Os seguintes exemplos não são intencionados a definir completamente ou de outro modo limitar o escopo da invenção.

Exemplo 1: Identificação de seqüências relacionadas com a SEQ ID NO:

7 e SEQ ID NO: 8

As seqüências (cDNA de tamanho natural, ESTs ou genômico) relacionadas com a SEQ ID NO: 7 e/ou seqüências de proteína relacionadas com a SEQ ID NO: 8 foram identificadas entre aquelas mantidas na base de dados da Entrez Nucleotides no National Center for Biotechnology Information (NCBI) usando as ferramentas de pesquisa de seqüência em base de dados, tais como o Basic Local Alignment Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215: 403-410; e Altschul et al. (1997) Nucleic Acid Res. 25: 3389-3402). O programa é usado para descobrir regiões de similaridade local entre as seqüências comparando-se as seqüências de ácido nucleico ou polipeptídeo com as bases de dados de seqüência e calculando-se a significância estatística de emparelhamentos. O polipeptídeo codificado pela SEQ ID NO: 8 foi usado para o algoritmo TBLASTN, com ajustes de referência e o filtro para ignorar seqüências de complexidade baixa desligado. Os resultados da análise foi visualizado pela comparação aos pares e classificados de acordo com a contagem de probabilidade (valor E), onde as contagens refletem a probabilidade d que um alinhamento particular ocorra ao acaso (quanto mais baixo o valor E, mais significante o acerto). Além dos valores E, as comparações foram também classificadas pela identidade percentual. A identidade percentual refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas seqüências de ácido nucleico (ou polipeptídeo) em um comprimento particular. Em alguns casos, os parâmetros de penalidade podem ser ajustados para modificar a severidade da procura.

Além das seqüências de ácido nucleico publicamente disponíveis na NCBI, as bases de dados de seqüência patenteada também são pesquisadas seguindo o mesmo procedimento como aqui descrito acima.

A Tabela A fornece uma lista de seqüências de ácido nucleico e proteína relacionadas com a seqüência de ácido nucleico como representada pela SEQ ID NO: 7 e a seqüência de proteína representada pela SEQ ID NO:

8, ver também a Fig. 4.

Tabela A: Seqüências de ácido nucleico relacionadas com a seqüência de ácido nucleico (SEQ ID NO: 7) úteis nos métodos da presente invenção e os polipeptídeos deduzidos correspondentes.

Nome	Organismo Fonte	Ácido Nucleico SEQ ID NO:	Poli-peptídeo SEQ ID NO:	Número de Acesso na Base de Dados	Situação
DEL1	Arabidopsis thaliana	7	8	NM 114685	Tamanho natural
OsDEL1	Oryza sativa	14	15	XM 467698	Tamanho natural
ZmDEL1	Zea mays	16	17	AY107996	Tamanho natural
OtDEL1	Ostreococcus tauri	18	19	AY675104	Tamanho natural
Equivalente a MtDEL1	Medicago truncatula	20	21	/	Tamanho natural
TaE2Fe	Triticum aestivum	22	23	DQ353854	Tamanho natural
Equivalente a OsDEL1	Oryza sativa	24	25	AK120032	Tamanho natural
Equivalente a SIDEL1	Silene latifolia	26	27	DV768235	parcial
Equivalente a HDEL1	Helianthus sp.	28	29	EE619641	parcial
Equivalente a GmDEL1	Glycine max	30	31	Gm59592851	Tamanho natural

Exemplo 2: Alinhamento de seqüências de polipeptídeo relevantes

AlignX do Vetor NTI (Invitrogen) está fundamentado no algoritmo popular Clustal de alinhamento progressivo (Thompson et al. (1997) Nucleic Acid Res 25: 4876-4882; Chenna et al. (2003). Nucleic Acid Res 31: 3497-3500). Os valores básicos são para a penalidade de abertura de intervalo de 10, para a penalidade de extensão de intervalo de 0,1 e a matriz de peso selecionada é Blosum 62 (se polipeptídeos são alinhados).

O resultado de um alinhamento de seqüência múltiplo usando polipeptídeos relevantes na identificação daqueles úteis na realização dos métodos da invenção é mostrado na Figura 2. Os dois domínios E2F TDP conservados podem ser facilmente identificados.

Exemplo 3: Cálculo da identidade percentual global entre seqüências de

polipeptídeo úteis na realização dos métodos da invenção

As porcentagens globais de similaridade e identidade entre seqüências de polipeptídeo de tamanho natural úteis na realização dos métodos da invenção foram determinadas usando um dos métodos disponíveis na técnica, o software MatGAT (Matrix Global Alignment Tool) (BMC Bioinformatics. 2003 4: 29. MatGAT: a application that generates similarity/identity matrices using protein ou DNA sequences. Campanella JJ, Bitincka L, Smalley J; software hospedado por Ledion Bitincka). O software MatGAT gera matrizes de similaridade/identidade para as seqüências de DNA ou proteína sem necessitar pré-alinhamento dos dados. O programa realiza uma série de alinhamentos aos pares usando o algoritmo de alinhamento global de Myers e Miller (com uma penalidade de abertura de intervalo de 12 e uma penalidade de extensão de intervalo de 2), calcula a similaridade e a identidade usando por exemplo Blosum 62 (para polipeptídeos) e depois coloca os resultados em uma matriz de distância. A similaridade de seqüência é mostrada na metade do fundo da linha divisória e a identidade de seqüência é mostrada na metade do topo da linha divisória diagonal.

Os parâmetros usados na comparação foram:

Matriz de Contagem: Blosum62

Primeiro Intervalo: 12

Intervalo de Extensão: 2

Os resultados da análise de software são mostrados na Tabela B para a similaridade e identidade globais em relação às seqüências de polipeptídeo de tamanho natural (excluindo as seqüências de polipeptídeo parciais). A identidade percentual é dada acima da diagonal e a similaridade percentual é dada abaixo da diagonal.

A identidade percentual entre as seqüências de polipeptídeo úteis na realização dos métodos da invenção pode ser tão baixa quanto 27,5 % de identidade de aminoácido comparada com a SEQ ID NO: 8.

Tabela B: Resultados de MatGAT para a similaridade e identidade globais em relação ás seqüências de polipeptídeo de tamanho natural.

	1	2	3	4	5	6	7	8	9
1. SEQ ID 8		42,9	44,2	27,5	51,5	45,6	43,4	36,2	35,7
2. SEQ ID 15	60,8		63,6	28,6	45,4	63,7	60,2	29,6	31,4
3. SEQ ID 17	63,1	76,4		29,5	45,1	74,8	69,4	30,7	31,8
4. SEQ ID 19	46,5	44,5	45,1		30,6	28,8	28,4	21,9	24,0
5. SEQ ID 21	69,7	60,8	61,5	44,1		46,8	43,1	36,4	37,3
6. SEQ ID 23	63,0	75,1	85,2	42,7	62,8		71,8	31,0	32,3
7. SEQ ID 25	61,9	75,3	82,6	43,1	60,5	82,4		30,4	32,9
8. SEQ ID 27	42,7	35,6	38,0	29,4	43,4	38,2	37,2		35,6
9. SEQ ID 29	47,6	41,7	43,2	35,6	49,6	42,4	43,1	53,3	

Exemplo 4: Identificação de domínios compreendidos nas seqüências de polipeptídeo úteis na realização dos métodos da invenção

A base de dados Integrated Resource of Protein Families, Domains and Sites (InterPro) é uma interface integrada para as bases de dados de assinatura habitualmente usadas para as pesquisas com base em texto e seqüência. A base de dados InterPro combina estas bases de dados, que usam metodologias diferentes e graus variáveis de informação biológica a cerca de proteínas bem caracterizadas para derivar assinaturas de proteína. As bases de dados colaboradoras incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. A Interpro é hospedada no European Bioinformatics Institute no Reino Unido.

Os resultados da varredura com InterPro da seqüência de polipeptídeo como representada pela SEQ ID NO: 8 são apresentadas na Tabela C.

Tabela C: Resultados de varredura com InterPro da seqüência de polipeptídeo como representada pela SEQ ID NO: 8

Base de dados	Número de Acesso	Nome do Acesso
InterPro	IPR003316	E2F TDP

Exemplo 5: Prognóstico de topologia das seqüências de polipeptídeo úteis na realização dos métodos da invenção (localização subcelular, transmembrana...)

TargetP 1.1 prognostica a localização subcelular de proteínas

eucarióticas. A designação da localização está fundamentada na presença prognosticada de qualquer uma das pré-seqüências de terminal N: peptídeo de trânsito de cloroplasto (cTP), peptídeo alvejador de mitocondria (mTP) ou peptídeo de sinal de caminho secretor (SP). As contagens sobre as quais o prognóstico final está fundamentado não são realmente probabilidades e as mesmas não necessariamente adicionam a uma. Entretanto, a localização com a contagem mais alta é o mais provável de acordo com TargetP e as relações entre as contagens (a classe de confiabilidade) podem ser uma indicação de quão certo o prognóstico é. A classe de confiabilidade (RC) varia de 1 a 5, onde 1 indica o prognóstico mais forte. TargetP é mantido no servidor da Technical University of Denmark.

Para as seqüências prognosticadas conterem uma pré seqüência de terminal N um sítio de clivagem potencial também pode ser prognosticado.

Vários parâmetros foram selecionados, tais como grupo de organismo (não vegetal ou vegetal), ajustes de corte (nenhum, ajuste pré definido de cortes, ou ajuste especificado pelo usuário de cortes) e o cálculo de prognóstico de sítios de clivagem (sim ou não).

Os resultados da análise de TargetP 1.1 da seqüência de polipeptídeo como representada pela SEQ ID NO: 8 são apresentados na Tabela D. O grupo de organismo "vegetal" foi selecionado, nenhum corte definido e o comprimento prognosticado do peptídeo de trânsito requisitado. Não existe nenhum prognóstico claro da localização subcelular, apenas um prognóstico fraco quanto a uma localização cloroplástica (classe de confiabilidade 3, que é razoavelmente baixo). As proteínas DEL1 portanto podem estar localizadas no citoplasma ou no núcleo.

Tabela D: Análise TargetP 1.1 da seqüência de polipeptídeo

como representada pela SEQ ID NO: 8

Comprimento (AA)	403
Peptídeo de Trânsito Cloroplástico	0,269
Peptídeo de Trânsito Mitocondrial	0,063
Peptídeo de sinal do caminho secretor	0,053
Outro alvejamento subcelular	0,721
Localização Prognosticado	/
Classe de confiabilidade	3
Comprimento do peptídeo de trânsito prognosticado	12

Muitos outros algoritmos podem ser usados para realizar tais análises, incluindo:

• ChloroP 1.1 hospedado no servidor da Technical University of Denmark;

• Protein Prowler Subcellular Localisation Predictor versão 1.2 hospedado no servidor do Institute for Molecular Bioscience, University of Queensland, Brisbane, Austrália;

• PENCE Proteome Analyst PA-GOSUB 2.5 hospedado no servidor da University of Alberta, Edmonton, Alberta, Canadá;

• PSORT (e em particular WOLF PSORT) para a localização subcelular, hospedado pelo Brinkman Laboratory na Simon Fraser University

Quando analisado com PSORT, usando os valores básicos, a proteína DEL1 da SEQ ID NO: 8 é prognosticado estar localizado no núcleo (dados de saída: nucl: 10, chlo: 1, cyto: 1, pero: 1, cysk: 1).

Exemplo 6: Ensaio de atividade DEL1

A atividade DEL1 pode ser ensaiada por exemplo pela determinação da atividade de ligação de DNA, diversos ensaios estão disponíveis e bem conhecidos na técnica (por exemplo o Ensaio de Mudança de Mobilidade Elétrica (EMSA) descrito em Kosugi & Ohashi (J. Biol. Chem. 19, 16553-16558, 2002a) ou em Kosugi & Ohashi (Plant J. 29, 45-59, 2002b). Um outro aspecto da atividade de DEL1, a saber a repressão de promotores regulados por E2F, pode ser ensaiado medindo-se a transcrição de genes repórteres sob o controle de promotores regulados por E2F na presença ou

ausência de proteínas DEL1 (por exemplo como descrito por Kosugi e Ohshi (2002a).

Além disso, a expressão de um ácido nucleico que codifica o polipeptídeo da SEQ ID NO: 2, representando um fragmento de um polipeptídeo DEL1, em plantas e em particular em arroz, tem o efeito de aumentar o rendimento da planta transgênica quando comparado com as plantas de controle, em que o rendimento aumentado compreende pelo menos o peso de sementes total ou número de sementes cheias. Além disso, a inibição da expressão de DEL1 em plantas transgênicas tem o efeito de aumentar a resistência ao estresse de plantas, medida como biomassa aumentada comparada com as plantas de controle no tratamento de radiação com UV-B (ver os Exemplos abaixo).

Exemplo 7: Clonagem de Gene de DEL1f

O fragmento de gene DEL1 (DELf1) de Arabidopsis thaliana foi amplificado pela PCR usando como padrão uma biblioteca de cDNA de mudas de Arabidopsis thaliana (Invitrogen, Paisley, UK). Depois da transcrição reversa de RNA extraído das mudas, os cDNAs foram clonados em pCMV Sport 6.0. O tamanho de inserto médio do banco foi de 1,5 kb e o número original de clones foi da ordem de $1,59 \times 10^7$ cfu. O título original foi determinado ser $9,6 \times 10^5$ cfu/ml depois da primeira amplificação de 6×10^{11} cfu/ml. Depois da extração de plasmídeo, 200 ng de padrão foram usados em uma mistura de PCR de 50 µl. Os iniciadores prm00536 (SEQ ID NO: 3; sentido, códon de partida em negrito, sítio AttB1 em itálico: 5'-ggggacaagtttgtacaaaaaa gcaggcttcacaatggttgggcttgatgatgc 3') e prm00324 (SEQ ID NO: 4; reverso, complementar, códon de parada em negrito, sítio AttB2 em itálico: 5' ggggaccactttgtacaagaaagctgggtctaacggtgttgtgatgtattag 3'), que incluem os sítios AttB para a recombinação de Gateway, foram usados para a amplificação pela PCR. A PCR foi realizada usando Hifi Taq DNA polimerase em condições padrão. Um fragmento de PCR de 1107 pares

de base (incluindo sítios attB) foi amplificado e purificado também usando métodos padrão. A primeira etapa do procedimento Gateway, a reação BP, foi depois realizada, durante o que o fragmento de PCR recombina in vivo com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada", p042. O plasmídeo pDONR201 foi adquirido da Invitrogen, como parte da tecnologia Gateway®.

Exemplo 8: Construção de Vetor

O clone de entrada p042 foi subseqüentemente usado em uma reação LR com p00831, um vetor de destinação usado para a transformação de Oryza sativa. Este vetor contém como elementos funcionais dentro das bordas de T-DNA: um marcador selecionável vegetal; um cassete de expressão de marcador triável; e um cassete Gateway intencionado para a recombinação LR in vivo com a seqüência de interesse já clonada no clone de entrada. Um promotor para embrião e expressão específica de aleurona (PRO0218, SEQ ID NO: 5) foi localizado a montante deste cassete Gateway. O cassete de expressão PRO0218: : fragmento DEL1 é dado na SEQ ID NO: 6

Depois da etapa de recombinação LR, o vetor de expressão resultante, p037 para DEL1f (Figura 3) foi transformado na cepa LBA4044 de Agrobacterium e subseqüentemente em plantas de Oryza sativa. As plantas de arroz transformadas foram deixadas crescer e foram depois examinadas quanto aos parâmetros descritos no Exemplo 3.

Exemplo 9: Avaliação e resultados de DEL1f sob o controle do promotor PRO0218

Aproximadamente 15 a 20 transformantes do arroz T0 independentes foram gerados. Os transformantes primários foram transferidos de uma câmara de cultura de tecido para uma estufa para crescimento e colheita de sementes T1. Quatro eventos, dos quais a progênie T1 segregou 3:1 quanto a presença/ausência do transgene, foram retidos. Para cada um

destes eventos, aproximadamente 10 mudas T1 contendo o transgene (hetero- e homo-zigotos) e aproximadamente 10 mudas T1 carecendo do transgene (nulizigotos) foram selecionados monitorando-se a expressão de marcador visual. As plantas T1 selecionadas foram transferidas para uma estufa. Cada planta recebeu um rótulo de código de barras único para ligar de modo não ambíguo os dados de fenotipagem à planta correspondente. As plantas T1 selecionadas foram cultivadas no solo em vasos de 10 cm de diâmetro sob os seguintes ajustes ambientais: fotoperíodo = 11,5 h, intensidade da luz do dia = 30.000 lux ou mais, temperatura diurna = 28° C ou mais alta, temperatura noturna = 22° C, umidade relativa = 60 a 70 %. Cuidado foi tomado para que as plantas fossem cultivadas sob condições ótimas. As plantas transgênicas e os nulizigotos correspondentes foram cultivados lado a lado em posições aleatórias. A partir do estágio de semeadura até o estágio de maturidade das plantas foram passadas diversas vezes através de uma cabina de formação de imagem digital. Em cada ponto de tempo imagens digitais (2048 x 1536 pixéis, 16 milhões de cores) foram tiradas de cada planta de pelo menos 6 ângulos diferentes.

Os panículos primários maduros foram colhidos, ensacados, rotulados com código de barras e depois secados por três dias na estufa a 37° C. Os panículos foram depois debulhados e todas as sementes coletadas. As cascas cheias foram separadas das vazias usando um dispositivo de sopragem a ar. Depois da separação, ambos os lotes de semente foram depois contados usando uma máquina de contagem comercialmente disponível. As cascas vazias foram descartadas. As cascas cheias foram pesadas em uma balança analítica e a área transversal das sementes foi medida usando a formação de imagem digital. Este procedimento resultou no conjunto dos seguintes parâmetros relacionados com semente:

O número de sementes cheias foi determinado contando-se o número de cascas cheias que permaneceram depois da etapa de separação. O

peso de semente total foi medido pesando-se todas as cascas cheias colhidas de uma planta. O número de semente total por planta foi medido contando-se o número de cascas colhidas de uma planta. O Peso de Mil Grãos (TKW) é extrapolado a partir do número de sementes cheias contadas e seu peso total. O índice de colheita é definido como a razão entre o peso de semente total e a área acima do solo (mm^2), multiplicado por um fator 10^6. Estes parâmetros foram derivados em uma maneira automatizada a partir das imagens digitais usando o software de análise de imagem e foram analisados estatisticamente. Os parâmetros de semente individuais (incluindo largura, comprimento, área, peso) foram medidos usando um dispositivo feito de encomenda consistindo de dois componentes principais, um dispositivo de pesagem e formação de imagem, ligados ao software para a análise de imagem.

Um fator ANOVA duplo (análise de variância) corrigido quanto ao planejamento desequilibrado foi usado como modelo estatístico para a avaliação global de características fenotípicas vegetais. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformados com este gene. O teste F foi realizado para checar quanto a um efeito do gene sobre todos os eventos de transformação e para verificar quanto a um efeito global do gene, também aqui chamado de "efeito de gene global". Se os valores do teste F mostram que os dados são significantes, então é concluído que existe um efeito de "gene", significando que não apenas a presença ou a posição do gene está causando o efeito. O limiar para significância para um efeito de gene global verdadeiro é ajustado ao nível de probabilidade de 5 % para o teste F.

Para checar quanto a um efeito dos genes dentro de um evento, isto é, quanto a um efeito específico de linhagem, um teste t foi realizado dentro de cada evento usando conjuntos de dados das plantas transgênicas e das plantas nulas correspondentes. "Plantas nulas" ou "segregantes nulos" ou "nulizigotos" são as plantas tratadas do mesmo modo como a planta

transgênica, mas da qual o transgene segregou. As plantas nulas também podem ser descritas como as plantas transformadas negativas homozigotas. O limiar para significância para o teste t é ajustado ao nível de 10 % de probabilidade. Os resultados para alguns eventos podem estar acima ou abaixo deste limiar. Isto está fundamentado na hipótese de que um gene teria apenas 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 também é chamado aqui de um "efeito de linhagem do gene". O valor p é obtido comparando-se o valor t com a distribuição t ou alternativamente, comparando-se o valor F com a distribuição F. O valor p depois dá a probabilidade de que a hipótese nula (isto é, que não existe nenhum efeito do transgene) esteja correta.

Os dados obtidos para DEL1f no primeiro experimento foram confirmados em um segundo experimento com plantas T2. Quatro linhagens que tiveram o padrão de expressão corrigido foram selecionadas para mais análise. Os lotes de semente das plantas positivas (tanto hetero- quanto homozigotas) em T1, foram triados monitorando-se a expressão de marcador. Para cada evento escolhido, os lotes de semente heterozigota foram depois retidos para a avaliação T2. Dentro de cada lote de semente um número igual de plantas positivas e negativas foram cultivadas na estufa para avaliação.

Um número total de 120 plantas transformadas por DEL1f foram avaliadas na geração T2, isto é 30 plantas por evento das quais 15 positivas para o transgene e 15 negativas.

Porque dois experimentos com eventos sobrepostos foram realizados, uma análise combinada foi realizada. Isto é útil para checar a compatibilidade dos efeitos em relação aos dois experimentos e se este for o caso, acumular evidência de ambos experimentos de modo a aumentar a confidência na conclusão. O método usado foi um método de modelo misto que leva em conta a estrutura de nível múltiplo dos dados (isto é, experimento

- evento - segregantes). Os valores p são obtidos comparando-se o teste de razão de probabilidade com as distribuições de qui quadrado.

Exemplo 10: Avaliação de transformantes DEL1f: medição de parâmetros relacionados com o rendimento

Na análise das sementes como descrita acima, os inventores verificaram que as plantas transformadas com a construção de gene DEL1f tiveram um rendimento de semente mais alto, expressado como número de sementes cheias, peso de sementes total e índice de colheita, comparado com plantas que carecem do transgene DEL1f. Deve ser mencionado que o aumento no rendimento foi obtido na ausência de condições de estresse.

Os resultados obtidos para as plantas na geração T1 estão resumidos na Tabela E:

Tabela E:

	% de diferença	Valor p
Nº de sementes cheias	+38	0,0055
Peso de sementes total	+42	0,0039

Estes resultados positivos foram mais uma vez obtidos na geração T2. Na Tabela F, os dados mostram os aumentos de % globais para o número de sementes cheias, peso de sementes total e índice de colheita, calculados a partir dos dados das linhagens individuais da geração T2 e dos respectivos valores p. Estes dados T2 foram re-avaliados em uma análise combinada com os resultados para a geração T1 e os valores p obtidos mostram que os efeitos observados foram altamente significantes.

Tabela F:

	Geração T2		Análise Combinada
	% de diferença	valor p	valor p
Nº de sementes cheias	+14	0,0006	0,0001
Peso sementes total	+12	0,0141	0,0002

Exemplo 11: Regulação negativa da expressão DEL1 leva à resistência ao estresse realçada

O alelo del1-1^{KO} de Arabidopsis thaliana (L.) Heynh. ecotipo Columbia (Col-0) foi obtido da coleção de mutante GABI-Kat T-DNA

(http://www.mpiz-koeln.mpg.de/GABI-Kat/GABIKat_homepage). Para a geração de linhagens $DEL1^{OE}$, a matriz de leitura de DEL1 aberta completa foi clonada no vetor pK2GW7 [Karimi et al. (2002) Trends Plant Sci. 7, 193-195] com a tecnologia GATEWAY, resultando no vetor pKDEL1, que foi mobilizado pelo plasmídeo auxiliar pRK2013 na cepa C58C1 RifR de Agrobacterium tumefaciens que abriga o plasmídeo pMP90 [Clough e Bent (1998) Plant J. 16, 735-743]. As plantas foram transformadas pelo método da imersão floral [S4]. As plantas $DEL1^{OE}$ transgênicas foram selecionadas em meio contendo canamicina.

As sementes de uma linhagem silenciada DEL1 ($del1\text{-}1^{KO}$, duas linhagens que super expressam DEL1 independentes ($DEL1^{OE}$ #2 e $DEL1^{OE}$ #4) e uma planta do tipo selvagem não transformada (controle) foram semeadas em uma mistura de solo e areia (4:1) e vernalizadas por 3 dias a 4° C por 3 dias. Subseqüentemente, as plantas foram submetidas a um regime de 70 $\mu mol.m^{-2}.s^{-1}$ de Radiação Fotossinteticamente Ativa, ciclo de 14 h/10 h dia/noite e uma temperatura de 20° C/16 a 18° C dia/noite em uma umidade relativa de 70 a 80 %. Depois de 18 dias, as plantas foram expostas à luz UV-B em um simulador solar (GSF; National Research Center for Enviroment and Health, Munich, Alemanha), que é uma câmara de crescimento que imita o espectro solar como o mesmo atinge a superfície da terra suplementado com uma dose adicional de UV-B (radiação: 172 $\mu mol.m^{-2}.s^{-1}$ (± 5 %), 3,27 Wm^{-2} (± 5 %). Depois de 0, 3, 5, 7 e 9 horas de exposição ao UV-B as plantas foram transferidas de volta para a câmara de crescimento normal e o seu desenvolvimento foi seguido com o tempo.

Os genótipos diferentes testados claramente responderam de modo diferente ao tratamento com UV. O tratamento com UV-B claramente afetou o crescimento das plantas de controle em uma maneira dependente da dose. As folhas de plantas tratadas por 5 horas com UV-B mostraram necrose e foram reduzidas em tamanho. Este fenótipo foi ainda mais exposto para as

plantas tratadas por 7 horas com UV-B. as plantas com níveis de DEL1 aumentados mostraram uma resposta ainda mais grave ao UV-B, demonstrando já necrose e inibição de crescimento de folha em uma dose de 3 horas de UV-B. Ao contrário, surpreendentemente, a linhagem del1-1^{KO} demonstrou tolerância aumentada contra a luz UV-B. Ao passo que o tamanho das plantas de controle foi claramente afetado por uma dose de 3 horas de UV-B, as plantas del1-1^{KO} tratadas por 3 horas foram indistingüíveis das plantas de controle. Também em uma dose de 5 horas de UV-B o nível de necrose foi severamente reduzida em comparação com as plantas de controle.

As observações visuais foram sustentadas pelas medições de biomassa (Fig. 5). A massa de 10 a 15 plantas foi medida e calculada em média por planta. Ao passo que as plantas de controle não transformadas mostraram uma diminuição dependente da dose de UV-B evidente na biomassa por planta, as linhagens del1-1^{KO} tratadas por menos do que 5 horas demonstraram uma biomassa igual àquela das plantas não tratadas. Também em uma dose de 5 horas de UV-B ou mais longo, as plantas del1-1^{KO} desempenharam melhor do que as plantas de controle. De acordo com as observações de que uma diminuição em DEL1 resulta em estabilidade de rendimento aumentada, o fenótipo oposto foi observado para as linhagens DEL1^{OE} #2 e DEL1^{OE} #4, que foram evidentemente mais afetadas do que as linhagens de controle em qualquer uma das doses de UV-B aplicadas.

Para identificar a base celular subjacente à estabilidade de rendimento aumentada das linhagens DEL1^{KO}, números e tamanhos de célula epidérmica foram determinados pela microscopia. A área de lâmina total da 5ª folha foi determinada a partir de quadros digitalizados tirados diretamente com uma câmara com dispositivo ligado à carga montado em um binocular (Stemi SV11, Zeiss, Jena, Alemanha). A área de lâmina de folha foi determinada a partir de imagens usando o programa de análise de imagem de domínio público ImageJ (versão 1.30; disponível no sítio da web do US

National Institutes of Health, NIH). A densidade de célula foi determinada a partir de imagens de tubo de desenho escaneadas do contorno de pelo menos 25 células da epiderme abaxial localizadas a 25 e 75 % da distância entre a ponta e a base da folha depois de depurar as folhas em metanol seguido pelo tratamento com ácido láctico. A partir das imagens a área de célula média e o número total de células por folha foram calculados dividindo-se a área das folhas pela área de célula média (calculada em média entre as posições apicais e basais). A análise comparativa de folhas tratadas por 5 horas com UV-B ilustrou que a área de lâmina de folha total foi diminuída em 52 % nas plantas de controle no tratamento com UV-B. Ao contrário, as linhagens del1-1^{KO} mostraram apenas uma diminuição de 15 %. Quando da contagem de números de célula, uma diminuição igual no número de célula total foi observada para ambos os genótipos (sendo de -28 % e -23 % para as plantas de controle e del1-1^{KO}, respectivamente). Ao contrário, ao passo que o tamanho de célula médio foi declinado com 33 % nas plantas de controle (plantas não tratadas versus plantas tratadas), as células de plantas del1-1^{KO} foram ampliadas em 11 % no tratamento com UV-B. Estes dados ilustram que a estabilidade no rendimento observada para as plantas del1-1^{KO} no tratamento de estresse é especificamente atribuído à ampliação de célula.

Exemplo 12: Plantas del1-1^{KO} demonstram níveis aumentados de expressão e atividade de fotoliase PHR1

Para revelar o mecanismo molecular que confere a resistência ao UV-B contra as linhagens deficientes em DEL1, uma análise de transcritoma foi realizada. As sementes silenciadas em DEL1 (del1-1^{KO}, superexpressão DEL1 #4 (DEL1^{OE} #4) e as sementes do tipo selvagem não transformadas (Col-0) foram semeadas em 1x meio Murashige e Skoog (Duchefa, Haarlem, Países Baixos) suplementado com 0,6 % de cultura de tecido vegetal em ágar (LabM, Bury, UK). Depois da vernalização por 2 dias a 4° C, as plantas foram transferidas para a câmara de crescimento e

cultivadas a 22° C e 65 $\mu E \cdot m^{-2} \cdot s^{-1}$ de irradiação em um fotoperíodo de 16 horas de luz/8 horas de escuro e uma umidade relativa de 50 a 60 %. Para a análise de microssérie os tecidos mais contrastantes foram analisados, sendo mudas de 8 dias de idade no total para o controle (Col-0) vs. del1-1^{KO} e o primeiro par de folhas a desenvolver para o controle vs. DEL1^{OE} #4, colhidos em 8, 15 e 22 dias depois da semeadura. Depois da análise estatística, 16 genes foram descobertos estar demonstrando um padrão de expressão complementar em plantas del1-1^{KO} vs. DEL1^{OE} #4. Um dos genes infra regulados na linhagem DEL1^{OE} #4 e super regulados em del1-1^{KO} codificados para a fotoliase CPD tipo 2, representando uma enzima capaz de reparar o dano ao DNA causado pelo UV nas células vegetais na presença de luz branca ou azul [Landry et al., 1997, PNAS 94, 328-332]. Estes dados mostraram que as plantas com níveis de DEL1 baixos ou nenhum demonstram uma atividade de reparo de DNA aumentada depois do tratamento de estresse com UV-B. Ao contrário, as plantas que super expressam DEL1, tendo um nível diminuído de fotoliase CPD tipo 2, não serão capazes de reparar o seu DNA no mesmo grau como as plantas de controle ou del1-1^{KO}.

Para testar esta hipótese, as sementes da linhagem del1-1^{KO}, duas linhagens que super expressam DEL1 independentes (DEL1^{OE} #2 e DEL1^{OE} #4) e uma planta do tipo selvagem não transformada (Col-0) foram semeadas em uma mistura de solo e areia (4:1) e vernalizadas a 4° C por 3 dias. Subseqüentemente, as plantas foram submetidas a um regime de 70 $\mu mol.m^{-2}.s^{-1}$ de PAR, ciclo de 14 h/10 h dia/noite e uma temperatura de 20° C/16 a 18° C dia/noite em uma umidade relativa de 70 a 80 %. Depois de 18 dias, as plantas foram expostas à luz UV-B (radiação: 172 $\mu mol.m^{-2}.s^{-1} \pm 5$ %) em um simulador solar (GSF; National Research Center for Environment and Health, Munich, Alemanha). Depois de 5 horas de irradiação de UV-B, as folhas foram colhidas imediatamente, ou deixadas recuperar em luz branca (70 $\mu mol.m^{-2}.s^{-1}$ de PAR) por mais 5 horas. A quantidade de atividade de

fotoliase foi medida determinando-se o número de fotoprodutos de dímero de ciclobutilpirimidina (CPD) pelo Elisa. O DNA genômico foi extraído usando um DNeasy Plant Mini Kit QIAGEN, Tóquio, Japão). Uma alíquota de 50 µl do DNA extraído a uma concentração de 0,02 $\mu g.ml^{-1}$ foi colocada em cada reservatório. Os CPDs foram detectados com anticorpos específicos (TDM-2) como anteriormente descrito (Tanaka et al., 2002, The Plant Journal 46, 317-326). Quando da comparação do nível de CPD das amostras colhidas diretamente depois do tratamento com UV-B, todos os genótipos demonstraram uma resposta de dano de DNA similar (Figura 6A). Ao contrário, quando da comparação das amostras deixadas recuperar em luz branca por 5 horas, quase 45 % dos sítios de dano de DNA foram reparados no mutante $del1\text{-}1^{KO}$. Contrastantemente, nenhum reparo de DNA pôde ser observado durante o mesmo período de tempo no controle e nas plantas que superexpressam $DEL1^{OE}$ (Figura 6B). Estes dados mostram que a linhagem $del1\text{-}1^{KO}$ desempenha melhor depois do tratamento com UV-B devido a um reparo de DNA mais rápido.

Exemplo 13: As linhagens $del1\text{-}1^{KO}$ demonstram um desempenho fotossintético melhor depois da irradiação com UV-B

As sementes de uma linhagem silenciada em DEL1 ($del1\text{-}1^{KO}$), duas linhagens que super expressam DEL1 independentes ($DEL1^{OE}$ #2 e $DEL1^{OE}$ #4) e uma planta do tipo selvagem não transformada (Col-0) foram semeadas em um composto com base em turfa comercial padrão, enriquecido com 1/3 de John Innes 3 (mistura com base e greda) e vernalizadas a 4° C por uma semana. Subseqüentemente, as plantas foram submetidas a um regime de 100 $\mu mol.m^{-2}.s^{-1}$, fotoperíodo de 12 horas e uma temperatura de 22° C em uma umidade relativa de 60 a 80 %. Depois de 25 a 30 dias, as plantas foram expostas a um pulso curto de luz UV-B. A eficiência do fotossistema II foi medido antes e 6 horas depois da irradiação com UV. Quando medido antes da irradiação, nenhuma diferença significante pôde ser encontrada entre os

genótipos diferentes. Isto indicou que todas as linhagens foram saudáveis e não estressadas antes do tratamento com UV e que a irradiação afetou todas as linhagens ao mesmo grau. Ao contrário, quando as plantas foram submetidas a um tratamento com UV-B longo de 48 horas ligado com PAR, o mutante del1-1^{KO} teve uma eficiência fotossintética levemente mais alta do que o controle, ao passo que as linhagens que super expressam tiveram um desempenho fotossintético reduzido em 12 % comparado com Col-0. Estes dados mostram que as linhagens del1-1^{KO} excederam em desempenho as plantas de controle na sua atividade fotossintética depois do tratamento com UV-B.

Exemplo 14: As plantas DEL1-1^{KO} também demonstram resistência realçada ao estresse sob irradiação UV-B constante

Para testar se a resistência ao estresse de del1-1^{KO} é verdadeira sob um nível baixo constante de UV-B (ao invés de em resposta a uma dose aguda), as plantas foram cultivadas por uma semana sob um fluxo baixo constante de luz UV-B. No início do experimento, nenhuma diferença significante no diâmetro da roseta foi observada para os genótipos distintos. Ao contrário, depois de uma semana de irradiação UV-B, o diâmetro da roseta das plantas del1-1^{KO} (1,6 cm) foi próximo àquele das plantas não tratadas (1,8 cm), ao passo que aquelas do controle (1,4 cm) e DEL1^{OE} (1,2 cm) foram significantemente mais afetadas pelo tratamento com UV-B.

LISTAGEM DA SEQÜÊNCIA

<110> CropDesign N.V.
Universiteit Gent
VIB vzw

<120> "plantas tendo características de crescimento melhoradas e um método para fabricar as mesmas"

<130> PF57951

<160> 31

<170> PatentIn versão 3.3

<210> 1
<211> 1044
<212> DNA
<213> Arabidopsis thaliana

<400> 1
atggttgggc ttgatgatgc tgcctcgaaa ttaggagtgg agagacgaag aatctatgat 60
attgttaatg ttctggagag tgttggggtt ttaacaagaa gagcaaagaa tcagtatacg 120
tggaaagggt tttccgcaat tccaggagca ttgaaggagc tacaagaaga gggggttaag 180
gatacttttc atcgtttcta tgtcaatgag aatgttaaag gatctgatga tgaggatgat 240
gatgaagagt cttctcagcc tcactctagt agccagactg atagttcaaa acctggttct 300
cttccccaat cttcagatcc ctccaaaata gataaccgac gagagaaatc tttaggattg 360
cttactcaga actttatcaa actctttatt tgctctgaag ctattaggat catctccctt 420
gatgacgctg caaaattact gcttggtgat gcccacaata catcaatcat gcgaactaaa 480
gtgaggcggc tttatgatat agcaaatgtc ttgtcgtcaa tgaatctcat agagaagact 540
cacaccttag attctaggaa accagctttc aagtggttag ggtacaatgg tgagcctact 600
ttcacactga gcagcgattt gttgcaattg gagtcaagaa aaagagcttt cggaactgat 660
attacaaacg tcaatgttaa gagaagcaaa tcatcatctt cgtccaaga aaacgctaca 720
gagagaaggc taaagatgaa aaagcactca acaccagaga gttcttataa caaaagcttt 780
gatgttcatg aatcaagaca tggatcaaga ggaggttacc attttggacc ttttgcacca 840
ggcactggta catatccaac tgctggttta gaggataact ctaggagagc ttttgatgtt 900
gagaatctgg attctgatta ccgtccctct taccaaaacc aagttttgaa agacctcttt 960
tcccattaca tggatgcttg gaagacatgg ttcagcgaag tcacccagga gaatccatta 1020
cctaatacat cacaacaccg ttag 1044

<210> 2
<211> 347
<212> PRT
<213> Arabidopsis thaliana

<400> 2
Met Val Gly Leu Asp Asp Ala Ala Ser Lys Leu Gly Val Glu Arg Arg
1 5 10 15
Arg Ile Tyr Asp Ile Val Asn Val Leu Glu Ser Val Gly Val Leu Thr
20 25 30
Arg Arg Ala Lys Asn Gln Tyr Thr Trp Lys Gly Phe Ser Ala Ile Pro
35 40 45
Gly Ala Leu Lys Glu Leu Gln Glu Glu Gly Val Lys Asp Thr Phe His
50 55 60

```
Arg Phe Tyr Val Asn Glu Asn Val Lys Gly Ser Asp Asp Glu Asp Asp
65                  70                  75                  80
Asp Glu Glu Ser Ser Gln Pro His Ser Ser Ser Gln Thr Asp Ser Ser
                85                  90                  95
Lys Pro Gly Ser Leu Pro Gln Ser Ser Asp Pro Ser Lys Ile Asp Asn
            100                 105                 110
Arg Arg Glu Lys Ser Leu Gly Leu Leu Thr Gln Asn Phe Ile Lys Leu
        115                 120                 125
Phe Ile Cys Ser Glu Ala Ile Arg Ile Ile Ser Leu Asp Asp Ala Ala
    130                 135                 140
Lys Leu Leu Leu Gly Asp Ala His Asn Thr Ser Ile Met Arg Thr Lys
145                 150                 155                 160
Val Arg Arg Leu Tyr Asp Ile Ala Asn Val Leu Ser Ser Met Asn Leu
                165                 170                 175
Ile Glu Lys Thr His Thr Leu Asp Ser Arg Lys Pro Ala Phe Lys Trp
            180                 185                 190
Leu Gly Tyr Asn Gly Glu Pro Thr Phe Thr Leu Ser Ser Asp Leu Leu
        195                 200                 205
Gln Leu Glu Ser Arg Lys Arg Ala Phe Gly Thr Asp Ile Thr Asn Val
    210                 215                 220
Asn Val Lys Arg Ser Lys Ser Ser Ser Ser Ser Gln Glu Asn Ala Thr
225                 230                 235                 240
Glu Arg Arg Leu Lys Met Lys Lys His Ser Thr Pro Glu Ser Ser Tyr
                245                 250                 255
Asn Lys Ser Phe Asp Val His Glu Ser Arg His Gly Ser Arg Gly Gly
            260                 265                 270
Tyr His Phe Gly Pro Phe Ala Pro Gly Thr Gly Thr Tyr Pro Thr Ala
        275                 280                 285
Gly Leu Glu Asp Asn Ser Arg Arg Ala Phe Asp Val Glu Asn Leu Asp
    290                 295                 300
Ser Asp Tyr Arg Pro Ser Tyr Gln Asn Gln Val Leu Lys Asp Leu Phe
305                 310                 315                 320
Ser His Tyr Met Asp Ala Trp Lys Thr Trp Phe Ser Glu Val Thr Gln
                325                 330                 335
Glu Asn Pro Leu Pro Asn Thr Ser Gln His Arg
            340                 345
```

<210> 3
<211> 54
<212> DNA
<213> Seqüência Artificial

<220>
<223> Iniciador : prm00536

<400> 3
ggggacaagt ttgtacaaaa aagcaggctt cacaatggtt gggcttgatg atgc 54

<210> 4
<211> 52
<212> DNA
<213> Seqüência Artificial

<220>

<223> Iniciador: prm00324

<400> 4
ggggaccact ttgtacaaga aagctgggtc taacggtgtt gtgatgtatt ag 52

<210> 5
<211> 1236
<212> DNA
<213> Oryza sativa

<400> 5
ggtcagccaa tacattgatc cgttgccaat catgcaaagt attttggctg tggccgagtg 60
ccggaattga taattgtgtt ctgactaaat taaatgacca gaagtcgcta tcttccaatg 120
tatccgaaac ctggattaaa caatcctgtt ctgttctcta gcccctcctg catggccgga 180
ttgtttttt gacatgtttt cttgactgag gcctgtttgt tctaaacttt ttcttcaaac 240
ttttaacttt ttcatcacat cagaactttt ctacacatat aaacttttaa cttttccgtc 300
acatcgttcc aatttcaatc aaactttcaa ttttggcgtg aactaaacac accctgagtc 360
ttttattgct cctccgtacg ggttggctgg ttgagaatag gtattttcag agagaaaatc 420
tagatattgg gaggaacttg gcatgaatgg ccactatatt tagagcaatt ctacggtcct 480
tgaggaggta ccatgaggta ccaaaatttt agtgtaaatt ttagtatctc attataacta 540
ggtattatga ggtaccaaat ttacaataga aaaaatagta cttcatggta ctttcttaag 600
taccgtaaaa ttgctcctat atttaagggg atgtttatat ctatccatat ccataatttg 660
attttgataa gaaaaaatgt gagcacacca agcatgtcca tgaccttgca ctcttggctc 720
actcgtcaac tgtgaagaac ctcaaaaatg ctcaatatag ctacaggtgc ctgaaaaaat 780
aactttaaag ttttgaacat cgatttcact aaacaacaat tattatctcc ctctgaaaga 840
tgatagttta gaactctaga atcattgtcg gcggagaaag taaattattt tccccaaatt 900
tccagctatg aaaaaaccct caccaaacac catcaaacaa gagttcacca aaccgcccat 960
gcggccatgc tgtcacgcaa cgcaccgcat tgcctgatgg ccgctcgatg catgcatgct 1020
tccccgtgca catatccgac agacgcgccg tgtcagcgag ctcctcgacc gacctgtgta 1080
gcccatgcaa gcatccaccc ccgccacgta caccccctcc tcctccctac gtgtcaccgc 1140
tctctccacc tatatatgcc cacctggccc ctctcctccc atctccactt caccccgatcg 1200
cttcttcttc ttcttcgttg cattcatctt gctagc 1236

<210> 6
<211> 2332
<212> DNA
<213> Seqüência Artificial

<220>
<223> Cassete de expressão

<400> 6
ggtcagccaa tacattgatc cgttgccaat catgcaaagt attttggctg tggccgagtg 60
ccggaattga taattgtgtt ctgactaaat taaatgacca gaagtcgcta tcttccaatg 120
tatccgaaac ctggattaaa caatcctgtt ctgttctcta gcccctcctg catggccgga 180
ttgtttttt gacatgtttt cttgactgag gcctgtttgt tctaaacttt ttcttcaaac 240
ttttaacttt ttcatcacat cagaactttt ctacacatat aaacttttaa cttttccgtc 300
acatcgttcc aatttcaatc aaactttcaa ttttggcgtg aactaaacac accctgagtc 360
ttttattgct cctccgtacg ggttggctgg ttgagaatag gtattttcag agagaaaatc 420
tagatattgg gaggaacttg gcatgaatgg ccactatatt tagagcaatt ctacggtcct 480
tgaggaggta ccatgaggta ccaaaatttt agtgtaaatt ttagtatctc attataacta 540
ggtattatga ggtaccaaat ttacaataga aaaaatagta cttcatggta ctttcttaag 600
taccgtaaaa ttgctcctat atttaagggg atgtttatat ctatccatat ccataatttg 660
attttgataa gaaaaaatgt gagcacacca agcatgtcca tgaccttgca ctcttggctc 720

actcgtcaac tgtgaagaac ctcaaaaatg ctcaatatag ctacaggtgc ctgaaaaaat 780
aactttaaag ttttgaacat cgatttcact aaacaacaat tattatctcc ctctgaaaga 840
tgatagttta gaactctaga atcattgtcg gcggagaaag taaattattt tccccaaatt 900
tccagctatg aaaaaaccct caccaaacac catcaaacaa gagttcacca aaccgcccat 960
gcggccatgc tgtcacgcaa cgcaccgcat tgcctgatgg ccgctcgatg catgcatgct 1020
tccccgtgca catatccgac agacgcgccg tgtcagcgag ctcctcgacc gacctgtgta 1080
gcccatgcaa gcatccaccc ccgccacgta cacccccctcc tcctccctac gtgtcaccgc 1140
tctctccacc tatatatgcc cacctggccc ctctcctccc atctccactt cacccgatcg 1200
cttcttcttc ttcttcgttg cattcatctt gctagcattt aaatcaacta gggatatcac 1260
aagtttgtac aaaaaagcag gcttcacaat ggttgggctt gatgatgctg cctcgaaatt 1320
aggagtggag agacgaagaa tctatgatat tgttaatgtt ctggagagtg ttggggtttt 1380
aacaagaaga gcaaagaatc agtatacgtg gaaagggttt tccgcaattc caggagcatt 1440
gaaggagcta caagaagagg gggttaagga cacttttcat cgtttctatg tcaatgagaa 1500
tgttaaagga tctgatgatg aggatgatga tgaagagtct tctcagcctc actctagtag 1560
ccagactgat agttcaaaac ctggttctct tccccaatct tcagatccct ccaaaatага 1620
taaccgacga gagaaatctt taggattgct tactcagaac tttatcaaac tctttatttg 1680
ctctgaagct attaggatca tctcccttga tgacgctgca aaattactgc ttggtgatgc 1740
ccacaataca tcaataatgc gaactaaagt gaggcggctt tatgatatag caaatgtctt 1800
gtcgtcaatg aatctcatag agaagactca caccttagat tctaggaaac cagctttcaa 1860
gtggttaggg tacaatggtg agcctacttt cacactgagc agcgatttgt tgcaattgga 1920
gtcaagaaaa agagctttcg gaactgatat tacaaacgtc aatgttaaga gaagcaaatc 1980
atcatcttcg tcccaagaaa acgctacaga gagaaggcta aagatgaaaa agcactcaac 2040
accagagagt tcttataaca aaagctttga tgttcatgaa tcaagacatg gatcaagagg 2100
aggttaccat tttggacctt ttgcaccagg cactggtaca tatccaactg ctggtttaga 2160
ggataactct aggagagctt ttgatgttga gaatctggat tctgattacc gtccctctta 2220
ccaaaaccaa gttttgaaag acctcttttc ccattacatg gatgcttgga agacatggtt 2280
cagcgaagtc acccaggaga atccattacc taatacatca caacaccgtt ag 2332

<210> 7
<211> 1212
<212> DNA
<213> Arabidopsis thaliana

<400> 7
atgtcagatc tatcgccaga aagattcaaa cttgccgtta cttctccttc ttccataccg 60
gaatcttctt cggctttaca attacaccat tcctatagtc gcaaacagaa atctctcgga 120
cttctttgta ccaatttctt agctttgtat aatcgagaag ggattgaaat ggttgggctt 180
gatgatgctg cctcgaaatt aggagtggag agacgaagaa tctatgatat tgttaatgtt 240
ctggagagtg ttggggtttt aacaagaaga gcaaagaatc agtatacgtg gaaagggttt 300
tccgcaattc caggagcatt gaaggagcta caagaagagg gggttaagga cacttttcat 360
cgtttctatg tcaatgagaa tgttaaagga tctgatgatg aggatgatga tgaagagtct 420
tctcagcctc actctagtag ccagactgat agttcaaaac ctggttctct tccccaatct 480
tcagatccct ccaaaatага taaccgacga gagaaatctt taggattgct tactcagaac 540
tttatcaaac tctttatttg ctctgaagct attaggatca tctcccttga tgacgctgca 600
aaattactgc ttggtgatgc ccacaataca tcaataatgc gaactaaagt gaggcggctt 660
tatgatatag caaatgtctt gtcgtcaatg aatctcatag agaagactca caccttagat 720
tctaggaaac cagctttcaa gtggttaggg tacaatggtg agcctacttt cacactgagc 780
agcgatttgt tgcaattgga gtcaagaaaa agagctttcg gaactgatat tacaaacgtc 840
aatgttaaga gaagcaaatc atcatcttcg tcccaagaaa acgctacaga gagaaggcta 900
aagatgaaaa agcactcaac accagagagt tcttataaca aaagctttga tgttcatgaa 960
tcaagacatg gatcaagagg aggttaccat tttggacctt ttgcaccagg cactggtaca 1020
tatccaactg ctggtttaga ggataactct aggagagctt ttgatgttga gaatctggat 1080
tctgattacc gtccctctta ccaaaaccaa gttttgaaag acctcttttc ccattacatg 1140
gatgcttgga agacatggtt cagcgaagtc acccaggaga atccattacc taatacatca 1200

caacaccgtt ag 1212

<210> 8
<211> 403
<212> PRT
<213> Arabidopsis thaliana

<400> 8

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

```
Asn Gln Val Leu Lys Asp Leu Phe Ser His Tyr Met Asp Ala Trp Lys
    370                 375                 380
Thr Trp Phe Ser Glu Val Thr Gln Glu Asn Pro Leu Pro Asn Thr Ser
385                 390                 395                 400
Gln His Arg
```

<210> 9
<211> 20
<212> PRT
<213> Seqüência Artificial

<220>
<223> Seqüência de Assinatura 1

<220>
<221> Variante
<222> (1)..(1)
<223> /:Substituir="Thr" / Substituir "Pro"

<220>
<221> Variante
<222> (5)..(5)
<223> / Substituir="Asp"

<220>
<221> Variante
<222> (9)..(9)
<223> / Substituir="Trp"

<220>
<221> Variante
<222> (10)..(10)
<223> / Substituir="Thr"

<220>
<221> Variante
<222> (13)..(13)
<223> / Substituir="Glu" / Substituir="Gln" / Substituir="Ser"

<220>
<221> Variante
<222> (14)..(14)
<223> / Substituir="Arg" / Substituir="Lys"

<220>
<221> Variante
<222> (16)..(16)
<223> / Substituir="Val"

<220>
<221> Variante
<222> (17)..(17)
<223> / Substituir="Thr"

<220>
<221> Variante
<222> (18)..(18)
<223> /replace="Ile" /replace="Arg"

<220>
<221> Variante
<222> (20)..(20)
<223> /replace="Gly" /replace="Pro" /replace="Asp"

<400> 9
Tyr Ser Arg Lys Gln Lys Ser Leu Gly Leu Leu Cys Thr Asn Phe Leu
1 5 10 15
Ala Leu Tyr Asn
20

<210> 10
<211> 12
<212> PRT
<213> Seqüência Artificial

<220>
<223> Seqüência de Assinatura 2

<220>
<221> Variante
<222> (1)..(1)
<223> /Substituir="Ser"

<220>
<221> Variante
<222> (4)..(4)
<223> /Substituir="Glu"

<220>
<221> Variante
<222> (5)..(5)
<223> /Substituir="Val"

<220>
<221> Variante
<222> (7)..(7)
<223> /Substituir="Thr" /Substituir="Val" /Substituir="Arg" /Substituir="Lys"

<220>
<221> Variante
<222> (8)..(8)
<223> /Substituir="Arg" /Substituir="Ser"

<400> 10
Gly Leu Asp Asp Ala Ala Ser Lys Leu Gly Val Glu
1 5 10

<210> 11
<211> 20
<212> PRT
<213> Seqüência Artificial

<220>
<223> Seqüência de Assinatura 3

<220>
<221> Variante
<222> (2)..(2)
<223> /Substituir="Lys"

<220>
<221> Variante
<222> (3)..(3)
<223> /Substituir="Asp"

<220>
<221> Variante
<222> (7)..(7)
<223> /Substituir="Arg"

<220>
<221> Variante
<222> (8)..(8)
<223> /Substituir="Leu"

<220>
<221> Variante
<222> (9)..(9)
<223> /Substituir="Met"

<220>
<221> Variante
<222> (10)..(10)
<223> /Substituir="Ser"

<220>
<221> Variante
<222> (12)..(12)
<223> /Substituir="Asn"

<220>
<221> Variante
<222> (14)..(14)
<223> /Substituir="Val"

<220>
<221> Variante
<222> (15)..(15)
<223> /Substituir="Gln" /Substituir="Met"

<220>

<221> Variante
<222> (18)..(18)
<223> /Substituir="Leu" /Substituir="Thr"

<220>
<221> Variante
<222> (19)..(19)
<223> /Substituir="Ala" /Substituir="Val" /Substituir="Met" /Substituir="Asn"

<220>
<221> Variante
<222> (20)..(20)
<223> /Substituir="Glu"

<400> 11
Arg Arg Glu Lys Ser Leu Gly Ile Leu Thr Gln Lys Phe Ile Lys Leu
1 5 10 15
Phe Ile Cys Ser
20

<210> 12
<211> 12
<212> PRT
<213> Seqüência Artificial

<220>
<223> Seqüência de Assinatura 4

<220>
<221> Variante
<222> (1)..(1)
<223> /Substituir="Val"

<220>
<221> Variante
<222> (2)..(2)
<223> /Substituir="Thr"

<220>
<221> Variante
<222> (4)..(4)
<223> /Substituir="Glu"

<220>
<221> Variante
<222> (5)..(5)
<223> /Substituir="Thr" /Substituir="Val" /Substituir="Glu"

<220>
<221> Variante
<222> (8)..(8)
<223> /Substituir="Arg"

<220>

<221> Variante
<222> (9)..(9)
<223> /substituir="Ile"

<220>
<221> Variante
<222> (10)..(10)
<223> /substituir="Ile"

<220>
<221> Variante
<222> (11)..(11)
<223> /substituir="Met" /substituir="Ile"

<220>
<221> Variante
<222> (12)..(12)
<223> /substituir="Glu"

<400> 12
Ile Ser Leu Asp Asp Ala Ala Lys Leu Leu Leu Gly
1 5 10

<210> 13
<211> 35
<212> PRT
<213> Seqüência Artificial

<220>
<223> Seqüência de assinatura 5

<220>
<221> Variante
<222> (1)..(1)
<223> /substituir="Ala"

<220>
<221> Variante
<222> (3)..(3)
<223> /substituir="Ile"

<220>
<221> Variante
<222> (12)..(12)
<223> /substituir="Ile"

<220>
<221> Variante
<222> (14)..(14)
<223> /substituir="Cys" /substituir="Thr"

<220>
<221> Variante
<222> (16)..(16)

<223> /Substituir="Leu"

<220>
<221> Variante
<222> (17)..(17)
<223> /Substituir="Arg" /Substituir="Ala"

<220>
<221> Variante
<222> (18)..(18)
<223> /Substituir="Phe"

<220>
<221> Variante
<222> (20)..(20)
<223> /Substituir="Arg" /Substituir="Lys" /Substituir="Asp"

<220>
<221> Variante
<222> (22)..(22)
<223> /Substituir="Val" /Substituir="Ile"

<220>
<221> Variante
<222> (23)..(23)
<223> /Substituir="Gln"

<220>
<221> VARIANT
<222> (24)..(24)
<223> /Substituir="Val" /Substituir="Gln"

<220>
<221> Variante
<222> (25)..(25)
<223> /Substituir="Gly" /Substituir="Ala" /Substituir="Thr"

<220>
<221> Variante
<222> (26)..(26)
<223> /Substituir="Glu"

<220>
<221> VARIANT
<222> (27)..(27)
<223> /Substituir="Thr" /Substituir="Glu"

<220>
<221> Variante
<222> (29)..(29)
<223> /Substituir="Gly"

<220>
<221> Variante

<222> (30)..(30)
<223> /Substituir="Arg"

<220>
<221> Variante
<222> (31)..(31)
<223> /Substituir="Lys"

<220>
<221> Variante
<222> (32)..(32)
<223> /Substituir="Pro"

<220>
<221> Variante
<222> (33)..(33)
<223> /Substituir="Leu" /Substituir="Ala" /Substituir="Arg"

<220>
<221> Variante
<222> (35)..(35)
<223> /Substituir="Phe"

<220>
<221> Variante
<222> (35)..(35)
<223> /Substituir="Lys"

<400> 13
Thr Lys Val Arg Arg Leu Tyr Asp Ile Ala Asn Val Leu Ser Ser Met
1 5 10 15
Asn Leu Ile Glu Lys Thr His Thr Leu Asp Ser Arg Lys Pro Ala Phe
20 25 30
Lys Trp Leu
35

<210> 14
<211> 1581
<212> DNA
<213> Oryza sativa

<400> 14
cgcggacgca ccaagatcca atcctcgcga gcagaatcga tggcgacggc ggcggtgatg 60
gctgctgtac cctcgtcttc gccggccgac gcggcggagg ctgtggttat gacggaggcg 120
gtgccttctc tcccgcaacg ccagcagccg gtgttcgtcg agggcagagg cgggaagctg 180
cgtgaccacg cctacagccg caagcagaag tcgctcggcc tcctctgctc caatttcgtc 240
gctctgtaca accgcgacga cgtggagtct atcgggctgg acgacgcggc taggaggctc 300
ggcgtggaga ggcgccggat ctacgacatc gtcaacgtgc tcgagagcgt agggatcctc 360
gtgaggaagg ccaagaatcg ttattcttgg ataggcttcg gcggcgtccc aatggcattg 420
cgagaactca aggagagggc attgagagag aagtctggat tggctcctct gcccgtggag 480
gagccgtctg cagccattat gtcggatgac gaagatgaag ataagatggg tgacgctgat 540
ggtgataccg agagcgagaa gctgagtcaa ccagttgaca accettctga caacaagcca 600
ggcgcacctc gctgccggct tagatctgac cataggaagg agaagtcgct tgggttgctc 660
acgcagaatt tcgtgaaact cttcctgacc atggaggttg acacaatctc acttgatgaa 720

```
gctgcaaagc tgctacttgg agaaggtcac gcagagaaca gtatgagaac taaagtccgg    780
agactgtatg acattgctaa cgtgctgtca tcactgaatt tcattgataa gatacaacag    840
gcagactcaa ggaaacctgc attccggtgg ttgggctcag cggggaaacc aaaagctgaa    900
aatggtgtca caatcgcagt acctccacca gggaagacca tatcgaacaa gagagcattt    960
gggactgaac tcactaacat tgacataaac agaagcagac tggactcaac aatcccaaag   1020
aaagcaaagc tgacactgag tggtggtgaa attttgaaga actgcaaatt gtcagtgcag   1080
aaacagctcg ggcagggtag caagggtggt tttgtttatg ggcctttcca ccctgctggt   1140
gcaagaaaac aagagcttga caatggtaat aaaggacaca cagataatgt tcaaaactgg   1200
gagagccttg ctgcttcatt tcgaccacaa taccagaacc aagcattggg cgatcttttt   1260
gctcattatg tggaagcctg gaaatcgtgg tactctgaat ttgcgcaagg cagcagcatg   1320
atgcagcagc actttggcat gcctgtcatt aaccagtttt tgtagtcaaa tcattaatct   1380
caagcactgt atttcttacc gcgtgcttgg gattttacct gtacaatttg tcaagaaaat   1440
gagggaacaa agacatgacc taggcctagg atccattgta tttactcaat gcaggaaata   1500
tccaggctat tcagtccaaa agataaacta gcctcattgt taggcatgta actgttaatt   1560
catctagcta taagctttcc t                                             1581
```

<210> 15
<211> 441
<212> PRT
<213> Oryza sativa

<400> 15

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

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

<210> 16
<211> 1461
<212> DNA
<213> Zea mays

<400> 16

```
gcacgagccc tcccagctcg cgcgccctcg atggatgcct ccgccgccac ccccgccccc     60
gggccttcct tctctggcgc cgagtcctcc gcggccgccg cccagccgcc ggcggaggct    120
ccgcagttgc gcgtccacgg cgccggcagc ggcagcggcg tcgcgcgagc ctgccgccac    180
cacgcgtaca gccgcaagca gaagtcgctc ggccttctct gctccaactt cgtggcgctg    240
tacgaccggg aggacgtgga ggtgattggg ctggacgacg cggccaagcg tctcggcgtc    300
gagcgacgcc ggatctacga catagtcaac gttctcgaga gcgtcgggat tcttgtgcgg    360
agggccaaga atcggtatac atggctcgga ttcgggggag tccctgctgc gctgaaagaa    420
ctcaaggaga gggcgctaag gaagatgtcc ggatcaccgg tgttactgtc aatggaggac    480
tcgtctactg ccaacttatc agatgatgag gatgatgaaa aattgggcga tgctgatgaa    540
gatgctgaga gcgagaagct cagccaacct gttgacaata cgtctgataa gcctgacgca    600
cccagctgcc gccttagatc tgatcatcgg aaggagaagt cccttgggct cctcactcag    660
aattttgtca agctcttcct caacatggag gttgggacaa tctcacttga cgaagctgca    720
aggcttctcc ttggagaggg acatgcagac agcaacatga gaacagccaa agttcgtcga    780
ttgtatgaca ttgccaatgt gctgtcttct ttgaacctca ttgagaagac gcagcaagca    840
gacacaagaa aacctgcatt ccggtggcta ggccaggcaa agcgaaagca agataacaat    900
gtcatggttt ctgtacctcc atccatgaag gcaatgccca ataagagatc atttggtact    960
gatcttacaa acattgacaa taagcgaggc aagttagact cagcagcgga gaacaaagtc   1020
aagctcatgc agggtgctgg taacatagtg aagacttttg agaggcagct ggtgcaaggg   1080
aaaaggaatg actttgttta tgggcccttc caccctgctg gtgcaaagaa acacgaaact   1140
gatgatcaaa ctgttaagca gcaggagagg aagaacattc aggactggga aaaccttgct   1200
gtgtccttcc gtccacaata tcagaatcaa gcactgaatg atcttttttgg tcattatgtg   1260
gaagcatgga aatcatggta cgtggatctt acccaggaaa cggcatcatg aagcagaatg   1320
ttggcaggtc agttgtaact tgtaagccgt ttccagtagt tagatcaact gagttgttga   1380
```

tgcttgctta gggaaccacc ctgtacaaga tggaataaaa acacgaaggc aattttgtct 1440
ttcgactgta tttactcgat g 1461

<210> 17
<211> 426
<212> PRT
<213> Zea mays

<400> 17

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

```
        355                 360                 365
Glu Thr Asp Asp Gln Thr Val Lys Gln Gln Glu Arg Lys Asn Ile Gln
    370                 375                 380
Asp Trp Glu Asn Leu Ala Val Ser Phe Arg Pro Gln Tyr Gln Asn Gln
385                 390                 395                 400
Ala Leu Asn Asp Leu Phe Gly His Tyr Val Glu Ala Trp Lys Ser Trp
                405                 410                 415
Tyr Val Asp Leu Thr Gln Glu Thr Ala Ser
            420                 425
```

<210> 18
<211> 1512
<212> DNA
<213> Ostreococcus tauri

<400> 18

```
atgaacgggg atgggggggt ggagagctcg cgcgcgtcgg gcgggacgga tgtcccgggc     60
gaagacgcgg cgcggacgcg cgcgacgccg agcgcgacgt ctcggaaaga taagtcgttg    120
tggacgctgt gtgaacggtt tttgacgatt tatggagatg gatcgaagga gagcgtgtcg    180
ctggacgatg cggcgacgcg gctcggggtg gagcgtcgaa ggatttacga cgtcgcgaac    240
gtgttggaga gcgtggaggt gctcgagcgc aaggcaaaga atcagtacac gtggcacggc    300
gtgcgaagac tcccggagtg cctgaaacgg ctcaaggaga gtggattgcg ggaatttggg    360
acggatgtcg agctcgatgg aagcacgagc gagggacgtg acggggaaaa ggaggatgga    420
acggctcgag gcggcgacgc ctccgatcga agctctccca caaactcgtc gacgatcaac    480
ttggacgcga aacagcgggg tgcgaaggtt actgggacta agttctttgg acaggggcga    540
ttcgccgtct cgagcgccag ttacgacagc cggcgggaga agagtcttgg cttgctgtcc    600
caaaaattcg ttcaactctt cctcgcgtca aagatgaatg ttgtcagtct ggaaacggcc    660
gccaggatta ttatgggaga agacgacgac gatgaagcca agttgaagac taaaattcgt    720
cgattgtacg atatcgccaa catcttgtgc tccttgcgcc tgattcgaaa ggtacacgtg    780
ggcgagacta gaaaaccagc cttcctttgg ctacagcgag aaaactccat cgcggagctc    840
atcgcgcagg gcaaggggct gatgtggttc gataaattga acgaagagga agagatgcgt    900
ttacaggcgt ccatcattga atcaaaactc gacgggaacg aggtgttgat ggttgacgcg    960
gagaataaac gacgaggcgc gttttacgat tcgcagacta agtcgggatc taagcgtcct   1020
cgaggacgtc caagactgcc gggtggtgac gtcgacgcgt tgccttcatc cgtacctcca   1080
ttaggcgcgt cttcgatgac gcccgaagaa gcggcgaggt tcaatcagtt gttcgcgttc   1140
accacgtcac aactcatggc acagtacccg ctggacgtgc gcgcggatgt gaacagcatc   1200
atggcgcaga gcgcgatggg taacgcaaac taccttcatc ttctcgcgca gtcttccgtg   1260
gcccaagcgc acgcggcgcg acggatggag aacaccgcaa aaggaatgag ctctgattcc   1320
agcatcgaca acacggctat gttcgtgccg ccacttcctg cgtttcccgc gatgcttccg   1380
gcgtggggca tgtctggctt gtcttaccac agcaaccaca tggaacacat gatgcggatg   1440
tacgaaactt ctctggcaac gattagcgac aacgaccccc cgcatccaga tggaaagatt   1500
gccgaacaat aa                                                       1512
```

<210> 19
<211> 503
<212> PRT
<213> Ostreococcus tauri

<400> 19

```
Met Asn Gly Asp Gly Gly Val Glu Ser Ser Arg Ala Ser Gly Gly Thr
1               5                  10                  15
Asp Val Pro Gly Glu Asp Ala Ala Arg Thr Arg Ala Thr Pro Ser Ala
            20                  25                  30
Thr Ser Arg Lys Asp Lys Ser Leu Trp Thr Leu Cys Glu Arg Phe Leu
```

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

```
465                 470                 475                 480
Tyr Glu Thr Ser Leu Ala Thr Ile Ser Asp Asn Asp Pro Ser His Pro
                485                 490                 495
Asp Gly Lys Ile Ala Glu Gln
            500
```

<210> 20
<211> 1161
<212> DNA
<213> Medicago truncatula

<400> 20

```
atggctgctt cctcctccga tcctccttca cgacaccaca cttacgaccg taaacaaaaa     60
tcccttggcc tcttatgcac caagttcttg agcttgtata acaaagatga tgttcgttta    120
attggtctcg acgatgctgc cgccaaatta ggtgttgaga gaagacggat ctatgatatt    180
gtcaatgttc tcgaaagcat cggggttctt gcaagaaaag ccaagaatca gtatacctgg    240
aaaggttttg cggcaattcc tgttgctcta caggagctta aggaagaggg tttatggcag    300
aatctcaatt cttcacaaga aggtgccaat gaagatgtga aggtatcgga tgaagaggat    360
gaggatgaat tgttatccca aaccactgga agtcagggtg aatcattatc ccaacccact    420
ggaagtcaga atgacaatct aaaccctaat tccgcttttc ccagatcttt gaaaaatgac    480
agaagggaaa aatctctggc gctgcttact cagaattttg tcaagctctt tgtctgttcc    540
aacctggaaa tgatatcgct tgatgatgca gcaaggttgt tgcttggaga tgcatataat    600
tcatcaacaa tgagaacaaa agtcaggcgc ctttatgata ttgcaaacgt gctaacctcc    660
atgaacctta ttgagaagac ccataccaca gatacaagaa aaccagcatt caggtggcta    720
ggcttaaaag cgaagacatt gaatgaggca tcactttaca attcaaaaca aaatgagtct    780
aggaaaaggg cgtttggaaa tgatgtcaca aacataagct ttgccaggaa tagaatggac    840
ttgttcatgg gcggggactt taagaagcaa aagacaatgg aaaatgatag tggactatgt    900
caggaagatg tgaaacaagg cataaaacag acttcagcag ctaactatca atttggtcct    960
tttgctcctg cctttgtatc caaagctgga agctctgaga ataaagtgaa gcaggtgcat   1020
gactgggaga gtctcgctac tgaacattgc cctcagtatc aaaaccaagc tttgaaagaa   1080
ctttactctc attcatggga agcatggaaa tcttggtact ctgaagttgc tgggaagagg   1140
tcaacgcaag ttttgtagta a                                             1161
```

<210> 21
<211> 385
<212> PRT
<213> Medicago truncatula

<400> 21

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

```
        115                 120                 125
Thr Gly Ser Gln Gly Glu Ser Leu Ser Gln Pro Thr Gly Ser Gln Asn
    130                 135                 140
Asp Asn Leu Asn Pro Asn Ser Ala Phe Pro Arg Ser Leu Lys Asn Asp
145                 150                 155                 160
Arg Arg Glu Lys Ser Leu Ala Leu Leu Thr Gln Asn Phe Val Lys Leu
                165                 170                 175
Phe Val Cys Ser Asn Leu Glu Met Ile Ser Leu Asp Asp Ala Ala Arg
            180                 185                 190
Leu Leu Leu Gly Asp Ala Tyr Asn Ser Ser Thr Met Arg Thr Lys Val
        195                 200                 205
Arg Arg Leu Tyr Asp Ile Ala Asn Val Leu Thr Ser Met Asn Leu Ile
    210                 215                 220
Glu Lys Thr His Thr Thr Asp Thr Arg Lys Pro Ala Phe Arg Trp Leu
225                 230                 235                 240
Gly Leu Lys Gly Lys Thr Leu Asn Glu Ala Ser Leu Tyr Asn Ser Lys
                245                 250                 255
Gln Asn Glu Ser Arg Lys Arg Ala Phe Gly Asn Asp Val Thr Asn Ile
            260                 265                 270
Ser Phe Ala Arg Asn Arg Met Asp Leu Phe Met Gly Gly Asp Phe Lys
        275                 280                 285
Lys Gln Lys Thr Met Glu Asn Asp Ser Gly Leu Cys Gln Glu Asp Val
    290                 295                 300
Lys Gln Gly Ile Lys Gln Thr Ser Ala Ala Asn Tyr Gln Phe Gly Pro
305                 310                 315                 320
Phe Ala Pro Ala Phe Val Ser Lys Ala Gly Ser Ser Glu Asn Lys Val
                325                 330                 335
Lys Gln Val His Asp Trp Glu Ser Leu Ala Thr Glu His Cys Pro Gln
            340                 345                 350
Tyr Gln Asn Gln Ala Leu Lys Glu Leu Tyr Ser His Tyr Met Glu Ala
        355                 360                 365
Trp Lys Ser Trp Tyr Ser Glu Val Ala Gly Lys Arg Ser Thr Gln Val
    370                 375                 380
Leu
385
```

<210> 22
<211> 1510
<212> DNA
<213> Triticum aestivum

<400> 22

```
ccctcctcca aggaccacaa aattcccccc tccctaccct ttccacctcg acatccaccg     60
gaaatggacc ccgccgccgc cgccgggggt ccggcggccg ctgcgcagcc accgccccct    120
cctcctcctc cttacctgcc cccgcggctg gtcatcgacg gcgccggagg cggctcgggc    180
gccgccgtga gggcctgccg gcaccacgcc tacagccgca agcagaagtc cctcggcctc    240
ctctgctcca acttcgtggc gctgtacgac cgggacgacg tggagacggt ggggctggac    300
gacgccgcca ggcggctcgg cgtcgagagg cgccggatct acgacatcgt caacgtgctc    360
gagagcgtcg ggattctcgt gaggagggcc aagaatcggt acacatggat cggatttgag    420
ggcgtccctg ccgcgctcaa ggagcttaag gagaggacac tgagagagat gtctggatta    480
gctccgccac cggaggaatc atctgctgcc aatgtgtcgg acgatgaaga cgacgatgat    540
aaattgggcg atgcagatgg ggacgccgac agcgagaagc ttagccagtc cctcgacaat    600
gcttctgata agcctaacgt gcccatgtgc ccacctagat ctgtagacca taggaaggag    660
aagtcgcttg cgctgctcac gcagaatttt gtcaagctct tcctcaccat ggaggttgag    720
```

```
acagtctcac ttgacgaggc tgcaaggctg ctccttggag agagacatgc cgagagcaat    780
atgagaacta aggttcgtcg actgtatgac atcgccaatg tgctatcttc tttgaacctc    840
attgagaaga cacagcaggt ggactccaga aaacctgcat tccggtggct cggtcaggca    900
aagcgaaagg aaggtgccac tgtcacggtt gctttaccac catccaggaa gattatgtct    960
agcaagaggg catttggtac cgacatcaca aacattgaca ataagagggg caagttagtc   1020
ttggaaacag agaacaaacc caaactcatg cagggtggca gcagcatgtt gaaaactttc   1080
gagagtcagc tcgggcaagg gaagagtagt ggctttgttt atgggccctt ccaccctgct   1140
ggtgcaagga aacatgaagt tgatgatcag acagtgaggg agaatgagat gaagaacatt   1200
caagactggg agagtctcgc tgtttcattc cgtccacagt accaaaatca cgcgctgaat   1260
gatcttttg gccattatgt tgaagcatgg aaatcatggt acttggatct tacacgggat   1320
tcgacctcat gaaaggatgg tttattttag gcaggtctgt tgtaatcagt tcttgtaggc   1380
agatcaactc atctgccgag gaccctagaa aggattcgac ctgtacaaga tgtcataaaa   1440
gcacaaaggc tcatttcgcc tttgctcgat gtatttactc aatgcaaaaa agaaacggaa   1500
taatgccaaa                                                          1510
```

<210> 23
<211> 422
<212> PRT
<213> Triticum aestivum

<400> 23

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

```
            260                 265                 270
Phe Arg Trp Leu Gly Gln Ala Lys Arg Lys Glu Gly Ala Thr Val Thr
        275                 280                 285
Val Ala Leu Pro Pro Ser Arg Lys Ile Met Ser Ser Lys Arg Ala Phe
    290                 295                 300
Gly Thr Asp Ile Thr Asn Ile Asp Asn Lys Arg Gly Lys Leu Val Leu
305                 310                 315                 320
Glu Thr Glu Asn Lys Pro Lys Leu Met Gln Gly Gly Ser Ser Met Leu
                325                 330                 335
Lys Thr Phe Glu Ser Gln Leu Gly Gln Gly Lys Ser Ser Gly Phe Val
            340                 345                 350
Tyr Gly Pro Phe His Pro Ala Gly Ala Arg Lys His Glu Val Asp Asp
        355                 360                 365
Gln Thr Val Arg Glu Asn Glu Met Lys Asn Ile Gln Asp Trp Glu Ser
    370                 375                 380
Leu Ala Val Ser Phe Arg Pro Gln Tyr Gln Asn His Ala Leu Asn Asp
385                 390                 395                 400
Leu Phe Gly His Tyr Val Glu Ala Trp Lys Ser Trp Tyr Leu Asp Leu
                405                 410                 415
Thr Arg Asp Ser Thr Ser
            420
```

<210> 24
<211> 1530
<212> DNA
<213> Oryza sativa

<400> 24

```
aaatctctcc tcccgctcgc cgaaacccta gccgtcgccg cgccatggcc gccgccgccg     60
atgctccccc gcctcctccg gaggtcgccc cgcccgcgcc cgcgcccgcc cccgcgccgg    120
cgccatatca gccgccgcgg ctggcggtgg ccgacggagc gggaggtggc ggcggcggcg    180
gcgggaagcc gtgcaggcac cacgcgtaca gccgcaagca gaagtcgctc ggcctcctct    240
gcaccaactt cgtggcgctg tacgaccggg aggacgtgga gtcggtgggg ctggacgacg    300
cggcgaggcg gctgggcgtc gagaggcgcc ggatctacga catcgtcaac gtgctcgaga    360
gcatcgggat gctcgtgagg agggccaaga atcggtatac gtggatcggc ttcggtggag    420
tccctgcggc gctcgcgaaa ctcaaggaga tgtcactgag ggcggtgtca agcgtggcgt    480
caccgtcgct ggatgaaaca tctgctgcta atgtctcgga tgatgaggat gatgacaagt    540
tagatgatgc tgagggcgat gcggagagcg agaagctcag cctcagccag tccattgaca    600
atccttctga taagcctgat gcacccccct gcaagcttcg atccgagcat aggaaggaga    660
agtcccttgg gctgctcact cagaattttg tcaagctctt cctcaccatg gagattgaga    720
cgatctcact tgatgaagcc gcaaagcggc tccttggaga gggacatgcg gcgaacaata    780
tgagaaccaa agttcggcga ttgtacgata ttgccaatgt gctgtcttct ttgaatctta    840
ttgagaagac acaacaggcg gactcaagaa aacctgcatt ccggtggctg ggccaggcaa    900
aacggaatga aggcgtcacg gttgctttac ccccaaccaa gacgttgcct aacaagagag    960
catttggtac tgatctgact aacattgaca ataagagggg taagttggac tccacaatgg   1020
agaacagagg caagcccacg caggatggtg gcaacctatt caataatttg cagaggcaat   1080
tagggcaaga gaacaggagc gattttgctt atggcccctt ccaccctgct gttgcaagga   1140
aacaagaaca tggtaatcgc acagtacaag agaaggagag gaagagcatt caagactggg   1200
agaaccttgc ttcttctttc cgtccgcaat atcaaaatcc aggactgaac gatctttttg   1260
gccactacat ggaagcaagg aggtcatggt actcggatct caggcgagac agcgcatcat   1320
aaaaccagat tcaggccggt ccgttgtaaa caaaattctt gtaggcagat aaactcattt   1380
gttgatgcat ggttagggat ttttgagtta tacaagatgt cacaaaaaca aaggctagtg   1440
ctggctttg attgattgta tttactcagt gcaagaaatt ggaaagtcgt ttcgtcaata   1500
atactaggat acctgattgt cgatttgttt                                   1530
```

<210> 25
<211> 425
<212> PRT
<213> Oryza sativa

<400> 25

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

```
    370                 375                 380
Trp Glu Asn Leu Ala Ser Ser Phe Arg Pro Gln Tyr Gln Asn Pro Gly
385                 390                 395                 400
Leu Asn Asp Leu Phe Gly His Tyr Met Glu Ala Arg Arg Ser Trp Tyr
                405                 410                 415
Ser Asp Leu Arg Arg Asp Arg Ala Ser
            420                 425
```

<210> 26
<211> 645
<212> DNA
<213> Silene latifolia

<400> 26

```
gcacgagtca ccatcagaaa atcgatcttt gcatcatggt tatagtagaa agcagaaatc     60
tttgggtctt ctttgttcaa atttcttgag attgtataat cgagacgatg tcgatttgat    120
tggtctagat gatgctgcca gtaaattagg agtcgagcga agaaggatct atgatattgt    180
taatgtgttg gagagtgtag gggtgttagc aagaaaagcg aaaaatcagt acacatggaa    240
gggctacaag gcaattccta aggctcttgc cttgttaaag gaagatggtt tgaaggagaa    300
ttttggtact gcagagggaa gaagtagagt aaagtttcag gttttagact ttgatgatga    360
tgtagcttct aaccctgata ccggaagcca gtcggaggca ccaacacata gtggtgcttc    420
taagcctgat aacagaaggg aaaagtcatt ggggcttcta acacaaaact ttgtgaagct    480
attcctttgt tctgaggctg agctgatctc tctggaggaa gctgcaaagt gcttgcttgg    540
agatggtctt aatgcacaag tgatgcgaac taaagtgaga cgattgtatg atattgccaa    600
tgtattatcc tccatgcaac tcattgaaaa gactactcaa acaga                    645
```

<210> 27
<211> 214
<212> PRT
<213> Silene latifolia

<400> 27

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

```
Cys Leu Leu Gly Asp Gly Leu Asn Ala Gln Val Met Arg Thr Lys Val
            180                 185                 190
Arg Arg Leu Tyr Asp Ile Ala Asn Val Leu Ser Ser Met Gln Leu Ile
        195                 200                 205
Glu Lys Thr Thr Gln Thr
    210
```

<210> 28
<211> 785
<212> DNA
<213> Helianthus

<400> 28
cgggggacg gatcgaactg ttggattgga tgacgccgca acgaggctag gtgttgagag 60
acggcggatt tatgacattg ttaatgtttt ggaaagcgtt ggcgtccttg tgaaaaaggc 120
aaaaaacacg tatcattggt taggattggg ggcaatacct aaggctttag agcagctaaa 180
ggaagaaggt tttaggaata atgatgaaca tgttgatgtt aaatctggaa aggtttctga 240
tgatgaagaa gatgaaagag tttctagcca tagtgtttct ttacaggaaa aatcagatct 300
agattcaatg cataaaactt cagggccgtt taaatcgggt agtgtaacgg aaaatcggaa 360
ggaaaaatct ttggggcttc ttaccaagaa tttcatcaag cttttcctat gcactaattc 420
ggatatgctt tcgcttgatg acgctgcaaa aatattgctt ggagatgctc agaatccgtc 480
attgacccga actaaagtca gacgcctata tgatattgct aatgtcttgt cttccatgca 540
cttcatcgag aagatccatc acccagaaac ccgaaagccc gcctttaggt ggttgggaat 600
ggcaagtcac ccaaatagcc aaacgagatc agccactggt gtagctcata tagagtccaa 660
aaaaagggca tttgggactg agcttacaaa catttgtttc aaaagaagca agttgggtga 720
ccaggaggtg aagctgtttg accttcagag tcaacgcccg cctatcaaac ccagcactca 780
taatc 785

<210> 29
<211> 261
<212> PRT
<213> Helianthus

<400> 29
```
Gly Gly Thr Asp Arg Thr Val Gly Leu Asp Asp Ala Ala Thr Arg Leu
1               5                   10                  15
Gly Val Glu Arg Arg Arg Ile Tyr Asp Ile Val Asn Val Leu Glu Ser
            20                  25                  30
Val Gly Val Leu Val Lys Lys Ala Lys Asn Thr Tyr His Trp Leu Gly
        35                  40                  45
Leu Gly Ala Ile Pro Lys Ala Leu Glu Gln Leu Lys Glu Glu Gly Phe
    50                  55                  60
Arg Asn Asn Asp Glu His Val Asp Val Lys Ser Gly Lys Val Ser Asp
65                  70                  75                  80
Asp Glu Glu Asp Glu Arg Val Ser Ser His Ser Val Ser Leu Gln Glu
                85                  90                  95
Lys Ser Asp Leu Asp Ser Met His Lys Thr Ser Gly Pro Phe Lys Ser
            100                 105                 110
Gly Ser Val Thr Glu Asn Arg Lys Glu Lys Ser Leu Gly Leu Leu Thr
        115                 120                 125
Lys Asn Phe Ile Lys Leu Phe Leu Cys Thr Asn Ser Asp Met Leu Ser
    130                 135                 140
Leu Asp Asp Ala Ala Lys Ile Leu Leu Gly Asp Ala Gln Asn Pro Ser
145                 150                 155                 160
```

```
Leu Thr Arg Thr Lys Val Arg Arg Leu Tyr Asp Ile Ala Asn Val Leu
                165                 170                 175
Ser Ser Met His Phe Ile Glu Lys Ile His His Pro Glu Thr Arg Lys
            180                 185                 190
Pro Ala Phe Arg Trp Leu Gly Met Ala Ser His Pro Asn Ser Gln Thr
        195                 200                 205
Arg Ser Ala Thr Gly Val Ala His Ile Glu Ser Lys Lys Arg Ala Phe
    210                 215                 220
Gly Thr Glu Leu Thr Asn Ile Cys Phe Lys Arg Ser Lys Leu Gly Asp
225                 230                 235                 240
Gln Glu Val Lys Leu Phe Asp Leu Gln Ser Gln Arg Pro Pro Ile Lys
                245                 250                 255
Pro Ser Thr His Asn
            260
```

<210> 30
<211> 1125
<212> DNA
<213> Glycine max

<400> 30

```
atgacttctt tgtctaccca tcacacctat agccgaaagc aaaaatctct cggcctcctc     60
tgcaccaatt ttctgagttt gtacaacaaa gaaggtgtgc gcctggtcgg tctcgatgac    120
gcagcttcgc ggttaggtgt agaaagacgt cggatctacg acatcgttaa tgttttagag    180
agtgtcggtg tgctaaccag aaaagctaag aatcagtata cttggaaagg attttgtgca    240
attcctgctg ccttacaaga acttaaggat gaaggtttga aggagaatcc tggatcattt    300
gatgactcca atgacaatgc aaaagtgtct gatgatgagg aggaggaaga agaaacattt    360
cctaatatta acattgggag tcagagtgag aaggaaaatc ctgattacac tgctactgtc    420
aaatcctcca agaatgaaaa cagaagggaa aaatctctgg cattgcttac tcagaatttt    480
gtcaagcttt ttgtctgctc taattttgaa atgatctccc ttgacgaagc tgcaaagttg    540
ttgcttggaa atgccaacaa tagaacaaaa gtcagacgtc tatatgatat tgcgaatgtg    600
ttatcctcca tgaacctaat cgagaagacc catacaacaa acacaagaaa accagcattc    660
aggtggttgg gtgtgagggg aaaaacatgg ggtggatcag ttgacttggc tcaaaattca    720
aatgtcaaag agtcacggaa aaggatgttt ggaactgaca ttggaaacat aagttttaaa    780
aggaacaaag ttgatttgtc catggatggc cagaattcta aaatgcaaaa tcaacacgaa    840
aatattagtc cacgagctca gttagagaaa atcatcataa aaaaagatgc taagcagact    900
tcagagagct atcagtttgg ccctttgct ccagcttatg tgcccaaagt tggaacctct    960
gaaaataata gtgtgaaaaa ggtgcatgac tgggacagcc tcgcacagga gcatcgtcct   1020
caatatcaaa accaaccttt aaaagatcta ttttcccact acatggaagc atggaaatct   1080
tggtactctg aagctgctgg gaaaaagtca acgaaaactt tataa                   1125
```

<210> 31
<211> 374
<212> PRT
<213> Glycine max

<400> 31

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

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

REIVINDICAÇÕES

1. Método para aumentar o rendimento de planta e/ou resistência ao estresse em relação às plantas de controle, caracterizado pelo fato de que compreende diminuir a expressão em uma planta de um ácido nucleico que codifica um polipeptídeo DEL1 e/ou diminuir a atividade de um polipeptídeo DEL1 e opcionalmente selecionar quanto a plantas tendo rendimento aumentado.

2. Método de acordo com a reivindicação 1, caracterizado pelo fato de que a dita proteína DEL1 compreende qualquer um ou mais dos motivos da: SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 e SEQ ID NO: 13.

3. Método para aumentar o rendimento de planta e/ou resistência ao estresse em relação às plantas de controle, caracterizado pelo fato de que compreende introduzir e expressar em uma planta um ácido nucleico DEL1 ou uma variante do mesmo.

4. Método de acordo com a reivindicação 3, caracterizado pelo fato de que o dito ácido nucleico ou variante codifica um ortólogo ou parálogo da proteína DEL1 da SEQ ID NO: 8.

5. Método de acordo com as reivindicações 3 ou 4, caracterizado pelo fato de que a dita introdução resulta no silenciamento mediado por RNA de um gene DEL1 endógeno.

6. Método de acordo com a reivindicação 3 ou 4, caracterizado pelo fato de que a dita variante é uma porção de um ácido nucleico DEL1 ou uma seqüência capaz de hibridizar a um ácido nucleico DEL1, porção ou seqüência hibridizadora estas que codificam um polipeptídeo que compreende a seqüência que corresponde à seqüência localizada no terminal C em relação aos domínios de ligação de DNA de uma proteína DEL1 e opcionalmente também uma ou mais das seqüências de assinatura da SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 ou SEQ ID NO: 13.

7. Método de acordo com a reivindicação 6, caracterizado pelo fato de que a dita porção codifica um polipeptídeo DEL1 que carece de parte ou toda da primeira e/ou parte ou toda da segunda região de ligação de DNA.

8. Método de acordo com a reivindicação 6 ou 7, caracterizado pelo fato de que a dita porção codifica um polipeptídeo que é representado pela SEQ ID NO: 2.

9. Método de acordo com qualquer uma das reivindicações 3, 4 ou 6 a 8, caracterizado pelo fato de que o dito ácido nucleico DEL1 ou variante do mesmo é superexpressado em uma planta.

10. Método de acordo com qualquer uma das reivindicações 3, 4 ou 6 a 9, caracterizado pelo fato de que o dito ácido nucleico DEL1 ou variante do mesmo é operavelmente ligado a um promotor específico de semente.

11. Método de acordo com qualquer uma das reivindicações de 1 a 10, caracterizado pelo fato de que o dito rendimento aumentado é rendimento de semente aumentado.

12. Método de acordo com qualquer uma das reivindicações de 1 a 10, caracterizado pelo fato de que a dita resistência ao estresse aumentada é a resistência aumentada ao estresse abiótico, preferivelmente resistência aumentada à radiação UV.

13. Construção, caracterizada pelo fato de que compreende:

(i) um ácido nucleico DEL1 f ou variante do mesmo;

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

(iii) uma seqüência de terminação de transcrição.

14. Construção de acordo com a reivindicação 13, caracterizada pelo fato de que a dita seqüência de controle é um promotor específico de semente.

15. Construção para o silenciamento mediado pelo RNA de

um gene DEL1 endógeno, caracterizada pelo fato de que compreende

(i) um ácido nucleico DEL1 ou variante do mesmo;

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

(iii) uma seqüência de terminação de transcrição.

16. Uso de uma construção como definida em qualquer uma das reivindicações 13 a 15, caracterizado pelo fato de ser para aumentar o rendimento ou aumentar a resistência ao estresse de uma planta.

17. Método para produzir uma planta transgênica tendo rendimento aumentado e/ou resistência ao estresse aumentada, caracterizado pelo fato de que compreende:

(i) introduzir e expressar em uma planta ou célula de planta um ácido nucleico DEL1 ou variante do mesmo;

(ii) cultivar a célula de planta sob condições que promovam o crescimento e o desenvolvimento vegetais.

18. Polipeptídeo DEL1 isolado, caracterizado pelo fato de que é selecionado do grupo que consiste de:

a) um polipeptídeo como dado na SEQ ID NO 21,

b) um polipeptídeo com uma seqüência de aminoácido que tem pelo menos 60 % de identidade de seqüência, preferivelmente 70 % de identidade de seqüência, mais preferivelmente 80 % ou 90 % de identidade de seqüência, o mais preferivelmente 95 %, 96 %, 97 %, 98 % ou 99 % de identidade de seqüência com a seqüência de aminoácido como dada na SEQ ID NO 21,

c) um homólogo, um derivado e/ou fragmento funcional de uma proteína como definida em (a).

19. Seqüência de ácido nucleico isolada, caracterizada pelo fato de que compreende:

(i) uma seqüência de ácido nucleico representada pela SEQ ID

NO: 20, ou o seu filamento complementar;

(ii) uma seqüência de ácido nucleico que codifica uma seqüência de aminoácido como definida na reivindicação 27;

(iii) uma seqüência de ácido nucleico capaz de hibridizar (preferivelmente sob condições severas) com uma seqüência de ácido nucleico de (i) ou (ii) acima, seqüência de hibridização esta que preferivelmente codifica uma proteína tendo atividade que aumenta o rendimento;

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

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

(vi) uma seqüência de ácido nucleico que tem em ordem crescente de preferência 60 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % ou 99 % de identidade de seqüência com a SEQ ID NO: 20 ou o complemento desta; e

(vii) uma porção de uma seqüência de ácido nucleico de acordo com qualquer uma das reivindicações de (i) a (vi) acima, porção esta que preferivelmente codifica uma proteína tendo atividade que aumenta o rendimento.

20. Uso de um ácido nucleico DEL1 ou variante do mesmo, ou de um polipeptídeo DEL1 ou homólogo do mesmo, caracterizado pelo fato de ser para (i) melhorar o rendimento, especialmente rendimento de semente, em relação às plantas de controle, ou (ii) para aumentar resistência ao estresse, especialmente resistência ao estresse abiótico, em relação às plantas de controle.

MSDLSPERFKLAVTSPSSIPESSSALQLHHSYS**RKQKSLGLLCTNFLALYNREGIEMVGLDD**
AASKLGVERRRIYDIVNVLESVGVLTRRAKNQYTWKGFSAIPGALKELQEEGVKDTFHRFYV
NENVKGSDDEDDDEESSQPHSSSQTDSSKPGSLPQSSDPSKIDN**RREKSLGLLTQNFIKLFI**
CSEAIRIISLDDAAKLLLGDAHNTSIMRTKVRRLYDIANVLSSMNLIEKTHTLDSRKPAFKW
LGYNGEPTFTLSSDLLQLESRKRAFGTDITNVNVKRSKSSSSSQENATERRLKMKKHSTPES
SYNKSFDVHESRHGSRGGYHFGPFAPGTGTYPTAGLEDNSRRAFDVENLDSDYRPSYQNQVL
KDLFSHYMDAWKTWFSEVTQENPLPNTSQHR

Figura 1

Alinhamento de sequência multipla CLUSTAL W (1.83)

```
SEQID17         -MDASAATPAPGPSFSGAESSAAAAQPPAEAPQLRVH-GAGSGSG-VARACRHHAYSRKQ
SEQID23         -MDPAAAAGAP----AAAAQPPPPPPPPYLPPRLVID-GAGGGSGAAVRACRHHAYSRKQ
SEQID25         -MAAAADAPPPPPEVAPPAPAPAPAPAPYQPPRLAVADGAGGGGGGGGKPCRHHAYSRKQ
SEQID15         MATAAVMAAVPSSSPADAAEAVVMTEAVPSLPQRQQPVFVEGRGG----KLRDHAYSRKQ
SEQID8          ------MSDLSPERFKLAVTSPSSIPESSSALQ------------------LHHSYSRKQ
SEQID27         ------------------HESPS---ENRS----------------------LHHGYSRKQ
SEQID21         -----------------MAASSS---DPPS----------------------RHHTYDRKQ
SEQID29         ------------------------------------------------------------
SEQID19         ------MNGDGGVESSRASGGTDVPGEDAARTR-----------------ATPSATSRKD


SEQID17         KSLGLLCSNFVALYDREDVEVIGLDDAAKRLGVERRRIYDIVNVLESVGILVRRAKNRYT
SEQID23         KSLGLLCSNFVALYDRDDVETVGLDDAARRLGVERRRIYDIVNVLESVGILVRRAKNRYT
SEQID25         KSLGLLCTNFVALYDREDVESVGLDDAARRLGVERRRIYDIVNVLESIGMLVRRAKNRYT
SEQID15         KSLGLLCSNFVALYNRDDVESIGLDDAARRLGVERRRIYDIVNVLESVGILVRKAKNRYS
SEQID8          KSLGLLCTNFLALYNREGIEMVGLDDAASKLGVERRRIYDIVNVLESVGVLTRRAKNQYT
SEQID27         KSLGLLCSNFLRLYNRDDVDLIGLDDAASKLGVERRRIYDIVNVLESVGVLARKAKNQYT
SEQID21         KSLGLLCTKFLSLYNKDDVRLIGLDDAAAKLGVERRRIYDIVNVLESIGVLARKAKNQYT
SEQID29         ---------------GGTDRTVGLDDAATRLGVERRRIYDIVNVLESVGVLVKKAKNTYH
SEQID19         KSLWTLCERFLTIYGDGSKESVSLDDAATRLGVERRRIYDVANVLESVEVLERKAKNQYT
                                :.***** :**********:.*****: :* ::*** *

SEQID17         WLGFGGVPAALKELKERALRKMSGSPVLLSMEDSSTANLSDDEDDE-KLGDADEDAESEK
SEQID23         WIGFEGVPAALKELKERTLREMSG--LAPPPEESSAANVSDDEDDDDKLGDADGDADSEK
SEQID25         WIGFGGVPAALAKLKEMSLRAVSS-VASPSLDETSAANVSDDEDDD-KLDDAEGDAESEK
SEQID15         WIGFGGVPMALRELKERALREKSG-LAPLPVEEPSAAIMSDDEDED-KMGDADGDTESEK
SEQID8          WKGFSAIPGALKELQEEGVKDTFH---RFY--VNENVKGSDDEDDD---EESSQPHSS--
SEQID27         WKGYKAIPKALALLKEDGLKENFG---TAE--GRSRVKFQVLDFDD---DVASNPDTG--
SEQID21         WKGFAAIPVALQELKEEGLWQNLN---SSQEGANEDVKVSDEEDED---ELLSQTTGSQG
SEQID29         WLGLGAIPKALEQLKEEGFRNNDE------HVDVKSGKVSDDEEDE----RVSSHSVS--
SEQID19         WHGVRRLPECLKRLKESGLREFGT--DVELDGSTSEGRDGEKEDGTARGGDASDRSSPTN
                * *   :* .*  *:*  .               .       :         .

SEQID17         ----------LSQPVDNTSD-KPDAPSCRLRS-DHRKEKSLGLLTQNFVKLFLNMEVG-T
SEQID23         ----------LSQSLDNASD-KPNVPMCPPRSVDHRKEKSLGLLTQNFVKLFLTMEVE-T
SEQID25         L---------SLSQSIDNPSD-KPDAPPCKLRS-EHRKEKSLGLLTQNFVKLFLTMEIE-T
SEQID15         ----------LSQPVDNPSDNKPGAPRCRLRS-DHRKEKSLGLLTQNFVKLFLTMEVD-T
SEQID8          -----------SQTDSSKPGSLPQSSDP--SKIDNRREKSLGLLTQNFIKLFICSEAIRI
SEQID27         -----------SQSE------APTHSGA--SKPDNRREKSLGLLTQNFVKLFLCSEAE-L
SEQID21         E---------SLSQPTGSQNDNLNPNSAFPRSLKNDRREKSLALLTQNFVKLFVCSNLE-M
SEQID29         ----------LQEKSDLDSMHKTSGPFKSGSVTENRKEKSLGLLTKNFIKLFLCTNSD-M
SEQID19         SSTINLDAKQRGEKVTGTKFFGQGRFAVSSASYDSRREKSLGLLSQKFVQLFLASKMN-V
                          :                    : *:****.**:::*::**:  :
```

Figura 2

```
SEQID17         ISLDEAARLLLGEGHADSNMRTAKVRRLYDIANVLSSLNLIEKTQQADTRKPAFRWLG--
SEQID23         VSLDEAARLLLGERHAESNMRT-KVRRLYDIANVLSSLNLIEKTQQVDSRKPAFRWLG--
SEQID25         ISLDEAAKRLLGEGHAANNMRT-KVRRLYDIANVLSSLNLIEKTQQADSRKPAFRWLG--
SEQID15         ISLDEAAKLLLGEGHAENSMRT-KVRRLYDIANVLSSLNFIDKIQQADSRKPAFRWLG--
SEQID8          ISLDDAAKLLLGDAHNTSIMRT-KVRRLYDIANVLSSMNLIEKTHTLDSRKPAFKWLG--
SEQID27         ISLEEAAKCLLGDGLNAQVMRT-KVRRLYDIANVLSSMQLIEKTTQT-------------
SEQID21         ISLDDAARLLLGDAYNSSTMRT-KVRRLYDIANVLTSMNLIEKTHTTDTRKPAFRWLG--
SEQID29         LSLDDAAKILLGDAQNPSLTRT-KVRRLYDIANVLSSMHFIEKIHHPETRKPAFRWLG--
SEQID19         VSLETAARIIMGEDDDDEAKLKTKIRRLYDIANILCSLRLIRKVHVGETRKPAFLWLQRE
                :**: **: ::*:     .    . *:********:* *:.:* *

SEQID17         --------------------------QAKR-KQDNNVMVSVPPSMKAMPNKRSFGTDLTN
SEQID23         --------------------------QAKR-KEGATVTVALPPSRKIMSSKRAFGTDITN
SEQID25         --------------------------QAKR-NEG--VTVALPPT-KTLPNKRAFGTDLTN
SEQID15         --------------------------SAGKPKAENGVTIAVPPPGKTISNKRAFGTELTN
SEQID8          --------------------------YNGE----PTFTLSS-DLLQLESRKRAFGTDITN
SEQID27         ------------------------------------------------------------
SEQID21         --------------------------LKGK----TLNEASLYNSKQNESRKRAFGNDVTN
SEQID29         --------------------------MASH---PNSQTRSATGVAHIESKKRAFGTELTN
SEQID19         NSIAELIAQGKGLMWFDKLNEEEEMRLQASIIESKLDGNEVLMVDAENKRRGAFYDSQTK

SEQID17         IDNKRGK-------LDSAAENKVKLMQGAGNIVK----TFERQLVQGKRNDFVYGPFHPA
SEQID23         IDNKRGK-------LVLETENKPKLMQGGSSMLK----TFESQLGQGKSSGFVYGPFHPA
SEQID25         IDNKRGK-------LDSTMENRGKPTQDGGNLFN----NLQRQLGQENRSDFAYGPFHPA
SEQID15         IDINRSR-------LDSTIPKKAKLTLSGGEILKNCKLSVQKQLGQGSKGGFVYGPFHPA
SEQID8          VNVKRSKSSS--SSQENATERRLKMKKHSTPESSYNKSFDVHESRHGSRGGYHFGPFAPG
SEQID27         ------------------------------------------------------------
SEQID21         ISFARNR-------MDLFMGGDFKKQKTMENDSGLCQEDVKQGIKQTSAANYQFGPFAPA
SEQID29         ICFKRSK-----------------------------------------------------
SEQID19         SGSKRPRGRPRLPGGDVDALPSSVPPLGASSMTPEEAARFNQLFAFTTSQLMAQYPLDVR

SEQID17         GAKKHETDDQTVKQQERKNIQDWENLAVSFRPQYQNQALN--------------------
SEQID23         GARKHEVDDQTVRENEMKNIQDWESLAVSFRPQYQNHALN--------------------
SEQID25         VARKQEHGNRTVQEKERKSIQDWENLASSFRPQYQNPGLN--------------------
SEQID15         GARKQELDNG--NKGHTDNVQNWESLAASFRPQYQNQALG--------------------
SEQID8          TGTY----PTAGLEDNSRRAFDVENLDSDYRPSYQNQVLK--------------------
SEQID27         ------------------------------------------------------------
SEQID21         FVSK----AGS-SENKVKQVHDWESLATEHCPQYQNQALK--------------------
SEQID29         ------------LGDQEVKLFDLQSQRPPIKPSTHN------------------------
SEQID19         ADVNSIMAQSAMGNANYLHLLAQSSVAQAHAARRMENTAKGMSSDSSIDNTAMFVPPLPA

SEQID17         -----------------DLFGHYVEAWKSWYVDLTQETAS---------------
SEQID23         -----------------DLFGHYVEAWKSWYLDLTRDSTS---------------
SEQID25         -----------------DLFGHYMEARRSWYSDLRRDRAS---------------
SEQID15         -----------------DLFAHYVEAWKSWYSEFAQGSSMMQQHFGMPVINQFL
SEQID8          -----------------DLFSHYMDAWKTWFSEVTQENPLPNTSQHR-------
SEQID27         ------------------------------------------------------
SEQID21         -----------------ELYSHYMEAWKSWYSEVAGKRSTQVL------------
SEQID29         ------------------------------------------------------
SEQID19         FPAMLPAWGMSGLSYHSNHMEHMMRMYETSLATISDNDPSHPDGKIAEQ-----
```

Figura 2 (cont.)

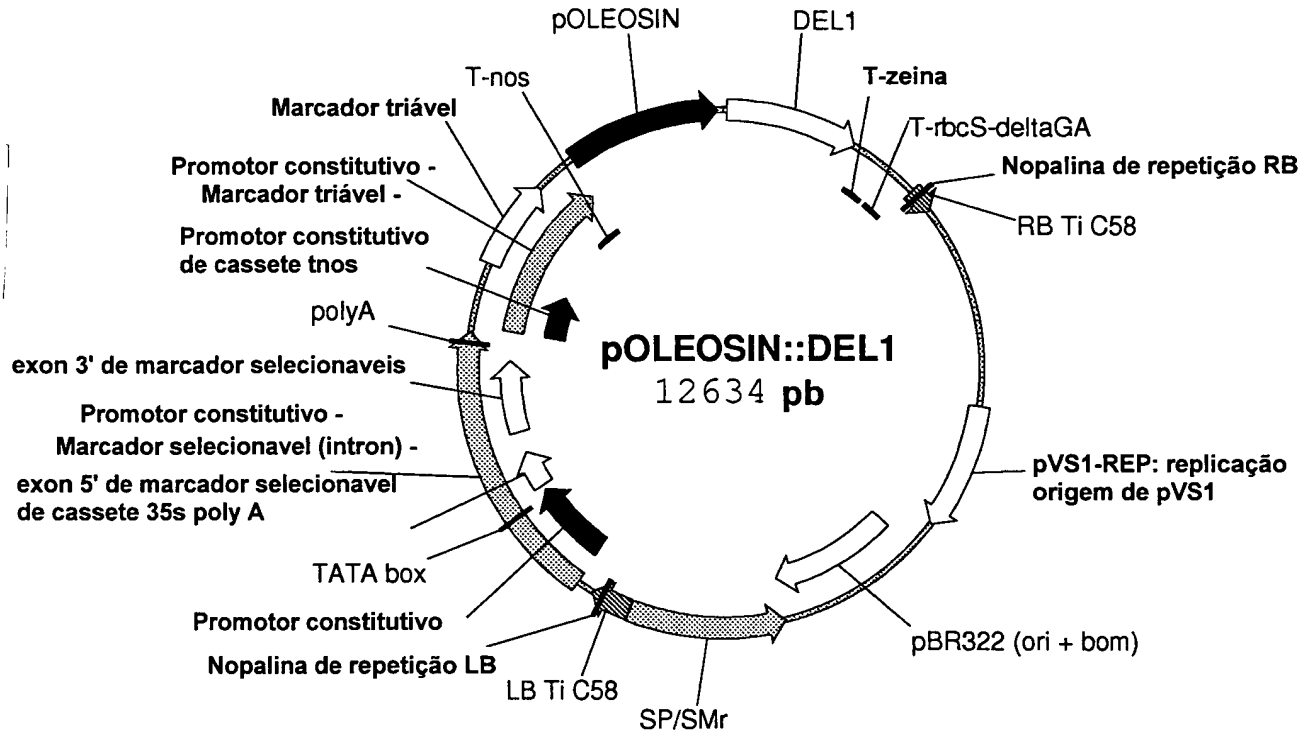
pOLEOSIN
DEL1
T-nos
T-zeina
Marcador triável
T-rbcS-deltaGA
Promotor constitutivo -
Marcador triável -
Nopalina de repetição RB
RB Ti C58
Promotor constitutivo
de cassete tnos
polyA
pOLEOSIN::DEL1
12634 pb
exon 3' de marcador selecionaveis
Promotor constitutivo -
Marcador selecionavel (intron) -
exon 5' de marcador selecionavel
de cassete 35s poly A
pVS1-REP: replicação
origem de pVS1
TATA box
Promotor constitutivo
pBR322 (ori + bom)
Nopalina de repetição LB
LB Ti C58
SP/SMr

Figura 3

SEQ ID NO: 1, fragmento de sequência de codificação DEL1

ATGGTTGGGCTTGATGATGCTGCCTCGAAATTAGGAGTGGAGAGACGAAGAATCTATGATAT
TGTTAATGTTCTGGAGAGTGTTGGGGTTTTAACAAGAAGAGCAAAGAATCAGTATACGTGGA
AAGGGTTTTCCGCAATTCCAGGAGCATTGAAGGAGCTACAAGAAGAGGGGGTTAAGGACACT
TTTCATCGTTTCTATGTCAATGAGAATGTTAAAGGATCTGATGATGAGGATGATGATGAAGA
GTCTTCTCAGCCTCACTCTAGTAGCCAGACTGATAGTTCAAAACCTGGTTCTCTTCCCCAAT
CTTCAGATCCCTCCAAAATAGATAACCGACGAGAGAAATCTTTAGGATTGCTTACTCAGAAC
TTTATCAAACTCTTTATTTGCTCTGAAGCTATTAGGATCATCTCCCTTGATGACGCTGCAAA
ATTACTGCTTGGTGATGCCCACAATACATCAATAATGCGAACTAAAGTGAGGCGGCTTTATG
ATATAGCAAATGTCTTGTCGTCAATGAATCTCATAGAGAAGACTCACACCTTAGATTCTAGG
AAACCAGCTTTCAAGTGGTTAGGGTACAATGGTGAGCCTACTTTCACACTGAGCAGCGATTT
GTTGCAATTGGAGTCAAGAAAAAGAGCTTTCGGAACTGATATTACAAACGTCAATGTTAAGA
GAAGCAAATCATCATCTTCGTCCCAAGAAAACGCTACAGAGAGAAGGCTAAAGATGAAAAAG
CACTCAACACCAGAGAGTTCTTATAACAAAAGCTTTGATGTTCATGAATCAAGACATGGATC
AAGAGGAGGTTACCATTTTGGACCTTTTGCACCAGGCACTGGTACATATCCAACTGCTGGTT
TAGAGGATAACTCTAGGAGAGCTTTTGATGTTGAGAATCTGGATTCTGATTACCGTCCCTCT
TACCAAAACCAAGTTTTGAAAGACCTCTTTTCCCATTACATGGATGCTTGGAAGACATGGTT
CAGCGAAGTCACCCAGGAGAATCCATTACCTAATACATCACAACACCGTTAG

SEQ ID NO: 2, fragmento de proteína DEL1

MVGLDDAASKLGVERRRIYDIVNVLESVGVLTRRAKNQYTWKGFSAIPGALKELQEEGVKDT
FHRFYVNENVKGSDDEDDDEESSQPHSSSQTDSSKPGSLPQSSDPSKIDNRREKSLGLLTQN
FIKLFICSEAIRIISLDDAAKLLLGDAHNTSIMRTKVRRLYDIANVLSSMNLIEKTHTLDSR
KPAFKWLGYNGEPTFTLSSDLLQLESRKRAFGTDITNVNVKRSKSSSSSQENATERRLKMKK
HSTPESSYNKSFDVHESRHGSRGGYHFGPFAPGTGTYPTAGLEDNSRRAFDVENLDSDYRPS
YQNQVLKDLFSHYMDAWKTWFSEVTQENPLPNTSQHR

SEQ ID NO: 3, prm00536

GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACAATGGTTGGGCTTGATGATGC

SEQ ID NO: 4, prm00324

GGGGACCACTTTGTACAAGAAAGCTGGGTCTAACGGTGTTGTGATGTATTAG

SEQ ID NO: 5, sequência promotora para expressão específica de embrião e aleurona

ggtcagccaatacattgatccgttgccaatcatgcaaagtattttggctgtggccgagtgcc
ggaattgataattgtgttctgactaaattaaatgaccagaagtcgctatcttccaatgtatc
cgaaacctggattaaacaatcctgttctgttctctagcccctcctgcatggccggattgttt
ttttgacatgttttcttgactgaggcctgtttgttctaaacttttttcttcaaacttttaact
ttttcatcacatcagaacttttctacacatataaacttttaacttttccgtcacatcgttcc
aatttcaatcaaactttcaattttggcgtgaactaaacacaccctgagtcttttattgctcc
tccgtacgggttggctggttgagaataggtattttcagagagaaaatctagatattgggagg
aacttggcatgaatggccactatatttagagcaattctacggtccttgaggaggtaccatga
ggtaccaaaattttagtgtaaattttagtatctcattataactaggtattatgaggtaccaa
atttacaatagaaaaaatagtacttcatggtactttcttaagtaccgtaaaattgctcctat
atttaaggggatgtttatatctatccatatccataatttgattttgataagaaaaaatgtga

Figura 4

GCACACCAAGCATGTCCATGACCTTGCACTCTTGGCTCACTCGTCAACTGTGAAGAACCTCA
AAAATGCTCAATATAGCTACAGGTGCCTGAAAAAATAACTTTAAAGTTTTGAACATCGATTT
CACTAAACAACAATTATTATCTCCCTCTGAAAGATGATAGTTTAGAACTCTAGAATCATTGT
CGGCGGAGAAAGTAAATTATTTTCCCCAAATTTCCAGCTATGAAAAAACCCTCACCAAACAC
CATCAAACAAGAGTTCACCAAACCGCCCATGCGGCCATGCTGTCACGCAACGCACCGCATTG
CCTGATGGCCGCTCGATGCATGCATGCTTCCCCGTGCACATATCCGACAGACGCGCCGTGTC
AGCGAGCTCCTCGACCGACCTGTGTAGCCCATGCAAGCATCCACCCCCGCCACGTACACCCC
CTCCTCCTCCCTACGTGTCACCGCTCTCTCCACCTATATATGCCCACCTGGCCCCTCTCCTC
CCATCTCCACTTCACCCGATCGCTTCTTCTTCTTCTTCGTTGCATTCATCTTGCTAGC

SEQ ID NO 6, Cassete de expressão com combinação de gene promotor, códon de códon de partida e parada do ORF indicado em negrito

GGTCAGCCAATACATTGATCCGTTGCCAATCATGCAAAGTATTTTGGCTGTGGCCGAGTGCC
GGAATTGATAATTGTGTTCTGACTAAATTAAATGACCAGAAGTCGCTATCTTCCAATGTATC
CGAAACCTGGATTAAACAATCCTGTTCTGTTCTCTAGCCCCTCCTGCATGGCCGGATTGTTT
TTTTGACATGTTTTCTTGACTGAGGCCTGTTTGTTCTAAACTTTTTCTTCAAACTTTTAACT
TTTTCATCACATCAGAACTTTTCTACACATATAAACTTTTAACTTTTCCGTCACATCGTTCC
AATTTCAATCAAACTTTCAATTTTGGCGTGAACTAAACACACCCTGAGTCTTTTATTGCTCC
TCCGTACGGGTTGGCTGGTTGAGAATAGGTATTTTCAGAGAGAAAATCTAGATATTGGGAGG
AACTTGGCATGAATGGCCACTATATTTAGAGCAATTCTACGGTCCTTGAGGAGGTACCATGA
GGTACCAAAATTTTAGTGTAAATTTTAGTATCTCATTATAACTAGGTATTATGAGGTACCAA
ATTTACAATAGAAAAAATAGTACTTCATGGTACTTTCTTAAGTACCGTAAAATTGCTCCTAT
ATTTAAGGGGATGTTTATATCTATCCATATCCATAATTTGATTTTGATAAGAAAAAATGTGA
GCACACCAAGCATGTCCATGACCTTGCACTCTTGGCTCACTCGTCAACTGTGAAGAACCTCA
AAAATGCTCAATATAGCTACAGGTGCCTGAAAAAATAACTTTAAAGTTTTGAACATCGATTT
CACTAAACAACAATTATTATCTCCCTCTGAAAGATGATAGTTTAGAACTCTAGAATCATTGT
CGGCGGAGAAAGTAAATTATTTTCCCCAAATTTCCAGCTATGAAAAAACCCTCACCAAACAC
CATCAAACAAGAGTTCACCAAACCGCCCATGCGGCCATGCTGTCACGCAACGCACCGCATTG
CCTGATGGCCGCTCGATGCATGCATGCTTCCCCGTGCACATATCCGACAGACGCGCCGTGTC
AGCGAGCTCCTCGACCGACCTGTGTAGCCCATGCAAGCATCCACCCCCGCCACGTACACCCC
CTCCTCCTCCCTACGTGTCACCGCTCTCTCCACCTATATATGCCCACCTGGCCCCTCTCCTC
CCATCTCCACTTCACCCGATCGCTTCTTCTTCTTCTTCGTTGCATTCATCTTGCTAGCATTT
AAATCAACTAGGGATATCACAAGTTTGTACAAAAAAGCAGGCTTCACA**ATG**GTTGGGCTTGA
TGATGCTGCCTCGAAATTAGGAGTGGAGAGACGAAGAATCTATGATATTGTTAATGTTCTGG
AGAGTGTTGGGGTTTTAACAAGAAGAGCAAAGAATCAGTATACGTGGAAAGGGTTTTCCGCA
ATTCCAGGAGCATTGAAGGAGCTACAAGAAGAGGGGGTTAAGGACACTTTTCATCGTTTCTA
TGTCAATGAGAATGTTAAAGGATCTGATGATGAGGATGATGATGAAGAGTCTTCTCAGCCTC
ACTCTAGTAGCCAGACTGATAGTTCAAAACCTGGTTCTCTTCCCCAATCTTCAGATCCCTCC
AAAATAGATAACCGACGAGAGAAATCTTTAGGATTGCTTACTCAGAACTTTATCAAACTCTT
TATTTGCTCTGAAGCTATTAGGATCATCTCCCTTGATGACGCTGCAAAATTACTGCTTGGTG
ATGCCCACAATACATCAATAATGCGAACTAAAGTGAGGCGGCTTTATGATATAGCAAATGTC
TTGTCGTCAATGAATCTCATAGAGAAGACTCACACCTTAGATTCTAGGAAACCAGCTTTCAA
GTGGTTAGGGTACAATGGTGAGCCTACTTTCACACTGAGCAGCGATTTGTTGCAATTGGAGT
CAAGAAAAAGAGCTTTCGGAACTGATATTACAAACGTCAATGTTAAGAGAAGCAAATCATCA
TCTTCGTCCCAAGAAAACGCTACAGAGAGAAGGCTAAAGATGAAAAAGCACTCAACACCAGA
GAGTTCTTATAACAAAAGCTTTGATGTTCATGAATCAAGACATGGATCAAGAGGAGGTTACC

Figura 4 (cont.)

ATTTTGGACCTTTTGCACCAGGCACTGGTACATATCCAACTGCTGGTTTAGAGGATAACTCT
AGGAGAGCTTTTGATGTTGAGAATCTGGATTCTGATTACCGTCCCTCTTACCAAAACCAAGT
TTTGAAAGACCTCTTTTCCCATTACATGGATGCTTGGAAGACATGGTTCAGCGAAGTCACCC
AGGAGAATCCATTACCTAATACATCACAACACCGT**TAG**

SEQ ID NO: 7, CDS completa DEL1

ATGTCAGATCTATCGCCAGAAAGATTCAAACTTGCCGTTACTTCTCCTTCTTCCATACCGGA
ATCTTCTTCGGCTTTACAATTACACCATTCCTATAGTCGCAAACAGAAATCTCTCGGACTTC
TTTGTACCAATTTCTTAGCTTTGTATAATCGAGAAGGGATTGAAATGGTTGGGCTTGATGAT
GCTGCCTCGAAATTAGGAGTGGAGAGACGAAGAATCTATGATATTGTTAATGTTCTGGAGAG
TGTTGGGGTTTTAACAAGAAGAGCAAAGAATCAGTATACGTGGAAAGGGTTTTCCGCAATTC
CAGGAGCATTGAAGGAGCTACAAGAAGAGGGGGTTAAGGACACTTTTCATCGTTTCTATGTC
AATGAGAATGTTAAAGGATCTGATGATGAGGATGATGATGAAGAGTCTTCTCAGCCTCACTC
TAGTAGCCAGACTGATAGTTCAAAACCTGGTTCTCTTCCCCAATCTTCAGATCCCTCCAAAA
TAGATAACCGACGAGAGAAATCTTTAGGATTGCTTACTCAGAACTTTATCAAACTCTTTATT
TGCTCTGAAGCTATTAGGATCATCTCCCTTGATGACGCTGCAAAATTACTGCTTGGTGATGC
CCACAATACATCAATAATGCGAACTAAAGTGAGGCGGCTTTATGATATAGCAAATGTCTTGT
CGTCAATGAATCTCATAGAGAAGACTCACACCTTAGATTCTAGGAAACCAGCTTTCAAGTGG
TTAGGGTACAATGGTGAGCCTACTTTCACACTGAGCAGCGATTTGTTGCAATTGGAGTCAAG
AAAAAGAGCTTTCGGAACTGATATTACAAACGTCAATGTTAAGAGAAGCAAATCATCATCTT
CGTCCCAAGAAAACGCTACAGAGAGAAGGCTAAAGATGAAAAAGCACTCAACACCAGAGAGT
TCTTATAACAAAAGCTTTGATGTTCATGAATCAAGACATGGATCAAGAGGAGGTTACCATTT
TGGACCTTTTGCACCAGGCACTGGTACATATCCAACTGCTGGTTTAGAGGATAACTCTAGGA
GAGCTTTTGATGTTGAGAATCTGGATTCTGATTACCGTCCCTCTTACCAAAACCAAGTTTTG
AAAGACCTCTTTTCCCATTACATGGATGCTTGGAAGACATGGTTCAGCGAAGTCACCCAGGA
GAATCCATTACCTAATACATCACAACACCGTTAG

SEQ ID NO: 8, DEL1, sequência de proteína de comprimento total

MSDLSPERFKLAVTSPSSIPESSSALQLHHSYSRKQKSLGLLCTNFLALYNREGIEMVGLDD
AASKLGVERRRIYDIVNVLESVGVLTRRAKNQYTWKGFSAIPGALKELQEEGVKDTFHRFYV
NENVKGSDDEDDDEESSQPHSSSQTDSSKPGSLPQSSDPSKIDNRREKSLGLLTQNFIKLFI
CSEAIRIISLDDAAKLLLGDAHNTSIMRTKVRRLYDIANVLSSMNLIEKTHTLDSRKPAFKW
LGYNGEPTFTLSSDLLQLESRKRAFGTDITNVNVKRSKSSSSSQENATERRLKMKKHSTPES
SYNKSFDVHESRHGSRGGYHFGPFAPGTGTYPTAGLEDNSRRAFDVENLDSDYRPSYQNQVL
KDLFSHYMDAWKTWFSEVTQENPLPNTSQHR

SEQ ID NO: 9, sequência de assinatura 1

(Y/T/P)SRK(Q/D)KSL(G/W)(L/T)LC(T/E/Q/S)(N/R/K)F(L/V)(A/T)(L/I/R)Y(N/G/P/D)

SEQ ID NO: 10, sequência de assinatura 2

(G/S)LD(D/E)(A/V)A(S/T/V/R/K)(K/R/S)LGVE

SEQ ID NO: 11, sequência de assinatura 3

R(R/K)(E/D)KSL(G/R)(I/L)(L/M)(T/S)Q(K/N)F(I/V)(K/Q/M)LF(I/L/T)(C/A/V/M/N)(S/E)

Figura 4 (cont.)

SEQ ID NO: 12, sequência de assinatura 4
(I/V) (S/T)L(D/E) (D/T/V/E)AA(K/R) (L/I) (L/I) (L/M/I) (G/E)

SEQ ID NO: 13, sequência de assinatura 5
(T/A)K(V/I)RRLYDIAN(V/I)L(S/C/T)S(M/L) (N/R/A) (L/F)I(E/R/K/D)K (T/V/I) (H/Q) (T/V/Q) (L/G/A/T) (D/E) (S/T/E)R(K/G) (P/R) (A/K) (F/P) (K/L/A/R) (W/F) (L/K)

SEQ ID NO: 14, sequência de codificação DEL1 de arroz
CGCGGACGCACCAAGATCCAATCCTCGCGAGCAGAATCGATGGCGACGGCGGCGGTGATGGC
TGCTGTACCCTCGTCTTCGCCGGCCGACGCGGCGGAGGCTGTGGTTATGACGGAGGCGGTGC
CTTCTCTCCCGCAACGCCAGCAGCCGGTGTTCGTCGAGGGCAGAGGCGGGAAGCTGCGTGAC
CACGCCTACAGCCGCAAGCAGAAGTCGCTCGGCCTCCTCTGCTCCAATTTCGTCGCTCTGTA
CAACCGCGACGACGTGGAGTCTATCGGGCTGGACGACGCGGCTAGGAGGCTCGGCGTGGAGA
GGCGCCGGATCTACGACATCGTCAACGTGCTCGAGAGCGTAGGGATCCTCGTGAGGAAGGCC
AAGAATCGTTATTCTTGGATAGGCTTCGGCGGCGTCCCAATGGCATTGCGAGAACTCAAGGA
GAGGGCATTGAGAGAGAAGTCTGGATTGGCTCCTCTGCCCGTGGAGGAGCCGTCTGCAGCCA
TTATGTCGGATGACGAAGATGAAGATAAGATGGGTGACGCTGATGGTGATACCGAGAGCGAG
AAGCTGAGTCAACCAGTTGACAACCCTTCTGACAACAAGCCAGGCGCACCTCGCTGCCGGCT
TAGATCTGACCATAGGAAGGAGAAGTCGCTTGGGTTGCTCACGCAGAATTTCGTGAAACTCT
TCCTGACCATGGAGGTTGACACAATCTCACTTGATGAAGCTGCAAAGCTGCTACTTGGAGAA
GGTCACGCAGAGAACAGTATGAGAACTAAAGTCCGGAGACTGTATGACATTGCTAACGTGCT
GTCATCACTGAATTTCATTGATAAGATACAACAGGCAGACTCAAGGAAACCTGCATTCCGGT
GGTTGGGCTCAGCGGGGAAACCAAAAGCTGAAAATGGTGTCACAATCGCAGTACCTCCACCA
GGGAAGACCATATCGAACAAGAGAGCATTTGGGACTGAACTCACTAACATTGACATAAACAG
AAGCAGACTGGACTCAACAATCCCAAAGAAAGCAAAGCTGACACTGAGTGGTGGTGAAATTT
TGAAGAACTGCAAATTGTCAGTGCAGAAACAGCTCGGGCAGGGTAGCAAGGGTGGTTTTGTT
TATGGGCCTTTCCACCCTGCTGGTGCAAGAAAACAAGAGCTTGACAATGGTAATAAAGGACA
CACAGATAATGTTCAAAACTGGGAGAGCCTTGCTGCTTCATTTCGACCACAATACCAGAACC
AAGCATTGGGCGATCTTTTTGCTCATTATGTGGAAGCCTGGAAATCGTGGTACTCTGAATTT
GCGCAAGGCAGCAGCATGATGCAGCAGCACTTTGGCATGCCTGTCATTAACCAGTTTTTGTA
GTCAAATCATTAATCTCAAGCACTGTATTTCTTACCGCGTGCTTGGGATTTTACCTGTACAA
TTTGTCAAGAAAATGAGGGAACAAAGACATGACCTAGGCCTAGGATCCATTGTATTTACTCA
ATGCAGGAAATATCCAGGCTATTCAGTCCAAAAGATAAACTCGCCTCATTGTTAGGCATGTA
ACTGTTAATTCATCTAGCTATAAGCTTTCCT

SEQ ID NO: 15, DEL1 de arroz (XP_467698)

MATAAVMAAVPSSSPADAAEAVVMTEAVPSLPQRQQPVFVEGRGGKLRDHAYSRKQKSLGLL
CSNFVALYNRDDVESIGLDDAARRLGVERRRIYDIVNVLESVGILVRKAKNRYSWIGFGGVP
MALRELKERALREKSGLAPLPVEEPSAAIMSDDEDEDKMGDADGDTESEKLSQPVDNPSDNK
PGAPRCRLRSDHRKEKSLGLLTQNFVKLFLTMEVDTISLDEAAKLLLGEGHAENSMRTKVRR
LYDIANVLSSLNFIDKIQQADSRKPAFRWLGSAGKPKAENGVTIAVPPPGKTISNKRAFGTE
LTNIDINRSRLDSTIPKKAKLTLSGGEILKNCKLSVQKQLGQGSKGGFVYGPFHPAGARKQE
LDNGNKGHTDNVQNWESLAASFRPQYQNQALGDLFAHYVEAWKSWYSEFAQGSSMMQQHFGM
PVINQFL

Figura 4 (cont.)

SEQ ID NO: 16, sequência de codificação de DEL1 de milho

GCACGAGCCCTCCCAGCTCGCGCGCCCTCGATGGATGCCTCCGCCGCCACCCCCGCCCCCGG
GCCTTCCTTCTCTGGCGCCGAGTCCTCCGCGGCCGCCGCCCAGCCGCCGGCGGAGGCTCCGC
AGTTGCGCGTCCACGGCGCCGGCAGCGGCAGCGGCGTCGCGCGAGCCTGCCGCCACCACGCG
TACAGCCGCAAGCAGAAGTCGCTCGGCCTTCTCTGCTCCAACTTCGTGGCGCTGTACGACCG
GGAGGACGTGGAGGTGATTGGGCTGGACGACGCGGCCAAGCGTCTCGGCGTCGAGCGACGCC
GGATCTACGACATAGTCAACGTTCTCGAGAGCGTCGGGATTCTTGTGCGGAGGGCCAAGAAT
CGGTATACATGGCTCGGATTCGGGGGAGTCCCTGCTGCGCTGAAAGAACTCAAGGAGAGGGC
GCTAAGGAAGATGTCCGGATCACCGGTGTTACTGTCAATGGAGGACTCGTCTACTGCCAACT
TATCAGATGATGAGGATGATGAAAAATTGGGCGATGCTGATGAAGATGCTGAGAGCGAGAAG
CTCAGCCAACCTGTTGACAATACGTCTGATAAGCCTGACGCACCCAGCTGCCGCCTTAGATC
TGATCATCGGAAGGAGAAGTCCCTTGGGCTCCTCACTCAGAATTTTGTCAAGCTCTTCCTCA
ACATGGAGGTTGGGACAATCTCACTTGACGAAGCTGCAAGGCTTCTCCTTGGAGAGGGACAT
GCAGACAGCAACATGAGAACAGCCAAAGTTCGTCGATTGTATGACATTGCCAATGTGCTGTC
TTCTTTGAACCTCATTGAGAAGACGCAGCAAGCAGACACAAGAAAACCTGCATTCCGGTGGC
TAGGCCAGGCAAAGCGAAAGCAAGATAACAATGTCATGGTTTCTGTACCTCCATCCATGAAG
GCAATGCCCAATAAGAGATCATTTGGTACTGATCTTACAAACATTGACAATAAGCGAGGCAA
GTTAGACTCAGCAGCGGAGAACAAAGTCAAGCTCATGCAGGGTGCTGGTAACATAGTGAAGA
CTTTTGAGAGGCAGCTGGTGCAAGGGAAAAGGAATGACTTTGTTTATGGGCCCTTCCACCCT
GCTGGTGCAAAGAAACACGAAACTGATGATCAAACTGTTAAGCAGCAGGAGAGGAAGAACAT
TCAGGACTGGGAAAACCTTGCTGTGTCCTTCCGTCCACAATATCAGAATCAAGCACTGAATG
ATCTTTTTGGTCATTATGTGGAAGCATGGAAATCATGGTACGTGGATCTTACCCAGGAAACG
GCATCATGAAGCAGAATGTTGGCAGGTCAGTTGTAACTTGTAAGCCGTTTCCAGTAGTTAGA
TCAACTGAGTTGTTGATGCTTGCTTAGGGAACCACCCTGTACAAGATGGAATAAAAACACGA
AGGCAATTTTGTCTTTCGACTGTATTTACTCGATG

SEQ ID NO: 17, DEL1 de milho, transladado de AY107996

MDASAATPAPGPSFSGAESSAAAAQPPAEAPQLRVHGAGSGSGVARACRHHAYSRKQKSLGL
LCSNFVALYDREDVEVIGLDDAAKRLGVERRRIYDIVNVLESVGILVRRAKNRYTWLGFGGV
PAALKELKERALRKMSGSPVLLSMEDSSTANLSDDEDDEKLGDADEDAESEKLSQPVDNTSD
KPDAPSCRLRSDHRKEKSLGLLTQNFVKLFLNMEVGTISLDEAARLLLGEGHADSNMRTAKV
RRLYDIANVLSSLNLIEKTQQADTRKPAFRWLGQAKRKQDNNVMVSVPPSMKAMPNKRSFGT
DLTNIDNKRGKLDSAAENKVKLMQGAGNIVKTFERQLVQGKRNDFVYGPFHPAGAKKHETDD
QTVKQQERKNIQDWENLAVSFRPQYQNQALNDLFGHYVEAWKSWYVDLTQETAS

SEQ ID 18, sequência decodificado DEL1 de Ostreococcus tauri (AY675104)

ATGAACGGGGATGGGGGGGTGGAGAGCTCGCGCGCGTCGGGCGGGACGGATGTCCCGGGCGA
AGACGCGGCGCGGACGCGCGCGACGCCGAGCGCGACGTCTCGGAAAGATAAGTCGTTGTGGA
CGCTGTGTGAACGGTTTTTGACGATTTATGGAGATGGATCGAAGGAGAGCGTGTCGCTGGAC
GATGCGGCGACGCGGCTCGGGGTGGAGCGTCGAAGGATTTACGACGTCGCGAACGTGTTGGA
GAGCGTGGAGGTGCTCGAGCGCAAGGCAAAGAATCAGTACACGTGGCACGGCGTGCGAAGAC
TCCCGGAGTGCCTGAAACGGCTCAAGGAGAGTGGATTGCGGGAATTTGGGACGGATGTCGAG
CTCGATGGAAGCACGAGCGAGGGACGTGACGGGGAAAAGGAGGATGGAACGGCTCGAGGCGG
CGACGCCTCCGATCGAAGCTCTCCCACAAACTCGTCGACGATCAACTTGGACGCGAAACAGC
GGGGTGAGAAGGTTACTGGGACTAAGTTCTTTGGACAGGGGCGATTCGCCGTCTCGAGCGCC
AGTTACGACAGCCGGCGGGAGAAGAGTCTTGGCTTGCTGTCCCAAAAATTCGTTCAACTCTT

Figura 4 (cont.)

CCTCGCGTCAAAGATGAATGTTGTCAGTCTGGAAACGGCCGCCAGGATTATTATGGGAGAAG
ACGACGACGATGAAGCCAAGTTGAAGACTAAAATTCGTCGATTGTACGATATCGCCAACATC
TTGTGCTCCTTGCGCCTGATTCGAAAGGTACACGTGGGCGAGACTAGAAAACCAGCCTTCCT
TTGGCTACAGCGAGAAAACTCCATCGCGGAGCTCATCGCGCAGGGCAAGGGGCTGATGTGGT
TCGATAAATTGAACGAAGAGGAAGAGATGCGTTTACAGGCGTCCATCATTGAATCAAAACTC
GACGGGAACGAGGTGTTGATGGTTGACGCGGAGAATAAACGACGAGGCGCGTTTTACGATTC
GCAGACTAAGTCGGGATCTAAGCGTCCTCGAGGACGTCCAAGACTGCCGGGTGGTGACGTCG
ACGCGTTGCCTTCATCCGTACCTCCATTAGGCGCGTCTTCGATGACGCCCGAAGAAGCGGCG
AGGTTCAATCAGTTGTTCGCGTTCACCACGTCACAACTCATGGCACAGTACCCGCTGGACGT
GCGCGCGGATGTGAACAGCATCATGGCGCAGAGCGCGATGGGTAACGCAAACTACCTTCATC
TTCTCGCGCAGTCTTCCGTGGCCCAAGCGCACGCGGCGCGACGGATGGAGAACACCGCAAAA
GGAATGAGCTCTGATTCCAGCATCGACAACACGGCTATGTTCGTGCCGCCACTTCCTGCGTT
TCCCGCGATGCTTCCGGCGTGGGGCATGTCTGGCTTGTCTTACCACAGCAACCACATGGAAC
ACATGATGCGGATGTACGAAACTTCTCTGGCAACGATTAGCGACAACGACCCCTCGCATCCA
GATGGAAAGATTGCCGAACAATAA

SEQ ID NO: 19, *Ostreococcus tauri* DEL1 (AAV68606)

MNGDGGVESSRASGGTDVPGEDAARTRATPSATSRKDKSLWTLCERFLTIYGDGSKESVSLD
DAATRLGVERRRIYDVANVLESVEVLERKAKNQYTWHGVRRLPECLKRLKESGLREFGTDVE
LDGSTSEGRDGEKEDGTARGGDASDRSSPTNSSTINLDAKQRGEKVTGTKFFGQGRFAVSSA
SYDSRREKSLGLLSQKFVQLFLASKMNVVSLETAARIIMGEDDDDEAKLKTKIRRLYDIANI
LCSLRLIRKVHVGETRKPAFLWLQRENSIAELIAQGKGLMWFDKLNEEEEMRLQASIIESKL
DGNEVLMVDAENKRRGAFYDSQTKSGSKRPRGRPRLPGGDVDALPSSVPPLGASSMTPEEAA
RFNQLFAFTTSQLMAQYPLDVRADVNSIMAQSAMGNANYLHLLAQSSVAQAHAARRMENTAK
GMSSDSSIDNTAMFVPPLPAFPAMLPAWGMSGLSYHSNHMEHMMRMYETSLATISDNDPSHP
DGKIAEQ

SEQ ID NO: 20, sequência de codificação tipo DEL1 de Medicago truncatula

atggctgcttcctcctccgatcctccttcacgacaccacacttacgaccgtaaacaaaaatc
ccttggcctcttatgcaccaagttcttgagcttgtataacaaagatgatgttcgtttaattg
gtctcgacgatgctgccgccaaattaggtgttgagagaagacggatctatgatattgtcaat
gttctcgaaagcatcggggttcttgcaagaaaagccaagaatcagtatacctggaaaggttt
tgcggcaattcctgttgctctacaggagcttaaggaagagggtttatggcagaatctcaatt
cttcacaagaaggtgccaatgaagatgtgaaggtatcggatgaagaggatgaggatgaattg
ttatcccaaaccactggaagtcagggtgaatcattatcccaacccactggaagtcagaatga
caatctaaaccctaattccgcttttcccagatctttgaaaaatgacagaagggaaaaatctc
tggcgctgcttactcagaattttgtcaagctctttgtctgttccaacctggaaatgatatcg
cttgatgatgcagcaaggttgttgcttggagatgcatataattcatcaacaatgagaacaaa
agtcaggcgcctttatgatattgcaaacgtgctaacctccatgaaccttattgagaagaccc
ataccacagatacaagaaaaccagcattcaggtggctaggcttaaaagggaagacattgaat
gaggcatcactttacaattcaaaacaaaatgagtctaggaaaagggcgtttggaaatgatgt
cacaaacataagctttgccaggaatagaatggacttgttcatgggcggggactttaagaagc
aaaagacaatggaaaatgatagtggactatgtcaggaagatgtgaaacaaggcataaaacag
acttcagcagctaactatcaatttggtccttttgctcctgcctttgtatccaaagctggaag
ctctgagaataaagtgaagcaggtgcatgactgggagagtctcgctactgaacattgccctc
agtatcaaaaccaagctttgaaagaactttactctcattacatggaagcatggaaatcttgg
tactctgaagttgctgggaagaggtcaacgcaagttttgtagtaa

Figura 4 (cont.)

SEQ ID NO: 21, sequência de proteína tipo DEL1 de Medicago truncatula

MAASSSDPPSRHHTYDRKQKSLGLLCTKFLSLYNKDDVRLIGLDDAAAKLGVERRRIYDIVN
VLESIGVLARKAKNQYTWKGFAAIPVALQELKEEGLWQNLNSSQEGANEDVKVSDEEDEDEL
LSQTTGSQGESLSQPTGSQNDNLNPNSAFPRSLKNDRREKSLALLTQNFVKLFVCSNLEMIS
LDDAARLLLGDAYNSSTMRTKVRRLYDIANVLTSMNLIEKTHTTDTRKPAFRWLGLKGKTLN
EASLYNSKQNESRKRAFGNDVTNISFARNRMDLFMGGDFKKQKTMENDSGLCQEDVKQGIKQ
TSAANYQFGPFAPAFVSKAGSSENKVKQVHDWESLATEHCPQYQNQALKELYSHYMEAWKSW
YSEVAGKRSTQVL

SEQ ID NO: 22, regulador de transcrição de triticum aestivum DQ353854 do ciclo de célula TaE2Fe

CCCTCCTCCAAGGACCACAAAATTCCCCCCTCCCTACCCTTTCCACCTCGACATCCACCGGA
AATGGACCCCGCCGCCGCCGCCGGGGCTCCGGCGGCCGCTGCGCAGCCACCGCCCCCTCCTC
CTCCTCCTTACCTGCCCCCGCGGCTGGTCATCGACGGCGCCGGAGGCGGCTCGGGCGCCGCC
GTGAGGGCCTGCCGGCACCACGCCTACAGCCGCAAGCAGAAGTCCCTCGGCCTCCTCTGCTC
CAACTTCGTGGCGCTGTACGACCGGGACGACGTGGAGACGGTGGGGCTGGACGACGCCGCCA
GGCGGCTCGGCGTCGAGAGGCGCCGGATCTACGACATCGTCAACGTGCTCGAGAGCGTCGGG
ATTCTCGTGAGGAGGGCCAAGAATCGGTACACATGGATCGGATTTGAGGGCGTCCCTGCCGC
GCTCAAGGAGCTTAAGGAGAGGACACTGAGAGAGATGTCTGGATTAGCTCCGCCACCGGAGG
AATCATCTGCTGCCAATGTGTCGGACGATGAAGACGACGATGATAAATTGGGCGATGCAGAT
GGGGACGCCGACAGCGAGAAGCTTAGCCAGTCCCTCGACAATGCTTCTGATAAGCCTAACGT
GCCCATGTGCCCACCTAGATCTGTAGACCATAGGAAGGAGAAGTCGCTTGGGCTGCTCACGC
AGAATTTTGTCAAGCTCTTCCTCACCATGGAGGTTGAGACAGTCTCACTTGACGAGGCTGCA
AGGCTGCTCCTTGGAGAGAGACATGCCGAGAGCAATATGAGAACTAAGGTTCGTCGACTGTA
TGACATCGCCAATGTGCTATCTTCTTTGAACCTCATTGAGAAGACACAGCAGGTGGACTCCA
GAAAACCTGCATTCCGGTGGCTCGGTCAGGCAAAGCGAAAGGAAGGTGCCACTGTCACGGTT
GCTTTACCACCATCCAGGAAGATTATGTCTAGCAAGAGGGCATTTGGTACCGACATCACAAA
CATTGACAATAAGAGGGGCAAGTTAGTCTTGGAAACAGAGAACAAACCCAAACTCATGCAGG
GTGGCAGCAGCATGTTGAAAACTTTCGAGAGTCAGCTCGGGCAAGGGAAGAGTAGTGGCTTT
GTTTATGGGCCCTTCCACCCTGCTGGTGCAAGGAAACATGAAGTTGATGATCAGACAGTGAG
GGAGAATGAGATGAAGAACATTCAAGACTGGGAGAGTCTCGCTGTTTCATTCCGTCCACAGT
ACCAAAATCACGCGCTGAATGATCTTTTTGGCCATTATGTTGAAGCATGGAAATCATGGTAC
TTGGATCTTACACGGGATTCGACCTCATGAAAGGATGGTTTATTTTAGGCAGGTCTGTTGTA
ATCAGTTCTTGTAGGCAGATCAACTCATCTGCCGAGGACCCTAGAAAGGATTCGACCTGTAC
AAGATGTCATAAAAGCACAAAGGCTCATTTCGCCTTTGCTCGATGTATTTACTCAATGCAAA
AAAGAAACGGAATAATGCCAAA

SEQ ID NO: 23, regulador de transcrição de trigo Q27W79 do ciclo de célula TaE2Fe

MDPAAAAGAPAAAAQPPPPPPPPYLPPRLVIDGAGGGSGAAVRACRHHAYSRKQKSLGLLCS
NFVALYDRDDVETVGLDDAARRLGVERRRIYDIVNVLESVGILVRRAKNRYTWIGFEGVPAA
LKELKERTLREMSGLAPPPEESSAANVSDDEDDDDKLGDADGDADSEKLSQSLDNASDKPNV
PMCPPRSVDHRKEKSLGLLTQNFVKLFLTMEVETVSLDEAARLLLGERHAESNMRTKVRRLY
DIANVLSSLNLIEKTQQVDSRKPAFRWLGQAKRKEGATVTVALPPSRKIMSSKRAFGTDITN
IDNKRGKLVLETENKPKLMQGGSSMLKTFESQLGQGKSSGFVYGPFHPAGARKHEVDDQTVR
ENEMKNIQDWESLAVSFRPQYQNHALNDLFGHYVEAWKSWYLDLTRDSTS

Figura 4 (cont.)

SEQ ID NO: 24, clone de cDNA 002-186-D11 (grupo de cultivar japônica) de Oryza sativa AK120032

AAATCTCTCCTCCCGCTCGCCGAAACCCTCGCCGTCGCCGCGCCATGGCCGCCGCCGCCGAT
GCTCCCCCGCCTCCTCCGGAGGTCGCCCCGCCCGCGCCCGCGCCCGCCCCCGCGCCGGCGCC
ATATCAGCCGCCGCGGCTGGCGGTGGCCGACGGAGCGGGAGGTGGCGGCGGCGGCGGCGGGA
AGCCGTGCAGGCACCACGCGTACAGCCGCAAGCAGAAGTCGCTCGGCCTCCTCTGCACCAAC
TTCGTGGCGCTGTACGACCGGGAGGACGTGGAGTCGGTGGGGCTGGACGACGCGGCGAGGCG
GCTGGGCGTCGAGAGGCGCCGGATCTACGACATCGTCAACGTGCTCGAGAGCATCGGGATGC
TCGTGAGGAGGGCCAAGAATCGGTATACGTGGATCGGCTTCGGTGGAGTCCCTGCGGCGCTC
GCGAAACTCAAGGAGATGTCACTGAGGGCGGTGTCAAGCGTGGCGTCACCGTCGCTGGATGA
AACATCTGCTGCTAATGTCTCGGATGATGAGGATGATGACAAGTTAGATGATGCTGAGGGCG
ATGCGGAGAGCGAGAAGCTCAGCCTCAGCCAGTCCATTGACAATCCTTCTGATAAGCCTGAT
GCACCCCCTTGCAAGCTTCGATCCGAGCATAGGAAGGAGAAGTCCCTTGGGCTGCTCACTCA
GAATTTTGTCAAGCTCTTCCTCACCATGGAGATTGAGACGATCTCACTTGATGAAGCCGCAA
AGCGGCTCCTTGGAGAGGGACATGCGGCGAACAATATGAGAACCAAAGTTCGGCGATTGTAC
GATATTGCCAATGTGCTGTCTTCTTTGAATCTTATTGAGAAGACACAACAGGCGGACTCAAG
AAAACCTGCATTCCGGTGGCTGGGCCAGGCAAAACGGAATGAAGGCGTCACGGTTGCTTTAC
CCCCAACCAAGACGTTGCCTAACAAGAGAGCATTTGGTACTGATCTGACTAACATTGACAAT
AAGAGGGGTAAGTTGGACTCCACAATGGAGAACAGAGGCAAGCCCACGCAGGATGGTGGCAA
CCTATTCAATAATTTGCAGAGGCAATTAGGGCAAGAGAACAGGAGCGATTTTGCTTATGGCC
CCTTCCACCCTGCTGTTGCAAGGAAACAAGAACATGGTAATCGCACAGTACAAGAGAAGGAG
AGGAAGAGCATTCAAGACTGGGAGAACCTTGCTTCTTCTTTCCGTCCGCAATATCAAAATCC
AGGACTGAACGATCTTTTTGGCCACTACATGGAAGCAAGGAGGTCATGGTACTCGGATCTCA
GGCGAGACAGAGCATCATAAAACCAGATTCAGGCCGGTCCGTTGTAAACAAAATTCTTGTAG
GCAGATAAACTCATTTGTTGATGCATGGTTAGGGATTTTTGAGTTATACAAGATGTCACAAA
AACAAAGGCTAGTGCTGGCCTTTGATTGATTGTATTTACTCAGTGCAAGAAATTGGAAAGTC
GTTTCGTCAATAATACTAGGATACCTGATTGTCGATTTGTTT

SEQ ID NO: 25, clone de cDNA 002-186-D11 [45 - 1319] (grupo cultivar japônica) de Oriza sativa

MAAAADAPPPPPEVAPPAPAPAPAPAPYQPPRLAVADGAGGGGGGGGKPCRHHAYSRKQK
SLGLLCTNFVALYDREDVESVGLDDAARRLGVERRRIYDIVNVLESIGMLVRRAKNRYTW
IGFGGVPAALAKLKEMSLRAVSSVASPSLDETSAANVSDDEDDDKLDDAEGDAESEKLSL
SQSIDNPSDKPDAPPCKLRSEHRKEKSLGLLTQNFVKLFLTMEIETISLDEAAKRLLGEG
HAANNMRTKVRRLYDIANVLSSLNLIEKTQQADSRKPAFRWLGQAKRNEGVTVALPPTKT
LPNKRAFGTDLTNIDNKRGKLDSTMENRGKPTQDGGNLFNNLQRQLGQENRSDFAYGPFH
PAVARKQEHGNRTVQEKERKSIQDWENLASSFRPQYQNPGLNDLFGHYMEARRSWYSDLR
RDRAS

SEQ ID NO: 26, Silene de biblioteca de flor macho de silene latifolia DV768235 S1_S1_09F04_SAC (parcial)

GCACGAGTCACCATCAGAAAATCGATCTTTGCATCATGGTTATAGTAGAAAGCAGAAATCTT
TGGGTCTTCTTTGTTCAAATTTCTTGAGATTGTATAATCGAGACGATGTCGATTTGATTGGT
CTAGATGATGCTGCCAGTAAATTAGGAGTCGAGCGAAGAAGGATCTATGATATTGTTAATGT
GTTGGAGAGTGTAGGGGTGTTAGCAAGAAAAGCGAAAAATCAGTACACATGGAAGGGCTACA
AGGCAATTCCTAAGGCTCTTGCCTTGTTAAAGGAAGATGGTTTGAAGGAGAATTTTGGTACT

Figura 4 (cont.)

GCAGAGGGAAGAAGTAGAGTAAAGTTTCAGGTTTTAGACTTTGATGATGATGTAGCTTCTAA
CCCTGATACCGGAAGCCAGTCGGAGGCACCAACACATAGTGGTGCTTCTAAGCCTGATAACA
GAAGGGAAAAGTCATTGGGGCTTCTAACACAAAACTTTGTGAAGCTATTCCTTTGTTCTGAG
GCTGAGCTGATCTCTCTGGAGGAAGCTGCAAAGTGCTTGCTTGGAGATGGTCTTAATGCACA
AGTGATGCGAACTAAAGTGAGACGATTGTATGATATTGCCAATGTATTATCCTCCATGCAAC
TCATTGAAAAGACTACTCAAACAGA

SEQ ID NO: 27, Silene de biblioteca de flor macho de silene latifolia DV768235 S1_S1_09F04_SAC sequência parcial [2-643]

HESPSENRSLHHGYSRKQKSLGLLCSNFLRLYNRDDVDLIGLDDAASKLGVERRRIYDIVNV
LESVGVLARKAKNQYTWKGYKAIPKALALLKEDGLKENFGTAEGRSRVKFQVLDFDDDVASN
PDTGSQSEAPTHSGASKPDNRREKSLGLLTQNFVKLFLCSEAELISLEEAAKCLLGDGLNAQ
VMRTKVRRLYDIANVLSSMQLIEKTTQT

SEQ ID NO: 28, Heliantuhus de girassol de folha prateada EE619641_18 CHWM6882.b1_D18.ab1 CHW (LMS), sequência parcial

CGGGGGGACGGATCGAACTGTTGGATTGGATGACGCCGCAACGAGGCTAGGTGTTGAGAGAC
GGCGGATTTATGACATTGTTAATGTTTTGGAAAGCGTTGGCGTCCTTGTGAAAAAGGCAAAA
AACACGTATCATTGGTTAGGATTGGGGGCAATACCTAAGGCTTTAGAGCAGCTAAAGGAAGA
AGGTTTTAGGAATAATGATGAACATGTTGATGTTAAATCTGGAAAGGTTTCTGATGATGAAG
AAGATGAAAGAGTTTCTAGCCATAGTGTTTCTTTACAGGAAAAATCAGATCTAGATTCAATG
CATAAAACTTCAGGGCCGTTTAAATCGGGTAGTGTAACGGAAAATCGGAAGGAAAAATCTTT
GGGGCTTCTTACCAAGAATTTCATCAAGCTTTTCCTATGCACTAATTCGGATATGCTTTCGC
TTGATGACGCTGCAAAAATATTGCTTGGAGATGCTCAGAATCCGTCATTGACCCGAACTAAA
GTCAGACGCCTATATGATATTGCTAATGTCTTGTCTTCCATGCACTTCATCGAGAAGATCCA
TCACCCAGAAACCCGAAAGCCCGCCTTTAGGTGGTTGGGAATGGCAAGTCACCCAAATAGCC
AAACGAGATCAGCCACTGGTGTAGCTCATATAGAGTCCAAAAAAAGGGCATTTGGGACTGAG
CTTACAAACATTTGTTTCAAAAGAAGCAAGTTGGGTGACCAGGAGGTGAAGCTGTTTGACCT
TCAGAGTCAACGCCCGCCTATCAAACCCAGCACTCATAATC

SEQ ID NO: 29, Heliantuhus de girassol de folha prateada EE619641 _18 CHWM6882.b1_D18.ab1 CHW (LMS), sequência parcial [2 - 784]

GGTDRTVGLDDAATRLGVERRRIYDIVNVLESVGVLVKKAKNTYHWLGLGAIPKALEQLKEE
GFRNNDEHVDVKSGKVSDDEEDERVSSHSVSLQEKSDLDSMHKTSGPFKSGSVTENRKEKSL
GLLTKNFIKLFLCTNSDMLSLDDAAKILLGDAQNPSLTRTKVRRLYDIANVLSSMHFIEKIH
HPETRKPAFRWLGMASHPNSQTRSATGVAHIESKKRAFGTELTNICFKRSKLGDQEVKLFDL
QSQRPPIKPSTHN

Figura 4 (cont.)

SEQ ID NO: 30, sequência de DNA de Glycine max

ATGACTTCTTTGTCTACCCATCACACCTATAGCCGAAAGCAAAAATCTCTCGGCCTCCTCTG
CACCAATTTTCTGAGTTTGTACAACAAAGAAGGTGTGCGCCTGGTCGGTCTCGATGACGCAG
CTTCGCGGTTAGGTGTAGAAAGACGTCGGATCTACGACATCGTTAATGTTTTAGAGAGTGTC
GGTGTGCTAACCAGAAAAGCTAAGAATCAGTATACTTGGAAAGGATTTTGTGCAATTCCTGC
TGCCTTACAAGAACTTAAGGATGAAGGTTTGAAGGAGAATCCTGGATCATTTGATGACTCCA
ATGACAATGCAAAAGTGTCTGATGATGAGGAGGAGGAAGAAGAAACATTTCCTAATATTAAC
ATTGGGAGTCAGAGTGAGAAGGAAAATCCTGATTACACTGCTACTGTCAAATCCTCCAAGAA
TGAAAACAGAAGGGAAAAATCTCTGGCATTGCTTACTCAGAATTTTGTCAAGCTTTTTGTCT
GCTCTAATTTTGAAATGATCTCCCTTGACGAAGCTGCAAAGTTGTTGCTTGGAAATGCCAAC
AATAGAACAAAAGTCAGACGTCTATATGATATTGCGAATGTGTTATCCTCCATGAACCTAAT
CGAGAAGACCCATACAACAAACACAAGAAAACCAGCATTCAGGTGGTTGGGTGTGAGGGGAA
AAACATGGGGTGGATCAGTTGACTTGGCTCAAAATTCAAATGTCAAAGAGTCACGGAAAAGG
ATGTTTGGAACTGACATTGGAAACATAAGTTTTAAAAGGAACAAAGTTGATTTGTCCATGGA
TGGCCAGAATTCTAAAATGCAAAATCAACACGAAAATATTAGTCCACGAGCTCAGTTAGAGA
AAATCATCATAAAAAAAGATGCTAAGCAGACTTCAGAGAGCTATCAGTTTGGCCCTTTTGCT
CCAGCTTATGTGCCCAAAGTTGGAACCTCTGAAAATAATAGTGTGAAAAAGGTGCATGACTG
GGACAGCCTCGCACAGGAGCATCGTCCTCAATATCAAAACCAACCTTTAAAAGATCTATTTT
CCCACTACATGGAAGCATGGAAATCTTGGTACTCTGAAGCTGCTGGGAAAAAGTCAACGAAA
ACTTTATAA

SEQ ID NO: 31, sequência de proteína de Glycine max

MTSLSTHHTYSRKQKSLGLLCTNFLSLYNKEGVRLVGLDDAASRLGVERRRIYDIVNVLESV
GVLTRKAKNQYTWKGFCAIPAALQELKDEGLKENPGSFDDSNDNAKVSDDEEEEEETFPNIN
IGSQSEKENPDYTATVKSSKNENRREKSLALLTQNFVKLFVCSNFEMISLDEAAKLLLGNAN
NRTKVRRLYDIANVLSSMNLIEKTHTTNTRKPAFRWLGVRGKTWGGSVDLAQNSNVKESRKR
MFGTDIGNISFKRNKVDLSMDGQNSKMQNQHENISPRAQLEKIIIKKDAKQTSESYQFGPFA
PAYVPKVGTSENNSVKKVHDWDSLAQEHRPQYQNQPLKDLFSHYMEAWKSWYSEAAGKKSTK
TL

Figura 4 (cont.)

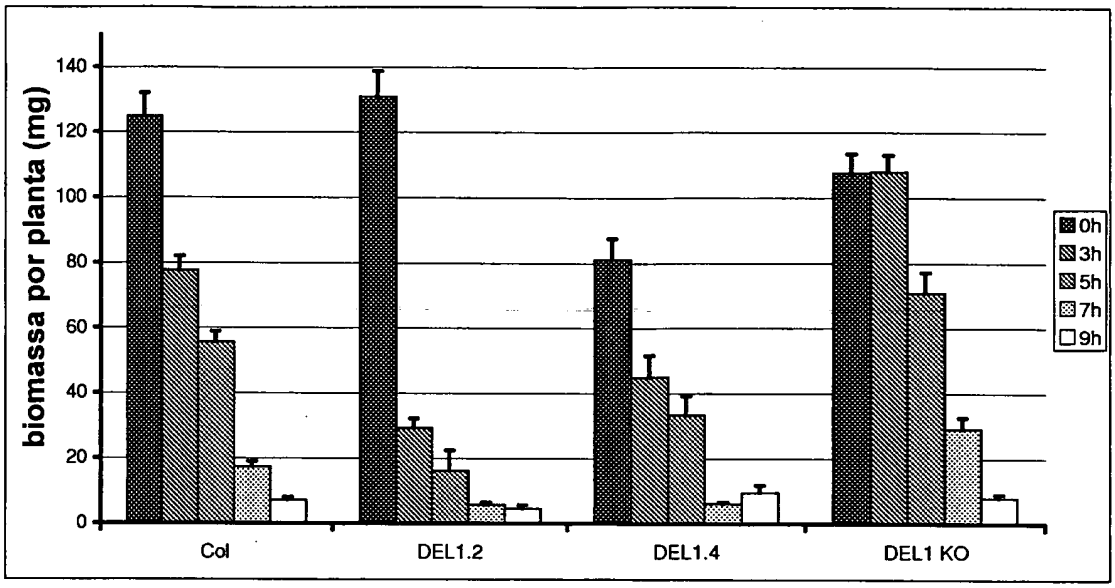

Figura 5

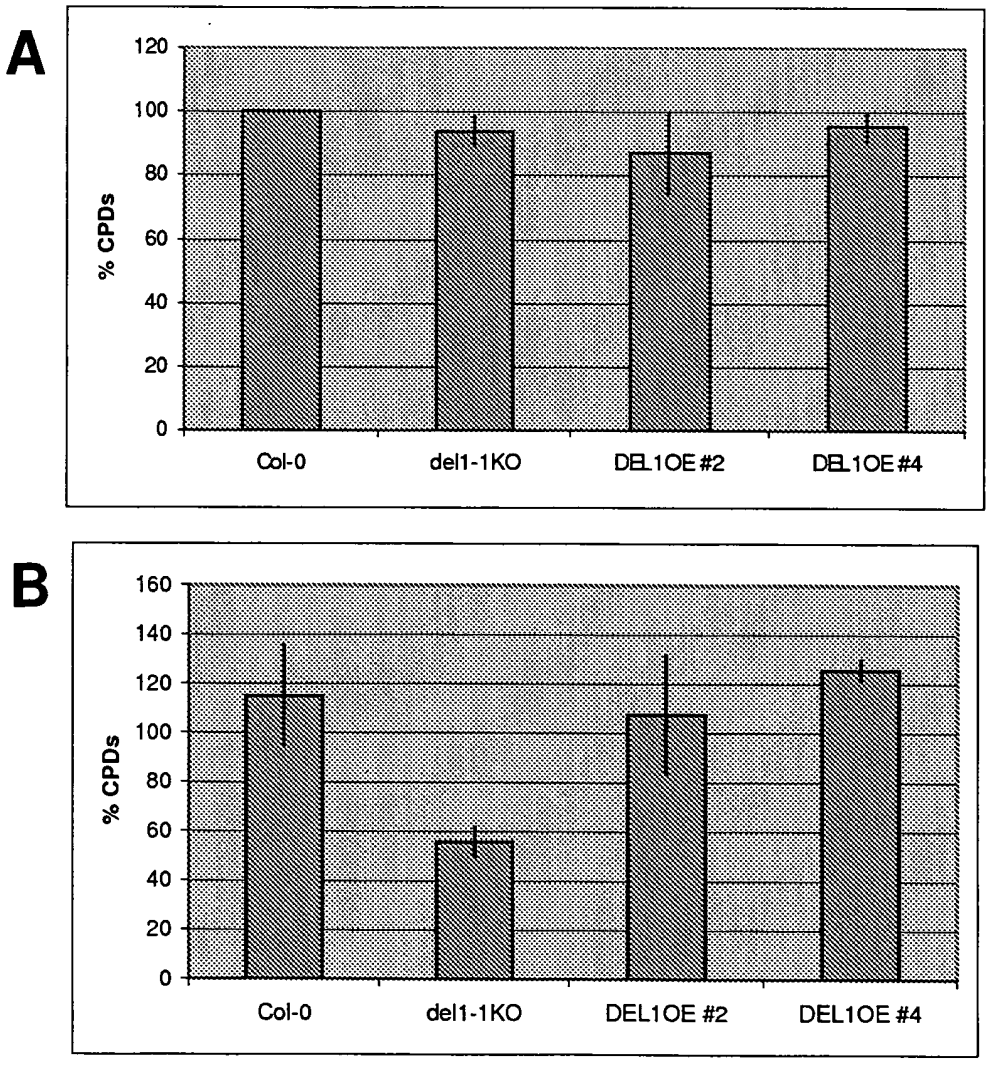

Figura 6

PI 0618378-6

RESUMO

"MÉTODOS PARA AUMENTAR O RENDIMENTO DE PLANTA E/OU RESISTÊNCIA AO ESTRESSE EM RELAÇÃO ÀS PLANTAS DE CONTROLE E PARA PRODUZIR UMA PLANTA TRANSGÊNICA TENDO RENDIMENTO AUMENTADO E/OU RESISTÊNCIA AO ESTRESSE AUMENTADA, CONSTRUÇÃO, USOS DE UMA CONSTRUÇÃO E DE UM ÁCIDO NUCLEICO DEL1 OU VARIANTE DO MESMO, OU DE UM POLIPEPTÍDEO DEL1 OU HOMÓLOGO DO MESMO, POLIPEPTÍDEO DEL1 ISOLADO, E, SEQÜÊNCIA DE ÁCIDO NUCLEICO ISOLADA"

A presente invenção diz respeito a um método para aumentar o rendimento de planta e/ou resistência ao estresse aumentada pela diminuição do nível de atividade de um polipeptídeo DEL1. Um tal método compreende introduzir em uma planta um ácido nucleico que compreende um ácido nucleico DEL1 variante. Um outro método compreende infra regular a expressão de um gene DEL1. A invenção também diz respeito a plantas transgênicas tendo introduzido nelas um ácido nucleico DEL1 variante destas, plantas estas que têm rendimento aumentado e/ou resistência ao estresse aumentada em relação às plantas de controle. A presente invenção também diz respeito a uma novo proteína DEL1 e sua seqüência codificadora e construções úteis nos métodos da invenção.