



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

(21) **PI 0922553-6 A2**

(22) Data de Depósito: 26/11/2009
(43) Data da Publicação: 07/05/2013
(RPI 2209)



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

(54) Título: MÉTODO PARA AUMENTAR A TOLERÂNCIA A ESTRESSE ABIÓTICO E/OU PARA AUMENTAR UMA CARACTERÍSTICA RELACIONADA AO RENDIMENTO EM UMA PLANTA, CONSTRUTO, PLANTA, CÉLULA DE PLANTA OU PARTE DA MESMA, INCLUINDO SEMENTES, PARTES CULTIVÁVEIS DA DITA PLANTA, PRODUTOS DERIVADOS DE DITA PLANTA E/OU DE DITAS PARTES CULTIVÁVEIS, MÉTODO PARA A PRODUÇÃO DE PLANTAS TRANSGÊNICAS, USO DE UM CONSTRUTO OU DE UM ÁCIDO NUCLEICO, MOLÉCULA DE ÁCIDO NUCLEICO ISOLADA, E, POLIPEPTÍDEO ISOLADO

(30) Prioridade Unionista: 03/12/2008 EP 08170613.7, 03/12/2008 EP 08170618.6, 10/12/2008 EP 08171212.7, 10/12/2008 EP 08171237.4, 17/12/2008 EP 081720336, 17/12/2008 EP 08172034.4, 17/12/2008 EP 08172038.5, 11/12/2008 US 61/121570, 11/12/2008 US 61/121571, 11/12/2008 US 61/121572, 11/12/2008 US 61/121577, 19/12/2008 US 61/138965, 19/12/2008 US 61/138970, 19/12/2008 US 61/138972, 11/12/2008 US 61/121570, 17/12/2008 EP 081720336, 17/12/2008 EP 08172034.4, 17/12/2008 EP 08172038.5, 11/12/2008 US 61/121570, 11/12/2008 US 61/121571, 11/12/2008 US 61/121577, 19/12/2008 US 61/138965, 19/12/2008 US 61/138970, 19/12/2008 US 61/138972, 17/12/2008 EP 08172038.5, 11/12/2008 US 61/121570, 11/12/2008 US 61/121571, 11/12/2008 US 61/121572, 11/12/2008 US 61/121577, 19/12/2008 US 61/138970, 19/12/2008 US 61/138972, 17/12/2008 EP 08172034.4, 17/12/2008 EP 08172038.5, 19/12/2008 US 61/138965, 19/12/2008 US 61/138972

(73) Titular(es): BASF PLANT SCIENCE GMBH

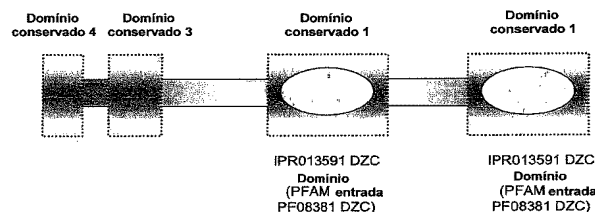
(72) Inventor(es): Ana Isabel Sanz Molinero, Christophe Reuzeau, Koen Bruynseels, Valerie Frankard, Yves Hatzfeld

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

(86) Pedido Internacional: PCT EP2009065906 de 26/11/2009

(87) Publicação Internacional: WO 2010/063637 de 10/06/2010

(57) Resumo: MÉTODO PARA DETECTARA PELO MENOS UM ANALITO EM PELO MENOS UMA AMOSTRA.A presente invenção refere-se a um método de determinação da concentração de um analito em uma amostra e/ou á cinética de ligação de um analito a uma molécula de sensor de analito.Para este propósito, a invenção é baseada na dectecção da interação entre o anatito e a molécula de sensor de analito, a última sendo fisicamente adsorvida em uma suporte poroso sólido, que está sob a forma de um microarranjo.Em uma modalidade preferida, o método pode ser operado em ciclos repetidos.



“MÉTODO PARA AUMENTAR A TOLERÂNCIA A ESTRESSE ABIÓTICO E/OU PARA AUMENTAR UMA CARACTERÍSTICA RELACIONADA AO RENDIMENTO EM UMA PLANTA, CONSTRUTO, PLANTA, CÉLULA DE PLANTA OU PARTE DA MESMA, INCLUINDO
5 SEMENTES, PARTES CULTIVÁVEIS DA DITA PLANTA, PRODUTOS DERIVADOS DE DITA PLANTA E/OU DE DITAS PARTES CULTIVÁVEIS, MÉTODO PARA A PRODUÇÃO DE PLANTAS TRANSGÊNICAS, USO DE UM CONSTRUTO OU DE UM ÁCIDO NUCLEICO, MOLÉCULA DE ÁCIDO NUCLEICO ISOLADA, E,
10 POLIPEPTÍDEO ISOLADO”

A presente invenção diz respeito em geral ao campo da biologia molecular e está relacionada a um método para melhorar a tolerância ao estresse abiótico em plantas modulando a expressão em uma planta com um ácido nucléico que codifica um tipo de alfin. A presente invenção também
15 está relacionada a plantas com expressão modulada de um ácido nucléico que codifica um polipeptídeo do tipo alfin, cujas plantas apresentam melhor tolerância ao estresse abiótico com relação às plantas tipo selvagem correspondentes ou outras plantas controle. A invenção também fornece construtos usados nos métodos da invenção.

Além disso, a presente invenção diz respeito em geral ao campo da biologia molecular e está relacionada a um método para melhorar a tolerância ao estresse abiótico em plantas modulando a expressão em uma planta com um ácido nucléico que codifica um YRP. A presente invenção também está relacionada a plantas com expressão modulada de um ácido
25 nucléico que codifica um YRP, cujas plantas apresentam melhor tolerância ao estresse abiótico com relação às plantas tipo selvagem correspondentes ou outras plantas controle. A invenção também fornece construtos usados nos métodos da invenção.

Além disso, a presente invenção diz respeito em geral ao campo da biologia molecular e está relacionada a um método para aumentar
30 várias características da planta relacionadas ao rendimento aumentando a

expressão em uma planta com uma sequência de ácidos nucleicos que codifica um polipeptídeo do tipo de Brevis Radix (BRXL). A presente invenção também está relacionada a plantas com melhor expressão de uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL, cujas
5 plantas apresentam melhores características relacionadas ao rendimento com relação às plantas controle. A invenção diz respeito adicionalmente a sequências de ácidos nucleicos, construtos de ácido nucléico, vetores e plantas contendo as ditas sequências de ácidos nucleicos.

Além disso, a presente invenção diz respeito em geral ao
10 campo da biologia molecular e está relacionada a um método para melhorar a tolerância ao estresse abiótico em plantas modulando a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo do tipo silky-1. A presente invenção também está relacionada a plantas com expressão modulada de um ácido nucléico que codifica um polipeptídeo do tipo silky-1,
15 cujas plantas apresentam melhor tolerância ao estresse abiótico com relação às plantas tipo selvagem correspondentes ou outras plantas controle. A invenção também fornece construtos usados nos métodos da invenção.

Além disso, a presente invenção diz respeito em geral ao campo da biologia molecular e está relacionada a um método para melhorar
20 várias características de crescimento da planta modulando a expressão em uma planta com um ácido nucléico que codifica um ARP6 (proteína 6 relacionada à actina). A presente invenção também está relacionada a plantas com expressão modulada de um ácido nucléico que codifica um ARP, cujas plantas apresentam melhores características de crescimento com relação às
25 plantas tipo selvagem correspondentes ou outras plantas controle. A invenção também fornece construtos usados nos métodos da invenção.

Além disso, a presente invenção diz respeito em geral ao campo da biologia molecular e está relacionada a um método para melhorar características relacionadas ao rendimento em plantas modulando a expressão

em uma planta com um ácido nucléico que codifica um polipeptídeo de POP (Prolil-oligopeptidase). A presente invenção também está relacionada a plantas com expressão modulada de um ácido nucléico que codifica um polipeptídeo de POP, cujas plantas apresentam melhores características relacionadas ao rendimento em plantas com relação às plantas tipo selvagem correspondentes ou outras plantas controle. A invenção também fornece construtos usados nos métodos da invenção.

Além disso, a presente invenção diz respeito em geral ao campo da biologia molecular e está relacionada a um método para melhorar características relacionadas ao rendimento modulando a expressão em uma planta com um ácido nucléico que codifica uma CRL (Folha amarrotada). A presente invenção também está relacionada a plantas com expressão modulada de um ácido nucléico que codifica uma CRL, cujas plantas apresentam melhores características relacionadas ao rendimento com relação às plantas tipo selvagem correspondentes ou outras plantas controle. A invenção também fornece construtos usados nos métodos da invenção.

A população mundial cada vez maior e a diminuição de suprimentos de terra cultivável disponíveis para pesquisa de combustíveis agrícolas no sentido de aumentar a eficiência de agricultura. Os meios convencionais para melhorias no cultivo e horticultura utilizam técnicas de cultivo seletivo para identificar plantas com características desejáveis. Entretanto, tais técnicas de cultivo seletivo apresentam obstáculos, o que quer dizer que estas técnicas são tipicamente trabalhosas e resultam em plantas que contêm frequentemente componentes genéticos heterogêneos que podem frequentemente não resultar na característica desejável que é passada para as plantas mãe ou a partir das mesmas. Os avanços na biologia molecular permitiram que os pesquisadores modificassem o germoplasma de animais e plantas. A engenharia genética de plantas implica no isolamento e manipulação de material genético (tipicamente na forma de DNA ou RNA) e

na introdução subsequente deste material genético em uma planta. Tal tecnologia apresenta a capacidade de distribuir cultivos ou plantas com várias características econômicas, agronômicas ou de horticulturas melhoradas.

Uma característica de interesse econômico particular é o maior rendimento. O rendimento é definido em geral como a produção mensurável de valor econômico de um cultivo. Isto pode ser definido nos termos de quantidade e/ou qualidade. O rendimento depende diretamente de vários fatores, por exemplo, o número e tamanho dos órgãos, arquitetura da planta (por exemplo, o número de ramificações), produção de semente, senescência da folha e mais. O desenvolvimento da raiz, absorção de nutrientes, tolerância ao estresse e vigor inicial também podem ser fatores importantes na determinação do rendimento. Portanto, a otimização dos fatores anteriormente mencionados pode contribuir para aumentar o rendimento do cultivo.

O rendimento da semente é uma característica particularmente importante, uma vez que as sementes de muitas plantas são importantes para nutrição humana e animal. As culturas tais como milho, arroz, trigo, canola e soja são responsáveis pela metade da absorção calórica humana total, quer por meio do consumo direto das próprias sementes ou por meio do consumo de produtos a base de carne constituídos por 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 ramos e raízes) e um endosperma (a fonte de nutrientes para o crescimento do embrião durante a germinação e durante o crescimento inicial de mudas). O desenvolvimento de uma semente envolve muitos genes, e exige a transferência de metabólitos a partir das raízes, folhas e caules na semente em crescimento. O endosperma, em particular, assimila os precursores metabólicos de carboidratos, óleos e proteínas e sintetiza-os em macromoléculas de armazenamento para preencher o grão.

A biomassa da planta é o rendimento para culturas de

forragem do tipo alfafa, milho para silagem e feno. Muitos substitutos para o rendimento foram usados em culturas de grão. Os principais entre estes são estimativas do tamanho da planta. O tamanho da planta pode ser medido de muitas maneiras dependendo da espécie e do estágio de desenvolvimento, mas incluem peso seco da planta total, peso seco do crescimento aéreo, peso fresco do crescimento aéreo, área da folha, volume do caule, altura da planta, diâmetro da roseta, comprimento da folha, comprimento da raiz, massa da raiz, número de cana e número de folha. Muitas espécies mantêm uma razão conservativa entre o tamanho de diferentes partes da planta em um dado estágio de desenvolvimento. Estas relações aliométricas são usadas para extrapolar a partir de uma das mesmas medidas de tamanho para uma outra (por exemplo, Tittonell et al 2005 *Agric Ecosys & Environ* 105: 213). O tamanho da planta em um estágio inicial de desenvolvimento será tipicamente correlacionado com o tamanho posterior da planta em desenvolvimento. Uma planta maior com uma maior área da folha pode absorver tipicamente mais luz e dióxido de carbono do que uma planta menor e, portanto, ganharão provavelmente um maior peso durante o mesmo período (Fasoula & Tollenaar 2005 *Maydica* 50:39). Isto é, além da continuação potencial da vantagem micro-ambiental ou genética que a planta apresentou para atingir o maior tamanho inicialmente. Existe um componente genético forte para o tamanho da planta e taxa de crescimento (por exemplo, ver Steege et al 2005 *Plant Physiology* 139:1078) e, assim, para uma faixa de tamanho da planta de genótipos diversos em uma condição ambiental, é provável de correlacionar com o tamanho em uma outra (Hittalmani et al 2003 *Theoretical Applied Genetics* 107:679). Da mesma maneira, um ambiente padrão é usado como um substituto para os ambientes diversos e dinâmicos encontrados em diferentes locais e períodos por culturas no campo.

Uma outra característica importante para muitas culturas é o vigor inicial. Melhorar o vigor inicial é um objetivo importante de programas

modernos de cultivo de arroz tanto em cultivares de arroz temperado quanto tropical. As raízes longas são importantes para a ancoragem adequado no solo em arroz semeado em água. Quando o arroz é semeado diretamente em campos alagados, e quando as plantas têm que emergir rapidamente por meio da água, os ramos mais longos são associados ao vigor. Quando a semeadeira é utilizada, mesocótilos mais longos e coleóptiles são importantes para boa emergência da muda. A capacidade de modificar por engenharia o vigor inicial nas plantas é de maior importância na agricultura. Por exemplo, pouco vigor inicial tem sido uma limitação na introdução de híbridos de milho (*Zea mays* L.) baseados no germoplasma do cinturão de milho no Atlântico europeu.

O índice da colheita, a razão do rendimento da semente para peso seco do crescimento aéreo, é relativamente estável em muitas condições ambientais e, assim, uma correlação consistente entre o tamanho da planta e rendimento do grão pode ser frequentemente obtida (por exemplo, Rebetzke et al 2002 Crop Science 42:739). Estes processos são ligados de maneira intrínseca em decorrência de a maioria da biomassa do grão ser dependente na produtividade fotossintética atual ou armazenada pela folhas e caule da planta (Gardener et al. (1985) Physiology of Crop Plants. Iowa State University Press, pp 68-73). Portanto, a seleção por tamanho da planta, mesmo em estágios iniciais de desenvolvimento, é usada como um indicador para o rendimento potencial futuro (por exemplo Tittonell et al 2005 Agric Ecosys & Environ 105: 213). Durante o teste com relação ao impacto de diferenças genéticas na tolerância ao estresse, a capacidade de padronizar propriedades do solo, temperatura, disponibilidade de água e nutriente e intensidade da luz é uma vantagem intrínseca dos ambientes de estufa ou câmara de crescimento comparados ao campo. Entretanto, limitações artificiais no rendimento em virtude de pouca polinização por causa da ausência do vento ou insetos, ou espaço insuficiente para raiz madura ou crescimento da copa, podem

restringir o uso destes ambientes controlados para testar as diferenças no rendimento. Portanto, medições de tamanho da planta em desenvolvimento inicial, em condições padronizadas em uma câmara de crescimento ou estufa, são práticas padrões para fornecer a indicação de vantagens de rendimento genético potencial.

Uma outra característica de particular importância é aquela de maior tolerância ao estresse abiótico. O estresse abiótico é uma causa principal de perda de cultivo em todo o mundo, reduzindo os rendimentos médios para a maioria das principais plantas de cultivo em mais de 50 % (Wang et al. (2003) Plant 218: 1 -14). Os estresses abióticos podem ser causados por seca, salinidade, extremos de temperatura, toxicidade química, excesso ou deficiência de nutrientes (macro elementos e/ou micro elementos), radiação e estresse oxidativo. A capacidade de aumentar tolerância à planta ao estresse abiótico seria de melhor vantagem econômica para os fazendeiros em todo o mundo e permitiria o cultivo de culturas durante condições adversas e em territórios quando o cultivo de culturas não for de outra forma possível.

Portanto, o rendimento do cultivo pode ser maior otimizando um dos fatores anteriormente mencionados.

Dependendo da finalidade do uso, a modificação de certas características de rendimento pode ser favorecida com relação a outras. Por exemplo, para aplicações tais como produção de forragem ou madeira, ou recursos de bio-combustível um aumento nas partes vegetativas de uma planta pode ser desejável, e para aplicações tais como produção de farinha, amido ou óleo, um aumento nos parâmetros da semente pode ser particularmente desejável. Ainda entre os parâmetros da semente, alguns podem ser favorecidos com relação a outros, dependendo da aplicação. Várias mecanismos podem contribuir para aumentar o rendimento da semente, quer aquele que está na forma de maior tamanho da semente ou maior número de semente.

Uma abordagem para aumentar o rendimento e/ou características relacionadas ao rendimento (o rendimento da semente e/ou da biomassa) em plantas pode ser por meio de modificação dos mecanismos de crescimento inerentes de uma planta, tal como o ciclo celular ou várias vias de sinalização envolvidas no crescimento da planta ou nos mecanismos de defesa.

Observa-se agora que a tolerância a vários estresses abióticos pode ser melhorada em plantas modulando a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo do tipo alfin.

10 Observa-se agora que a tolerância a vários estresses abióticos pode ser melhorada em plantas modulando a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo YRP.

15 Observa-se agora que várias características relacionadas ao rendimento podem ser melhores em plantas com relação às plantas controle, aumentando a expressão em uma planta com uma sequência de ácidos nucleicos que codifica um polipeptídeo do tipo de Brevis Radix (BRXL). As melhores características relacionadas ao rendimento compreendem um ou mais de: maior altura da planta e aumento no peso de mil grãos (TKW).

20 Observa-se agora que a tolerância a vários estresses abióticos pode ser melhorada em plantas modulando a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo do tipo silky-1.

Observa-se agora que várias características de crescimento podem ser melhoradas em plantas modulando a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo ARP6 em uma planta.

25 Observa-se agora que várias características de crescimento podem ser melhoradas em plantas modulando a expressão em uma planta com um ácido nucléico que codifica um POP em uma planta.

Observa-se agora que várias características relacionadas ao rendimento podem ser melhores em plantas modulando a expressão em uma

planta com um ácido nucléico que codifica uma CRL (Folha amarrotada) em uma planta.

FUNDAMENTOS DA INVENÇÃO

1. Polipeptídeos do tipo alfin

5 O dedo PHD, um dedo de zinco Cys₄-His-Cys₃, é observado em muitas proteínas regulatórias de plantas a animais e que são frequentemente associados à regulação transcricional mediada por cromatina. O dedo PHD mostrou ativar a transcrição em células de levedura, planta e animal (Halbach et al., Nucleic Acids Res. 2000 September 15; 28(18); 3542-10 3550).

Os domínios de dedo de zinco derivados do tipo alfin pertencem à família do domínio de dedo PHD (R. Aasland, et al., Trends Biochem Sci (1995) 20:56-9). Especula-se que o domínio PHD do tipo alfin desempenha o papel de se ligar ao DNA de uma maneira sensível ao EDTA 15 que interfere na necessidade do zinco se ligar em um motivo de núcleo hexâmero tanto de GNGGTG quanto de GTGGNG (D. Bastola, et al., Plant Mol Biol. (1998) 38:1123-35). Oito genes do fator do tipo alfin (ALF) foram identificados em *Arabidopsis* (J. L. Riechmann, et al., Science (2000) 290:2105-10). A expressão de uma versão anti-sentido do tipo alfin fez com 20 que a alfafa transgênica crescesse menos, ao passo que a superexpressão constitutiva por um promotor constitutivo aumentou o crescimento da raiz tanto em condições normais quanto naquelas estressadas pelo sal (I. Winicov Plant (2000) 210:416-22).

2. Polipeptídeos YRP

25 Os YRP de *Hordeum vulgare* (SEQ ID NO: 11 e 13) são fatores de transcrição que codificam elementos da classe GARP de fatores de transcrição.

3. Polipeptídeos do tipo de Brevis Radix (BRXL)

O polipeptídeo de Brevis Radix (BRX) foi identificado por

meio de um alelo de perda de função natural no acesso Umkirch-1 (Uk-1) de *Arabidopsis*, e resulta em menor tamanho do meristema da raiz, menor tamanho da célula madura e, assim, menor crescimento da raiz (Mouchel et al. (2004) Genes Dev 18: 700-714), devido à interrupção do hormônio brassinoesteróide da planta e vias de sinalização de auxina.

BRX e parólogos do tipo BRX (BRXL) pertencem a uma família de gene específica de planta conservada (coletivamente denominada BRXL) que codifica proteínas que são previstas para regular a transcrição direta ou indiretamente. Os genes BRXL são observados em todas as plantas superiores para as quais os dados são disponíveis, mas não em organismos unicelulares e animais. Nos genomas de planta totalmente sequenciados de *Arabidopsis*, álamo (*Populus trichocarpa*) e arroz (*Oryza sativa*), cinco genes BRXL podem ser observados (Briggs et al. (2006) Plant Physiol 140: 1307-1316).

Quatro domínios de alta conservação podem ser distinguidos nas proteínas da família BRX. A homologia entre as proteínas da família BRX nas espécies e entre estas é especialmente conservada nestas regiões:

1. Nas terminações N, dois domínios curtos de aproximadamente 10 e 25 aminoácidos respectivamente são conservados, contendo Cys conservada, cujo espaçamento é indicativo de um motivo de ligação de zinco potencial.

2. A região média de proteínas da família BRX contém um domínio altamente conservado de aproximadamente 65 aminoácidos.

3. Um segundo domínio altamente conservado de aproximadamente 60 aminoácidos está presente, homólogo ao primeiro domínio médio, que constitui um tipo inédito de repetição em tandem, que é a principal característica de proteínas da família BRX (portanto, denominado o domínio BRX).

As regiões em alfa hélice, que são características de domínios

de interação de ligação de DNA e proteína-proteína, foram identificadas nos domínios BRX conservados. Experimentos de dois híbridos em levedura demonstraram que o domínio BRX é um domínio de interação proteína-proteína inédito, que provavelmente medeia a homodimerização e heterodimerização no BRXL e/ou entre este e também famílias de proteína do tipo PRAF (PH, RCC1, e FYVE) (Briggs et al. (2006) Plant Physiol 140: 1307-1316; van Leeuwen et al. (2004) Trends Plant Sci 9: 378-384). As proteínas do tipo PRAF também contêm repetições de regulador de condensação de cromossomo 1 (RCC1), que fornecem frequentemente atividade de troca do nucleotídeo guanina.

Na patente U.S. 7214786 "Moléculas de ácido nucléico e outras moléculas associadas a plantas e usos das mesmas para melhoria da planta", são descritas sequências de ácidos nucleicos que codificam polipeptídeos BRXL (SEQ IDs NO: 35674, 30290, 17003) e construtos que compreendem estes. Os polinucleotídeos recombinantes revelados e polipeptídeos recombinantes são usados na produção de plantas transgênicas para produzir plantas com melhores propriedades. Na patente U.S. 7365185, "Sequências de planta genômica e usos da mesma", são descritas sequências promotoras de ácido nucléico de arroz de polipeptídeos BRXL (SEQ IDs NO: 59178, 70484, 70442, 37078, 78410, 64873) e construtos que compreendem estas. A invenção revela adicionalmente composições, células hospedeiras transformadas, plantas transgênicas e sementes contendo as sequências promotoras genômicas de arroz, e métodos para preparar e usar as mesmas.

4. Polipeptídeos do tipo silky-1

Silky-1 é um elemento da família de fatores de transcrição MADS e que está envolvido no desenvolvimento da flor.

5. Polipeptídeos ARP6

As proteínas relacionadas à actina (ARPs) constituem uma família de proteínas eucarióticas cujas sequências principais desempenham

homologia com as actinas convencionais. Ao passo que as actinas desempenham papéis de citoesqueleto bem caracterizados, as ARPs são implicadas em várias funções celulares tanto no citoplasma quanto no núcleo. As ARPs citoplasmáticas, por exemplo, são conhecidas por participar da

5 montagem de filamentos ramificados de actina e movimento mediado por dineína de vesículas em muitos eucariotos. As ARPs nucleares são componentes de vários complexos modificados por cromatina envolvidos na regulação transcricional. Em plantas, por exemplo, foi recentemente descrita a

10 existência de um complexo do tipo SWR1/SRCAP. Estes complexos parecem desestabilizar interações proteína-proteína e proteína-DNA no nucleossomo, permitindo o remodelamento da cromatina e, portanto, influenciado a expressão de gene. Proteínas ARP6 de levedura e mamífero funcionam com os complexos SWR1 e SRCAP, respectivamente, que depositam o variante H2A.Z de histona na cromatina. A ARP6 de *Arabidopsis thaliana* interage

15 com as proteínas ARP6, PIE1 e SEF indicando que a função de ARP6 é conservada também no reino das plantas (March-Diaz, R. et al. 2007. Plant Physiol. 143, 893- 901). Mutações neutras em AtARP6 resultam em misregulação da expressão de um número de genes e na floração inicial e

20 fenótipos anões em *Arabidopsis thaliana* (Deal 2007, The Plant Cell, Vol. 19: 74-83).

As ARPs e actinas possuem uma estrutura terciária comum centrada na bolsa de ligação de nucleotídeo conhecida como o dobramento de actina (Kabach et al. 1995. FASEB J. 9,167-1745). As ARPs são agrupadas em várias classes ou subfamílias que são altamente conservadas em uma

25 ampla faixa de eucariotos. Cada classe ou subfamília é distinguida por seu grau de similaridade com a actina convencional (Kandasamy et al. 2004 Trends Plant Sci 9: 196-202).

6. Polipeptídeos POP

As proteases catalisam a hidrólise de ligações de peptídeo,

tanto no final de uma cadeia de polipeptídeo (exopeptidases) quanto na cadeia de polipeptídeo (endopeptidases). Elas são classificadas de acordo com sua conservação estrutural, ver, por exemplo, a base de dados MEROPS (Rawlings et al., Nucleic Acids Research 34, D270-272, 2006). As serina proteases apresentam em seu sítio ativo uma serina que desempenha um papel na hidrólise do substrato, e formam nas plantas o maior grupos de proteinases. No grupo das serina proteases, vários subgrupos são discriminados, tais como a família subtilisina, a família quimotripsina, família B de carboxipeptidase D-Ala-D-Ala ou a família prolil oligopeptidase.

As prolil-oligopeptidases são postuladas para desempenhar um papel na formação, processamento e degradação de peptídeos biologicamente ativos, e foram descritas em bactérias, arqueias bem como em eucariotos. No grupo de prolil-oligopeptidases, 4 subfamílias podem ser discriminadas: prolil-oligopeptidase (S9A), dipeptidil-peptidase IV (S9B), aminoacilpeptidase (S9C) e glutamil endopeptidase (S9D).

Apesar do fato de que muitas proteases foram identificadas, pouco se sabe sobre os substratos da mesmas enzimas. Portanto, também a função e a regulação de proteases são dificilmente caracterizadas.

7. Polipeptídeos do tipo folha amarrotada (CRL)

Folha amarrotada (crl) é o nome dado a um mutante de *Arabidopsis thaliana* com morfogênese anormal de todos os órgãos da planta e divisão de plastídeos. A análise histológica revelou que os planos de divisão celular foram distorcidos em meristemas apicais do ramo (SAMs), radículas e embriões em plantas que possuem a mutação crl. Além disso, padrões de diferenciação do córtex e de células endodérmicas na inflorescência, caules e células endodérmicas da raiz foram interrompidos no mutante crl. Estes resultados sugerem que anormalidades morfológicas observadas no mutante crl foram em decorrência da divisão celular e diferenciação anormais. Além do mais, células do mutante crl continham um número menor de plastídeos

maiores, indicando que a divisão de plastídeos foi inibida no *crl*. O gene mutado e responsável pelo fenótipo foi denominado CRL (Folha amarrotada). O gene CRL codifica uma proteína com uma massa molecular de 30 kDa que está localizada no envelope do plastídeo. A proteína CRL é conservada em várias espécies de planta, incluindo uma samambaia, e em cianobactérias. A proteína CRL de *Arabidopsis* apresenta um domínio de membrana putativa localizado entre os aminoácidos 19-36 e um domínio conservado entre resíduos de aminoácido 42-236. Este domínio é altamente conservado entre as proteínas CRL presentes em outras espécies de planta.

SUMÁRIO DA INVENÇÃO

1. Polipeptídeos do tipo alfin

De maneira surpreendente, observa-se agora que a modulação da expressão de um ácido nucléico que codifica um polipeptídeo do tipo alfin fornece plantas com melhor tolerância a vários estresses abióticos com relação às plantas controle.

De acordo com uma modalidade, é fornecido um método para melhorar a tolerância em plantas a vários estresses abióticos, com relação a tolerância em plantas controle, compreendendo modular a expressão de um ácido nucléico que codifica um polipeptídeo do tipo alfin em uma planta.

2. Polipeptídeos YRP

De maneira surpreendente, observa-se agora que a modulação da expressão de um ácido nucléico que codifica um polipeptídeo YRP fornece plantas com melhor tolerância a vários estresses abióticos com relação às plantas controle.

De acordo com uma modalidade, é fornecido um método para melhorar a tolerância em plantas a vários estresses abióticos, com relação a tolerância em plantas controle, compreendendo modular a expressão de um ácido nucléico que codifica um polipeptídeo YRP em uma planta.

3. Polipeptídeos do tipo de Brevis Radix (BRXL)

De maneira surpreendente, observa-se agora que o aumento da expressão em uma planta com uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL da maneira aqui definida, fornece plantas com mais características relacionadas ao rendimento com relação às plantas controle.

De acordo com uma modalidade, é fornecido um método para aumentar as características relacionadas ao rendimento em plantas com relação às plantas controle, compreendendo aumentar a expressão em uma planta com uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL da maneira aqui definida. As características relacionadas ao maior rendimento compreendem um ou mais de: maior altura da planta e aumento no peso de mil grãos (TKW).

4. Polipeptídeos do tipo silky-1

De maneira surpreendente, observa-se agora que a modulação da expressão de um ácido nucléico que codifica um polipeptídeo do tipo silky-1 fornece plantas com melhor tolerância a vários estresses abióticos com relação às plantas controle.

De acordo com uma modalidade, é fornecido um método para melhorar a tolerância em plantas a vários estresses abióticos, com relação a tolerância em plantas controle, compreendendo modular a expressão de um ácido nucléico que codifica um polipeptídeo do tipo silky-1 em uma planta.

5. Polipeptídeos ARP6

De maneira surpreendente, observa-se agora que a modulação da expressão de um ácido nucléico que codifica um polipeptídeo ARP6 fornece plantas com melhores características relacionadas ao rendimento com relação às plantas controle.

De acordo com uma modalidade, é fornecido um método para melhorar características relacionadas ao rendimento de uma planta com relação às plantas controle, compreendendo modular a expressão de um ácido

nucléico que codifica um polipeptídeo ARP6 em uma planta.

6. Polipeptídeos POP

De maneira surpreendente, observa-se agora que a modulação da expressão de um ácido nucléico que codifica um polipeptídeo de POP fornece plantas com melhores características relacionadas ao rendimento com relação às plantas controle.

De acordo com uma modalidade, é fornecido um método para melhores características relacionadas ao rendimento de uma planta com relação às plantas controle, compreendendo modular a expressão de um ácido nucléico que codifica um polipeptídeo de POP em uma planta.

7. Polipeptídeos do tipo folha amarrotada (CRL)

De maneira surpreendente, observa-se agora que a modulação da expressão de um ácido nucléico que codifica um polipeptídeo CRL fornece plantas com melhores características relacionadas ao rendimento com relação às plantas controle.

De acordo com uma modalidade, é fornecido um método para melhorar características relacionadas ao rendimento de uma planta com relação às plantas controle, compreendendo modular a expressão de um ácido nucléico que codifica um polipeptídeo CRL em uma planta.

Definições

Polipeptídeo(s)/Proteína(s)

As palavras "polipeptídeo" e "proteína" são aqui usadas indiferentemente e referem-se a aminoácidos em uma forma polimérica de qualquer tamanho, ligados juntos por ligações de peptídeo.

Polinucleotídeo(s)/Ácido nucléico(s)/Sequência de ácidos nucleicos(s)/sequência de nucleotídeos(s)

Os termos "polinucleotídeo(s)", "sequência de ácidos nucleicos(s)", "sequência de nucleotídeos(s)", "ácido nucléico(s)", "molécula de ácido nucléico" são aqui usados indiferentemente e referem-se a

nucleotídeos, tanto ribonucleotídeos ou desoxirribonucleotídeos quanto uma combinação de ambos, em uma forma polimérica não ramificada de qualquer tamanho.

Planta controle(s)

- 5 A escolha de plantas controle adequadas é uma parte rotineira de uma plataforma experimental e pode incluir plantas tipo selvagem correspondentes ou plantas correspondentes sem o gene de interesse. A planta controle é tipicamente da mesma espécie de planta ou ainda da mesma variedade da planta a ser avaliada. A planta controle também pode ser um
- 10 nulizigoto da planta a ser avaliada. Os nulizigotos são indivíduos que perdem o transgene por segregação. Uma "planta controle", da maneira aqui usada, refere-se não apenas a plantas completas, mas também à partes da planta, incluindo sementes e partes da semente.

Homólogo(s)

- 15 "Homólogos" de uma proteína incluem peptídeos, oligopeptídeos, polipeptídeos, proteínas e enzimas com substituições, deleções e/ou inserções de aminoácido com relação à proteína não modificada em questão e com atividade biológica e funcional similar à proteína não modificada a partir da qual eles são derivados.

- 20 Uma deleção refere-se a 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-determinando em uma proteína. As inserções podem compreender fusões N-terminais e/ou C-terminais, bem
- 25 como inserções intra-sequência de aminoácidos únicos ou múltiplos. Em geral, inserções na sequência de aminoácidos serão menores que Fusões N ou C-terminais, da ordem de cerca de 1 a 10 resíduos. Exemplos de proteínas ou peptídeos de fusão N ou C-terminal incluem o domínio de ligação ou domínio de ativação de um ativador transcricional da maneira usada no sistema de dois

híbridos em levedura, proteínas de revestimento de fago, (histidina)-6-tag, glutathione S-transferase-tag, proteína A, proteína de ligação de maltose, dihidrofolato redutase, epítipo Tag•100, epítipo c-myc, epítipo FLAG[®], lacZ, CMP (peptídeo de ligação de calmodulina), epítipo de HA, epítipo de proteína C e epítipo de VSV.

Uma substituição refere-se a troca de aminoácidos da proteína por outros aminoácidos com propriedades similares (tais como hidrofobicidade, hidrofilicidade, antigenicidade, propensão similares para formar ou quebrar estruturas em α -hélice ou estruturas em folha β). As substituições de aminoácido são tipicamente de resíduos únicos, mas podem ser agrupados dependendo das restrições funcionais colocadas no polipeptídeo; as inserções serão em geral da ordem de cerca de 1 a 10 resíduos de aminoácido. As substituições de aminoácido são preferivelmente substituições conservativas de aminoácido. As tabelas de substituição conservativa são bem conhecidas na técnica (ver, por exemplo, Creighton (1984) Proteins. W. H. Freeman e Company (Eds) e tabela 1 a seguir).

Tabela 1: Exemplos de substituições conservadas de aminoácido

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 realizadas facilmente usando técnicas sintéticas de peptídeo bem conhecidas na técnica, tais como síntese de peptídeo de fase sólida e similares, ou por manipulação do DNA recombinante. Métodos para a manipulação de sequências de DNA para produzir variantes de substituição,

inserção ou deleção de uma proteína são conhecidos na técnica. Por exemplo, técnicas para preparar mutações por substituição em sítio pré-determinados em DNA são bem conhecidas pelos versados na técnica e incluem mutagênese M13, mutagênese *in vitro* T7-Gen (USB, Cleveland, OH),
5 mutagênese sítio direcionada de mudança rápida (Stratagene, San Diego, CA), mutagênese sítio direcionada mediada por PCR ou outros protocolos de mutagênese sítio direcionada.

Derivados

"Derivados" incluem peptídeos, oligopeptídeos, polipeptídeos
10 que podem compreender, comparados à sequência de aminoácidos da forma da proteína de ocorrência natural, tal como a proteína de interesse, substituições de aminoácidos com resíduos de aminoácido de ocorrência não natural, ou adições de resíduos de aminoácido de ocorrência não natural. "Derivados" de uma proteína também incluem peptídeos, oligopeptídeos,
15 polipeptídeos que compreendem resíduos de aminoácido alterados de ocorrência natural (glicosilados, acilados, prenilados, fosforilados, miristoilados, sulfatados etc.) ou alterados de ocorrência não natural comparados com a sequência de aminoácidos de uma forma de ocorrência não natural do polipeptídeo. Um derivado também pode compreender um ou mais
20 substituintes ou adições de não aminoácidos comparados com a sequência de aminoácidos a partir da qual ele é derivado, por exemplo, uma molécula repórter, ou outro ligante, ligada covalente ou não covalentemente à sequência de aminoácidos, tal como uma molécula repórter que é ligada para facilitar sua deleção, e resíduos de aminoácido de ocorrência não natural com relação
25 à sequência de aminoácidos de uma proteína de ocorrência natural. Além disso, "derivados" também incluem fusões da forma da proteína de ocorrência natural com peptídeos de marcação tal como FLAG, HIS6 ou tioredoxina (para uma revisão de peptídeos de marcação, ver Terpe, Appl. Microbiol. Biotechnol. 60, 523-533, 2003).

Ortólogo(s)/Parólogo(s)

Ortólogos e parólogos incluem conceitos evolucionários usados para descrever as relações ancestrais de genes. Os parólogos são genes na mesma espécie que originaram por meio de duplicação de um gene ancestral; os ortólogos são genes de diferentes organismos que originaram por meio de especiação, e são também derivados de um gene ancestral comum.

Domínio

A palavra "domínio" refere-se a um conjunto de aminoácidos conservados em posições específicas junto com alinhamento de sequências de proteínas evolucionariamente relacionadas. Embora os aminoácidos em outras posições possam variar entre homólogos, os aminoácidos que são altamente conservados nas posições específicas indicam aminoácidos que são provavelmente essenciais em estrutura, estabilidade ou função de uma proteína. Identificadas por seu alto grau de conservação em sequências alinhadas de uma família de proteína homólogas, elas podem ser usadas como identificadores para determinar se qualquer polipeptídeo em questão pertence a uma família de polipeptídeo previamente identificada.

Motivo/Sequência consensual/assinatura

O termo "motivo" ou "sequência consensual" ou "assinatura" refere-se a uma pequena região conservada na sequência de proteínas evolucionariamente relacionadas. Os motivos são em geral partes altamente conservadas de domínios, mas também podem incluir apenas parte do domínio, ou estar localizados fora do domínio conservado (se todos os aminoácidos do motivo estiverem fora de um domínio definido).

Hibridização

A palavra "hibridização", da maneira aqui definida, é um processo em que sequências de nucleotídeos complementares substancialmente homólogas anelam-se umas às outras. 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 em uma matriz tais como microesferas magnéticas, microesferas de sefarose ou qualquer outra resina. O processo de hibridização pode, além
5 disso, ocorrer com um dos ácidos nucleicos complementares imobilizados em um suporte sólido, tal como uma membrana nitro-celulose ou de náilon ou imobilizados, por exemplo, por fotolitografia, por exemplo, um em suporte de vidro silicoso (o último conhecido como arranjos ou microarranjos de ácido
10 nucléico ou como chipes de ácido nucléico). A fim de permitir que a hibridização ocorra, as moléculas de ácido nucléico são em geral térmica ou quimicamente desnaturadas para fundir uma fita dupla em duas fitas simples e/ou para remover estruturas em forma de grampo ou outras secundárias a partir de ácidos nucleicos de fita simples.

A palavra "severidade" refere-se às condições nas quais uma
15 hibridização ocorre. A severidade da hibridização é influenciada por condições tais como temperatura, concentração de sal, concentração iônica e composição de tampão de hibridização. Em geral, baixas condições de severidade são selecionadas para estar cerca de 30 °C abaixo do ponto de fusão térmica (T_m) para a sequência específica em uma concentração iônica e
20 pH definidos. São condições de severidade do meio quando a temperatura está 20 °C abaixo da T_m , e são condições de alta severidade quando a temperatura está 10 °C abaixo da T_m . As condições de hibridização de alta severidade são tipicamente usadas para isolar sequências de hibridização que apresentam alta similaridade de sequência com a sequência alvo de ácidos nucleicos.
25 Entretanto, ácidos nucleicos podem desviar em sequência e ainda codificam um polipeptídeo substancialmente idêntico, em virtude da degeneração do código genético. Portanto, as condições de hibridização de severidade média podem algumas vezes ser necessárias para identificar tais moléculas de ácido nucléico.

A T_m é a temperatura em concentração iônica e pH definidos, em que 50 % da sequência alvo hibridiza em uma sonda perfeitamente combinada. A T_m é dependente das condições de solução e da composição base e tamanho da sonda. Por exemplo, sequências mais longas hibridizam especificamente em temperaturas maiores. A taxa máxima de hibridização 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 as duas fitas de ácido nucléico, promovendo por meio disso a formação de híbrido; este efeito é visível em concentrações de sódio de até 0,4M (para concentrações mais elevadas, este efeito pode ser ignorado). A formamida reduz a temperatura de fusão de duplexes 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 a hibridização a ser realizada em 30 a 45 °C, embora a taxa de hibridização seja menor. As combinações de par de base reduzem 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 combinação de base. A T_m pode ser calculada usando as equações a seguir, dependendo dos tipos de híbridos:

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

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

2) Híbridos de DNA-RNA ou RNA-RNA:

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

3) Híbridos de oligo-DNA ou oligo-RNA^d:

Para <20 nucleotídeos: $T_m = 2 (I_n)$

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

^a ou para outro cátion monovalente, mas exato apenas na faixa

0,01-0,4 M.

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

^c L = tamanho de duplex nos pares de base.

^d oligo, oligonucleotídeo; I_n = comprimento eficiente de

5 oligonucleotídeo iniciador = $2 \cdot (\text{número de G/C}) + (\text{número de A/T})$.

A ligação não específica pode ser controlada usando qualquer uma de inúmeras técnicas conhecidas tais como, por exemplo, bloquear a membrana com proteína contendo soluções, adições de RNA heterólogo, DNA e SDS ao tampão de hibridização, e tratamento com Rnase. Para sondas
10 não homólogas, uma série de hibridizações pode ser realizada variando um de (i) reduzir progressivamente a temperatura de anelamento (por exemplo, de 68 °C para 42 °C) ou (ii) reduzir progressivamente a concentração de formamida (por exemplo, de 50 % a 0 %). Os versados na técnica estão cientes de vários parâmetros que podem ser alterados durante a hibridização e
15 que tanto manterão quanto mudarão as condições de severidade.

Além das condições de hibridização, a especificidade de hibridização também depende tipicamente da função de lavagens pós-hibridização. Para remover a parte inferior resultante da hibridização não específica, as amostras são lavadas com soluções salinas diluídas. Fatores
20 importantes de tais lavagens incluem a concentração iônica e temperatura da solução final de lavagem: quanto menor a concentração de sal e quanto maior a temperatura da lavagem, maior a severidade da lavagem. As condições de lavagem são tipicamente realizadas com severidade de hibridização ou abaixo da mesmas. Uma hibridização positiva fornece um sinal que é pelo menos
25 duas vezes o daquele da parte inferior. Em geral, condições severas adequadas para ensaios de hibridização de ácido nucléico ou procedimentos de detecção para amplificação de gene são da maneira apresentada anteriormente. Condições mais ou menos severas também podem ser selecionadas. Os versados na técnica estão cientes de vários parâmetros que podem ser

alterados durante a lavagem e que tanto manterão quanto mudarão as condições de severidade.

Por exemplo, condições de hibridização de alta severidade típicas para híbridos de DNA com mais de 50 nucleotídeos incluem hibridização a 65 °C em SSC 1x ou a 42 °C em SSC 1x e 50 % de formamida, seguido por lavagem a 65 °C em SSC 0,3x. Exemplos de condições de hibridização de severidade média para híbridos de DNA com mais de 50 nucleotídeos incluem hibridização a 50 °C em SSC 4x ou a 40 °C em SSC e 50 % de formamida, seguido por lavagem a 50 °C em SSC 2x. O tamanho do híbrido é o tamanho antecipado pelo ácido nucléico de hibridização. Quando os ácidos nucleicos de sequência conhecida são hibridizados, o tamanho do híbrido pode ser determinado alinhando as sequências e identificando as regiões conservadas aqui descritas. SSC 1x é NaCl 0,15M e citrato de sódio 15mM; a solução de hibridização e soluções de lavagem podem incluir adicionalmente reagente de Denhardt 5x, SDS 0,5-1,0 %, 100 µg/mL de DNA de esperma de salmão fragmentado desnaturado, pirofosfato de sódio 0,5 %.

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

Variante de união

O termo "variante de união", da maneira aqui usada, inclui variantes de uma sequência de ácidos nucleicos na qual introns e/ou exons selecionados foram excisados, substituídos, deslocados ou adicionados, ou na qual os introns foram diminuídos ou aumentados. Tais variantes serão aqueles nos quais a atividade biológica da proteína é substancialmente mantida; isto pode ser atingido mantendo seletivamente segmentos funcionais da proteína. Tais variantes de união podem ser encontrados na natureza ou podem ser

preparados pelo homem. Os métodos para prever e isolar tais variantes de união são bem conhecidos na técnica (ver, por exemplo, Foissac e Schiex (2005) BMC Bioinformatics 6: 25).

Variante alélico

5 Os alelos ou variantes alélicos são formas alternativas de um dado gene, localizados na mesma posição cromossomal. Os variantes alélicos incluem polimorfismos de nucleotídeo único (SNPs), bem como polimorfismos de pequena inserção/deleção (INDELs). O tamanho de INDELs é em geral menor que 100 bp. SNPs e INDELs formam o maior
10 conjunto de variantes de sequência em cepas polimórficas de ocorrência natural da maioria dos organismos.

Embaralhamento de gene/Evolução direcionada

O embaralhamento de gene ou evolução direcionada consiste em repetições de embaralhamento de DNA seguido por seleção apropriada
15 e/ou seleção para gerar variantes de ácidos nucleicos, ou porções destes, que codificam proteínas com uma atividade biológica modificada (Castle et al., (2004) Science 304(5674): 1151-4; patentes U.S. 5.811.238 e 6.395.547).

Elemento regulatório/Sequência controle/Promotor

Os termos "elemento regulatório", "sequência controle" e
20 "promotor" são também usados aqui indiferentemente e são obtidos em um amplo contexto para se referir às sequências regulatórias de ácidos nucleicos capazes de realizar a expressão das sequências nas quais eles são ligados. O termo "promotor" refere-se tipicamente a uma sequência de ácidos nucleicos controle localizada à montante do início transcricional de um gene e que está
25 envolvido no reconhecimento e ligação de RNA polimerase e outras proteínas, direcionando por meio disso a transcrição de um ácido nucléico operavelmente ligado. São incluídas pelos termos anteriormente mencionados sequências regulatórias transcricionais derivadas de um gene genômico eucariótico clássico (incluindo a caixa TATA, que é exigido para início exato

da transcrição, com ou sem uma sequência de caixa CCAAT) e elementos regulatórios adicionais (isto é, sequências que ativam à montante, melhoradores e silenciadores) que alteram a expressão de gene em resposta ao desenvolvimento e/ou estímulos externos, ou de uma maneira específica para tecido. É também incluída no termo uma sequência regulatória transcricional de um gene procarioto clássico, em que no caso ele pode incluir uma sequência em caixa 35 e/ou sequências regulatórias transcricionais em caixa 10. O termo "elemento regulatório" também inclui uma molécula de fusão sintética ou derivado que confere, ativa ou melhora a expressão de uma molécula de ácido nucléico em uma célula, tecido ou órgão.

Um "promotor de planta" compreende elementos regulatórios, que medeiam a expressão de um segmento de sequência codificante em células de planta. Dessa maneira, um promotor de planta não precisa ser de origem vegetal, mas pode ser originado de vírus ou microrganismos, por exemplo, de vírus que atacam células de planta. O "promotor de planta" também pode ser originado de uma célula de planta, por exemplo, da planta que é transformada com a sequência de ácidos nucleicos a ser expressa no processo inventivo e aqui descrita. Isto também se aplica a outros sinais regulatórios de "planta", tais como exterminadores de "planta". Os promotores à montante das sequências de nucleotídeos usadas nos métodos da presente invenção podem ser modificados por uma ou mais substituição(s), inserção(s) e/ou deleção(s) de nucleotídeo sem interferir com a funcionalidade ou atividade de ambos os promotores, o quadro de leitura aberto (ORF) ou a região regulatória 3', tais como exterminadores ou outras regiões regulatórias 3' que estão localizadas longe do ORF. Além disso, é possível que a atividade dos promotores seja maior pela modificação de sua sequência, ou que eles sejam substituídos completamente por promotores mais ativos, mesmo promotores de organismos heterólogos. Para a expressão em plantas, a molécula de ácido nucléico, da maneira descrita anteriormente, tem que estar

ligada ou compreende um promotor adequado que expressa o gene no ponto exato de tempo e com o padrão de expressão espacial exigido.

Para a identificação de promotores funcionalmente equivalentes, a resistência do promotor e/ou padrão de expressão de um promotor candidato podem ser analisados, por exemplo, ligando operavelmente o promotor a um gene repórter e ensaiando o nível e padrão de expressão do gene repórter em vários tecidos da planta. Genes repórteres bem conhecidos e adequados incluem, por exemplo, beta-glucuronidase ou beta-galactosidase. A atividade do promotor é ensaiada medindo a atividade enzimática da beta-glucuronidase ou beta-galactosidase. A resistência do promotor e/ou padrão de expressão podem ser comparadas àquela de um promotor de referência (tal como aqueles usado nos métodos da presente invenção). Alternativamente, a resistência do promotor pode ser ensaiada quantificando níveis de RNAm ou comparando os níveis de RNAm do ácido nucléico usado nos métodos da presente invenção, com níveis de RNAm de genes próprios, tal como RNAr 18S, usando métodos conhecidos na técnica, tais como Northern blotting com análise densitométrica de autoradiogramas, PCR em tempo real quantitativa ou RT-PCR (Heid et al., 1996 Genome Methods 6: 986-994). Pretende-se, em geral, que "promotor fraco" seja um promotor que direciona a expressão de uma sequência codificante em um nível baixo. Pretende-se que "nível baixo" esteja em níveis de cerca de 1/10.000 transcritos a cerca de 1/100.000 transcritos, a cerca de 1/500.0000 transcritos per célula. De modo contrário, um "promotor forte" direciona a expressão de uma sequência codificante em alto nível, ou em cerca de 1/10 transcritos a cerca de 1/100 transcritos a cerca de 1/1.000 transcritos per célula. Em geral, pretende-se que "promotor de média resistência" seja um promotor que direciona a expressão de uma sequência codificante em um nível menor que um promotor forte, em particular, em um nível que está em todos os exemplos a seguir que foram obtidos durante o controle de um

promotor 35S CaMV.

Operavelmente ligado

O termo "operavelmente ligado", da maneira aqui usada, refere-se a uma ligação funcional entre a sequência promotora e o gene de interesse, de maneira tal que a sequência promotora seja capaz de iniciar a transcrição do gene de interesse.

Promotor constitutivo

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

Tabela 2a: Exemplos de promotores constitutivos

Fonte do Gene	Referência
Actina	McElroy et al, Plant Cell, 2: 163-171, 1990
HMGP	WO 2004/070039
CAMV 35S	Odell et al, Nature, 313: 810-812, 1985
CaMV 19S	Nilsson et al., Physiol. Plant. 100:456-462, 1997
GOS2	de Pater et al, Plant J Nov;2(6):837-44, 1992, WO 2004/065596
Ubiquitina	Christensen et al, Plant Mol. Biol. 18: 675-689, 1992
Ciclofilina de arroz	Buchholz et al, Plant Mol Biol. 25(5): 837-43, 1994
Histona H3 de milho	Lepetit et al, Mol. Gen. Genet. 231:276-285, 1992
Histona H3 de alfafa	Wu et al. Plant Mol. Biol. 11:641-649, 1988
Actina 2	An et al, Plant J. 10(1); 107-121, 1996
34S FMV	Sanger et al., Plant. Mol. Biol., 14, 1990: 433-443
Subunidade pequena de rubisco	U.S. 4.962.028
OCS	Leisner (1988) Proc Natl Acad Sci USA 85(5): 2553
SAD1	Jain et al., Crop Science, 39 (6), 1999: 1696
SAD2	Jain et al., Crop Science, 39 (6), 1999: 1696
Nos	Shaw et al. (1984) Nucleic Acids Res. 12(20):7831-7846
V-ATPase	WO 01/14572
Super promotor	WO 95/14098
Proteínas de caixa G	WO 94/12015

Promotor ubíquo

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

Promotor regulado pelo desenvolvimento

Um promotor regulado pelo desenvolvimento é ativo durante certos estágios de desenvolvimento ou em partes da planta que passam por mudanças no desenvolvimento.

Promotor indutível

5 Um promotor indutível induziu ou aumentou o início da transcrição em resposta a uma substância química (para uma revisão, ver Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48:89- 108), estímulos ambientais e físicos, ou pode ser "indutível pelo estresse", isto é, ativado quando uma planta é exposta a várias condições de estresse, ou um "indutível por patógeno", isto é, ativado quando uma planta é exposta a vários patógenos.

Promotor específico de órgão/específico de tecido

Um promotor específico de órgão e específico de tecido é aquele que é capaz de iniciar preferivelmente a transcrição em certos órgãos ou tecidos, tais como as folhas, raízes, tecido de semente, etc. Por exemplo, 15 um "promotor específico de raiz" é um promotor que é ativo transcricionalmente de maneira predominante em raízes de planta, substancialmente para a exclusão de quaisquer outras partes de uma planta, apesar de permitir ainda qualquer expressão com perdas nestas outras partes da planta. Os promotores capazes de iniciar a transcrição apenas em certas 20 células são aqui referidos como "específicos de célula".

Exemplos de promotores específicos de raiz são listados na tabela 2b a seguir:

Tabela 2b: Exemplos de promotores específicos de raiz

Fonte do gene	Referência
RCc3	Plant Mol Biol. 1995 Jan;27(2):237-48
PHT1 de <i>Arabidopsis</i>	Kovama et al., 2005; Mudge et al. (2002, Plant J. 31:341)
Transportador de fosfato de <i>Medicago</i>	Xiao et al., 2006
Pyk10 de <i>Arabidopsis</i>	Nitz et al. (2001) Plant Sci 161(2): 337-346
genes expressíveis em raiz	Tingey et al., EMBO J. 6: 1, 1987.
Gene indutível por auxina em tabaco	Van der Zaal et al., Plant Mol. Biol. 16, 983, 1991.

β -tubulina	Oppenheimer, et al., Gene 63: 87, 1988.
genes específicos de raiz de tabaco	Conkling, et al., Plant Physiol. 93: 1203, 1990.
Gene G1-3b de <i>B. napus</i>	patente dos Estados Unidos 5.401.836
SbPRP1	Suzuki et al., Plant Mol. Biol. 21: 109-119, 1993.
LRX1	Baumberger et al. 2001, Genes & Dev. 15:1128
BTG-26 de <i>Brassica napus</i>	U.S. 20050044585
LeAMT1 (tomate)	Lauter et al. (1996, PNAS 3:8139)
The LeNRT1-1 (tomate)	Lauter et al. (1996, PNAS 3:8139)
gene patatina classe I (potato)	Liu et al., Plant Mol. Biol. 153:386-395, 1991.
KDC1 (<i>Daucus carota</i>)	Downey et al. (2000, J. Biol. Chem. 275:39420)
Gene TobRB7	W Song (1997) PhD Thesis, North Carolina State University, Raleigh, NC USA
OsRAB5a (arroz)	Wang et al. 2002, Plant Sci. 163:273
ALF5 (<i>Arabidopsis</i>)	Diener et al. (2001, Plant Cell 13:1625)
NRT2;1Np (<i>N. plumbaginifolia</i>)	Quesada et al. (1997, Plant Mol. Biol. 34:265)

Um promotor específico de semente é transcricionalmente ativo de maneira predominante em tecido de semente, mas não necessariamente de maneira exclusiva em tecido de semente (em casos de expressão com perdas). O promotor específico de semente pode ser ativo durante o desenvolvimento da semente e/ou durante a germinação. O promotor específico de semente pode ser específico de endosperma/aleurona/embrião. Exemplos de promotores específicos de semente (específicos de endosperma/aleurona/embrião) são mostrados na tabela 2c até a tabela 2f a seguir. Exemplos adicionais de promotores específicos de semente são fornecidos em Qing Qu e Takaiwa (Plant Biotechnol. J. 2, 113-125, 2004), cuja revelação é aqui incorporada pela referência a medida que é totalmente apresentada.

Tabela 2c: exemplos de promotores específicos de semente

Fonte do 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 noz brasileira	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.

LMW de trigo e glutenina-1 de HMW	Mol Gen Genet 216:81-90, 1989; NAR 17:461-2, 1989
SPA de trigo	Albani et al, Plant Cell, 9: 171-184, 1997
α , β , γ -gliadinas de trigo	EMBO J. 3:1409-15, 1984
promotor Itr1 de cevada	Diaz et al. (1995) Mol Gen Genet 248(5):592-8
hordeína B1, C, D, de 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, The Plant Journal, 116(1): 53-62, 1998
blz2	EP99106056.7
promotor sintético	Vicente-Carbajosa et al., Plant J. 13: 629-640, 1998.
prolamina NRP33 de arroz	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
a-globulina Glb-1 de arroz	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
OSH1 de arroz	Sato et al, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996
α -globulina REB/OHP-1 de arroz	Nakase et al. Plant Mol. Biol. 33: 513-522, 1997
ADP-glicose pirofosforilase de arroz	Trans Res 6:157-68, 1997
família do gene ESR de milho	Plant J 12:235-46, 1997
α -cafirina de sorgo	DeRose et al., Plant Mol. Biol 32:1029-35, 1996
KNOX	Postma-Haarsma et al, Plant Mol. Biol. 39:257-71, 1999
oleosina de arroz	Wu et al, J. Biochem. 123:386, 1998
Oleosina de girassol	Cummins et al., Plant Mol. Biol. 19: 873-876, 1992
PRO0117, proteína ribossomal 40S putativa de arroz	WO 2004/070039
PRO0136, alanina aminotransferase de arroz	não publicado
PRO0147, inibidor ITR1 de tripsina (cevada)	não publicado
PRO0151, WSI18 de arroz	WO 2004/070039
PRO0175, RAB21 de arroz	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 do tipo catepsina β	Cejudo et al, Plant Mol Biol 20:849-856, 1992
Ltp2 de cevada	Kalla et al., Plant J. 6:849-60, 1994
Chi26	Leah et al., Plant J. 4:579-89, 1994
B-Peru de milho	Selinger et al., Genetics 149:1125-38, 1998

Tabela 2d: exemplos de promotores específicos de endosperma

Fonte do gene	Referência
glutelina (arroz)	Takaiwa et al. (1986) Mol Gen Genet 208:15-22; Takaiwa et al. (1987) FEBS Letts. 221:43-47
Zeína	Matzke et al., (1990) Plant Mol Biol 14(3): 323-32
LMW de trigo e glutenina-1 de HMW	Colot et al. (1989) Mol Gen Genet 216:81-90, Anderson et al. (1989) NAR 17:461-2
SPA de trigo	Albani et al. (1997) Plant Cell 9:171-184
Gliadinas de trigo	Rafalski et al. (1984) EMBO 3:1409-15

Promotor Itr1 de cevada	Diaz et al. (1995) Mol Gen Genet 248(5):592-8
Hordeína B1, C, D, de cevada	Cho et al. (1999) Theor Appl Genet 98:1253-62; Muller et al. (1993) Plant J 4:343-55; Sorenson et al. (1996) Mol Gen Genet 250:750-60
DOF de cevada	Mena et al, (1998) Plant J 116(1): 53-62
blz2	Onate et al. (1999) J Biol Chem 274(14):9175-82
promotor sintético	Vicente-Carbajosa et al. (1998) Plant J 13:629-640
prolamina NRP33 de arroz	Wu et al, (1998) Plant Cell Physiol 39(8) 885-889
globulina Glb-1 de arroz	Wu et al. (1998) Plant Cell Physiol 39(8) 885-889
arroz globulin REB/OHP-1	Nakase et al. (1997) Plant Molec Biol 33: 513-522
ADP-glicose pirofosforilase de arroz	Russell et al. (1997) Trans Res 6:157-68
Família do gene ESR de milho	Opsahl-Ferstad et al. (1997) Plant J 12:235-46
Cafirina de sorgo	DeRose et al. (1996) Plant Mol Biol 32:1029-35

Tabela 2e: Exemplos de promotores específicos de embrião:

Fonte do gene	Referência
OSH1 de arroz	Sato et al, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996
KNOX	Postma-Haarsma et al, Plant Mol. Biol. 39:257-71, 1999
PRO0151	WO 2004/070039
PRO0175	WO 2004/070039
PRO005	WO 2004/070039
PRO0095	WO 2004/070039

Tabela 2f: Exemplos de promotores específicos de aleurona:

Fonte do gene	Referência
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 do tipo catepsina β	Cejudo et al, Plant Mol Biol 20:849-856, 1992
Ltp2 de cevada	Kalla et al., Plant J. 6:849-60, 1994
Chi26	Leah et al., Plant J. 4:579-89, 1994
B-Peru de milho	Selinger et al., Genetics 149:1125-38,1998

Um promotor específico de tecido verde da maneira aqui definida é um promotor que é ativo transcricionalmente de maneira predominante em tecido verde, substancialmente para a exclusão de quaisquer outras partes de uma planta, embora permitindo ainda qualquer expressão com perdas nestas outras partes da planta.

Exemplos de promotores específicos de tecido verde que podem ser usados para realizar os métodos da invenção são mostrados na tabela 2g a seguir.

Tabela 2g: Exemplos de promotores específicos de tecido verde

Gene	Expressão	Referência
Ortofosfato diquinase de milho	Específico de folha	Fukavama et al., 2001

Fosfoenolpiruvato carboxilase de milho	Específico de folha	Kausch et al., 2001
Fosfoenolpiruvato carboxilase de arroz	Específico de folha	Liu et al., 2003
Rubisco de subunidade pequena de arroz	Específico de folha	Nomura et al., 2000
beta expansina EXBP9 de arroz	Específico de raiz	WO 2004/070039
Rubisco de subunidade pequena de <i>Pigeonpea</i>	Específico de folha	Panguluri et al., 2005
RBCS3A de ervilha	Específico de folha	

Um outro exemplo de um promotor específico de tecido é um promotor específico de meristema, que é transcricionalmente ativo de maneira predominante em tecido meristemático, substancialmente para a exclusão de quaisquer outras partes de uma planta, embora permitindo ainda qualquer expressão com perdas nestas outras partes da planta. Exemplos de promotores específicos de meristema verde que podem ser usados para realizar os métodos da invenção são mostrados na tabela 2h a seguir.

Tabela 2h: Exemplos de promotores específicos de meristema

Fonte do gene	Padrão de expressão	Referência
OSH1 de arroz	Meristema apical de ramo, a partir do estágio globular do embrião até o estágio de muda	Sato <i>et al.</i> (1996) Proc. Natl. Acad. Sci. USA, 93: 8117-8122
Metalotioneína de arroz	Específico de meristema	BAD87835.1
WAK1 & WAK 2	Meristemas apicais de ramo e raiz, e em folhas que se expandem e sépalas	Wagner & Kohorn (2001) Plant Cell 13(2): 303-318

Finalizador

A palavra "finalizador" inclui uma sequência controle que é uma sequência de DNA na extremidade de uma unidade transcricional que sinaliza 3' em processo e a poliadenilação de um transcrito primário e finalização de transcrição. O finalizador pode ser derivado do gene natural, de uma variedade de outros genes de planta, ou de DNA-T. O finalizador a ser adicionado pode ser derivado, por exemplo, dos genes nopalina sintase ou octopina sintase ou, alternativamente, de um outro gene de planta ou, menos preferivelmente, de qualquer outro gene de eucarioto.

Modulação

A palavra "modulação" significa, com relação a expressão ou

expressão de gene, um processo no qual o nível de expressão é mudado pela dita expressão de gene em comparação à planta controle, o nível de expressão pode ser menor ou maior. A expressão original não modulada pode ser de qualquer tipo de expressão de um RNA estrutural (RNAr, RNAt) ou RNAm com tradução subsequente. O termo "modular a atividade" pode significar qualquer mudança da expressão das sequências inventivas de ácidos nucleicos ou proteínas codificadas, que levam ao melhor rendimento e/ou maior crescimento da plantas.

Expressão

O termo "expressão" ou "expressão de gene" significa a transcrição de um gene específico ou genes específicos ou construto genético específico. O termo "expressão" ou "expressão de gene" em particular significa a transcrição de um gene ou genes ou construto genético no RNA estrutural (RNAr, RNAt) ou RNAm com ou sem tradução subsequente do último em uma proteína. O processo inclui transcrição de DNA e processamento do produto do RNAm resultante.

Maior expressão/superexpressão

O termo "melhor expressão" ou "superexpressão", da maneira aqui usada, significa qualquer forma de expressão que é adicional ao nível de expressão tipo selvagem original.

Os métodos para aumentar a expressão de genes ou produtos de gene são bem documentados na técnica e incluem, por exemplo, superexpressão dirigida por promotores apropriados, o uso de melhoradores de transcrição ou melhoradores de tradução. Os ácidos nucleicos isolados que funcionam como elementos promotores ou melhoradores podem ser introduzidos em uma posição apropriada (tipicamente à montante) de uma forma não heteróloga de um polinucleotídeo, de maneira a supra-regular a expressão de um ácido nucléico que codifica o polipeptídeo de interesse. Por exemplo, promotores endógenos podem ser alterados *in vivo* por mutação,

deleção e/ou substituição (ver Kmiec, U.S. 5.565.350; Zarling et al., WO9322443), promotores isolados podem ser introduzidos em uma célula de planta na orientação e distância adequadas de um gene da presente invenção, de maneira a controlar a expressão do gene.

5 Se a expressão do polipeptídeo for desejada, em geral, é desejável incluir uma região de poliadenilação na extremidade 3' de uma região codificante de polinucleotídeo. A região de poliadenilação pode ser derivada do gene natural, de uma variedade de outros genes de planta, ou de DNA-T. A sequência da extremidade 3' a ser adicionada pode ser derivada,
10 por exemplo, dos genes nopalina sintase ou octopina sintase ou, alternativamente, de um outro gene de planta ou, menos preferivelmente, de qualquer outro gene de eucarioto.

 Uma sequência de intron também pode ser adicionada na região não traduzida 5' (UTR) ou na sequência codificante da sequência
15 codificante parcial para aumentar a quantidade da mensagem madura que se acumula no citossol. A inclusão de um intron divisível na unidade de transcrição tanto nos construtos de expressão de planta quanto de animal mostrou aumentar a expressão de gene nos níveis tanto de RNAm quanto de proteína até 1.000 vezes (Buchman e Berg (1988) MoI. Célul biol. 8: 4395-
20 4405; Callis et al. (1987) Genes Dev 1 :1183-1200). Tal melhoria de intron da expressão de gene é tipicamente maior quando colocado próximo da extremidade 5' da unidade de transcrição. O uso dos introns de milho, intron AdM-S 1, 2 e 6, o intron Bronze-1, é conhecido na técnica. Para informação
25 geral ver: The Maize Handbook, capítulo 116, Freeling e Walbot, Eds., Springer, N.Y. (1994).

Gene endógeno

 A referência aqui a um gene "endógeno" não apenas refere-se ao gene em questão da maneira observada em uma planta em sua forma natural (isto é, sem nenhuma intervenção humana), mas também refere-se

àquele mesmo gene (ou um ácido nucléico/gene substancialmente homólogo) em uma forma subsequentemente isolada (re)introduzida em uma planta (um transgene). Por exemplo, uma planta transgênica contendo um transgene como este pode encontrar uma redução substancial da expressão de transgene e/ou redução substancial de expressão do gene endógeno. O gene isolado pode ser isolado de um organismo ou pode ser produzido em laboratório, por exemplo, por síntese química.

Menor expressão

A referência aqui a "menor expressão" ou "redução ou eliminação substancial" de expressão é obtida para significar uma diminuição na expressão de gene endógeno e/ou níveis de polipeptídeo e/ou atividade de polipeptídeo com relação às plantas controle. A redução ou eliminação substancial em ordem crescente de preferência está pelo menos 10 %, 20 %, 30 %, 40 % ou 50 %, 60 %, 70 %, 80 %, 85 %, 90 %, ou 95 %, 96 %, 97 %, 98 %, 99 % ou mais reduzida, comparada àquela de plantas controle. Os métodos para diminuir a expressão são conhecidos na técnica e os versados na técnica serão facilmente capazes de adaptar-se a métodos conhecidos para silenciar, de maneira a atingir a redução de expressão de um gene endógeno em uma planta completa, ou em partes da mesma, por meio do uso de um promotor apropriado, por exemplo.

Para a redução ou eliminação substancial de expressão um gene endógeno em uma planta, um tamanho suficiente de nucleotídeos substancialmente contíguos de uma sequência de ácidos nucleicos é exigido. A fim de realizar o silenciamento de gene, este pode ser tão pequeno quanto 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10 ou menos nucleotídeos, alternativamente, este pode apresentar o gene inteiro (incluindo a UTR 5' e/ou 3', tanto em parte quanto por completo). O intervalo dos nucleotídeos substancialmente contíguos pode ser derivado do ácido nucléico que codifica a proteína de interesse (gene alvo), ou de qualquer ácido nucléico capaz de

codificar um ortólogo, parólogo ou homólogo da proteína de interesse. Preferivelmente, o intervalo de nucleotídeos substancialmente contíguos é capaz de formar ligações de hidrogênio com o gene alvo (tanto na fita sentido quanto anti-sentido), mais preferivelmente, o intervalo de nucleotídeos substancialmente contíguos apresenta, em uma ordem crescente de preferência, 50 %, 60 %, 70 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 %, 100 % de identidade de sequência com o gene alvo (tanto fita sentido quanto anti-sentido). Uma sequência de ácidos nucleicos que codifica um polipeptídeo (funcional) não é uma exigência para os vários métodos aqui discutidos para a redução ou eliminação substancial de expressão de um gene endógeno.

Exemplos de vários métodos para a redução ou eliminação substancial de expressão em uma planta de um gene endógeno, ou para diminuir níveis e/ou atividade de uma proteína, são conhecidos pelos versados na técnica. Os versados na técnica serão facilmente capazes de se adaptar a métodos conhecidos para silenciar, de maneira a atingir redução de expressão de um gene endógeno em uma planta completa, ou em parte da mesmas, por meio do uso de um promotor apropriado, por exemplo.

Esta redução ou eliminação substancial de expressão pode ser atingida usando ferramentas e técnicas de rotina. Um método preferido para a redução ou eliminação substancial de expressão de gene endógeno é introduzir e expressar em uma planta um construto genético no qual o ácido nucléico (neste caso, um intervalo de nucleotídeos substancialmente contíguos derivados do gene de interesse, ou de qualquer ácido nucléico capaz de codificar um ortólogo, parólogo ou homólogo de qualquer uma da proteína de interesse) é clonado como uma repetição invertida (em parte ou completamente), separado por um espaçador (DNA não codificante).

Em um método preferido como este, a expressão do gene endógeno é reduzida ou substancialmente eliminada por meio de

silenciamento mediado por RNA usando uma repetição invertida de um ácido nucléico ou uma parte deste (neste caso, um intervalo de nucleotídeos substancialmente contíguos derivados do gene de interesse, ou de qualquer ácido nucléico capaz de codificar um ortólogo, parólogo ou homólogo da proteína de interesse), preferivelmente capaz de formar uma estrutura em forma de grampo. A repetição invertida é clonada em um vetor de expressão compreendendo sequências controle. Uma sequência de ácidos nucleicos de DNA não codificantes (um espaçador, por exemplo, um fragmento de região de anexação de matriz (MAR), um intron, um poliligador, etc.) está localizada entre os dois ácidos nucleicos invertidos que formam a repetição invertida. Após a transcrição da repetição invertida, um RNA quimérico com uma estrutura alto complementar é formado (parcial ou completo). Esta estrutura de RNA de fita dupla é referida como o RNA em grampo (RNAhp). O RNAhp é processado pela planta em RNAssi que são incorporados em um complexo de silenciamento induzido por RNA (RISC). O RISC cliva adicionalmente os transcritos de RNAm, reduzindo substancialmente por meio disso o número de transcritos de RNAm a ser traduzido em polipeptídeos. Para detalhes mais 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 depende da introdução e expressão em uma planta de um construto genético no qual o ácido nucléico é clonado como uma repetição invertida, mas qualquer um ou mais de vários métodos bem conhecidos de "silenciamento de gene" podem ser usados para atingir os mesmos efeitos.

Um método como este para a redução de expressão de gene endógeno é o silenciamento mediado por RNA de expressão de gene (infra-regulação). O silenciamento neste caso é ativado em uma planta por uma sequência de RNA de fita dupla (RNAds), que é substancialmente similar ao gene alvo endógeno. Este RNAds é processado adicionalmente pela planta em

cerca de 20 a cerca de 26 nucleotídeos, denominados RNAs de pequena interferência (RNAssi). Os RNAssi são incorporados em um complexo de silenciamento induzido por RNA (RISC) que cliva o RNAm transcrito do gene endógeno alvo, reduzindo substancialmente por meio disso o número de transcritos de RNAm a ser traduzido em um polipeptídeo. Preferivelmente, a sequência de RNA de fita dupla corresponde a um gene alvo.

Um outro exemplo de um método de silenciamento de RNA envolve a introdução de sequências de ácidos nucleicos, ou partes das mesmas, (neste caso, um intervalo de nucleotídeos substancialmente contíguos derivados do gene de interesse, ou de qualquer ácido nucléico capaz de codificar um ortólogo, parólogo ou homólogo da proteína de interesse) em uma orientação sentido em uma planta. "Orientação sentido" refere-se a uma sequência de DNA que é homóloga a um transcrito de RNAm da mesma. Portanto, pode ser introduzida em uma planta menos uma cópia da sequência de ácidos nucleicos. A sequência adicional de ácidos nucleicos reduzirá a expressão do gene endógeno, dando origem a um fenômeno conhecido como co-supressão. A redução de expressão de gene será mais evidenciada se várias cópias adicionais de uma sequência de ácidos nucleicos forem introduzidas na planta, a medida que existe uma correlação positiva entre altos níveis de transcrito e a ativação da co-supressão.

Um outro exemplo de um método de silenciamento de RNA envolve o uso de sequências anti-sentido de ácidos nucleicos. Uma sequência "anti-sentido" de ácidos nucleicos compreende uma sequência de nucleotídeos que é complementar a uma sequência "sentido" de ácidos nucleicos que codifica uma proteína, isto é, complementar à fita codificante de uma molécula de DNAc de fita dupla ou complementar a uma sequência de transcrito de RNAm. A sequência anti-sentido de ácidos nucleicos é preferivelmente complementar ao gene endógeno a ser silenciado. A complementaridade pode estar localizada na "região codificante" e/ou na

"região não codificante" de um gene. O termo "região codificante" refere-se a uma região da sequência de nucleotídeos compreendendo códons que são traduzidos em resíduos de aminoácido. O termo "região não codificante" refere-se às sequências 5' e 3' que flanqueiam a região codificante que é transcrita, mas não é traduzida em aminoácidos (também referidas como as regiões não traduzidas 5' e 3').

As sequências anti-sentido de ácidos nucleicos podem ser desenhadas de acordo com as regras de pareamento de bases Watson e Crick. A sequência anti-sentido de ácidos nucleicos pode ser complementar à sequência total de ácidos nucleicos (neste caso, um intervalo de nucleotídeos substancialmente contíguos derivados do gene de interesse, ou de qualquer ácido nucléico capaz de codificar um ortólogo, parólogo ou homólogo da proteína de interesse), mas também pode ser um oligonucleotídeo que é anti-sentido a apenas uma parte da sequência de ácidos nucleicos (incluindo a UTR de RNAm 5' e 3'). Por exemplo, a sequência de oligonucleotídeo anti-sentido pode ser complementar à região ao redor do sítio de início de tradução de um transcrito de RNAm que codifica um polipeptídeo. O tamanho de uma sequência de oligonucleotídeo anti-sentido adequada é conhecido na técnica e pode iniciar com cerca de 50, 45, 40, 35, 30, 25, 20, 15 ou 10 nucleotídeos em comprimento ou menos. Uma sequência anti-sentido de ácidos nucleicos de acordo com a invenção pode ser construída usando síntese química e reações de ligação enzimática usando métodos conhecidos na técnica. Por exemplo, uma sequência anti-sentido de ácidos nucleicos (por exemplo, uma sequência de oligonucleotídeo anti-sentido) pode ser sintetizada quimicamente usando nucleotídeos de ocorrência natural ou nucleotídeos modificados de várias maneiras desenhados para aumentar a estabilidade biológica das moléculas ou para aumentar a estabilidade física do duplex formado entre as sequências anti-sentido e sentido de ácidos nucleicos, por exemplo, nucleotídeos substituídos por derivados de fosforotioato e acridina podem ser usados.

Exemplos de nucleotídeos modificados que podem ser usados para gerar as sequências anti-sentido de ácidos nucleicos são bem conhecidos na técnica. Modificações de nucleotídeo conhecidas incluem metilação, ciclização e “cobertura” e substituição de um ou mais dos nucleotídeos de ocorrência natural com um análogo tal como inosina. Outras modificações de nucleotídeos são bem conhecidas na técnica.

A sequência anti-sentido de ácidos nucleicos pode ser produzida biologicamente usando um vetor de expressão, no qual uma sequência de ácidos nucleicos foi subclonada em uma orientação anti-sentido (isto é, RNA transcrito a partir do ácido nucléico inserido apresentarão uma orientação anti-sentido em um ácido nucléico alvo de interesse). Preferivelmente, a produção de sequências anti-sentido de ácidos nucleicos em plantas ocorre por meio de um construto de ácido nucléico estavelmente integrado compreendendo um promotor, um oligonucleotídeo anti-sentido operavelmente ligado e um finalizador.

As moléculas de ácido nucléico usadas para silenciar nos métodos da invenção (quer introduzidas em uma planta ou geradas *in situ*) hibridizam ou se ligam a transcritos de RNAm e/ou DNA genômico que codifica um polipeptídeo para inibir por meio disso a expressão da proteína, por exemplo, inibindo a transcrição e/ou tradução. A hibridização pode ser por complementaridade de nucleotídeo convencional para formar um duplex estável ou, por exemplo, no caso de uma sequência anti-sentido de ácidos nucleicos que se liga a duplexes de DNA, por meio de interações específicas no encaixe principal da dupla hélice. As sequências anti-sentido de ácidos nucleicos podem ser introduzidas em uma planta por transformação ou injeção direta em um sítio de tecido específico. Alternativamente, as sequências anti-sentido de ácidos nucleicos podem ser modificadas para alvejar células selecionadas e então administradas sistemicamente. Por exemplo, para administração sistêmica, as sequências anti-sentido de ácidos

nucleicos podem ser modificadas de maneira tal que elas se ligam especificamente aos receptores ou antígenos expressos em uma superfície de célula selecionada, por exemplo, ligando a sequência anti-sentido de ácidos nucleicos aos peptídeos ou anticorpos que se ligam aos receptores ou antígenos de superfície celular. As sequências anti-sentido de ácidos nucleicos também podem ser distribuídas nas células usando os vetores aqui descritos.

De acordo com um aspecto adicional, a sequência anti-sentido de ácidos nucleicos é uma sequência a-anomérica de ácidos nucleicos. Uma sequência a-anomérica de ácidos nucleicos forma híbridos de fita dupla específicos com RNA de complementaridade no qual, ao contrário das unidades b usuais, as fitas correm em paralelo umas as outras (Gaultier et al. (1987) Nucl Ac Res 15: 6625-6641). A sequência anti-sentido de ácidos nucleicos também pode compreender um 2'-o-metilribonucleotídeo (Inoue et al. (1987) Nucl Ac Res 15, 6131 -6148) ou um análogo quimérico de RNA-DNA (Inoue et al. (1987) FEBS Lett. 215, 327-330).

A redução ou eliminação substancial de expressão de gene endógeno também pode ser realizada usando ribozimas. As ribozimas são moléculas catalíticas de RNA com atividade de ribonuclease que são capazes de clivar uma sequência de fita simples de ácidos nucleicos, tal como um RNAm, no qual elas apresentam uma região de complementaridade. 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 RNAm que codificam um polipeptídeo, reduzindo substancialmente por meio disso o número de transcritos de RNAm a ser traduzido em um polipeptídeo. Uma ribozima com especificidade para uma sequência de ácidos nucleicos pode ser desenhada (ver, por exemplo: Cech et al. patente U.S. 4.987.071; e Cech et al. patente U.S. 5.116.742). Alternativamente, os transcritos de RNAm que correspondem a uma sequência de ácidos nucleicos podem ser usados para selecionar um RNA

catalítico com uma atividade específica de ribonuclease a partir de um agrupamento de moléculas de RNA (Bartel e Szostak (1993) Science 261, 1411-1418). O uso de ribozimas para 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 97/381 16).

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

O silenciamento de gene também pode ocorrer se existir uma mutação em um gene endógeno e/ou uma mutação em um gene isolado/ácido nucléico introduzido subsequentemente em uma planta. A redução ou eliminação substancial pode ser causada por um polipeptídeo não funcional. Por exemplo, o polipeptídeo pode ser ligar a várias proteínas de interação; uma ou mais mutação(s) e/ou truncação(s) pode, portanto, fornecer um polipeptídeo que ainda é capaz de se ligar às proteínas de interação (tais como proteínas receptoras) mas que não exhibe sua função normal (tal como ligante de sinalização).

Uma abordagem adicional do silenciamento de gene é o alvejamento de sequências de ácidos nucleicos complementares à região regulatória do gene (por exemplo, o promotor e/ou melhoradores) para formar estruturas de três hélices que previnem 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 inibir sua função em planta, ou interferência na via de sinalização na qual um polipeptídeo está envolvido, serão

conhecidos pelos versados na técnica. Em particular, pode-se contemplar que moléculas sintetizadas em laboratório podem ser usadas para inibir a função biológica de um polipeptídeo alvo, ou para interferir com a via de sinalização na qual o polipeptídeo alvo está envolvido.

5 Alternativamente, um programa de seleção pode ser estabelecido para identificar em uma população de planta variantes naturais de um gene, cujos variantes codificam polipeptídeos com menor atividade. Tais variantes naturais também podem ser usados, por exemplo, para realizar recombinação homóloga.

10 RNAsmicro artificiais e/ou naturais (RNAsmi) podem ser usados para neutralizar a expressão de gene e/ou tradução de RNAm. RNAsmi endógenos são RNAs pequenos de fita simples com tipicamente 19-24 nucleotídeos de comprimento. Eles funcionam principalmente para regular a expressão de gene e/ ou tradução de RNAm. A maioria dos RNAsmicro de
15 planta (RNAsmi) apresentam complementaridade perfeita ou quase perfeita com suas sequências alvo. Entretanto, existem alvos naturais com até cinco combinações. Eles são processados a partir de RNAs mais longos não codificantes com estruturas que se dobram sobre elas mesmas características por RNases específicas de dita dupla da família Dicer. Mediante
20 processamento, eles são incorporados no complexo de silenciamento induzido por RNA (RISC) se ligando a seu principal componente, uma proteína argonauta. RNAsmi funcionam como os componentes de especificidade de RISC, uma vez que eles pareiam bases para alvejar ácidos nucleicos, principalmente RNAs, no citoplasma. Eventos subsequentes regulatórios
25 incluem a clivagem do RNAm alvo e a destruição e/ou inibição translacional. Os efeitos da superexpressão de RNAmi são, assim, frequentemente refletidos em menores níveis de RNAm de gene alvos.

 RNAsmicro artificiais (RNAsmia), que apresentam tipicamente 21 nucleotídeos de comprimento, podem ser geneticamente

modificados especificamente para regular negativamente a expressão de gene de genes simples ou múltiplos de interesse. Os determinantes de seleção alvo de RNAmicro de planta são bem conhecidos na técnica. Parâmetros empíricos para reconhecimento alvo foram definidos e podem ser usados para auxiliar no desenho de RNAsmias específicos, (Schwab et al., Dev. Cell 8, 517-527, 2005). Ferramentas convenientes para desenhar e gerar RNAsmia e seus precursores são também disponíveis para o público (Schwab et al, Plant Cell 18, 1121-1133, 2006).

Para desempenho ideal, as técnicas de silenciamento de gene usadas para reduzir a expressão em uma planta de um gene endógeno exigem o uso de sequências de ácidos nucleicos de plantas monocotiledôneas para a transformação de plantas monocotiledôneas, e de plantas dicotiledôneas para a transformação de plantas dicotiledôneas. Preferivelmente, uma sequência de ácidos nucleicos de qualquer dada espécie de planta é introduzida em uma da mesma espécie. Por exemplo, uma sequência de ácidos nucleicos de arroz é transformada em uma planta de arroz. Entretanto, não é uma exigência absoluta que a sequência de ácidos nucleicos a ser introduzida se origine da mesma espécie de planta, como a planta na qual ela será introduzida. É suficiente que exista homologia substancial entre o gene endógeno alvo e o ácido nucléico a ser introduzido.

São descritos anteriormente exemplos de vários métodos para a redução ou eliminação substancial de expressão em uma planta de um gene endógeno. Os versados na técnica serão facilmente capazes de se adaptar aos métodos anteriormente mencionados para silenciar, de maneira a atingir a redução de expressão de um gene endógeno em uma planta completa, ou em parte da mesmas, por meio do uso de um promotor apropriado, por exemplo.

Marcador selecionável (gene)/Gene repórter

"Marcador selecionável", "gene de marcador selecionável" ou "gene repórter" incluem qualquer gene que confere um fenótipo em uma

célula na qual é expresso para facilitar a identificação e/ou seleção de células que são transfectadas ou transformadas com um construto de ácido nucléico da invenção. Estes genes marcadores possibilitam a identificação de uma transferência satisfatória das moléculas de ácido nucléico por meio de uma

5 série de princípios diferentes. Os marcadores adequados podem ser selecionados a partir de marcadores que conferem resistência à antibiótico ou herbicida, que introduzem uma nova característica metabólica ou que permitem seleção visual. Exemplos de genes de marcadores selecionáveis incluem genes que conferem resistência a antibióticos (tal como nptII, que

10 fosforila neomicina e canamicina, ou hpt, que fosforila higromicina, ou genes que conferem resistência a, por exemplo, bleomicina, estreptomicina, tetraciclina, cloranfenicol, ampicilina, gentamicina, geneticina (G418), espectinomomicina ou blasticidina), a herbicidas (por exemplo, barra que fornece resistência a Basta[®]; aroA ou gox que fornece resistência contra glifosato, ou

15 os genes que conferem resistência a, por exemplo, imidazolinona, fosfinotricina ou sulfoniluréia), ou genes que fornecem uma característica metabólica (tal como manA que permite que as plantas usem manose como única fonte de carbono ou xilose isomerase para a utilização de xilose, ou marcadores anti-nutritivos tal como a resistência a 2-desoxiglicose). A

20 expressão de genes marcadores visuais resulta na formação de cor (por exemplo, β -glucuronidase, GUS ou β -galactosidase com seus substratos coloridos, por exemplo, X-Gal), luminescência (tal como o sistema luciferina/luciferase) ou fluorescência (proteína verde fluorescente, GFP e derivados destes). Esta lista representa apenas um pequeno número de

25 possíveis marcadores. Os versados na técnica são familiares com tais marcadores. Diferentes marcadores são preferidos, dependendo do organismo e do método de seleção.

Sabe-se que, mediante integração estável ou transiente de ácidos nucleicos em células de planta, apenas uma minoria das células obtém

o DNA estrangeiro e, se desejado, integra-o em seu genoma, dependendo do vetor de expressão usado e da técnica de transfecção usada. Para identificar e selecionar estes integrantes, um gene que codifica um marcador selecionável (tais como aqueles descritos anteriormente) é em geral introduzido nas células hospedeiras junto com o gene de interesse. Estes marcadores podem, por exemplo, ser usados em mutantes nos quais estes genes não são funcionais, por exemplo, pela deleção por métodos convencionais. Além disso, as moléculas de ácido nucléico que codificam um marcador selecionável podem ser introduzidas em uma célula hospedeira no mesmo vetor que compreende a sequência 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 nucléico introduzido podem ser identificadas, por exemplo, por seleção (por exemplo, células que integraram o marcador selecionável sobrevivem, ao passo que as outras células morrem). Os genes marcadores podem ser removidos ou excisados das célula transgênica, uma vez que eles já não são necessários. As técnicas para a remoção de gene marcador são conhecidas na tecnologia, as técnicas usadas são descritas anteriormente na seção de definições.

Uma vez que os genes marcadores, particularmente genes para a resistência a antibióticos e herbicidas, não são mais exigidos ou são indesejados na célula transgênica hospedeira, uma vez que os ácidos nucleicos foram introduzidos satisfatoriamente, o processo de acordo com a invenção para introduzir os ácidos nucleicos vantajosamente emprega técnicas que possibilitam a remoção ou excisão destes genes marcadores. Um método como este é o que é conhecido como co-transformação. O método de co-transformação emprega dois vetores simultaneamente para a transformação, um vetor que carrega o ácido nucléico de acordo com a invenção e um segundo que carrega o(s) gene(s) marcador(s). Uma grande proporção de transformantes recebe ou, no case de plantas, compreende (até 40 % ou mais

dos transformantes), ambos os vetores. No caso de transformação com agrobactérias, os transformantes recebem geral apenas uma parte do vetor, isto é, a sequência flanqueada pelo DNA-T, que em geral representa o cassete de expressão. Os genes marcadores podem ser subsequentemente removidos da planta transformada realizando cruzamentos. Em um outro método, os genes marcadores integrados em um transposon são usados para a transformação junto com o ácido nucléico desejado (conhecido como a tecnologia Ac/Ds tecnologia). Os transformantes podem ser cruzados com uma fonte de transposase ou os transformantes são transformados com um construto de ácido nucléico que confere a expressão de uma transposase, de maneira transiente ou estável. Em alguns casos (aproximadamente 10 %), o transposon pula do genoma da célula hospedeira, uma vez que a transformação ocorreu satisfatoriamente e é perdido. Em inúmeros casos adicionais, o transposon salta um local diferente. Nestes casos, o gene marcador tem que ser eliminado realizando cruzamentos. Na microbiologia, as técnicas que foram desenvolvidas tornam possível, ou facilitam, a deleção de tais eventos. Um método vantajoso adicional depende do que é conhecido como sistemas de recombinação; cuja vantagem é que a eliminação por 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 sequências localizadas entre as sequências loxP. Se o gene marcador for integrado entre as sequências loxP, ele é removido uma vez que a transformação ocorreu satisfatoriamente, por expressão da recombinase. Os sistemas de recombinação adicionais 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 sítio específica no genoma da planta com as sequências de ácidos nucleicos de acordo com a invenção é possível. Naturalmente, estes métodos também podem ser aplicados a microrganismos tais como levedura, fungos ou

bactérias.

Transgênico/Transgene/Recombinante

Com propósitos da invenção, "transgênico", "transgene" ou "recombinante" significa com relação a, por exemplo, uma sequência de ácidos nucleicos, um cassete de expressão, construto de gene ou um vetor compreendendo a sequência de ácidos nucleicos ou um organismo transformado com as sequências de ácidos nucleicos, cassetes de expressão ou vetores de acordo com a invenção, todas aquelas construções acarretam em métodos recombinantes nos quais cada um:

10 (a) a sequências de ácidos nucleicos que codificam proteínas usadas nos métodos da invenção, ou

(b) sequência(s) genética(s) controle que é(são) operavelmente ligada(s) com a sequência de ácidos nucleicos de acordo com a invenção, por exemplo um promotor, ou

15 (c) a) e b) não estão localizadas em seu ambiente genético natural ou foram modificadas por métodos recombinantes, que são possíveis para a modificação obter a forma de, por exemplo, uma substituição, adição, deleção, inversão ou inserção de um ou mais resíduos de nucleotídeo. Entende-se que o ambiente genético natural significa o locus genômico ou cromosômico natural na planta original ou a presença em uma biblioteca genômica. No caso de uma biblioteca genômica, o ambiente genético natural da sequência de ácidos nucleicos é preferivelmente mantido, pelo menos em parte. O ambiente flanqueia a sequência de ácidos nucleicos pelo menos em um lado e apresenta um tamanho de sequência de pelo menos 50 bp, 20 preferivelmente pelo menos 500 bp, especialmente preferível pelo menos 1.000 bp, acima de tudo preferivelmente pelo menos 5.000 bp. Um cassete de expressão de ocorrência natural, por exemplo, a combinação de ocorrência natural do promotor natural das sequências de ácidos nucleicos com a sequência correspondente de ácidos nucleicos que codifica um polipeptídeo

usado nos métodos da presente invenção, da maneira definida anteriormente, se torna um cassete de expressão transgênico quando este cassete de expressão é modificado por métodos não naturais, sintéticos ("artificiais") tal como, por exemplo, tratamento mutagênico. Os métodos adequados são descritos, por exemplo, em US 5.565,350 ou WO 00/15815.

Entende-se assim que uma planta transgênica com propósitos da invenção significa, da maneira anterior, que os ácidos nucleicos usados no método da invenção não estão em seu locus natural no genoma da dita planta, sendo possível que os ácidos nucleicos sejam expressos homóloga ou heterologamente. Entretanto, da maneira mencionada, transgênico também significa que, embora os ácidos nucleicos de acordo com a invenção ou usados no método inventivo estejam em sua posição natural no genoma de uma planta, a sequência foi modificada com relação à sequência natural, e/ou que as sequências regulatórias das sequências naturais foram modificadas. Entende-se preferivelmente que transgênico significa a expressão dos ácidos nucleicos de acordo com a invenção em um locus não natural no genoma, isto é, expressão homóloga ou, preferivelmente, heteróloga dos ácidos nucleicos ocorreu. As plantas transgênicas preferidas são aqui mencionadas.

Transformação

O termo "introdução" ou "transformação", da maneira aqui referida, inclui a transferência de um polinucleotídeo exógeno em uma célula hospedeira, independente do método usado para a transferência. O tecido de planta capaz de propagação clonal subsequente, que por organogênese ou embriogênese, pode ser transformado com um construto genético da presente invenção e em uma planta completa regenerada dele derivado. O tecido particular escolhido variará dependendo dos sistemas de propagação clonal disponíveis para, e melhores adequados a, as espécies particulares que são transformadas. Os alvos de tecido exemplar incluem discos de folha, pólen, embriões, cotilédones, hipocótilos, megagametófitos, tecido do calo, tecido

meristemático existente (por exemplo, meristema apical, brotos auxiliares e meristemas de raiz), e tecido de meristema induzido (por exemplo, meristema de cotilédone e meristema de hipocótilo). O polinucleotídeo pode ser introduzido transientemente ou estavelmente em uma célula hospedeira e pode ser mantido não integrado, por exemplo, como um plasmídeo. Alternativamente, ele pode ser integrado no genoma hospedeiro. As células de planta transformadas resultantes podem então ser usadas para regenerar uma planta transformada de uma maneira conhecida pelos versados na técnica.

A transferência de genes estrangeiros no genoma de uma planta é denominada transformação. A transformação de espécie de planta é agora uma técnica bastante rotineira. Vantajosamente, qualquer de vários 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 de planta ou células de planta podem ser utilizados para transformação transiente ou estável. Os métodos de transformação incluem o uso de lipossomos, eletroporação, produtos químicos que aumentam a absorção de DNA livre, injeção do DNA diretamente na planta, arma de bombardeamento de partícula, transformação usando vírus ou pólen e microinjeção. Os métodos podem ser selecionados a partir do método de cálcio/polietileno glicol para protoplastos (Krens, F.A. et al., (1982) Nature 296, 72-74; Negrutiu I et al. (1987) Plant Mol Biol 8: 363-373); eletroporação de protoplastos (Shillito R.D. et al. (1985) Bio/Technol 3, 1099-1102); microinjeção no material da planta (Crossway A et al., (1986) Mol. Gen Genet 202: 179- 185); bombardeamento de partícula revestida de DNA ou RNA (Klein TM et al., (1987) Nature 327: 70) infecção com (não integrativa) vírus e similares. As plantas transgênicas, incluindo as plantas transgênicas de cultivo, são preferivelmente produzidas por meio da transformação ativada por *Agrobacterium*. Um método vantajoso de transformação é a transformação em planta. Com esta finalidade, é possível, por exemplo,

permitir que as agrobactérias possam agir nas sementes da planta ou inocular o meristema da planta com agrobactérias. Foi fornecido particularmente expediente de acordo com a invenção para permitir uma suspensão de agrobactérias transformadas para agir na planta intacta ou pelo menos no

5 primórdio da flor. A planta cresce subsequentemente 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 quaisquer dos seguintes: pedido de patente européia E.P.

10 1198985 A1, Aldemita e Hodges (Plant 199: 612-617, 1996); Chan et al. (Plant Mol Biol 22 (3): 491-506, 1993), Hiei et al. (Plant J 6 (2): 271-282, 1994), cujas revelações são aqui incorporadas pela referência a medida que são completamente apresentadas. No caso da transformação do milho, o método preferido é descrito tato em Ishida et al. (Nat. Biotechnol 14(6): 745-

15 50, 1996) quanto em Frame et al. (Plant Physiol 129(1): 13-22, 2002), cujas revelações são aqui incorporadas pela referência a medida que são completamente apresentadas. Os ditos métodos são descritos adicionalmente a título de exemplo em B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S. D. Kung e R.

20 Wu, Academic Press (1993) 128-143 e em Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205-225). Os ácidos nucleicos ou o construto a ser expresso 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

25 transformadas por um vetor como este podem estão ser usadas em maneiras conhecidas para a transformação de plantas, tais como as plantas usadas como um modelo, como *Arabidopsis* (*Arabidopsis thaliana* está no escopo da presente invenção e não é considerada como uma planta de cultivo), ou plantas de cultivo tais como, a título de exemplo, plantas de tabaco, por

exemplo, imergindo folhas machucadas ou folhas rasgadas em uma solução de agrobactérias e então cultivando-as em meios adequados. 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 conhecida *inter alia* a partir de F. F. White, Vectors for Gene Transfer in Higher Plants; in 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 são então regeneradas em plantas intactas, também é possível transformar as células de meristemas de planta e, em particular, aquelas células que se desenvolvem gametas. Neste caso, os gametas transformados seguem o desenvolvimento natural da planta, dando origem à plantas transgênicas. Assim, por exemplo, as sementes de *Arabidopsis* são tratadas com agrobactérias e as sementes são obtidas a partir de plantas em desenvolvimento, a partir das quais uma certa proporção é transformada e, assim, são transgênicas [Feldman, KA e Marks MD (1987). MoI Gen Genet 208:274-289; Feldmann K (1992). Em: C Koncz, N-H Chua e J Shell, eds, Methods in *Arabidopsis* Research. Word Scientific, Singapura, pp. 274-289]. Os métodos alternativos baseiam-se na remoção repetida das inflorescências e da incubação do sítio de excisão no centro da roseta com agrobactérias transformadas, por meio do qual as sementes transformadas podem ser provavelmente obtidas em um ponto posterior do tempo (Chang (1994). Plant J. 5: 551-558; Katavic (1994). MoI Gen Genet, 245: 363-370). Entretanto, um método especialmente efetivo é o método de infiltração a vácuo com suas modificações tal como o método "imersão floral". No caso de infiltração a vácuo de *Arabidopsis*, as plantas intactas em pressão reduzida são tratadas com uma suspensão de agrobactérias [Bechthold, N (1993). C R Acad Sci Paris Life Sci, 316: 1 194-1 199], embora no caso do método de "imersão floral" o tecido floral em desenvolvimento seja incubado rapidamente com uma suspensão de

agrobactérias tratadas com agente tensoativo [Clough, SJ e Bent AF (1998) *The plant J.* 16, 735-743]. Uma certa proporção de sementes transgênicas é coletada em ambos os casos, e estas sementes podem ser distinguidas de sementes não transgênicas crescendo nas condições seletivas descritas anteriormente. Além do mais, a transformação estável de plastídeos é vantajosa em decorrência de os plastídeos serem maternalmente inerentes na maioria das culturas, reduzindo ou eliminado o risco do transgene fluir por meio do pólen. A transformação do genoma do cloroplasto é atingida em geral por um processo que foi esquematicamente exibido em Klaus et al., 2004 [Nature Biotechnology 22 (2), 225-229]. Resumidamente, as sequências a serem transformadas são clonadas junto com um gene de marcador selecionável entre sequências homólogas que flanqueiam no genoma do cloroplasto. Estas sequências flaqueantes homólogas direcionam a integração sítio específica no plastoma. A transformação plastidial foi descrita por muitas espécies de planta diferentes e uma perspectiva é fornecida em Bock (2001) *Transgenic plastids in basic research and plant biotechnology. J Mol Biol.* 2001 Sep 21; 312 (3):425-38 ou Maliga, P (2003) *Progress towards commercialization of plastid transformation technology. Trends Biotechnol.* 21, 20-28. O progresso biotecnológico adicional foi recentemente relatado na forma de transformantes de plastídeo sem marcador, que pode ser produzido por um gene marcador co-integrado transiente (Klaus et al., 2004, *Nature Biotechnology* 22(2), 225- 229).

Marcação da ativação do DNA-T

A marcação da ativação do DNA-T (Hayashi et al. *Science* (1992) 1350-1353), envolve inserção de DNA-T, contendo em geral um promotor (também pode ser um melhorador de tradução ou um intron), na região genômica do gene de interesse ou 10 kb à montante ou à jusante da região codificante de um gene em uma configuração, de maneira tal que o promotor direcione a expressão do gene alvejado. Tipicamente, a regulação de

expressão do gene alvejado por seu promotor natural é interrompida e o gene perde o controle do promotor recém introduzido. O promotor é tipicamente embebido em um DNA-T. Este DNA-T é inserido aleatoriamente no genoma da planta, por exemplo, por meio de infecção por *Agrobacterium* infecção e
 5 leva a expressão modificada de genes próximos ao DNA-T inserido. As plantas transgênicas resultantes mostram fenótipos dominantes em virtude da expressão modificada de genes próximos ao promotor introduzido.

TILLING

O termo "TILLING" é uma abreviação de "Lesões locias induzidas alvejadas em genomas" e refere-se a uma tecnologia de mutagênese usada para gerar e/ou identificar ácidos nucleicos que codificam proteínas com expressão e/ou atividade modificadas. TILLING também permite a seleção de plantas que carregam tais variantes mutantes. Estes variantes mutantes podem exibir expressão modificada, tanto no tamanho ou local
 10 quanto na cronometragem (se as mutações afetarem o promotor, por exemplo). Estes variantes mutantes podem exibir maior atividade do que aquela exibida pelo gene na sua forma natural. TILLING combina mutagênese de alta densidade com métodos de seleção de alto rendimento. As etapas seguidas tipicamente em TILLING são: (a) mutagênese EMS (Redei
 15 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,
 20 eds, *Methods on Molecular Biology*, Vol. 82. Human Press, Totowa, NJ, pp 91 -104); (b) preparação de DNA e agrupamento de indivíduos; (c) amplificação por PCR de uma região de interesse; (d) desnaturação e anelamento para permitir a formação de heteroduplexes; (e) DHPLC, quando a presença de um heteroduplex em um agrupamento é detectada como um

pico extra no cromatograma; (f) identificação do mutante individual; e (g) sequenciamento do produto de PCR mutante. Os métodos para TILLING são bem conhecidos na técnica (McCalium et al., (2000) Nat Biotechnol 18: 455-457; revisado por Stemple (2004) Nat Rev Genet 5(2): 145-50).

5 Recombinação homóloga

A recombinação homóloga permite introdução em um genoma de um ácido nucléico selecionado em uma posição selecionada definida. A recombinação homóloga é uma tecnologia padrão usada rotineiramente em ciências biológicas para organismos inferiores tal como levedura ou o musgo
10 *Physcomitrella*. Os métodos para realizar a recombinação homóloga em plantas foram descritos não apenas para plantas modelo (Offringa et al. (1990) EMBO J 9(10): 3077-84), mas também para plantas de cultivo, por exemplo, arroz (Terada et al. (2002) Nat Biotech 20(10): 1030-4; lida e Terada (2004) Curr Opin Biotech 15(2): 132-8), e existem abordagens de que são em geral
15 aplicáveis com relação ao organismo alvo (Miller et al, Nature Biotechnol. 25, 778-785, 2007).

Rendimento

O termo "rendimento" significa em geral uma produção mensurável de valor econômico, tipicamente relacionado a um cultivo
20 específico, a uma área e a um período de tempo. Partes da planta individual contribuem diretamente para o rendimento com base nos seus números, tamanho e/ou peso, ou o rendimento atual é o rendimento por metro quadrado para um cultivo e ano, que é determinado dividindo a produção total (inclui tanto produção coletada quanto avaliada) por metros quadrados plantados. O
25 termo "rendimento" de uma planta pode relacionar-se a biomassa vegetativa (biomassa da raiz e/ou ramo), aos órgãos reprodutivos e/ou às propágulas (tais como sementes) da mesma planta.

Vigor inicial

"Vigor inicial" refere-se a crescimento ativo de maneira

saudável bem balanceado, especialmente durante os estágios iniciais de crescimento da planta, e pode resultar da maior aptidão da planta em virtude de, por exemplo, as plantas serem melhor adaptadas ao seu ambiente (isto é, otimizar o uso de recursos de energia e repartição entre ramo e raiz). As

5 plantas com vigor inicial também mostram maior sobrevivência da muda e um melhor estabelecimento do cultivo, que resulta frequentemente em campos altamente uniformes (com o cultivo crescendo de maneira uniforme, isto é, com a maioria de plantas atingindo os vários estágios de desenvolvimento substancialmente ao mesmo tempo), e frequentemente

10 melhor e maior rendimento. Portanto, o vigor inicial pode ser determinado medindo várias fatores, tais como peso de mil grãos, germinação percentual, emergência percentual, crescimento da muda, altura da muda, comprimento da raiz, biomassa da raiz e ramo e muito mais.

Aumento/Ampliação/Melhoria

15 Os termos "aumento", "ampliação" ou "melhoria" são indiferentes e podem significar, no sentido do pedido, pelo menos 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 % ou 10 %, preferivelmente pelo menos 15 % ou 20 %, mais preferivelmente 25 %, 30 %, 35 % ou 40 % mais rendimento e/ou crescimento em comparação com as plantas controle da maneira aqui

20 definida.

Rendimento da semente

O maior rendimento da semente pode ser manifestado como um ou mais dos seguintes: a) um aumento na biomassa da semente (peso total da semente) que pode ser em uma base individual de semente e/ou por planta

25 e/ou por metro quadrado; b) maior número de flores por planta; c) maior número de sementes (preenchidas); d) maior taxa de preenchimento da semente (que é expressa como a razão entre o número de sementes preenchidas dividido pelo número total de sementes); e) maior índice da colheita, que é expresso como uma razão do rendimento de Partes cultiváveis,

tais como sementes, dividido pela biomassa total; e f) aumento no peso de mil grãos (TKW), e g) maior número de panículas primárias, que é extrapolado do número de sementes preenchidas contadas e seu peso total. Um maior TKW pode resultar de um maior tamanho da semente e/ou peso da semente, e também pode resultar de um aumento no tamanho do embrião e/ou endosperma.

Um aumento no rendimento da semente também pode ser manifestado como um aumento no tamanho da semente e/ou volume da semente. Além disso, um aumento no rendimento da semente também pode se manifestar como um aumento na área da semente e/ou tamanho da semente e/ou largura da semente e/ou perímetro da semente. O maior rendimento da semente também pode resultar em arquitetura modificada, ou pode ocorrer em decorrência da arquitetura modificada.

Índice de vegetação

O "índice de vegetação", da maneira aqui usada, é calculado a partir de imagens digitais das plantas. Para cada pixel pertencente ao objeto da planta na imagem, a razão do valor verde versus o valor vermelho (no modelo RGB para codificar a cor) é calculada. O índice de vegetação é expresso como o percentual de pixels para o qual a razão verde para vermelho excede um dado limiar. Em condições normais de crescimento, em condições de crescimento em estresse salino, e em condições de crescimento com disponibilidade reduzida de nutrientes, o índice de vegetação de plantas é medido no último imageamento antes da floração. Ao contrário, em condições de crescimento de estresse por secura, o índice de vegetação de plantas é medido no primeiro imageamento após a secura.

Planta

O termo "planta", da maneira aqui usada, inclui as plantas completas, ancestrais e progênie das plantas e partes da planta, incluindo sementes, ramos, caules, folhas, raízes (incluindo tubérculos), flores e tecidos

e órgãos, em que cada um dos anteriormente mencionados compreendem o gene/ácido nucléico de interesse. O termo "planta" também inclui células de planta, culturas em suspensão, tecido do calo, embriões, regiões meristemáticas, gametófitos, esporófitos, pólen e microesporos, em que

5 novamente cada um dos anteriormente mencionados compreende o gene/ácido nucléico de interesse.

As plantas que são particularmente usadas nos métodos da invenção incluem todas as plantas que pertencem à superfamília *Viridiplantae*, em particular, plantas monocotiledôneas e dicotiledôneas

10 incluindo plantas de forragem ou legumes forrageiros, plantas ornamentais, culturas de alimento, árvores ou arbustos selecionados da lista compreendendo *Acer* spp., *Actinidia* spp., *Abelmoschus* spp., *Agave sisalana*, *Agropyron* spp., *Agrostis stolonifera*, *Allium* spp., *Amaranthus* spp., *Ammofila arenaria*, *Ananas comosus*, *Annona* spp., *Apium graveolens*, *Arachis* spp,

15 *Artocarpus* spp., *Asparagus officinalis*, *Avena* spp. (por exemplo, *Avena sativa*, *Avena fatua*, *Avena byzantina*, *Avena fatua* var. *sativa*, *Avena hybrida*), *Averrhoa carambola*, *Bambusa* sp., *Benincasa hispida*, *Bertholletia excelsea*, *Beta vulgaris*, *Brassica* spp. (por exemplo, *Brassica napus*, *Brassica rapa* ssp. [canola, óleo de semente de colza, nabita]), *Cadaba farinosa*,

20 *Camellia sinensis*, *Canna indica*, *Cannabis sativa*, *Capsicum* spp., *Carex elata*, *Carica papaya*, *Carissa macrocarpa*, *Carya* spp., *Carthamus tinctorius*, *Castanea* spp., *Ceiba pentandra*, *Cichorium endivia*, *Cinnamomum* spp., *Citrullus lanatus*, *Citrus* spp., *Cocos* spp., *Coffea* spp., *Colocasia esculenta*, *Cola* spp., *Corchorus* sp., *Coriandrum sativum*, *Corylus* spp.,

25 *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*, *Eragrostis tef*, *Erianthus* sp., *Eriobotrya japonica*, *Eucalyptus* sp., *Eugenia uniflora*, *Fagopyrum* spp., *Fagus* spp.,

Festuca arundinacea, *Ficus carica*, *Fortunella* spp., *Fragaria* spp., *Ginkgo biloba*, *Glycine* spp. (por exemplo, *Glycine max*, *Soja hispida* ou *Soja max*), *Gossypium hirsutum*, *Helianthus* spp. (por exemplo, *Helianthus annuus*), *Hemerocallis fulva*, *Hibiscus* spp., *Hordeum* spp. (por exemplo, *Hordeum vulgare*), *Ipomoea batatas*, *Juglans* spp., *Lactuca sativa*, *Lathyrus* spp., *Lens culinaris*, *Linum usitatissimum*, *Litchi chinensis*, *Lotus* spp., *Luffa acutangula*, *Lupinus* spp., *Luzula sylvatica*, *Lycopersicon* spp. (por exemplo, *Lycopersicon esculentum*, *Lycopersicon lycopersicum*, *Lycopersicon pyriforme*), *Macrotyloma* spp., *Malus* spp., *Malpighia emarginata*, *Mammea americana*, *Mangifera indica*, *Manihot* spp., *Manilkara zapota*, *Medicago sativa*, *Melilotus* spp., *Mentha* spp., *Miscanthus sinensis*, *Momordica* spp., *Morus nigra*, *Musa* spp., *Nicotiana* spp., *Olea* spp., *Opuntia* spp., *Ornithopus* spp., *Oryza* spp. (por exemplo, *Oryza sativa*, *Oryza latifolia*), *Panicum miliaceum*, *Panicum virgatum*, *Passiflora edulis*, *Pastinaca sativa*, *Pennisetum* sp., *Persea* spp., *Petroseium crispum*, *Phalaris arundinacea*, *Phaseolus* spp., *Phleum pratense*, *Foenix* spp., *Phragmites australis*, *Physalis* spp., *Pinus* spp., *Pistacia vera*, *Pisum* spp., *Poa* spp., *Populus* spp., *Prosopis* spp., *Prunus* spp., *Psidium* spp., *Punica granatum*, *Pyrus communis*, *Quercus* spp., *Raphanus sativus*, *Rheum rhabarbarum*, *Ribes* spp., *Ricinus communis*, *Rubus* spp., *Saccharum* spp., *Salix* sp., *Sambucus* spp., *Secale cereale*, *Sesamum* spp., *Sinapis* sp., *Solanum* spp. (por exemplo, *Solanum tuberosum*, *Solanum integrifolium* ou *Solanum lycopersica*), *Sorghum bicolor*, *Spinacia* spp., *Syzygium* spp., *Tagetes* spp., *Tamarindus indica*, *Theobroma cacao*, *Trifolium* spp., *Tripsacum dactyloides*, *Triticale* sp., *Triticosecale rimpaii*, *Triticum* spp. (por exemplo, *Triticum aestivum*, *Triticum durum*, *Triticum turgidum*, *Triticum hybernum*, *Triticum macha*, *Triticum sativum*, *Triticum monococcum* 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.

DESCRIÇÃO DETALHADA DA INVENÇÃO

De maneira surpreendente, observa-se agora que a modulação da expressão em uma planta com um ácido nucléico que codifica um polipeptídeo do tipo alfin fornece plantas com melhor tolerância ao estresse abiótico com relação às plantas controle. De acordo com uma primeira modalidade, a presente invenção fornece um método para melhorar a tolerância a vários estresses abióticos em plantas com relação às plantas controle, compreendendo modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo do tipo alfin e opcionalmente selecionar plantas com melhor tolerância ao estresse abiótico.

Além disso, observa-se surpreendente agora que modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo YRP fornece plantas com melhor tolerância ao estresse abiótico com relação às plantas controle. De acordo com uma primeira modalidade, a presente invenção fornece um método para melhorar a tolerância a vários estresses abióticos em plantas com relação às plantas controle, compreendendo modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo YRP e opcionalmente selecionar plantas com melhor tolerância ao estresse abiótico.

Além disso, observa-se surpreendente agora que aumentar a expressão em uma planta, com uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL da maneira aqui definida, fornece plantas com mais características relacionadas ao rendimento com relação às plantas controle. De acordo com uma primeira modalidade, a presente invenção fornece um método para aumentar as características relacionadas ao rendimento em plantas com relação às plantas controle, compreendendo aumentar a expressão em uma planta com uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL.

A invenção também fornece até aqui sequências de ácidos

nucleicos desconhecidas que codificam polipeptídeos BRXL, e polipeptídeos BRXL.

De acordo com uma modalidade da presente invenção, é fornecido, portanto, uma Molécula de ácido nucléico isolado selecionada de:

5 (i) uma sequência de ácidos nucleicos da maneira representada por qualquer uma de SEQ ID NO: 75, SEQ ID NO: 77, ou SEQ ID NO: 79;

(ii) o complemento de uma sequência de ácidos nucleicos da maneira representada por qualquer uma de SEQ ID NO: 75, SEQ ID NO: 77, ou SEQ ID NO: 79;

10 (iii) uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL com, em uma ordem crescente de preferência, pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com a sequência de polipeptídeos representada por qualquer uma de SEQ ID NO: 15 76, SEQ ID NO: 78, ou SEQ ID NO: 80.

De acordo com uma modalidade adicional da presente invenção, é também fornecido um polipeptídeo isolado selecionado de:

(i) uma sequência de polipeptídeos da maneira representada por qualquer uma de SEQ ID NO: 76, SEQ ID NO: 78, ou SEQ ID NO: 80;

20 (ii) uma sequência de polipeptídeos com, em uma ordem crescente de preferência, pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com uma sequência de polipeptídeos da maneira representada por qualquer uma de SEQ ID NO: 76, SEQ ID NO: 78, ou SEQ 25 ID NO: 80;

(iii) derivados de quaisquer das sequências de polipeptídeos fornecidas em (i) ou (ii) anteriormente.

Além disso, observa-se surpreendente agora que modular a expressão em uma planta com um ácido nucléico que codifica um

polipeptídeo do tipo silky-1 fornece plantas com melhor tolerância ao estresse abiótico com relação às plantas controle. De acordo com uma primeira modalidade, a presente invenção fornece um método para melhorar a tolerância a vários estresses abióticos em plantas com relação às plantas controle, compreendendo modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo do tipo silky-1 e opcionalmente selecionar plantas com melhor tolerância ao estresse abiótico.

Além disso, observa-se surpreendente agora que modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo ARP6 fornece plantas com melhores características relacionadas ao rendimento com relação às plantas controle. De acordo com uma primeira modalidade, a presente invenção fornece um método para melhorar características relacionadas ao rendimento em plantas com relação às plantas controle, compreendendo modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo ARP6 e opcionalmente selecionar plantas com melhores características relacionadas ao rendimento.

Além disso, observa-se surpreendentemente agora que modular a expressão em uma planta com um ácido nucléico que codifica uma prolil-oligopeptidase, aqui denominado "polipeptídeo de POP", fornece plantas com melhores características relacionadas ao rendimento com relação às plantas controle. De acordo com uma primeira modalidade, a presente invenção fornece um método para melhorar características relacionadas ao rendimento em plantas com relação às plantas controle, compreendendo modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo de POP e opcionalmente selecionar plantas com melhores características relacionadas ao rendimento.

Além disso, observa-se surpreendentemente agora que modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo CRL fornece plantas com melhores características relacionadas

ao rendimento com relação às plantas controle. De acordo com uma primeira modalidade, a presente invenção fornece um método para melhorar características relacionadas ao rendimento em plantas com relação às plantas controle, compreendendo modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo CRL e opcionalmente selecionar plantas com melhores características relacionadas ao rendimento, preferivelmente qualquer uma de o rendimento total da semente (peso total das sementes), número de sementes preenchidas (número de semente preenchidas), taxa de preenchimento (taxa de preenchimento) e índice da colheita (índice da colheita), e/ou identificar a planta transgênica selecionando a planta transgênica que superexpressa o gene CRL (introduzido ou endógeno) e/ou a proteína CRL.

Um método preferido para modular (preferivelmente aumentar) a expressão de um ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL, é introduzir e expressar em uma planta um ácido nucléico que codifica um polipeptídeo do tipo alfin, um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL.

Com relação aos polipeptídeos do tipo alfin, qualquer referência a partir daqui a uma "proteína usada nos métodos da invenção" é obtida para significar um polipeptídeo do tipo alfin da maneira aqui definida. Qualquer referência daqui por diante a um "ácido nucléico usado nos métodos da invenção" é obtida para significar um ácido nucléico capaz de codificar um polipeptídeo do tipo alfin como este. O ácido nucléico a ser introduzido em uma planta (e, portanto, usado na realização dos métodos da invenção) é qualquer ácido nucléico que codifica o tipo de proteína que será agora descrita, daqui por diante também denominado "ácido nucléico do tipo alfin"

ou "gene do tipo alfin".

Com relação aos polipeptídeos YRP, qualquer referência daqui por diante a uma "proteína usada nos métodos da invenção" é obtida para significar um polipeptídeo YRP da maneira aqui definida. Qualquer referência
5 daqui por diante a um "ácido nucléico usado nos métodos da invenção" é obtida para significar um ácido nucléico capaz de codificar um polipeptídeo YRP como este. O ácido nucléico a ser introduzido em uma planta (e, portanto, usado na realização dos métodos da invenção) é qualquer ácido nucléico que codifica o tipo de proteína que será agora descrita, daqui por
10 diante também denominado "ácido nucléico YRP " ou "gene YRP".

O YRP de *Hordeum vulgare* (SEQ ID NO: 11 e 13) é fator de transcrição que codifica elementos da classe GARP de fatores de transcrição.

Com relação aos polipeptídeos BRXL, qualquer referência daqui por diante a uma "proteína usada nos métodos da invenção" é obtida
15 para significar um polipeptídeo BRXL da maneira aqui definida. Qualquer referência daqui por diante a uma "sequência de ácidos nucleicos usada nos métodos da invenção" é obtida para significar uma sequência de ácidos nucleicos capaz de codificar um polipeptídeo BRXL como este. A sequência de ácidos nucleicos a ser introduzida em uma planta (e, portanto, usado na
20 realização dos métodos da invenção) é qualquer sequência de ácidos nucleicos que codifica o tipo de polipeptídeo, que será agora descrito, daqui por diante também denominada "sequência de ácidos nucleicos BRXL" ou "gene BRXL".

Com relação aos polipeptídeos do tipo silky, qualquer
25 referência daqui por diante a uma "proteína usada nos métodos da invenção" é obtida para significar um polipeptídeo do tipo silky-1 da maneira aqui definida. Qualquer referência daqui por diante a um "ácido nucléico usado nos métodos da invenção" é obtida para significar um ácido nucléico capaz de codificar um polipeptídeo do tipo silky-1 como este. O ácido nucléico a ser

introduzido em uma planta (e, portanto, usado na realização dos métodos da invenção) é qualquer ácido nucléico que codifica o tipo de proteína que será agora descrita, daqui por diante também denominado "ácido nucléico do tipo silky-1" ou "gene do tipo silky-1".

5 Com relação aos polipeptídeos ARP6, qualquer referência daqui por diante a uma "proteína usada nos métodos da invenção" é obtida para significar um polipeptídeo ARP6 da maneira aqui definida. Qualquer referência daqui por diante a um "ácido nucléico usado nos métodos da invenção" é obtida para significar um ácido nucléico capaz de codificar um
10 polipeptídeo ARP6 como este. O ácido nucléico a ser introduzido em uma planta (e, portanto, usado na realização dos métodos da invenção) é qualquer ácido nucléico que codifica o tipo de proteína que será agora descrita, daqui por diante também denominado "ácido nucléico ARP6" ou "gene ARP6".

Com relação aos polipeptídeos POP, qualquer referência daqui
15 por diante a uma "proteína usada nos métodos da invenção" é obtida para significar um polipeptídeo de POP da maneira aqui definida. Qualquer referência daqui por diante a um "ácido nucléico usado nos métodos da invenção" é obtida para significar um ácido nucléico capaz de codificar um polipeptídeo de POP como este. O ácido nucléico a ser introduzido em uma
20 planta (e, portanto, usado na realização dos métodos da invenção) é qualquer ácido nucléico que codifica o tipo de proteína que será agora descrita, daqui por diante também denominado "ácido nucléico POP" ou "gene POP".

Com relação aos polipeptídeos CRL, qualquer referência daqui
por diante a uma "proteína usada nos métodos da invenção" é obtida para
25 significar um polipeptídeo CRL da maneira aqui definida. Qualquer referência daqui por diante a um "ácido nucléico usado nos métodos da invenção" é obtida para significar um ácido nucléico capaz de codificar um polipeptídeo CRL como este. O ácido nucléico a ser introduzido em uma planta (e, portanto, usado na realização dos métodos da invenção) é qualquer ácido

nucléico que codifica o tipo de proteína que será agora descrita, daqui por diante também denominado "ácido nucléico CRL" ou "gene CRL".

Um "polipeptídeo do tipo alfin" da maneira aqui definida refere-se a qualquer polipeptídeo compreendendo um motivo de núcleo hexâmero tanto de GNGGTG quanto de GTGGNG.

Exemplos de tais polipeptídeos do tipo alfin incluem ortólogos e parólogos das sequências representados por qualquer de SEQ ID NO: 2 e SEQ ID NO: 4.

Os polipeptídeos do tipo alfin e ortólogos e parólogos destes apresentam tipicamente em uma ordem crescente de preferência pelo menos 25 %, 26 %, 27 %, 28 %, 29 %, 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade completa de sequência com o aminoácido representado por qualquer de SEQ ID NO: 2 e SEQ ID NO: 4.

A identidade completa de sequência é determinada usando um algoritmo de alinhamento global, tal como o algoritmo de Needleman Wunsch no programa GAP (GCG Wisconsin Package, Accelrys), preferivelmente com parâmetros padrões e preferivelmente com sequências de proteínas maduras (isto é, sem levar em consideração sinais de secreção ou peptídeos de trânsito). Comparada com a identidade completa de sequência, a identidade de sequência em geral será maior quando apenas domínios ou motivos conservados forem considerados.

Preferivelmente, a sequência de polipeptídeos que, quando usada na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos do tipo alfin compreendendo as sequências de aminoácidos

representadas por SEQ ID NO: 2 e SEQ ID NO: 4. em vez de com qualquer outro grupo. Ferramentas e técnicas para a construção e análise de árvores filogenéticas são bem conhecidas na técnica.

Um "polipeptídeo YRP", da maneira aqui definida, refere-se a qualquer polipeptídeo de acordo com SEQ ID NO: 11 e SEQ ID NO: 13 e ortólogos e parólogos das sequências representados por qualquer de SEQ ID NO: 11 e SEQ ID NO: 13.

Os polipeptídeos YRP e ortólogos e parólogos destes apresentam tipicamente em uma ordem crescente de preferência pelo menos 25 %, 26 %, 27 %, 28 %, 29 %, 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade completa de sequência com o aminoácido representado por qualquer de SEQ ID NO: 11 e SEQ ID NO: 13.

A identidade completa de sequência é determinada usando um algoritmo de alinhamento global, tal como o algoritmo de Needleman Wunsch no programa GAP (GCG Wisconsin Package, Accelrys), preferivelmente com parâmetros padrões e preferivelmente com sequências de proteínas maduras (isto é, sem levar em consideração sinais de secreção ou peptídeos de trânsito). Comparada com a identidade completa de sequência, a identidade de sequência em geral será maior quando apenas domínios ou motivos conservados forem considerados.

Preferivelmente, a sequência de polipeptídeos, que quando usada na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos YRP compreendendo as sequências de aminoácidos representadas por SEQ ID NO: 11 e SEQ ID NO: 13 em vez de com qualquer

outro grupo. Ferramentas e técnicas para a construção e análise de árvores filogenéticas são bem conhecidas na técnica.

Um "polipeptídeo BRXL", da maneira aqui definida, refere-se a qualquer polipeptídeo compreendendo pelo menos dois domínios BRX com um domínio IPR013591 DZC de entrada InterPro (PFAM entrada PF08381 DZC).

Alternativamente, ou adicionalmente, "polipeptídeo BRXL", da maneira aqui definida, refere-se a qualquer polipeptídeo compreendendo (i) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio BRX da maneira representada por SEQ ID NO: 81; e (ii) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio BRX da maneira representada por SEQ ID NO: 82.

Alternativamente, ou adicionalmente, "polipeptídeo BRXL", da maneira aqui definida, refere-se a qualquer polipeptídeo compreendendo (i) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 1 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 83; e (ii) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 2 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 84.

Adicionalmente, um "polipeptídeo BRXL" da maneira aqui definida compreende adicionalmente (1) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos

com um domínio conservado 3 da maneira representada por SEQ ID NO: 85; e (1) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 4 da maneira representada por SEQ ID NO: 86.

Alternativamente, ou adicionalmente, um "polipeptídeo BRXL", da maneira aqui definida, refere-se a qualquer sequência de polipeptídeos com, em uma ordem crescente de preferência, pelo menos 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um polipeptídeo da maneira representada por SEQ ID NO: 18.

Alternativamente, ou adicionalmente, um "polipeptídeo BRXL", da maneira aqui definida, refere-se a qualquer polipeptídeo com, em uma ordem crescente de preferência, pelo menos 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com quaisquer das sequências de polipeptídeos fornecidas na tabela A3 aqui.

Alternativamente, ou adicionalmente, um "polipeptídeo BRXL", da maneira aqui definida, refere-se a qualquer polipeptídeo que, em um ensaio de dois híbridos em levedura, interage com ele mesmo ou com um outro polipeptídeo BRLX.

Um "polipeptídeo do tipo silky-1", da maneira aqui definida, refere-se a qualquer polipeptídeo representado por qualquer de SEQ ID NO: 91, SEQ ID NO: 93 e SEQ ID NO: 95 e ortólogos e parólogos destes.

Os polipeptídeos do tipo silky e ortólogos e parólogos destes apresentam tipicamente em uma ordem crescente de preferência pelo menos 25 %, 26 %, 27 %, 28 %, 29 %, 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %,

61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade completa de sequência com o aminoácido representado por qualquer de SEQ ID NO: 91, SEQ ID NO: 93 e SEQ ID NO: 95.

A identidade completa de sequência é determinada usando um algoritmo de alinhamento global, tal como o algoritmo de Needleman Wunsch no programa GAP (GCG Wisconsin Package, Accelrys), preferivelmente com parâmetros padrões e preferivelmente com sequências de proteínas maduras (isto é, sem levar em consideração sinais de secreção ou peptídeos de trânsito). Comparada com a identidade completa de sequência, a identidade de sequência em geral será maior quando apenas domínios ou motivos conservados forem considerados.

Preferivelmente, a sequência de polipeptídeos, que quando usada na construção de uma árvore filogenética, agrupa-se com o grupo de sifky-1 -like polipeptídeos compreendendo as sequências de aminoácidos representadas por SEQ ID NO: 91, SEQ ID NO: 93 e SEQ ID NO: 95 em vez de com qualquer outro grupo. Ferramentas e técnicas para a construção e análise de árvores filogenéticas são bem conhecidas na técnica.

Um "polipeptídeo ARP6", da maneira aqui definida, refere-se a qualquer polipeptídeo compreendendo um domínio actina e com, em uma ordem crescente de preferência, pelo menos 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade completa de sequência com o

aminoácido representado por SEQ ID NO: 102 ou por qualquer polipeptídeo da tabela A5.

Os polipeptídeos ARP6 compreendem tipicamente um domínio actina. Os domínios de actina são bem conhecidos na técnica. Por exemplo, a base de dados Pfam dos domínios de proteína (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002)) refere-se ao domínio actina como o número de referência: PF00022.

O domínio Pfam PF00022 baseia-se em torno das pesquisas de modelo oculto de Markov (HMM) da maneira fornecida pelo pacote HMMER2. Em HMMER2, como BLAST, os valores E (valores de expectativas) são calculados. O valor E é o número de acertos que pode ser esperado para apresentar uma pontuação igual ou melhor que esta acidental. Um bom valor E é muito menor que 1. Em torno de 1 é o que espera-se exatamente ao acaso. A princípio, tudo o que se precisa para decidir na significância de uma combinação é o valor E. A seguir estão as pontuações de domínio que definem o domínio actina da maneira fornecida na base de dados Pfam.

Parâmetro	Modelo HMM			
	Modelo <i>ls</i>		Modelo <i>fs</i>	
	Sequência	Pontuação de domínio	Sequência	Pontuação de domínio
Ponto de corte da coleta	-144,0	-144,0	12,4	12,4
Ponto de corte de confiança	-144,0	-144,0	12,6	12,6
Ponto de corte de barulho	-145,0	-145,0	12,3	12,3

O modelo HMM usado para construir o domínio actina é indicado. A ordem que as combinações *ls* (global) e *fs* (fragmento) são alinhadas no modelo fornece o alinhamento completo. O método de construção pode ser primeiro global, quando as combinações *ls* são alinhadas primeiro, seguidas por combinações *fs* que não se sobrepõem, biclassificadas, quando as combinações são alinhadas a fim de classificar o valor *e*, ou

primeiro local, quando as combinações fs são alinhadas primeiro seguidas pelas combinações ls que não se sobrepõem. A classificação de um domínio único alinhado a um HMM é indicada. Se existir mais de um domínio, a classificação da sequência é a soma de todas as pontuações do domínio para a tal entrada Pfam. Se existir apenas um domínio único, a classificação da sequência e dos domínios com relação à proteína será idêntica.

O ponto de corte da coleta usado no domínio actina é indicado. Este valor é o limite de pesquisa usado para construir o alinhamento completo. O ponto de corte da coleta é a classificação mínima que uma sequência tem que manter a fim de pertencer ao alinhamento completo de uma entrada Pfam. Para cada Pfam HMM existem dois valores de ponto de corte, um ponto de corte de sequência e um ponto de corte de domínio.

O ponto de corte de confiança refere-se às classificações bit da menor combinação de classificação no alinhamento completo.

O ponto de corte de barulho (NC) refere-se às classificações bit da combinação de classificação mais alta no alinhamento não completo.

Alternativamente, o homólogo de uma proteína ARP6 apresenta em uma ordem crescente de preferência pelo menos 25 %, 26 %, 27 %, 28 %, 29 %, 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade completa de sequência com o aminoácido representado por SEQ ID NO: 102, contanto que a proteína homóloga compreenda os motivos conservados da maneira anteriormente representada. A identidade completa de sequência é determinada usando um algoritmo de alinhamento global, tal como o algoritmo de Needleman Wunsch no programa GAP (GCG

Wisconsin Package, Accelrys), preferivelmente com parâmetros padrões e preferivelmente com sequências de proteínas maduras (isto é, sem levar em consideração sinais de secreção ou peptídeos de trânsito). Comparada com a identidade completa de sequência, a identidade de sequência em geral será maior quando apenas domínios ou motivos conservados forem considerados.

Preferivelmente, a sequência de polipeptídeos que, quando usada na construção de uma árvore filogenética, tal como aquela representada na figura 1 de Kandasamy et al. 2004. Trends Plant Sci 9: 196-202, agrupa-se com o grupo de polipeptídeos ARP6 compreendendo a sequência de aminoácidos de AtARP6 da maneira representada por SEQ ID NO: 102 em vez de com qualquer outro grupo.

Uma "Prolil-oligopeptidase" ou "polipeptídeo de POP", da maneira aqui definida, refere-se a qualquer serina protease que classifica-se como uma família S9 peptidase na base de dados MEROPS. Nas serina proteases, as proteínas POP pertencem ao dobramento α/β hidrolase no clã SC (Tripathi & Sowdhamini, BMC Genomics, 7, 200, 2006). Preferivelmente, o polipeptídeo de POP pertence à subfamília S9B das prolil-oligopeptidases. O polipeptídeo de POP usado nos métodos da invenção compreende um domínio Peptidase_S9 (entrada Pfam PF00326) e preferivelmente também um domínio DPPIV_N (entrada Pfam PF00930).

Preferivelmente, o polipeptídeo de POP também compreende um ou mais dos seguintes motivos:

Motivo 1 (SEQ ID NO: 118):

(V/S/L)(Y/H)GGP

Motivo 2 (SEQ ID NO: 119):

(0/A)(YZF)(LZW)(RZTZS)(SZN)(RZQZKZI)G(IZWZY)(LZ
AZS)(VZFZQ)(WMAZL)(KZDZI)(LZMMI)(DZN
) (NZY)(RZG)G(SZT)(UM/S/T/L)(R/G)(R/Y)G(L/R/E)

Motivo 3 (SEQ ID NO: 120):

(RZH)(IZL)(GZCZT)(IZLZV)(YZCZSZTZL)G(WZGZR)S(Y
ZAZH)GG(YZF)(MZLZT)(AZSZT)

Motivo 4 (SEQ ID NO: 121):

(YZF)(DZE)(TZSZA)(YZHZFZR)(YZG)(TZIZD)(EZDZQ)(
5 KZNZS)(YZLZH)(MMY)(GZT)

Motivo 5 (SEQ ID NO: 122):

S(VZIZP)(MZI)(HZNZS)(HZF)(VZI)

Motivo 6 (SEQ ID NO: 123):

(HZQZL)G(MZLZT)(IZEZK)D(EZKZL)(NMR)V(HZTZP)(FZPZI)

10 Motivo 7 (SEQ ID NO: 124):

(FZY)(PZE)(DZGZN)(EZD)(RZQZN)H(MZGZP)(PZFZL)(R
ZD)(GZRZK).

Preferivelmente de maneira adicional, o polipeptídeo de POP
também compreende um ou mais dos seguintes motivos:

15 Motivo 8 (SEQ ID NO: 125):

KLRRERLR(QZE)RGLGVT(C/R)YEW

Motivo 9 (SEQ ID NO: 126):

HG(LZI)AEYIAQEEM(DZE)R(KZR)(NZTZM)G(YZF)WW
S(LZP)DS Motivo 10 (SEQ ID NO: 127):

20 GFIWASE(KZR)(SZT)GFRHL

Motivo 11 (SEQ ID NO: 128):

LR(SZN)(QZKZR)GILVWK(LZM)D

Motivo 12 (SEQ ID NO: 129):

IG(LMI)(CZY)GWSYGG(YZF)

25 Motivo 13 (SEQ ID NO: 130):

CAV(SZA)GAPVT(SZA)WDGYD

Motivo 14 (SEQ ID NO: 131):

HGMIDENVHFRHTARL

Mais preferivelmente, o polipeptídeo de POP compreende em

uma ordem crescente de preferência, pelo menos 2, pelo menos 3, pelo menos 4, pelo menos 5, pelo menos 6, pelo menos 7, pelo menos 8, pelo menos 9, pelo menos 10, pelo menos 11, ou todos os 12 motivos.

Alternativamente, o homólogo de uma proteína POP apresenta em uma ordem crescente de preferência pelo menos 25 %, 26 %, 27 %, 28 %, 29 %, 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade completa de sequência com o aminoácido representado por SEQ ID NO: 117, contanto que a proteína homóloga compreenda pelo menos um dos motivos conservados da maneira anteriormente representada. A identidade completa de sequência é determinada usando um algoritmo de alinhamento global, tal como o algoritmo de Needleman Wunsch no programa GAP (GCG Wisconsin Package, Accelrys), preferivelmente com parâmetros padrões e preferivelmente com sequências de proteínas maduras (isto é, sem levar em consideração sinais de secreção ou peptídeos de trânsito). Comparada com a identidade completa de sequência, a identidade de sequência em geral será maior quando apenas domínios ou motivos conservados forem considerados. Preferivelmente, os motivos em um polipeptídeo de POP apresentam, em uma ordem crescente de preferência, pelo menos 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade de sequência com os motivos representados por SEQ ID NO: 118 to SEQ ID NO: 129 (Motivos 1 a 12).

Preferivelmente, a sequência de polipeptídeos que, quando

usada na construção de uma árvore filogenética, tal como aquela representada na figura 13 (Tripathi & Sowdhamini 2006), agrupa-se com o grupo de polipeptídeos POP indicado pela seta compreendendo a sequência de aminoácidos representada por SEQ ID NO: 1 17 em vez de com qualquer outro grupo.

Um "polipeptídeo CRL", da maneira aqui definida, refere-se a qualquer polipeptídeo compreendendo um domínio de proteína com, em uma ordem crescente de preferência, pelo menos 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade completa de sequência com o domínio de aminoácido representado pela sequência de aminoácidos 42-236 de SEQ ID NO: 156 e opcionalmente com um domínio transmembrana, preferivelmente da maneira representada pela sequência de aminoácidos 19-36 de SEQ ID NO: SEQ ID NO: 156.

Os métodos para determinar a presença de um domínio transmembrana que estão em uma proteína são bem conhecidos na técnica (Os detalhes adicionais são fornecidos na seção de exemplo). Espera-se que uma proteína com um domínio transmembrana esteja localizada em uma estrutura membranosa. Preferivelmente, a proteína da invenção, quando presente em uma célula, localiza-se em uma membrana.

Uma proteína preferida da invenção refere-se a um polipeptídeo CRL com, em uma ordem crescente de preferência, pelo menos 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade completa de sequência com o domínio de aminoácido representado pela sequência de aminoácidos 42-236 de SEQ ID NO: 156 e opcionalmente com um domínio transmembrana, preferivelmente da maneira representada pela sequência de aminoácidos 19-36 de SEQ ID NO: SEQ ID NO: 156.

98 %, ou 99 % de identidade de sequência com qualquer um ou mais dos seguintes motivos:

(i) Motivo 1; EQAFWRxPXKPFRQR (SEQ ID NO: 207);

(ii) Motivo 2: NFCDR (SEQ ID NO: 208);

5 (iii) Motivo 3: RGKRCLYEGS (SEQ ID NO: 209);

(iv) Motivo 4: QVWGxKXGPYEFK (SEQ ID NO: 210);

em que X representa qualquer aminoácido.

Alternativamente, o homólogo de uma proteína CRL apresenta em uma ordem crescente de preferência pelo menos 25 %, 26 %, 27 %, 28 %, 10 29 %, 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 15 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade completa de sequência com qualquer um dos aminoácidos apresentados na tabela A7, preferivelmente com a sequência representada por SEQ ID NO: 156. A identidade completa de sequência é determinada usando um algoritmo de alinhamento global, tal como o algoritmo de Needleman 20 Wunsch no programa GAP (GCG Wisconsin Package, Accelrys), preferivelmente com parâmetros padrões e preferivelmente com sequências de proteínas maduras (isto é, sem levar em consideração sinais de secreção ou peptídeos de trânsito). Comparada com a identidade completa de sequência, a identidade de sequência em geral será maior quando apenas domínios ou 25 motivos conservados forem considerados. Para alinhamentos locais, o algoritmo de Smith-Waterman é particularmente usado (Smith TF, Waterman MS (1981) J. Mol. Biol 147(1);195-7). Um método alternativo para determinar a identidade de sequência é fornecido pelo método conhecido como MATGAT (seção de exemplos).

Preferivelmente, a sequência de polipeptídeos que, quando usada na construção de uma árvore filogenética, tal como aquela representada na figura 16, agrupa-se com o grupo de polipeptídeo CRLs compreendendo a sequência de aminoácidos codificada por uma planta monocotiledônea, mais preferivelmente por uma sequência codificada por uma planta dicotiledônea, acima de tudo preferivelmente pela sequência representada por SEQ ID NO: 156 em vez de com qualquer outro grupo.

Os termos "domínio", "assinatura" e "motivo" são definidos aqui na seção "definições". As bases de dados especialistas existem para a identificação de domínios, por exemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-244), InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318), Prosite (Bucher e Bairoch (1994), Uma sintaxe de perfil generalizado para motivos de sequência biomolecular e sua função na interpretação de sequência automática. (Em) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; HuIo et al., Nucl. Acids. Res. 32:D134-D137, (2004)), ou Pfam (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002)). Um conjunto de ferramentas para análise de sílica de sequências protéicas é disponível no servidor proteômico ExPASy (Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: o servidor proteômico para conhecimento aprofundado e análise de proteína, Nucleic Acids Res. 31 :3784-3788(2003)). Os domínios ou motivos também pode ser identificados usando técnicas de rotina, tal como por alinhamento de sequência.

Com relação aos polipeptídeos BRXL, um alinhamento dos polipeptídeos de tabela A3 aqui é mostrado na figura 5. Tais alinhamentos são usadas para identificar os principais domínios ou motivos conservados entre os polipeptídeos BRXL da maneira aqui definida. Quatro domínios como

estes são (i) um domínio conservado 1 que representa um domínio BRX, que compreende o domínio IPR013591 DZC (entrada PFAM PF08381 DZC; marcado por X na figura 5); (2) um domínio conservado 2 que representa um domínio BRX, que compreende um domínio C-terminal IPR013591 DZC (entrada PFAM PF08381 DZC; marcado com um X na figura 5); (3) um domínio conservado 3 e (4) um domínio conservado 4, ambos contendo Cys conservado, cujo espaçamento é indicativo de um motivo de ligação de zinco potencial. Todos os quatro domínios conservados são enquadrados na figura 5.

Os métodos para o alinhamento de sequências 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, que varre as sequências completas) de duas sequências que maximizam o número de combinações e minimiza o número de intervalos. O algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-10) calcula a identidade percentual de sequência e realiza uma análise estatística da similaridade entre as duas sequências. O software para realizar a análise BLAST é publicamente disponível por meio do National Centre for Biotechnology information (NCBI). Os homólogos podem ser facilmente identificados usando, por exemplo, o algoritmo de alinhamento de sequência múltipla ClustalW (versão 1.83), com os parâmetros de alinhamento aos pares padrões, e um método de classificação em percentual. Os percentuais globais de similaridade e identidade também podem ser determinados usando um dos métodos disponíveis no pacote do software MatGAT (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: a application that generates similarity/identity matrices using protein or DNA sequences). A edição de manual menor pode ser realizada para otimizar o alinhamento entre motivos conservados, como pode ser evidente para os versados na técnica. Além disso,

ao contrário, usando sequências de comprimento total para a identificação de homólogos, domínios específicos também podem ser usados. A identidade dos valores de sequência pode ser determinada com relação ao ácido nucléico total ou a sequência de aminoácidos ou com relação aos domínios selecionados ou motivo(s) conservado(s), usando os programas mencionados anteriormente usando os parâmetros padrões. Para alinhamentos locais, o algoritmo de Smith-Waterman é particularmente usado (Smith TF, Waterman MS (1981) J. Mol. Biol 147(1);195-7).

Com relação aos polipeptídeos BRXL, a seção de exemplos aqui descreve na tabela B1 a identidade percentual entre o polipeptídeo BRXL da maneira representada por SEQ ID NO: 18 e os polipeptídeos BRXL listados na tabela A3, que pode ser tão baixa quanto 43 % de identidade de sequência de aminoácido. Em alguns exemplos, os parâmetros padrões podem ser ajustados para modificar a severidade da pesquisa. Por exemplo, usando BLAST, o limite de significância estatística (denominado valor "esperado") para relatar combinações contra sequências de base de dados pode ser maior para mostrar menos combinações severas. Da mesma maneira, as combinações exatamente pequenas podem ser identificadas.

A tarefa da previsão de localização subcelular protéica é importante e bem estudada. Conhecer uma localização de proteína auxilia a elucidar sua função. Os métodos experimentais para a localização de proteína variam de imunolocalização para marcação de proteínas usando proteína verde fluorescente (GFP) ou beta-glucuronidase (GUS). Tais métodos são exatos embora trabalhosos comparados com métodos computacionais. Muito recentemente foi realizado progresso na previsão computacional da localização protéica dos dados da sequência. Entre os algoritmos bem conhecidos pelos versados na técnica, estão disponíveis nas ferramentas proteômicas ExPASy hospedadas pelo Swiss Institute for Bioinformatics, por exemplo, PSort, TargetP, ChloroP, LocTree, Predotar, LipoP, MITOPROT,

PATS, PTS1, SignalP, TMHMM e outros.

Com relação aos polipeptídeos BRXL, de acordo com TargetP, a localização subcelular predita de um polipeptídeo BRXL da maneira representada por SEQ ID NO: 18 é o núcleo (ver a seção de exemplo).

5 Os polipeptídeos do tipo alfin, quando expressos em plantas, em particular em plantas de arroz, conferem melhor tolerância aos estresses abióticos naquelas plantas.

Além disso, os polipeptídeos YRP, quando expressos em plantas, em particular em plantas de arroz, conferem melhor tolerância aos
10 estresses abióticos naquelas plantas.

Além disso, polipeptídeos do tipo silky-1, quando expressos em plantas, em particular em plantas de arroz, conferem melhor tolerância aos estresses abióticos naquelas plantas.

Além disso, ferramentas e técnicas para medir a atividade de
15 polipeptídeos ARP6 são bem conhecidos na tecnologia, por exemplo, da maneira descrita nas referências da literatura aqui incluídas. Além do mais, polipeptídeos ARP6, quando expressos em arroz de acordo com os métodos da presente invenção, da maneira apresentada na seção de exemplos, fornecem plantas com mais características relacionadas ao rendimento, em
20 particular, em qualquer uma ou mais de peso da semente, índice da colheita, taxa de preenchimento da semente e o número de sementes preenchidas por planta.

Além disso, polipeptídeos POP (pelo menos na sua forma nativa) apresentam tipicamente atividade de endopeptidase, que clivam após
25 os resíduos de prolina, e em uma extensão menor após os resíduos Ala. As ferramentas e técnicas para medir a atividade de prolil endopeptidase são bem conhecidas na tecnologia, ver, por exemplo Nomura (FEBS Letters 209, 235-237, 1986). Os detalhes adicionais são fornecidos na seção de exemplo. Além do mais, os polipeptídeos POP, quando expressos em arroz de acordo com o

métodos da presente invenção, da maneira apresentada na seção de exemplos, fornecem plantas com mais características relacionadas ao rendimento, em particular maior biomassa e/ou maior rendimento da semente.

Além disso, os polipeptídeo CRL apresentam tipicamente a
5 atividade de modular a fotossíntese quando presente em uma planta. As ferramentas e técnicas para medir modulação de fotossíntese são bem conhecidas na tecnologia, por exemplo, por determinação do teor de clorofila ou da fluorescência de clorofila (Asano et al 2004, The plant Journal, Volume 38, pp. 448-459(12). Além do mais, os polipeptídeo CRLs, quando expressos
10 em arroz de acordo com o métodos da presente invenção, da maneira apresentada na seção de exemplos, fornecem plantas com mais características relacionadas ao rendimento, em particular em qualquer uma ou mais selecionadas do rendimento total da semente (peso total da semente), número de sementes preenchidas (número de semente preenchida), taxa de
15 preenchimento (taxa de preenchimento) e índice da colheita (índice da colheita).

Com relação aos polipeptídeos do tipo alfin, a presente invenção pode ser realizada, por exemplo, transformando plantas com a sequência de ácidos nucleicos representada por qualquer de SEQ ID NO: 1
20 que codifica a sequência de polipeptídeos de SEQ ID NO: 2, ou SEQ ID NO: 3 que codifica a sequência de polipeptídeos de SEQ ID NO: 4. Entretanto, o desempenho da invenção não é restrito a estas sequências; os métodos da invenção podem ser vantajosamente realizados usando qualquer ácido nucléico codificante do tipo alfin ou polipeptídeo do tipo alfin da maneira
25 aqui definida.

Exemplos de ácidos nucleicos que codificam polipeptídeos do tipo alfin são fornecidos na tabela A1 da seção de exemplos aqui. Tais ácidos nucleicos são usados na realização dos métodos da invenção. Os ortólogos e parólogos das sequências de aminoácidos fornecidos na tabela A1 podem ser

facilmente obtidos usando ferramentas e técnicas de rotina, tal como uma pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envolve realizar o BLAST em uma sequência de consulta (por exemplo, usando quaisquer das sequências listadas na tabela A1 da seção de exemplos) contra qualquer base de dados de sequência, tal como a base de dados NCBI disponível publicamente. BLASTN ou TBLASTX (usando valores padrões) são em geral usados durante o início de uma sequência de nucleotídeos, e BLASTP ou TBLASTN (usando valores padrões) durante o início de uma sequência de proteína. Os resultados do BLAST podem ser opcionalmente filtrados. As sequências de comprimento total tanto dos resultados de filtrados quanto dos resultados de não filtrados são então novamente submetidas ao BLAST (segundo BLAST) contra sequências do organismo a partir do qual a sequência de consulta é derivada (quando a sequência de consulta é SEQ ID NO: 1 ou SEQ ID NO: 2, o segundo BLAST portanto, seria contra sequências de *Solanum lycopersicum*; quando a sequência de consulta é SEQ ID NO: 3 ou SEQ ID NO: 4, o segundo BLAST portanto, seria contra sequências de *Populus trichocarpa*). Os resultados do primeiro e segundo BLASTs são então comparados. Um parólogo é identificado se um acerto de alta classificação do primeiro blast for da mesma espécie a partir da qual a sequência de consulta é derivada, um novo BLAST então resulta de maneira ideal na sequência de consulta entre os maiores acertos; um ortólogo é identificado se um acerto de alta classificação no primeiro BLAST não for da mesma espécie a partir da qual a sequência de consulta é derivada, e preferivelmente resulta em novo BLAST na sequência de consulta que está entre os maiores acertos.

Com relação aos polipeptídeos YRP, a presente invenção pode ser realizada, por exemplo, transformando plantas com a sequência de ácidos nucleicos representada por qualquer de SEQ ID NO: 10 que codifica a sequência de polipeptídeos de SEQ ID NO: 11, ou SEQ ID NO: 12 que

codifica a sequência de polipeptídeos de SEQ ID NO: 13. Entretanto, o desempenho da invenção não é restrito a estas sequências; os métodos da invenção podem ser vantajosamente realizados usando qualquer ácido nucléico que codifica YRP ou polipeptídeo YRP da maneira aqui definida.

5 Exemplos de ácidos nucleicos que codificam polipeptídeos YRP são fornecidos na tabela A2 da seção de exemplos aqui. Tais ácidos nucleicos são usados na realização dos métodos da invenção. Os ortólogos e parólogos das sequências de aminoácidos fornecidos na tabela A2 podem ser facilmente obtidos usando ferramentas e técnicas de rotina, tal como uma
10 pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envolve realizar o BLAST em uma sequência de consulta (por exemplo, usando quaisquer das sequências listadas na tabela A2 da seção de exemplos) contra qualquer base de dados de sequência, tal como a base de dados NCBI disponível publicamente. BLASTN ou TBLASTX (usando valores padrões)
15 são em geral usados durante o início de uma sequência de nucleotídeos, e BLASTP ou TBLASTN (usando valores padrões) durante o início de uma sequência de proteína. Os resultados do BLAST podem ser opcionalmente filtrados. As sequências de comprimento total tanto dos resultados de filtrados quanto dos resultados de não filtrados são então novamente submetidas ao
20 BLAST (segundo BLAST) contra sequências do organismo a partir do qual a sequência de consulta é derivada (quando a sequência de consulta é SEQ ID NO: 10 ou SEQ ID NO: 11, o segundo BLAST portanto, seria contra sequências de *Hordeum vulgare*; quando a sequência de consulta é SEQ ID NO: 12 ou SEQ ID NO: 13, o segundo BLAST portanto, seria contra
25 sequências de *Hordeum vulgare*). Os resultados do primeiro e segundo BLASTs são então comparados. Um parólogo é identificado se um acerto de alta classificação do primeiro blast for da mesma espécie a partir da qual a sequência de consulta é derivada, um novo BLAST então resulta de maneira ideal na sequência de consulta entre os maiores acertos; um ortólogo é

identificado se um acerto de alta classificação no primeiro BLAST não for da mesma espécie a partir da qual a sequência de consulta é derivada, e preferivelmente resulta em novo BLAST na sequência de consulta que está entre os maiores acertos.

5 Com relação aos polipeptídeos BRXL, a presente invenção é ilustrada transformando plantas com a sequência de ácidos nucleicos representada por SEQ ID NO: 17, que codifica a sequência BRXL de polipeptídeos de SEQ ID NO: 18. Entretanto, o desempenho da invenção não é restrito a estas sequências; os métodos da invenção podem ser
10 vantajosamente realizados usando qualquer sequência de ácidos nucleicos que codifica um polipeptídeo BRXL da maneira aqui definida.

Exemplos de sequências de ácidos nucleicos que codificam polipeptídeos BRXL são fornecidos na tabela A3 do exemplo 1 aqui. Tais sequências de ácidos nucleicos são usadas na realização dos métodos da
15 invenção. A sequências de polipeptídeos fornecidas na tabela A3 do exemplo 1 são exemplos de sequências de ortólogos e parólogos do polipeptídeo BRXL representadas por SEQ ID NO: 18, os termos "ortólogos" e "parólogos" sendo da maneira aqui definida. Ortólogos e parólogos adicionais podem ser facilmente identificados realizando uma assim denominada
20 pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envolve realizar o BLAST em uma sequência de consulta (por exemplo, usando quaisquer das sequências listadas na tabela A3 do exemplo 1) contra qualquer base de dados de sequência, tal como a base de dados NCBI disponível publicamente. BLASTN ou TBLASTX (usando valores padrões)
25 são em geral usados durante o início de uma sequência de nucleotídeos, e BLASTP ou TBLASTN (usando valores padrões) durante o início de uma sequência de proteína. Os resultados do BLAST podem ser opcionalmente filtrados. As sequências de comprimento total tanto dos resultados de filtrados quanto dos resultados de não filtrados são então novamente submetidas ao

BLAST (segundo BLAST) contra sequências do organismo a partir do qual a sequência de consulta é derivada (quando a sequência de consulta é SEQ ID NO: 17 ou SEQ ID NO: 18, o segundo BLAST portanto, seria contra sequências de álamo). Os resultados do primeiro e segundo BLASTs são
5 então comparados. Um parólogo é identificado se um acerto de alta classificação do primeiro blast for da mesma espécie a partir da qual a sequência de consulta é derivada, um novo BLAST então resulta de maneira ideal na sequência de consulta entre os maiores acertos; um ortólogo é
10 identificado se um acerto de alta classificação no primeiro BLAST não for da mesma espécie a partir da qual a sequência de consulta é derivada, e preferivelmente resulta em novo BLAST na sequência de consulta que está entre os maiores acertos.

Com relação aos polipeptídeos do tipo silky-1, a presente invenção pode ser realizada, por exemplo, transformando plantas com a
15 sequência de ácidos nucleicos representada por qualquer de SEQ ID NO: 90 que codifica a sequência de polipeptídeos de SEQ ID NO: 91, SEQ ID NO: 92 que codifica a sequência de polipeptídeos de SEQ ID NO: 93, ou SEQ ID NO: 94 que codifica a sequência de polipeptídeos de SEQ ID NO: 95. Entretanto, o desempenho da invenção não é restrito a estas sequências; os
20 métodos da invenção podem ser vantajosamente realizados usando qualquer ácido nucléico codifica tipo silky-1 ou polipeptídeo do tipo silky-1 da maneira aqui definida.

Exemplos de ácidos nucleicos que codificam polipeptídeos do tipo silky-1 são fornecidos na tabela A4 da seção dos exemplos aqui. Tais
25 ácidos nucleicos são usados na realização dos métodos da invenção. Ortólogos e parólogos das sequências de aminoácidos fornecidos na tabela A4 podem ser facilmente obtidos usando ferramentas e técnicas de rotina, tal como uma pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envolve realizar o BLAST em uma sequência de consulta (por

exemplo, usando quaisquer das sequências listadas na tabela A4 da seção dos exemplos) contra qualquer base de dados de sequência, tal como a base de dados NCBI disponível publicamente. BLASTN ou TBLASTX (usando valores padrões) são em geral usados durante o início de uma sequência de nucleotídeos, e BLASTP ou TBLASTN (usando valores padrões) durante o início de uma sequência de proteína. Os resultados do BLAST podem ser opcionalmente filtrados. As sequências de comprimento total tanto dos resultados de filtrados quanto dos resultados de não filtrados são então novamente submetidas ao BLAST (segundo BLAST) contra sequências do organismo a partir do qual a sequência de consulta é derivada (quando a sequência de consulta é SEQ ID NO: 90 ou SEQ ID NO: 91, o segundo BLAST portanto, seria contra sequências de *Populus trichocarpa*; quando a sequência de consulta é SEQ ID NO: 92 ou SEQ ID NO: 93, o segundo BLAST portanto, seria contra sequências de *Solanum lycopersicum*; quando a sequência de consulta é SEQ ID NO: 94 ou SEQ ID NO: 95, o segundo BLAST portanto, seria contra sequências de *Triticum aestivum*). Os resultados do primeiro e segundo BLASTs são então comparados. Um parólogo é identificado se um acerto de alta classificação do primeiro blast for da mesma espécie a partir da qual a sequência de consulta é derivada, um novo BLAST então resulta de maneira ideal na sequência de consulta entre os maiores acertos; um ortólogo é identificado se um acerto de alta classificação no primeiro BLAST não for da mesma espécie a partir da qual a sequência de consulta é derivada, e preferivelmente resulta em novo BLAST na sequência de consulta que está entre os maiores acertos.

Com relação aos polipeptídeos ARP6, a presente invenção é ilustrada transformando plantas com a sequência de ácidos nucleicos representada por SEQ ID NO: 101, que codifica a sequência de polipeptídeos de SEQ ID NO: 102. Entretanto, o desempenho da invenção não é restrito a estas sequências; os métodos da invenção podem ser vantajosamente

realizados usando qualquer ácido nucléico que codifica ARTP6 ou polipeptídeo ARP6 da maneira aqui definida.

Exemplos de ácidos nucleicos que codificam polipeptídeos ARP6 são fornecidos na tabela A5 da seção dos exemplos aqui. Tais ácidos nucleicos são usados na realização dos métodos da invenção. As sequências de aminoácidos fornecidas na tabela A5 da seção dos exemplos são exemplos de sequências de ortólogos e parólogos do polipeptídeo ARP6 representadas por SEQ ID NO: 102, os termos "ortólogos" e "parólogos" sendo da maneira aqui definida. Ortólogos e parólogos adicionais podem ser facilmente identificados realizando uma assim denominada pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envolve realizar o BLAST em uma sequência de consulta (por exemplo, usando quaisquer das sequências listadas na tabela A5 da seção dos exemplos) contra qualquer base de dados de sequência, tal como a base de dados NCBI disponível publicamente. BLASTN ou TBLASTX (usando valores padrões) são em geral usados durante o início de uma sequência de nucleotídeos, e BLASTP ou TBLASTN (usando valores padrões) durante o início de uma sequência de proteína. Os resultados do BLAST podem ser opcionalmente filtrados. As sequências de comprimento total tanto dos resultados de filtrados quanto dos resultados de não filtrados são então novamente submetidas ao BLAST (segundo BLAST) contra sequências do organismo a partir do qual a sequência de consulta é derivada (quando a sequência de consulta é SEQ ID NO: 101 ou SEQ ID NO: 102, o segundo BLAST portanto, seria contra sequências de *Arabidopsis thaliana*). Os resultados do primeiro e segundo BLASTs são então comparados. Um parólogo é identificado se um acerto de alta classificação do primeiro blast for da mesma espécie a partir da qual a sequência de consulta é derivada, um novo BLAST então resulta de maneira ideal na sequência de consulta entre os maiores acertos; um ortólogo é identificado se um acerto de alta classificação no primeiro BLAST não for da

mesma espécie a partir da qual a sequência de consulta é derivada, e preferivelmente resulta em novo BLAST na sequência de consulta que está entre os maiores acertos.

Com relação aos polipeptídeos POP, a presente invenção é
5 ilustrada transformando plantas com a sequência de ácidos nucleicos representada por SEQ ID NO: 116, que codifica a sequência de polipeptídeos de SEQ ID NO: 1 17. Entretanto, o desempenho da invenção não é restrito a estas sequências; os métodos da invenção podem ser vantajosamente realizados usando qualquer ácido nucléico que codifica POP ou polipeptídeo
10 de POP da maneira aqui definida.

Exemplos de ácidos nucleicos que codificam polipeptídeos POP são fornecidos na tabela A6 da seção dos exemplos aqui. Tais ácidos nucleicos são usados na realização dos métodos da invenção. As sequências de aminoácidos fornecidas na tabela A6 da seção dos exemplos são exemplos
15 de sequências de ortólogos e parólogos do polipeptídeo de POP representadas por SEQ ID NO: 2, os termos "ortólogos" e "parólogos" sendo da maneira aqui definida. Ortólogos e parólogos adicionais podem ser facilmente identificados realizando uma assim denominada pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envolve realizar o
20 BLAST em uma sequência de consulta (por exemplo, usando quaisquer das sequências listadas na tabela A6 da seção dos exemplos) contra qualquer base de dados de sequência, tal como a base de dados NCBI disponível publicamente. BLASTN ou TBLASTX (usando valores padrões) são em geral usados durante o início de uma sequência de nucleotídeos, e BLASTP ou
25 TBLASTN (usando valores padrões) durante o início de uma sequência de proteína. Os resultados do BLAST podem ser opcionalmente filtrados. As sequências de comprimento total tanto dos resultados de filtrados quanto dos resultados de não filtrados são então novamente submetidas ao BLAST (segundo BLAST) contra sequências do organismo a partir do qual a

sequência de consulta é derivada (quando a sequência de consulta é SEQ ID NO: 116 ou SEQ ID NO: 117, o segundo BLAST portanto, seria contra sequências de *Arabidopsis*). Os resultados do primeiro e segundo BLASTs são então comparados. Um parólogo é identificado se um acerto de alta
5 classificação do primeiro blast for da mesma espécie a partir da qual a sequência de consulta é derivada, um novo BLAST então resulta de maneira ideal na sequência de consulta entre os maiores acertos; um ortólogo é identificado se um acerto de alta classificação no primeiro BLAST não for da mesma espécie a partir da qual a sequência de consulta é derivada, e
10 preferivelmente resulta em novo BLAST na sequência de consulta que está entre os maiores acertos.

Com relação aos polipeptídeos CRL, a presente invenção é ilustrada transformando plantas com a sequência de ácidos nucleicos representada por SEQ ID NO: 155, que codifica a sequência de polipeptídeos
15 de SEQ ID NO: 156. Entretanto, o desempenho da invenção não é restrito a estas sequências; os métodos da invenção podem ser vantajosamente realizados usando qualquer ácido nucléico que codifica CRL ou polipeptídeo CRL da maneira aqui definida.

Exemplos de ácidos nucleicos que codificam polipeptídeos
20 CRL são fornecidos na tabela A7 da seção dos exemplos aqui. Tais ácidos nucleicos são usados na realização dos métodos da invenção. As sequências de aminoácidos fornecidas na tabela A7 da seção dos exemplos são exemplos de sequências de ortólogos e parólogos do polipeptídeo CRL representadas por SEQ ID NO: 2, os termos "ortólogos" e "parólogos" sendo da maneira
25 aqui definida. Ortólogos e parólogos adicionais podem ser facilmente identificados realizando uma assim denominada pesquisa blast recíproca. Tipicamente, isto envolve um primeiro BLAST que envolve realizar o BLAST em uma sequência de consulta (por exemplo, usando quaisquer das sequências listadas na tabela A7 da seção dos exemplos) contra qualquer base

de dados de sequência, tal como a base de dados NCBI disponível publicamente. BLASTN ou TBLASTX (usando valores padrões) são em geral usados durante o início de uma sequência de nucleotídeos, e BLASTP ou TBLASTN (usando valores padrões) durante o início de uma sequência de proteína. Os resultados do BLAST podem ser opcionalmente filtrados. As sequências de comprimento total tanto dos resultados de filtrados quanto dos resultados de não filtrados são então novamente submetidas ao BLAST (segundo BLAST) contra sequências do organismo a partir do qual a sequência de consulta é derivada (quando a sequência de consulta é SEQ ID NO: 155 ou SEQ ID NO: 156, o segundo BLAST portanto, seria contra sequências de *Arabidopsis thaliana*). Os resultados do primeiro e segundo BLASTs são então comparados. Um parólogo é identificado se um acerto de alta classificação do primeiro blast for da mesma espécie a partir da qual a sequência de consulta é derivada, um novo BLAST então resulta de maneira ideal na sequência de consulta entre os maiores acertos; um ortólogo é identificado se um acerto de alta classificação no primeiro BLAST não for da mesma espécie a partir da qual a sequência de consulta é derivada, e preferivelmente resulta em novo BLAST na sequência de consulta que está entre os maiores acertos.

Os acertos de alta classificação são aqueles com um baixo valor E. Quanto menor o valor E, mais significativa é a classificação (ou em outras palavras, tanto menor o acaso que o acerto foi observado pelo acaso). A computação do valor E é bem conhecida na técnica. Além dos valores E, as comparações também são classificadas por identidade percentual. A identidade percentual refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as das sequências comparadas de ácido nucléico (ou polipeptídeo) como relação a um tamanho particular. No caso de famílias maiores, ClustalW pode ser usado, seguido por uma árvore de união ao redor, para auxiliar o agrupamento de visualização de genes relacionados e para

identificar ortólogos e parólogos.

Os variantes de ácido nucléico também podem ser usados para praticar os métodos da invenção. Exemplos de tais variantes incluem ácidos nucleicos que codificam homólogos e derivados de qualquer uma das sequências de aminoácidos fornecidas na tabela A1 a A7 da seção de
 5 exemplos, os termos "homólogo" e "derivado" sendo da maneira aqui definida. São também usados nos métodos da invenção ácidos nucleicos que codificam homólogos e derivados de ortólogos ou parólogos de qualquer uma das sequências de aminoácidos fornecidas na tabela A1 a A7 da seção dos
 10 exemplos. Homólogos e derivados usados nos métodos da presente invenção apresentam substancialmente a mesma atividade biológica e funcional da proteína não modificada a partir da qual eles são derivados. Os variantes adicionais usados na prática dos métodos da invenção são variantes em que o uso do códon é otimizado ou em que os sítios alvos de RNAmi são
 15 removidos.

Os variantes de ácido nucléico adicionais usados na prática dos métodos da invenção incluem porções de ácidos nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky-1, ou polipeptídeos ARP6, ou polipeptídeos
 20 POP, ou polipeptídeos CRL, ácidos nucleicos que hibridizam em ácidos nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky-1, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, variantes de união de ácidos nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos
 25 YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky-1, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, variantes alélicos de ácidos nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky-1, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, e

variantes de ácidos nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky-1, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, obtidos por embaralhamento de gene. Os termos sequência hibridizante, variante de
5 união, variante alélica e embaralhamento de gene são da maneira aqui descrita.

Os ácidos nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky-1, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos
10 CRL, não precisam ser ácidos nucleicos de comprimento total, uma vez que o desempenho dos métodos da invenção não dependem do uso de sequências de ácidos nucleicos de comprimento total. De acordo com a presente invenção, é fornecido um método para melhorar a tolerância ao estresse abiótico em plantas, compreendendo introduzir e expressar em uma planta uma porção de
15 qualquer uma das sequências de ácidos nucleicos fornecidas na tabela A1 a A7 da seção dos exemplos, ou uma porção de um ácido nucléico que codifica um ortólogo, parólogo ou homólogo de quaisquer das sequências de aminoácidos fornecidas na tabela A1 a A7 da seção dos exemplos.

Uma porção de um ácido nucléico pode ser preparada, por
20 exemplo, produzindo uma ou mais deleções no ácido nucléico. As porções podem ser usadas na forma isolada ou elas podem ser fundidas às sequências codificantes (ou não codificantes) a fim de, por exemplo, produzir uma proteína que combina várias atividades. Quando fundidas a outras sequências codificantes, o polipeptídeo resultante produzido na tradução pode ser maior
25 que aquele previsto para a porção de proteína.

Com relação aos polipeptídeos do tipo alfin, as porções usadas nos métodos da invenção codificam um polipeptídeo do tipo alfin da maneira aqui definida, e apresentam substancialmente a mesma atividade biológica das sequências de aminoácidos fornecidas na tabela A1 da seção dos exemplos.

Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos fornecidos na tabela A1 da seção dos exemplos, ou é uma porção de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A1 da seção dos exemplos.

- 5 Preferivelmente a porção apresenta pelo menos 650, 700, 750, 800, 850, 900, 950, 1.000, 1.050, 1.100, 1.150, 1.200, 1.250 ou mais nucleotídeos consecutivos de comprimento, os nucleotídeos consecutivos sendo de qualquer uma das sequências de ácidos nucleicos fornecidas na tabela A1 da seção dos exemplos, ou de um ácido nucléico que codifica um ortólogo ou
- 10 parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A1 da seção dos exemplos. Acima de tudo preferivelmente, a porção é uma porção do ácido nucléico de SEQ ID NO: 1 ou SEQ ID NO: 3. Preferivelmente, a porção codifica um fragmento de uma sequência de aminoácidos que, quando usado na construção de uma árvore filogenética,
- 15 agrupa-se com o grupo de polipeptídeos do tipo alfin compreendendo a sequência de aminoácidos representada por SEQ ID NO: 2 ou SEQ ID NO: 4, em vez de com qualquer outro grupo.

- Com relação aos polipeptídeos YRP, as porções usadas nos métodos da invenção codificam um polipeptídeo YRP da maneira aqui
- 20 definida, e apresentam substancialmente a mesma atividade biológica das sequências de aminoácidos fornecidas na tabela A2 da seção dos exemplos. Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos fornecidos na tabela A2 da seção dos exemplos, ou é uma porção de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das
- 25 sequências de aminoácidos fornecidas na tabela A2 da seção dos exemplos. Preferivelmente a porção apresenta pelo menos 700, 750, 800, 850, 900, 950, 1.000, 1.050, 1.100, 1.150, 1.200, 1.250, 1.300, 1.350, 1.350, 1.400, ou mais nucleotídeos consecutivos de comprimento, os nucleotídeos consecutivos sendo de qualquer uma das sequências de ácidos nucleicos fornecidas na

tabela A2 da seção dos exemplos, ou de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A2 da seção dos exemplos. Acima de tudo preferivelmente, a porção é uma porção do ácido nucléico de SEQ ID NO: 10 ou SEQ ID NO: 12. Preferivelmente, a porção codifica um fragmento de uma sequência de aminoácidos que, quando usado na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos YRP compreendendo a sequência de aminoácidos representada por SEQ ID NO: 11 ou SEQ ID NO: 13, em vez de com qualquer outro grupo.

Com relação aos polipeptídeos BRXL, as porções usadas nos métodos da invenção, codificam um polipeptídeo BRXL, da maneira aqui definida, e apresentam substancialmente a mesma atividade biológica das sequências de polipeptídeos fornecidas na tabela A3 do exemplo 1. Preferivelmente, a porção é uma porção de qualquer uma das sequências de ácidos nucleicos fornecidas na tabela A3 do exemplo 1, ou é uma porção de uma sequência de ácidos nucleicos que codifica um ortólogo ou parólogo de qualquer uma das sequências de polipeptídeos fornecidas na tabela A3 do exemplo 1. Preferivelmente, a porção apresenta, em uma ordem crescente de preferência pelo menos 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.050, 1.060, 1.070, 1.080 ou mais nucleotídeos consecutivos de comprimento, os nucleotídeos consecutivos sendo de qualquer uma das sequências de ácidos nucleicos fornecidas na tabela A3 do exemplo 1, ou de uma sequência de ácidos nucleicos que codifica um ortólogo ou parólogo de qualquer uma das sequências de polipeptídeos fornecidas na tabela A3 do exemplo 1. Preferivelmente, a porção é uma porção de uma sequência de ácido nucléico que codifica uma sequência de polipeptídeos compreendendo (i) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 1 (compreendendo

um domínio BRX) da maneira representada por SEQ ID NO: 83; e (ii) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 2 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 84. Mais preferivelmente, a porção é uma porção de uma sequência de ácido nucléico que codifica uma sequência de polipeptídeos com, em uma ordem crescente de preferência, pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com o polipeptídeo BRXL da maneira representada por SEQ ID NO: 18 ou com quaisquer das sequências de polipeptídeos fornecidas na tabela A3 aqui. Acima de tudo preferivelmente, a porção é uma porção da sequência de ácidos nucleicos de SEQ ID NO: 17.

Com relação aos polipeptídeos do tipo silky-1, as porções usadas nos métodos da invenção codificam um polipeptídeo do tipo silky-1 da maneira aqui definida, e apresentam substancialmente a mesma atividade biológica das sequências de aminoácidos fornecidas na tabela A4 da seção dos exemplos. Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos fornecidos na tabela A4 da seção dos exemplos, ou é uma porção de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A4 da seção dos exemplos. Preferivelmente a porção apresenta pelo menos 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.050, 1.100 ou mais nucleotídeos consecutivos de comprimento, os nucleotídeos consecutivos sendo de qualquer uma das sequências de ácidos nucleicos fornecidas na tabela A4 da seção dos exemplos, ou de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A4 da seção dos exemplos. Acima de tudo preferivelmente, a porção é uma porção do ácido nucléico de SEQ ID NO: 90, SEQ ID NO: 92 ou SEQ ID

NO: 94. Preferivelmente, a porção codifica um fragmento de uma sequência de aminoácidos que, quando usado na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos do tipo silky-1 compreendendo a sequência de aminoácidos representada por SEQ ID NO: 91, SEQ ID NO: 93 ou SEQ ID NO: 95, em vez de com qualquer outro grupo.

Com relação aos polipeptídeos ARP6, as porções usadas nos métodos da invenção codificam um polipeptídeo ARP6 da maneira aqui definida, e apresentam substancialmente a mesma atividade biológica das sequências de aminoácidos fornecidas na tabela A5 da seção dos exemplos.

Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos fornecidos na tabela A5 da seção dos exemplos, ou é uma porção de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A5 da seção dos exemplos.

Preferivelmente, a porção apresenta pelo menos 100, 200, 300, 400, 500, 550,

600, 650, 700, 750, 800, 850, 900, 950, 1.000 nucleotídeos consecutivos de comprimento, os nucleotídeos consecutivos sendo de qualquer uma das sequências de ácidos nucleicos fornecidas na tabela A5 da seção dos exemplos, ou de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A5 da

seção dos exemplos. Acima de tudo preferivelmente a porção é uma porção do ácido nucléico de SEQ ID NO: 101. Preferivelmente, a porção codifica um fragmento de uma sequência de aminoácidos que, quando usado na construção de uma árvore filogenética, tal como aquela representada na figura 1 de Kandasamy et al. 2004. Trends Plant Sci 9: 196-202, agrupa-se com o grupo de polipeptídeos ARP6 compreendendo a sequência de aminoácidos de AtARP6 da maneira representada por SEQ ID NO: 102 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos POP, as porções usadas nos métodos da invenção codificam um polipeptídeo de POP da maneira aqui

definida, e apresentam substancialmente a mesma atividade biológica das sequências de aminoácidos fornecidas na tabela A6 da seção dos exemplos. Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos fornecidos na tabela A6 da seção dos exemplos, ou é uma porção de um ácido

5 nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A6 da seção dos exemplos. Preferivelmente a porção apresenta pelo menos 500, 600, 700, 800, 900, 1.000, 1.100, 1.200, 1.300, 1.400, 1.500, 1.600, 1.700, 1.800, 1.900, 2.000, 2.100, 2.200, 2.300 nucleotídeos consecutivos de comprimento, os

10 nucleotídeos consecutivos sendo de qualquer uma das sequências de ácidos nucleicos fornecidas na tabela A6 da seção dos exemplos, ou de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A6 da seção dos exemplos. Acima de tudo preferivelmente, a porção é uma porção do ácido nucléico de

15 SEQ ID NO: 1. Preferivelmente, a porção codifica um fragmento de uma sequência de aminoácidos que, quando usado na construção de uma árvore filogenética, tal como aquela representada na figura 13 (Tripathi & Sowdhamini 2006), agrupa-se com o grupo de polipeptídeos POP indicado pela seta compreendendo a sequência de aminoácidos representada por SEQ

20 ID NO: 117 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos CRL, as porções usadas nos métodos da invenção codificam um polipeptídeo CRL da maneira aqui definida, e apresentam substancialmente a mesma atividade biológica das sequências de aminoácidos fornecidas na tabela A7 da seção dos exemplos.

25 Preferivelmente, a porção é uma porção de qualquer um dos ácidos nucleicos fornecidos na tabela A7 da seção dos exemplos, ou é uma porção de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A7 da seção dos exemplos. Preferivelmente a porção apresenta pelo menos 100, 200, 300, 400, 500, 550,

600, 650, 700, 750, 800, 850, 900, 950, 1.000 nucleotídeos consecutivos de comprimento, os nucleotídeos consecutivos sendo de qualquer uma das sequências de ácidos nucleicos fornecidas na tabela A7 da seção dos exemplos, ou de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A7 da
5 seção dos exemplos. Acima de tudo preferivelmente, a porção é uma porção do ácido nucléico de SEQ ID NO: 155. Preferivelmente, a porção codifica um fragmento de uma sequência de aminoácidos que, quando usado na construção de uma árvore filogenética, tal como aquela representada na figura
10 16, agrupa-se com o grupo de polipeptídeos CRL compreendendo a sequência de aminoácidos codificada por uma planta monocotiledônea, mais preferivelmente por uma sequência codificada por uma planta dicotiledônea, acima de tudo preferivelmente pela sequência representada por SEQ ID NO: 156 em vez de com qualquer outro grupo.

15 Um outro variante de ácido nucléico usado nos métodos da invenção é um ácido nucléico capaz de hibridizar, em menores condições de severidade, preferivelmente em condições severas, em um ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo
20 ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL, da maneira aqui definida, ou em uma porção da maneira aqui definida.

De acordo com a presente invenção, é fornecido um método para melhorar a tolerância ao estresse abiótico em plantas, compreendendo introduzir e expressar em uma planta um ácido nucléico capaz de hibridizar
25 em qualquer um dos ácidos nucleicos fornecidos na tabela A1 a A7 da seção dos exemplos, ou compreendendo introduzir e expressar em uma planta um ácido nucléico capaz de hibridizar em um ácido nucléico que codifica um ortólogo, parólogo ou homólogo de quaisquer das sequências de ácidos nucleicos fornecidas na tabela A1 a A7 da seção dos exemplos.

Com relação aos polipeptídeos do tipo alfin, as sequência hibridizantes usadas nos métodos da invenção codificam um polipeptídeo do tipo alfin, da maneira aqui definida, com substancialmente a mesma atividade biológica das sequências de aminoácidos fornecidas na tabela A1 da seção dos exemplos. Preferivelmente, a sequência hibridizante é capaz de hibridizar no complemento de qualquer um dos ácidos nucleicos fornecidos na tabela A1, ou em uma porção de quaisquer das mesmas sequências, uma porção que é definida da maneira anterior, ou a sequência hibridizante é capaz de hibridizar no complemento de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A1. Acima de tudo preferivelmente, a sequência hibridizante é capaz de hibridizar no complemento de um ácido nucléico da maneira representada por SEQ ID NO: 1 ou SEQ ID NO: 3 ou em uma porção deste.

Preferivelmente, a sequência hibridizante codifica um polipeptídeo com uma sequência de aminoácidos que, quando em comprimento total e usado na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos do tipo alfin compreendendo a sequência de aminoácidos representada por SEQ ID NO: 2 ou SEQ ID NO: 4 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos YRP, as sequência hibridizantes usadas nos métodos da invenção codificam um polipeptídeo YRP, da maneira aqui definida, com substancialmente a mesma atividade biológica das sequências de aminoácidos fornecidas na tabela A2 da seção dos exemplos. Preferivelmente, a sequência hibridizante é capaz de hibridizar no complemento de qualquer um dos ácidos nucleicos fornecidos na tabela A2, ou em uma porção de quaisquer das mesmas sequências, uma porção que é definida da maneira anterior, ou a sequência hibridizante é capaz de hibridizar no complemento de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A2.

Acima de tudo preferivelmente, a sequência hibridizante é capaz de hibridizar no complemento de um ácido nucleico da maneira representada por SEQ ID NO: 10 ou SEQ ID NO: 12 ou em uma porção deste.

Preferivelmente, a sequência hibridizante codifica um polipeptídeo com uma sequência de aminoácidos que, quando em comprimento total e usado na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos YRP compreendendo a sequência de aminoácidos representada por SEQ ID NO: 11 ou SEQ ID NO: 13 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos BRXL, as sequências hibridizantes usadas nos métodos da invenção codificam um polipeptídeo BRXL, da maneira aqui definida, e apresentam substancialmente a mesma atividade biológica das sequências de polipeptídeos fornecidas na tabela A3 do exemplo 1. Preferivelmente, a sequência hibridizante é capaz de hibridizar em qualquer uma das sequências de ácidos nucleicos fornecidas na tabela A3 do exemplo 1, ou em um complemento da mesmas, ou em uma porção de quaisquer da mesmas sequências, uma porção que é definida da maneira anterior, ou em que a sequência hibridizante é capaz de hibridizar em uma sequência de ácidos nucleicos que codifica um ortólogo ou parólogo de qualquer uma das sequências de polipeptídeos fornecidas na tabela A3 do exemplo 1, ou em um complemento da mesmas. Preferivelmente, a sequência hibridizante é capaz de hibridizar em uma sequência de ácidos nucleicos que codifica uma sequência de polipeptídeos compreendendo (i) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 1 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 83; e (ii) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de

aminoácidos com um domínio conservado 2 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 84. Mais preferivelmente, a sequência hibridizante é capaz de hibridizar em uma sequência de ácidos nucleicos que codifica uma sequência de polipeptídeos com, em uma ordem crescente de preferência, pelo menos 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos to o polipeptídeo BRXL da maneira representada por SEQ ID NO: 18 ou em quaisquer das sequências de polipeptídeos fornecidas na tabela A3 aqui. Acima de tudo preferivelmente, a sequência hibridizante é capaz de hibridizar em uma sequência de ácidos nucleicos da maneira representada por SEQ ID NO: 17 ou em uma porção deste.

Com relação aos polipeptídeos do tipo silky-1, as sequência hibridizantes usadas nos métodos da invenção codificam um polipeptídeo do tipo silky-1, da maneira aqui definida, com substancialmente a mesma atividade biológica das sequências de aminoácidos fornecidas na tabela A4 da seção dos exemplos. Preferivelmente, a sequência hibridizante é capaz de hibridizar no complemento de qualquer um dos ácidos nucleicos fornecidos na tabela A4, ou em uma porção de quaisquer da mesmas sequências, uma porção que é definida da maneira anterior, ou a sequência hibridizante é capaz de hibridizar no complemento de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A4. Acima de tudo preferivelmente, a sequência hibridizante é capaz de hibridizar no complemento de um ácido nucléico da maneira representada por SEQ ID NO: 90, SEQ ID NO: 92 ou SEQ ID NO: 94 ou em uma porção deste.

Preferivelmente, a sequência hibridizante codifica um polipeptídeo com uma sequência de aminoácidos que, quando em comprimento total e usado na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos do tipo silky compreendendo a sequência de

aminoácidos representada por SEQ ID NO: 91, SEQ ID NO: 93 ou SEQ ID NO: 95 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos ARP6, as sequência
5 hibridizantes usadas nos métodos da invenção codificam um polipeptídeo
ARP6, da maneira aqui definida, com substancialmente a mesma atividade
biológica das sequências de aminoácidos fornecidas na tabela A5 da seção
dos exemplos. Preferivelmente, a sequência hibridizante é capaz de hibridizar
no complemento de qualquer um dos ácidos nucleicos fornecidos na tabela
10 A5 da seção dos exemplos, ou em uma porção de quaisquer da mesmas
sequências, uma porção que é definida da maneira anterior, ou a sequência
hibridizante é capaz de hibridizar no complemento de um ácido nucléico que
codifica um ortólogo ou parólogo de qualquer uma das sequências de
aminoácidos fornecidas na tabela A5 da seção dos exemplos. Acima de tudo
preferivelmente, a sequência hibridizante é capaz de hibridizar no
15 complemento de um ácido nucléico da maneira representada por SEQ ID NO:
101 ou em uma porção deste.

Preferivelmente, a sequência hibridizante codifica um
polipeptídeo com uma sequência de aminoácidos que, quando em
comprimento total e usado na construção de uma árvore filogenética, tal como
20 aquela representada na figura 1 de Kandasamy et al. 2004. Trends Plant Sci 9:
196-202, agrupa-se com o grupo de polipeptídeos ARP6 compreendendo a
sequência de aminoácidos de A1ARP6 da maneira representada por SEQ ID
NO: 102 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos POP, as sequência
25 hibridizantes usadas nos métodos da invenção codificam um polipeptídeo de
POP, da maneira aqui definida, com substancialmente a mesma atividade
biológica das sequências de aminoácidos fornecidas na tabela A6 da seção
dos exemplos. Preferivelmente, a sequência hibridizante é capaz de hibridizar
no complemento de qualquer um dos ácidos nucleicos fornecidos na tabela

A6 da seção dos exemplos, ou em uma porção de quaisquer das mesmas sequências, uma porção que é definida da maneira anterior, ou a sequência hibridizante é capaz de hibridizar no complemento de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A6 da seção dos exemplos. Acima de tudo preferivelmente, a sequência hibridizante é capaz de hibridizar no complemento de um ácido nucléico da maneira representada por SEQ ID NO: 116 ou em uma porção deste.

Preferivelmente, a sequência hibridizante codifica um polipeptídeo com uma sequência de aminoácidos que, quando em comprimento total e usado na construção de uma árvore filogenética, tal como aquela representada na figura 13 (Tripathi & Sowdhamini 2006), agrupa-se com o grupo de polipeptídeos POP indicado pela seta compreendendo a sequência de aminoácidos representada por SEQ ID NO: 117 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos CRL, as sequências hibridizantes usadas nos métodos da invenção codificam um polipeptídeo CRL, da maneira aqui definida, com substancialmente a mesma atividade biológica das sequências de aminoácidos fornecidas na tabela A7 da seção dos exemplos. Preferivelmente, a sequência hibridizante é capaz de hibridizar no complemento de qualquer um dos ácidos nucleicos fornecidos na tabela A7 da seção dos exemplos, ou em uma porção de quaisquer das mesmas sequências, uma porção que é definida da maneira anterior, ou a sequência hibridizante é capaz de hibridizar no complemento de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer uma das sequências de aminoácidos fornecidas na tabela A7 da seção dos exemplos. Acima de tudo preferivelmente, a sequência hibridizante é capaz de hibridizar no complemento de um ácido nucléico da maneira representada por SEQ ID NO: 155 ou em uma porção deste.

Preferivelmente, a sequência hibridizante codifica um polipeptídeo com uma sequência de aminoácidos que, quando em comprimento total e usado na construção de uma árvore filogenética, tal como aquela representada na figura 16, agrupa-se com o grupo de polipeptídeos CRL compreendendo a sequência de aminoácidos codificada por uma planta monocotiledônea, mais preferivelmente por uma sequência codificada por uma planta dicotiledônea, acima de tudo preferivelmente pela sequência representada por SEQ ID NO: 156 em vez de com qualquer outro grupo.

Um outro variante de ácido nucléico usado nos métodos da invenção é um variante de união que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL, da maneira aqui definida anteriormente, um variante de união sendo da maneira aqui definida.

De acordo com a presente invenção, é fornecido um método para melhorar a tolerância ao estresse abiótico e/ou melhorar as características relacionadas ao rendimento em plantas, compreendendo introduzir e expressar em uma planta um variante de união de qualquer uma das sequências de ácidos nucleicos fornecidas na tabela A1 a A7 da seção dos exemplos, ou um variante de união de um ácido nucléico que codifica um ortólogo, parólogo ou homólogo de quaisquer das sequências de aminoácidos fornecidas na tabela A1 a A7 da seção dos exemplos.

Com relação aos polipeptídeos do tipo alfin, variantes de união preferidos são variantes de união de um ácido nucléico representados por qualquer de SEQ ID NO: 1 ou SEQ ID NO: 3, ou um variante de união de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer de SEQ ID NO: 2 ou SEQ ID NO: 4. Preferivelmente, a sequência de aminoácidos codificada pelo variante de união, quando usada na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos do tipo alfin

compreendendo a sequência de aminoácidos representada por SEQ ID NO: 2 ou SEQ ID NO: 4 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos YRP, variantes de união preferidos são variantes de união de um ácido nucleico representados por qualquer de SEQ ID NO: 10 ou SEQ ID NO: 12, ou um variante de união de um ácido nucleico que codifica um ortólogo ou parólogo de qualquer de SEQ ID NO: 11 ou SEQ ID NO: 13. Preferivelmente, a sequência de aminoácidos codificada pelo variante de união, quando usada na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos YRP compreendendo a sequência de aminoácidos representada por SEQ ID NO: 11 ou SEQ ID NO: 13 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos BRXL, variantes de união preferidos são variantes de união de uma sequência de ácidos nucleicos representada por SEQ ID NO: 17, ou um variante de união de uma sequência de ácidos nucleicos que codifica um ortólogo ou parólogo de SEQ ID NO: 18. Preferivelmente, o variante de união é um variante de união de uma sequência de ácidos nucleicos que codifica uma sequência de polipeptídeos compreendendo (i) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 1 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 83; e (ii) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 2 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 84. Mais preferivelmente, o variante de união é um variante de união de uma sequência de ácidos nucleicos que codifica uma sequência de polipeptídeos com, em uma ordem crescente de preferência, pelo menos 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98

%, 99 % ou mais de identidade de sequência de aminoácidos com o polipeptídeo BRXL da maneira representada por SEQ ID NO: 18 ou com quaisquer das sequências de polipeptídeos fornecidas na tabela A3 aqui. Acima de tudo preferivelmente, o variante de união é um variante de união de uma sequência de ácidos nucleicos da maneira representada por SEQ ID NO: 17, ou de uma sequência de ácidos nucleicos que codifica uma sequência de polipeptídeos da maneira representada por SEQ ID NO: 18.

Com relação aos polipeptídeos do tipo silky, variantes de união preferidos são variantes de união de um ácido nucléico representados por qualquer de SEQ ID NO: 90, SEQ ID NO: 92 ou SEQ ID NO: 94, ou um variante de união de um ácido nucléico que codifica um ortólogo ou parólogo de qualquer de SEQ ID NO: 91, SEQ ID NO: 93 ou SEQ ID NO: 95. Preferivelmente, a sequência de aminoácidos codificada pelo variante de união, quando usada na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos do tipo silky compreendendo a sequência de aminoácidos representada por SEQ ID NO: 91, SEQ ID NO: 93 ou SEQ ID NO: 95 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos ARP6, variantes de união preferidos são variantes de união de um ácido nucléico representados por SEQ ID NO: 101, ou um variante de união de um ácido nucléico que codifica um ortólogo ou parólogo de SEQ ID NO: 102. Preferivelmente, a sequência de aminoácidos codificada pelo variante de união, quando usada na construção de uma árvore filogenética, tal como aquela representada na figura 1 de Kandasamy et al. 2004. Trends Plant Sci 9: 196-202, agrupa-se com o grupo de polipeptídeos ARP6 compreendendo a sequência de aminoácidos de AtARP6 da maneira representada por SEQ ID NO: 102 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos POP, variantes de união preferidos são variantes de união de um ácido nucléico representados por

SEQ ID NO: 116, ou um variante de união de um ácido nucléico que codifica um ortólogo ou parólogo de SEQ ID NO: 117. Preferivelmente, a sequência de aminoácidos codificada pelo variante de união, quando usada na construção de uma árvore filogenética, tal como aquela representada na figura 13 (Tripathi & Sowdhamini 2006), agrupa-se com o grupo de polipeptídeos POP indicado pela seta compreendendo a sequência de aminoácidos representada por SEQ ID NO: 117 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos CRL, variantes de união preferidos são variantes de união de um ácido nucléico representados por SEQ ID NO: 155, ou um variante de união de um ácido nucléico que codifica um ortólogo ou parólogo de SEQ ID NO: 156. Preferivelmente, a sequência de aminoácidos codificada pelo variante de união, quando usada na construção de uma árvore filogenética, tal como aquela representada na figura 16, agrupa-se com o grupo de polipeptídeos CRL compreendendo a sequência de aminoácidos codificada por uma planta monocotiledônea, mais preferivelmente por uma sequência codificada por uma planta dicotiledônea, acima de tudo preferivelmente pela sequência representada por SEQ ID NO: 156 em vez de com qualquer outro grupo.

Um outro variante de ácido nucléico usado na realização dos métodos da invenção é um variante alélico de um ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL, da maneira aqui definida anteriormente, um variante alélico sendo da maneira aqui definida.

De acordo com a presente invenção, é fornecido um método para melhorar a tolerância ao estresse abiótico e/ou melhorar as características relacionadas ao rendimento em plantas, compreendendo introduzir e expressar em uma planta um variante alélico de qualquer um dos ácidos nucleicos fornecidos na tabela A1 a A7, ou compreendendo introduzir e expressar em

uma planta um variante alélico de um ácido nucléico que codifica um ortólogo, parólogo ou homólogo de quaisquer das sequências de aminoácidos fornecidas na tabela A1 a A7.

Com relação aos polipeptídeos do tipo alfin, os polipeptídeos codificados por variantes alélicos usados nos métodos da presente invenção apresentam substancialmente a mesma atividade biológica do polipeptídeo do tipo alfin de qualquer de SEQ ID NO: 2 ou quaisquer dos aminoácidos representados na tabela A1 da seção dos exemplos. Os variantes alélicos existem na natureza, e está incluído nos métodos da presente invenção o uso destes alelos naturais. Preferivelmente, o variante alélico é um variante alélico de qualquer de SEQ ID NO: 1 ou SEQ ID NO: 3 ou um variante alélico de um ácido nucléico que codifica um ortólogo ou parólogo de SEQ ID NO: 2 ou SEQ ID NO: 4. Preferivelmente, a sequência de aminoácidos codificada pelo variante alélico, agrupa-se com os polipeptídeos do tipo alfin compreendendo a sequência de aminoácidos representada por SEQ ID NO: 2 ou SEQ ID NO: 4 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos YRP, os polipeptídeos codificados por variantes alélicos usados nos métodos da presente invenção apresentam substancialmente a mesma atividade biológica do polipeptídeo YRP de qualquer de SEQ ID NO: 11, ou quaisquer dos aminoácidos representados na tabela A2 da seção dos exemplos. Os variantes alélicos existem na natureza, e está incluído nos métodos da presente invenção o uso destes alelos naturais. Preferivelmente, o variante alélico é um variante alélico de qualquer de SEQ ID NO: 10 ou SEQ ID NO: 12 ou um variante alélico de um ácido nucléico que codifica um ortólogo ou parólogo de SEQ ID NO: 11 ou SEQ ID NO: 13. Preferivelmente, a sequência de aminoácidos codificada pelo variante alélico, agrupa-se com os polipeptídeos YRP compreendendo a sequência de aminoácidos representada por SEQ ID NO: 11 ou SEQ ID NO: 13 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos BRXL, os variantes alélicos usados nos métodos da presente invenção apresentam substancialmente a mesma atividade biológica do polipeptídeo BRXL de SEQ ID NO: 18 e quaisquer das sequências de polipeptídeos representadas na tabela A3 do exemplo 1. Os variantes alélicos existem na natureza, e está incluído nos métodos da presente invenção o uso destes alelos naturais. Preferivelmente, o variante alélico é um variante alélico de uma sequência de polipeptídeos compreendendo (i) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 1 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 83; e (ii) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 2 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 84. Mais preferivelmente, o variante alélico é um variante alélico que codifica uma sequência de polipeptídeos com, em uma ordem crescente de preferência, pelo menos 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com o polipeptídeo BRXL da maneira representada por SEQ ID NO: 18 ou com quaisquer das sequências de polipeptídeos fornecidas na tabela A aqui. Acima de tudo preferivelmente, o variante alélico é um variante alélico de SEQ ID NO: 17 ou um variante alélico de uma sequência de ácidos nucleicos que codifica um ortólogo ou parólogo de SEQ ID NO: 18.

Com relação aos polipeptídeos do tipo silky-1, os polipeptídeos codificados por variantes alélicos usados nos métodos da presente invenção apresentam substancialmente a mesma atividade biológica do polipeptídeo do tipo silky-1 de qualquer de SEQ ID NO: 91, ou quaisquer dos aminoácidos representados na tabela A4 da seção dos exemplos. Os

variantes alélicas existem na natureza, e está incluído nos métodos da presente invenção o uso destes alelos naturais. Preferivelmente, o variante alélico é um variante alélico de qualquer de SEQ ID NO: 90, SEQ ID NO: 92 ou SEQ ID NO: 94 ou um variante alélico de um ácido nucleico que codifica um ortólogo ou parólogo de SEQ ID NO: 91, SEQ ID NO: 93 ou SEQ ID NO: 95. Preferivelmente, a sequência de aminoácidos codificada pelo variante alélico, agrupa-se com os polipeptídeos do tipo silky-1 compreendendo a sequência de aminoácidos representada por SEQ ID NO: 91, SEQ ID NO: 93 ou SEQ ID NO: 95 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos ARP6, os polipeptídeos codificados por variantes alélicas usados nos métodos da presente invenção apresentam substancialmente a mesma atividade biológica do polipeptídeo ARP6 de SEQ ID NO: 102 e quaisquer dos aminoácidos representados na tabela A5 da seção dos exemplos. Os variantes alélicos existem na natureza, e está incluído nos métodos da presente invenção o uso destes alelos naturais. Preferivelmente, o variante alélico é um variante alélico de SEQ ID NO: 101 ou um variante alélico de um ácido nucleico que codifica um ortólogo ou parólogo de SEQ ID NO: 102. Preferivelmente, a sequência de aminoácidos codificada pelo variante alélico, quando usada na construção de uma árvore filogenética, tal como aquela representada na figura 1 de Kandasamy et al. 2004. Trends Plant Sci 9: 196-202, agrupa-se com o grupo de polipeptídeos ARP6 compreendendo a sequência de aminoácidos de AtARP6, da maneira representada por SEQ ID NO: 102, em vez de com qualquer outro grupo.

Com relação aos polipeptídeos POP, os polipeptídeos codificados por variantes alélicas usados nos métodos da presente invenção apresentam substancialmente a mesma atividade biológica do polipeptídeo de POP de SEQ ID NO: 117, e quaisquer dos aminoácidos representados na tabela A6 da seção dos exemplos. Os variantes alélicos existem na natureza, e está incluído nos métodos da presente invenção o uso destes alelos naturais.

Preferivelmente, o variante alélico é um variante alélico de SEQ ID NO: 116 ou um variante alélico de um ácido nucléico que codifica um ortólogo ou parólogo de SEQ ID NO: 117. Preferivelmente, a sequência de aminoácidos codificada pelo variante alélico, quando usada na construção de uma árvore filogenética, tal como aquela representada na figura 13 (Tripathi & Sowdhamini 2006), agrupa-se com o grupo de polipeptídeos POP indicado pela seta compreendendo a sequência de aminoácidos representada por SEQ ID NO: 117 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos CRL, os polipeptídeos codificados por variantes alélicos usados nos métodos da presente invenção apresentam substancialmente a mesma atividade biológica do polipeptídeo CRL de SEQ ID NO: 156, e quaisquer dos aminoácidos representados na tabela A7 da seção dos exemplos. Os variantes alélicos existem na natureza, e está incluído nos métodos da presente invenção o uso destes alelos naturais.

Preferivelmente, o variante alélico é um variante alélico de SEQ ID NO: 155 ou um variante alélico de um ácido nucléico que codifica um ortólogo ou parólogo de SEQ ID NO: 156. Preferivelmente, a sequência de aminoácidos codificada pelo variante alélico, quando usada na construção de uma árvore filogenética, tal como aquela representada na figura 16, agrupa-se com o grupo de polipeptídeos CRL compreendendo a sequência de aminoácidos codificada por uma planta monocotiledônea, mais preferivelmente por uma sequência codificada por uma planta dicotiledônea, acima de tudo preferivelmente pela sequência representada por SEQ ID NO: 156 em vez de com qualquer outro grupo.

O embaralhamento de gene ou evolução direcionada também pode ser usada para gerar variantes de ácidos nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky-1, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, da maneira definida anteriormente; o termo

"embaralhamento de gene" sendo da maneira aqui definida.

De acordo com a presente invenção, é fornecido um método para melhorar a tolerância ao estresse abiótico e/ou características relacionadas ao rendimento em plantas, compreendendo introduzir e expressar em uma planta um variante de qualquer uma das sequências de ácidos nucleicos fornecidas na tabela A1 a A7 da seção dos exemplos, ou compreendendo introduzir e expressar em uma planta um variante de um ácido nucléico que codifica um ortólogo, parólogo ou homólogo de quaisquer das sequências de aminoácidos fornecidas na tabela A1 a A7 da seção dos exemplos, cujo variante ácido nucléico é obtido por embaralhamento de gene.

Com relação aos polipeptídeos do tipo alfin, preferivelmente, a sequência de aminoácidos codificada pelo ácido nucléico variante obtido por embaralhamento de gene, quando usada na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos do tipo alfin compreendendo a sequência de aminoácidos representada por SEQ ID NO: 2 ou SEQ ID NO: 4 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos YRP, preferivelmente, a sequência de aminoácidos codificada pelo ácido nucléico variante obtido por embaralhamento de gene, quando usada na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos YRP compreendendo a sequência de aminoácidos representada por SEQ ID NO: 11 ou SEQ ID NO: 13 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos BRXL, preferivelmente, a sequência de ácidos nucleicos variante obtida por embaralhamento de gene codifica uma sequência de polipeptídeos compreendendo (i) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 1 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 83; e (ii) em uma ordem

crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 2 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 84. Mais preferivelmente, a

5 sequência de ácidos nucleicos variante obtida por embaralhamento de gene codifica uma sequência de polipeptídeos com, em uma ordem crescente de preferência, pelo menos 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com o polipeptídeo BRXL da maneira representada por SEQ ID

10 NO: 18 ou com quaisquer das sequências de polipeptídeos fornecidas na tabela A3 aqui. Acima de tudo preferivelmente, a sequência de ácidos nucleicos obtida por embaralhamento de gene codifica uma sequência de polipeptídeos da maneira representada por SEQ ID NO: 18.

Com relação aos polipeptídeos do tipo silky-1,

15 preferivelmente, a sequência de aminoácidos codificada pelo ácido nucléico variante obtido por embaralhamento de gene, quando usada na construção de uma árvore filogenética, agrupa-se com o grupo de polipeptídeos do tipo silky-1 compreendendo a sequência de aminoácidos representada por SEQ ID NO: 91, SEQ ID NO: 93 ou SEQ ID NO: 95 em vez de com qualquer outro

20 grupo.

Com relação aos polipeptídeos ARP6, preferivelmente, a sequência de aminoácidos codificada pelo ácido nucléico variante obtido por embaralhamento de gene, quando usada na construção de uma árvore filogenética tal como aquela representada na figura 1 de Kandasamy et al.

25 2004. Trends Plant Sci 9: 196-202, agrupa-se com o grupo de polipeptídeos ARP6 compreendendo a sequência de aminoácidos de AtARP6, da maneira representada por SEQ ID NO: 102, em vez de com qualquer outro grupo.

Com relação aos polipeptídeos POP, preferivelmente, a sequência de aminoácidos codificada pelo ácido nucléico variante obtido por

embaralhamento de gene, quando usada na construção de uma árvore filogenética tal como aquela representada na figura 13 (Thpathi & Sowdhamini 2006), agrupa-se com o grupo de polipeptídeos POP indicado pela seta, compreendendo a sequência de aminoácidos representada por SEQ ID NO: 117 em vez de com qualquer outro grupo.

Com relação aos polipeptídeos CRL, preferivelmente, a sequência de aminoácidos codificada pelo ácido nucléico variante obtido por embaralhamento de gene, quando usada na construção de uma árvore filogenética tal como aquela representada na figura 16, agrupa-se com o grupo de polipeptídeos CRL compreendendo a sequência de aminoácidos codificada por uma planta monocotiledônea, mais preferivelmente por uma sequência codificada por uma planta dicotiledônea, acima de tudo preferivelmente pela sequência representada por SEQ ID NO: 156 em vez de com qualquer outro grupo.

Além disso, os variantes de ácido nucléico também podem ser obtidos por mutagênese sítio direcionada. Vários métodos estão disponíveis para atingir a mutagênese sítio direcionada, o mais comum sendo métodos a base de PCR (Current Protocols in Molecular Biology. Wiley Eds.).

Ácidos nucleicos que codificam polipeptídeos do tipo alfin podem ser derivados de qualquer fonte natural ou artificial. O ácido nucléico pode ser modificado a partir de sua forma nativa em composição e/ou ambiente genômico por meio de manipulação humana deliberada. Preferivelmente, o ácido nucléico que codifica polipeptídeo do tipo alfin é de uma planta, preferivelmente de maneira adicional, de uma planta monocotiledônea ou dicotiledônea, mais preferivelmente da família *Allium* ou *Hordeum*.

Ácidos nucleicos que codificam polipeptídeos YRP podem ser derivados de qualquer fonte natural ou artificial. O ácido nucléico pode ser modificado a partir de sua forma nativa em composição e/ou ambiente

genômico por meio de manipulação humana deliberada. Preferivelmente, o ácido nucléico que codifica o polipeptídeo YRP é de uma planta, preferivelmente de maneira adicional, de uma planta monocotiledônea ou dicotiledônea, mais preferivelmente da família *Populus* ou *Solanum*.

5 As sequências de ácidos nucleicos que codificam polipeptídeos BRXL podem ser derivadas de qualquer fonte natural ou artificial. A sequência de ácidos nucleicos pode ser modificada a partir de sua forma nativa em composição e/ou ambiente genômico por meio de manipulação humana deliberada. A sequência de ácidos nucleicos que codifica um
10 polipeptídeo BRXL é de uma planta, preferivelmente de maneira adicional, de uma planta dicotiledônea, mais preferivelmente da família *Salicaceae*, acima de tudo preferivelmente, a sequência de ácidos nucleicos é de *Populus trichocarpa*.

Ácidos nucleicos que codificam polipeptídeos do tipo silky-1
15 podem ser derivados de qualquer fonte natural ou artificial. O ácido nucléico pode ser modificado a partir de sua forma nativa em composição e/ou ambiente genômico por meio de manipulação humana deliberada. Preferivelmente, o ácido nucléico que codifica o polipeptídeo do tipo silky-1 é de uma planta, preferivelmente de maneira adicional, de uma planta
20 monocotiledônea ou dicotiledônea, mais preferivelmente da família *Allium* ou *Hordeum*.

Ácidos nucleicos que codificam polipeptídeos ARP6 podem ser derivados de qualquer fonte natural ou artificial. O ácido nucléico pode ser modificado a partir de sua forma nativa em composição e/ou ambiente
25 genômico por meio de manipulação humana deliberada. Preferivelmente, o ácido nucléico que codifica o polipeptídeo ARP6 é de uma planta, preferivelmente de maneira adicional, de uma planta dicotiledônea, mais preferivelmente da família *Brassicaceae*, acima de tudo preferivelmente, o ácido nucléico é de *Arabidopsis thaliana*.

Ácidos nucleicos que codificam polipeptídeos POP podem ser derivados de qualquer fonte natural ou artificial. O ácido nucléico pode ser modificado a partir de sua forma nativa em composição e/ou ambiente genômico por meio de manipulação humana deliberada. Preferivelmente, o

5 ácido nucléico que codifica o polipeptídeo de POP é de uma planta, preferivelmente de maneira adicional, de uma planta dicotiledônea, mais preferivelmente, da família *Brassicaceae*, acima de tudo preferivelmente, o ácido nucléico é de *Arabidopsis thaliana*.

Ácidos nucleicos que codificam polipeptídeos CRL podem ser

10 derivados de qualquer fonte natural ou artificial. O ácido nucléico pode ser modificado a partir de sua forma nativa em composição e/ou ambiente genômico por meio de manipulação humana deliberada. Preferivelmente, o ácido nucléico que codifica o polipeptídeo CRL é de uma planta, preferivelmente de maneira adicional, de uma planta dicotiledônea, mais

15 preferivelmente, da família *Brassicaceae*, acima de tudo preferivelmente o ácido nucléico é de *Arabidopsis thaliana*.

Vantajosamente, a invenção também fornece até aqui ácidos nucleicos que codificam CRL desconhecido e polipeptídeos CRL.

De acordo com uma modalidade adicional da presente

20 invenção, é fornecido, portanto, uma Molécula de ácido nucléico isolado selecionada de:

(i) um ácido nucléico representadas por SEQ ID NO: 41;

(ii) o complemento de um ácido nucléico representadas por
SEQ ID NO: 41 :

25 (iii) um ácido nucléico que codifica o polipeptídeo da maneira representada por SEQ ID NO: 42, preferivelmente em função da degeneração do código genético, o dito ácido nucléico isolado pode ser derivado de uma sequência de polipeptídeos da maneira representada por SEQ ID NO: 42 e, preferivelmente de maneira adicional, confere melhores características

relacionadas ao rendimento com relação às plantas controle;

(iv) um ácido nucléico com, em uma ordem crescente de preferência pelo menos 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade de sequência com quaisquer das sequências de ácidos nucleicos da tabela A, e preferivelmente de maneira adicional que confere melhores características relacionadas ao rendimento com relação às plantas controle;

(v) uma molécula de ácido nucléico que hibridiza em uma molécula de ácido nucléico de (i) a (iv) em condições de hibridização severa, e preferivelmente confere melhores características relacionadas ao rendimento com relação às plantas controle;

(vi) um ácido nucléico que codifica um polipeptídeo ASPAT com, em uma ordem crescente de preferência, pelo menos 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade de sequência com a sequência de aminoácidos representada por SEQ ID NO: 42 e quaisquer das outras sequências de aminoácido na tabela A, e preferivelmente que confere melhores características relacionadas ao rendimento com relação às plantas controle.

De acordo com uma modalidade adicional da presente invenção, é também fornecido um polipeptídeo isolado selecionado de:

(i) uma sequência de aminoácidos representada por SEQ ID

NO: 42;

(ii) uma sequência de aminoácidos com, em uma ordem crescente de preferência, pelo menos 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade de sequência com a sequência de aminoácidos representada por SEQ ID NO: 42, e quaisquer das outras sequências de aminoácido na tabela A, e preferivelmente que confere melhores características relacionadas ao rendimento com relação às plantas controle;

(iii) derivados de quaisquer das sequências de aminoácidos fornecidas em (i) ou (ii) anteriormente.

Com relação aos polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos do tipo silky, o desempenho dos métodos da invenção fornece plantas com melhor tolerância ao estresse abiótico.

Com relação aos polipeptídeos BRXL, o desempenho dos métodos da invenção fornece plantas com mais características relacionadas ao rendimento com relação às plantas controle. Os termos "rendimento" e "o rendimento da semente" são descritos com mais detalhes aqui na seção "definições".

Com relação aos polipeptídeos ARP6, ou polipeptídeos CRL, o desempenho dos métodos da invenção fornece plantas com melhores características relacionadas ao rendimento. Em particular, o desempenho dos métodos da invenção fornece plantas com maior rendimento, especialmente maior rendimento da semente com relação às plantas controle. Os termos "rendimento" e "o rendimento da semente" são descritos com mais detalhes aqui na seção "definições".

Com relação aos polipeptídeos POP, o desempenho dos

métodos da invenção fornece plantas com melhores características relacionadas ao rendimento. Em particular, o desempenho dos métodos da invenção fornece plantas com maior rendimento, especialmente maior biomassa (raiz e/ou ramo) e rendimento da semente com relação às plantas controle. Os termos "rendimento" e "o rendimento da semente" são descritos com mais detalhes aqui na seção "definições". O desempenho dos métodos da invenção fornece plantas com período de floração modificada (preferivelmente período de floração inicial).

A referência aqui a melhores características relacionadas ao rendimento é obtida para significar um aumento em biomassa (peso) de uma ou mais partes de uma planta, que podem incluir partes aéreas (coletáveis) e/ou partes (coletáveis) subterrâneas. Em particular, tais Partes cultiváveis são sementes, e o desempenho dos métodos da invenção resulta em plantas com maior rendimento da semente com relação ao rendimento da semente da plantas controle.

Tendo o milho como um exemplo, um aumento no rendimento pode ser manifestado como um ou mais dos seguintes: aumento no número de plantas estabelecidas por metro quadrado, um aumento no número de espigas por planta, um aumento no número de linhas, número de grãos por linha, peso do grão, peso de mil grãos, tamanho/diâmetro da espiga, aumento na taxa de preenchimento da semente (que é o número de sementes preenchidas dividido pelo número total de sementes e multiplicado por 100), entre outros. Tendo o arroz como um exemplo, um rendimento aumento pode se manifestar como um aumento em um ou mais dos seguintes; número de plantas por metro quadrado, número de panículas por planta, número de espiguetas por panícula, número de flores (flores pequenas) por panícula (que é expresso como uma razão do número de sementes preenchidas para o número de panículas primárias), aumento na taxa de preenchimento da semente (que é o número de sementes preenchidas dividido pelo número total de sementes e

multiplicado por 100), aumento em peso de mil grãos, entre outros.

Com relação ao estresse abiótico, a presente invenção fornece um método para melhorar tolerância ao estresse em plantas, com relação às plantas controle, cujo método compreende modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo do tipo silky-1, da maneira aqui definida.

As plantas respondem tipicamente a exposição aos estresse crescendo mais lentamente. Em condições de estresse grave, a planta pode ainda parar de crescer por completo. Pouco estresse, por outro lado, é definido aqui como sendo qualquer estresse ao qual uma planta é exposta que não resulta na suspensão do crescimento da planta completamente, sem a capacidade de retomar o crescimento. Pouco estresse, no sentido da invenção, leva a uma redução no crescimento das plantas estressadas em menos de 40 %, 35 %, 30 % ou 25 %, mais preferivelmente menos de 20 % ou 15 %, em comparação à planta controle em condições de não estresse. Em virtude dos avanços nas práticas agrícolas (irrigação, fertilização, tratamentos com pesticida) os estresses graves não são frequentemente encontrados em plantas de cultivo cultivadas. Como uma consequência, o crescimento comprometido induzido por pouco estresse é frequentemente uma característica não desejável para a agricultura. Poucos estresses são os estresses bióticos e/ou abióticos de cada dia (ambiental) aos quais uma planta é exposta. Os estresses abióticos podem ser em virtude da secura ou excesso água, estresse anaeróbico, estresse pelo sal, toxicidade química, estresse oxidativo e temperaturas de calor, frio ou congelamento. O estresse abiótico pode ser um estresse osmótico causado por um estresse pela água (particularmente em virtude da secura), estresse pelo sal, estresse oxidativo ou um estresse iônico. Os estresses bióticos são tipicamente aqueles estresses causados por patógenos, tais como bactérias, vírus, fungos, nematódeos e insetos.

Em particular, os métodos da presente invenção podem ser realizados em condições de (pouca) seca para fornecer plantas com melhor tolerância à seca com relação às plantas controle, que podem se manifestar como um maior rendimento com relação às plantas controle. Da maneira

5 relatada em Wang et al. (Plant (2003) 218: 1-14), o estresse abiótico leva a uma série de mudanças morfológicas, fisiológicas, bioquímicas e moleculares que afetam adversamente o crescimento e produtividade da planta. Sabe-se que a seca, salinidade, temperaturas extremas e estresse oxidativo estão interconectados e podem induzir o crescimento e dano celular por meio de

10 mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) descrevem um grau particularmente alto de "informação cruzada" entre estresse por seca e estresse por alta salinidade. Por exemplo, seca e/ou salinização são manifestadas principalmente como estresse osmótico, resultando na interrupção da homeostase e distribuição de íon na célula. O

15 estresse oxidativo, que frequentemente acompanha estresse por temperatura alta ou baixa, salinidade ou seca, pode causar a desnaturação de proteínas funcionais e estruturais. Como uma consequência, estes diversos estresses ambientais frequentemente ativam vias celulares similares de sinalização e resposta celulares, tais como a produção de proteínas de estresse, supra-

20 regulação de anti-oxidantes, acúmulo de solutos compatíveis e interrupção de crescimento. Segundo o termo condições de "não estresse", da maneira aqui usada, são aquelas condições ambientais que permitem crescimento ideal de plantas. Os versados na técnica estão cientes das condições normais do solo e condições climáticas para um dado local. As plantas com condições ideais de

25 crescimento, (crescidas em condições de não estresse) rendem tipicamente em uma ordem crescente de preferência pelo menos 97 %, 95 %, 92 %, 90 %, 87 %, 85 %, 83 %, 80 %, 77 % ou 75 % da produção média de tal planta em um dado ambiente. A produção média pode ser calculada com base na coleta e/ou estação. Os versados na técnica estão cientes das produções de rendimento

médio de um cultivo.

O desempenho dos métodos da invenção fornece plantas crescidas em condições de secura (pouca), melhor tolerância à secura com relação às plantas controle crescidas em condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para melhorar a tolerância à secura em plantas crescidas em condições de secura (pouca), cujo método compreende modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo do tipo silky-1.

O desempenho dos métodos da invenção fornece plantas crescidas em condições de deficiência de nutriente, particularmente em condições de deficiência de nitrogênio, melhor tolerância aos estresses causados por deficiência de nutriente com relação às plantas controle. Portanto, de acordo com a presente invenção, é fornecido um método para melhorar a tolerância aos estresses causados por deficiência de nutriente, cujo método compreende modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo do tipo silky-1. A deficiência de nutriente pode resultar de uma necessidade de nutrientes tais como nitrogênio, fosfatos e outros compostos contendo fósforo, potássio, cálcio, magnésio, manganês, ferro e boro, entre outros.

O desempenho dos métodos da invenção fornece plantas crescidas em condições de estresse pelo sal, melhor tolerância ao sal com relação às plantas controle crescidas em condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para melhorar tolerância ao sal em plantas crescidas em condições de estresse pelo sal, cujo método compreende modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo do tipo silky-1. O termo estresse pelo sal não é restrito ao

sal comum (NaCl), mas pode ser qualquer um ou mais de: NaCl, KCl, LiCl, MgCl₂, CaCl₂, entre outros.

Com relação às características relacionadas ao rendimento, a presente invenção fornece um método para aumentar as características relacionadas ao rendimento de plantas com relação às plantas controle, cujo método compreende aumentar a expressão em uma planta com uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL da maneira aqui definida.

A presente invenção fornece um método para aumentar o rendimento, especialmente o rendimento da semente de plantas, com relação às plantas controle, cujo método compreende modular a expressão em uma planta com um ácido nucléico que codifica ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL, da maneira aqui definida.

Uma vez que as plantas transgênicas, de acordo com a presente invenção, apresentam mais características relacionadas ao rendimento e/ou rendimento, é provável que estas plantas exibam uma maior taxa de crescimento (durante pelo menos parte de seu ciclo de vida), com relação a a taxa de crescimento de plantas controle em um estágio correspondente em seu ciclo de vida.

A maior taxa de crescimento pode ser específica para uma ou mais partes de uma planta (incluindo sementes), ou pode ser substancialmente em toda a planta. As plantas com uma maior taxa de crescimento podem apresentar um ciclo de vida menor. O ciclo de vida de uma planta pode ser obtido para significar o tempo necessário para crescer a partir de uma semente madura seca até o estágio quando a planta produziu sementes maduras secas, similar ao material e partida. Este ciclo de vida pode ser influenciado por fatores tais como velocidade de germinação, vigor inicial, taxa de crescimento, índice de vegetação, período de floração 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 todo o ciclo de vida da planta. A maior taxa de crescimento durante os estágios iniciais no ciclo de vida de uma planta pode refletir maior vigor. O aumento na taxa de crescimento pode alterar o ciclo de coleta de uma

5 planta, o que permite que as plantas sejam semeadas posteriormente e/ou coletadas mais cedo do que seria de outra forma possível (um efeito similar pode ser obtido com período de floração inicial). Se a taxa de crescimento for suficientemente maior, ela pode permitir o semeio adicional de sementes da mesma espécie de planta (por exemplo, semeio e coleta de plantas de arroz

10 após semeio e coleta de plantas de arroz adicionais, todas em um período de crescimento convencional). Similarmente, se a taxa de crescimento for suficientemente maior, ela pode permitir o semeio adicional de sementes de diferentes espécies de plantas (por exemplo, o semeio e coleta de plantas de milho seguido por, por exemplo, o semeio e coleta de soja, batata ou qualquer

15 outra planta adequada). Os períodos adicionais de coleta do mesmo rizoma no caso de algumas plantas de cultivo também podem ser possíveis. Alterar o ciclo de coleta de uma planta pode levar a um aumento na produção anual de biomassa por metro quadrado (em virtude de um aumento no número de períodos (dito em um ano) que qualquer planta particular pode ser crescida e

20 coletada). Um aumento na taxa de crescimento também pode permitir o cultivo de plantas transgênicas em uma área geográfica mais ampla do que suas duplicatas tipo selvagem, uma vez que as limitações territoriais para crescer um cultivo são frequentemente determinadas por condições ambientais adversas tanto no período de plantio (estação inicial) quanto no

25 período de coleta (estação final). Tais condições adversas podem ser evitadas se o ciclo de coleta for menor. A taxa de crescimento pode ser determinada derivando vários parâmetros a partir de curvas de crescimento, tais parâmetros podem ser: T-Mid (o tempo obtido para plantas alcançarem 50 % de seu tamanho máximo) e T-90 (tempo obtido para as plantas atingirem 90

% de seu tamanho máximo), entre outros.

De acordo com uma característica preferida da presente invenção, o desempenho dos métodos da invenção fornece plantas com uma maior taxa de crescimento com relação às plantas controle. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar a taxa de crescimento de plantas, cujo método compreende aumentar a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo BRXL, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL, da maneira aqui definida.

O desempenho dos métodos da invenção fornece plantas crescidas em condições de não estresse ou em condições de pouco estresse com mais características relacionadas ao rendimento, com relação às plantas controle crescidas em condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar as características relacionadas ao rendimento em plantas crescidas em condições de não estresse ou em condições de pouco estresse, cujo método compreende aumentar a expressão em uma planta com uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL.

O desempenho dos métodos da invenção fornece plantas crescidas em condições de menor disponibilidade de nutriente, particularmente em condições de menor disponibilidade de nitrogênio, com mais características relacionadas ao rendimento e/ou rendimento com relação às plantas controle crescidas em condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar as características relacionadas ao rendimento em plantas crescidas em condições de menor disponibilidade de nutriente, preferivelmente menor disponibilidade de nitrogênio, cujo método compreende aumentar a expressão em uma planta com uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL,

ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL. A menor disponibilidade de nutriente pode resultar de uma deficiência ou excesso de nutrientes tais como nitrogênio, fosfatos e outros compostos contendo fósforo, potássio, cálcio, cádmio, magnésio, manganês, ferro e boro, entre outros. Preferivelmente, a menor disponibilidade de nutriente é a menor disponibilidade de nitrogênio.

O desempenho dos métodos da invenção fornece plantas crescidas em condições de estresse pelo sal, maior rendimento com relação às plantas controle crescidas em condições comparáveis. Portanto, de acordo com a presente invenção, é fornecido um método para aumentar o rendimento em plantas crescidas em condições de estresse pelo sal, cujo método compreende modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL. O termo estresse pelo sal não é restrito ao sal comum (NaCl), mas pode ser qualquer um ou mais de: NaCl, KCl, LiCl, MgCl₂, CaCl₂, entre outros.

A presente invenção inclui plantas, ou partes da mesmas, (incluindo sementes) obteníveis pelos métodos de acordo com a presente invenção. As plantas, ou partes da mesmas, compreendem um transgene de ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL, da maneira definida anteriormente, operavelmente ligado a um promotor que funciona em plantas.

A invenção também fornece construtos e vetores genéticos para facilitar a introdução e/ou expressão em plantas de ácidos nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL. O construto de genes pode ser

inserido em vetores, que podem ser comercialmente disponíveis, adequado para transformar em plantas e adequado para a expressão do gene de interesse nas células transformadas. A invenção também fornece o uso de um construto de gene da maneira aqui definida nos métodos da invenção.

5 Mais especificamente, a presente invenção fornece um construto compreendendo:

(a) um ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL, da maneira definida anteriormente;

(b) uma ou mais sequências controle capazes de direcionar a expressão da sequência de ácidos nucleicos de (a) e, opcionalmente

(c) uma sequência de finalização de transcrição.

15 Preferivelmente, o ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL, é da maneira definida anteriormente. O termo "sequência controle" e "sequência de finalização" são da maneira aqui definida.

20 Com relação aos polipeptídeos BRXL, preferivelmente, uma das sequências controle de um construto é um promotor constitutivo isolado de um genoma de planta. Um exemplo de um promotor constitutivo é um promotor GOS2, preferivelmente um promotor GOS2 de arroz, acima de tudo preferivelmente uma sequência de GOS2 da maneira representada por SEQ
25 ID NO: 87.

As plantas são transformadas com um vetor compreendendo qualquer dos ácidos nucleicos descritos anteriormente. Os versados na técnica estão cientes dos elementos genéticos que têm que estar presentes no vetor, afim de transformar selecionar e propagar satisfatoriamente as células

hospedeiras contendo a sequência de interesse. A sequência de interesse está operavelmente ligada a uma ou mais sequências controle (pelo menos a um promotor).

5 Vantajosamente, qualquer tipo de promotor, seja natural ou sintético, pode ser usado para direcionar a expressão da sequência de ácidos nucleicos, mas preferivelmente o promotor é de origem vegetal. Um promotor constitutivo é particularmente usado nos métodos. Preferivelmente, o promotor constitutivo é também um promotor ubíquo de concentração de meio. Ver a seção de "definições" aqui para as definições dos vários tipos de
10 promotor.

Com relação aos polipeptídeos BRXL, vantajosamente, qualquer tipo de promotor, seja natural ou sintético, pode ser usado para aumenta a expressão da sequência de ácidos nucleicos. Um promotor constitutivo é particularmente usado nos métodos, preferivelmente um
15 promotor constitutivo isolado de um genoma de planta. O promotor de planta constitutivo direciona a expressão de uma sequência codificante em um nível que está em todos os exemplos a seguir, que foram obtidos no control de um promotor viral 35S CaMV. Um exemplo de um promotor como este é um promotor GOS2 da maneira representada por SEQ ID NO: 87.

20 No caso de genes BRXL, os promotores específicos de órgão, por exemplo, para a expressão preferida em folhas, caules, tubérculos, meristemas, sementes, são usados na realização dos métodos da invenção. Promotores regulados pelo desenvolvimento e indutíveis também são usados na realização dos métodos da invenção. Ver a seção "Definições" aqui para
25 definições dos vários tipos de promotores.

Com relação aos polipeptídeos do tipo alfin, pode ser evidente que a aplicabilidade da presente invenção não esteja restrita ao ácido nucléico que codifica polipeptídeo do tipo alfin representado por SEQ ID NO: 1 ou SEQ ID NO: 3, nem que a aplicabilidade da invenção esteja restrita à

expressão de um ácido nucléico que codifica o polipeptídeo do tipo alfin quando direcionada por um promotor constitutivo.

O promotor constitutivo é preferivelmente um promotor de média resistência, mais preferivelmente selecionado de um promotor derivado de planta, tal como um promotor GOS2, mais preferivelmente é o promotor GOS2 de arroz. Preferivelmente de maneira adicional, o promotor constitutivo é representado por uma sequência de ácidos nucleicos substancialmente similar a SEQ ID NO: 7, acima de tudo preferivelmente, o promotor constitutivo é da maneira representada por SEQ ID NO: 7. Ver a seção de “definições” aqui para exemplos adicionais de promotores constitutivos.

Opcionalmente, uma ou mais sequências finalizadoras pode ser usada no construto introduzido em uma planta. Preferivelmente, o construto compreende um cassete de expressão compreendendo um promotor (GOS2), substancialmente similar a SEQ ID NO: 7, e o ácido nucléico que codifica o polipeptídeo do tipo alfin.

Com relação aos polipeptídeos YRP, pode ser evidente que a aplicabilidade da presente invenção não esteja restrita ao ácido nucléico que codifica o polipeptídeo YRP representado por SEQ ID NO: 10 ou SEQ ID NO: 12, nem que a aplicabilidade da invenção esteja restrita a expressão de um ácido nucléico que codifica o polipeptídeo YRP quando direcionado por um promotor constitutivo.

O promotor constitutivo é preferivelmente um promotor de média resistência, mais preferivelmente selecionado de um promotor derivado de planta, tal como um promotor GOS2, mais preferivelmente é o promotor GOS2 de arroz. Preferivelmente de maneira adicional, o promotor constitutivo é representado por uma sequência de ácidos nucleicos substancialmente similar a SEQ ID NO: 16, acima de tudo preferivelmente, o promotor constitutivo é da maneira representada por SEQ ID NO: 16. Ver a

seção de “definições” aqui para exemplos adicionais de promotores constitutivos.

Opcionalmente, uma ou mais sequências finalizadoras podem ser usadas no construto introduzido em uma planta. Preferivelmente, o construto compreende um cassete de expressão compreendendo um promotor (GOS2), substancialmente similar a SEQ ID NO: 16, e o ácido nucleico que codifica o polipeptídeo YRP.

Com relação aos polipeptídeos BRXL, pode ser evidente que a aplicabilidade da presente invenção não esteja restrita a uma sequência de ácidos nucleicos que codifica o polipeptídeo BRXL, da maneira representadas por SEQ ID NO: 17, nem que a aplicabilidade da invenção esteja restrita a expressão de uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL quando direcionado por um promotor constitutivo.

Opcionalmente, uma ou mais sequências finalizadoras podem ser usadas no construto introduzido em uma planta.

Com relação aos polipeptídeos do tipo silky, pode ser evidente que a aplicabilidade da presente invenção não esteja restrita ao ácido nucleico que codifica polipeptídeo do tipo silky-1 representado por SEQ ID NO: 90, SEQ ID NO: 92 ou SEQ ID NO: 94, nem que a aplicabilidade da invenção esteja restrita a expressão de ácido nucleico que codifica um polipeptídeo do tipo silky-1 quando direcionado por um promotor constitutivo.

O promotor constitutivo é preferivelmente um promotor de média resistência, mais preferivelmente selecionado de um promotor derivado de planta, tal como um promotor GOS2, mais preferivelmente é o promotor GOS2 de arroz. Preferivelmente de maneira adicional, o promotor constitutivo é representado por uma sequência de ácidos nucleicos substancialmente similar a SEQ ID NO: 96, acima de tudo preferivelmente, o promotor constitutivo é da maneira representada por SEQ ID NO: 96. Ver a seção de “definições” aqui para exemplos adicionais de promotores

constitutivos.

Opcionalmente, uma ou mais sequências finalizadoras podem ser usadas no construto introduzido em uma planta. Preferivelmente, o construto compreende um cassete de expressão compreendendo um promotor (GOS2), substancialmente similar a SEQ ID NO: 96, e o ácido nucléico que codifica o polipeptídeo do tipo silky-1.

Com relação aos polipeptídeos ARP6, pode ser evidente que a aplicabilidade da presente invenção não esteja restrita ao ácido nucléico que codifica o polipeptídeo ARP6 representado por SEQ ID NO: 101, nem que a aplicabilidade da invenção esteja restrita a expressão de um ácido nucléico que codifica o polipeptídeo ARP6 quando direcionado por um promotor constitutivo.

O promotor constitutivo é preferivelmente um promotor de média resistência, mais preferivelmente selecionado de um promotor derivado de planta, tal como um promotor GOS2, mais preferivelmente é o promotor GOS2 de arroz. Preferivelmente de maneira adicional, o promotor constitutivo é representado por uma sequência de ácidos nucleicos substancialmente similar a SEQ ID NO: 115, acima de tudo preferivelmente, o promotor constitutivo é da maneira representada por SEQ ID NO: 115. Ver a seção de “definições” aqui para exemplos adicionais de promotores constitutivos.

Opcionalmente, uma ou mais sequências finalizadoras podem ser usadas no construto introduzido em uma planta. Preferivelmente, o construto compreende um cassete de expressão compreendendo um promotor GOS2, substancialmente similar a SEQ ID NO: 115, e o ácido nucléico que codifica o polipeptídeo ARP6.

Com relação aos polipeptídeos POP, pode ser evidente que a aplicabilidade da presente invenção não esteja restrita ao ácido nucléico que codifica o polipeptídeo de POP representadas por SEQ ID NO: 116, nem que

a aplicabilidade da invenção esteja restrita a expressão de um ácido nucléico que codifica o polipeptídeo de POP quando direcionado por um promotor constitutivo.

O promotor constitutivo é preferivelmente um promotor de
5 média resistência, mais preferivelmente selecionado de um promotor derivado de planta, tal como um promotor GOS2, mais preferivelmente é o promotor GOS2 de arroz. Preferivelmente de maneira adicional, o promotor constitutivo é representado por uma sequência de ácidos nucleicos substancialmente similar a SEQ ID NO: 134, acima de tudo preferivelmente,
10 o promotor constitutivo é da maneira representada por SEQ ID NO: 134. Ver a seção de “definições” aqui para exemplos adicionais de promotores constitutivos.

Opcionalmente, uma ou mais sequências finalizadoras podem ser usadas no construto introduzido em uma planta. Preferivelmente, o
15 construto compreende um cassete de expressão compreendendo um promotor GOS2, substancialmente similar a SEQ ID NO: 134, e o ácido nucléico que codifica o polipeptídeo de POP.

Com relação aos polipeptídeos CRL, pode ser evidente que a aplicabilidade da presente invenção não esteja restrita ao polipeptídeo CRL
20 que codifica ácido nucléico representadas por SEQ ID NO: 155, nem que a aplicabilidade da invenção esteja restrita a expressão de um ácido nucléico que codifica o polipeptídeo CRL quando direcionado por um promotor constitutivo, ou quando direcionado por um promotor específico de raiz.

O promotor constitutivo é preferivelmente um promotor de
25 média resistência, mais preferivelmente selecionado de um promotor derivado de planta, tal como um promotor GOS2, mais preferivelmente é o promotor GOS2 de arroz. Preferivelmente de maneira adicional, o promotor constitutivo é representado por uma sequência de ácidos nucleicos substancialmente similar a SEQ ID NO: 213, acima de tudo preferivelmente,

o promotor constitutivo é da maneira representada por SEQ ID NO: 213. Ver a seção de “definições” aqui para exemplos adicionais de promotores constitutivos.

Opcionalmente, uma ou mais sequências finalizadoras podem ser usadas no construto introduzido em uma planta. Preferivelmente, o construto compreende um cassete de expressão compreendendo um promotor (GOS2), substancialmente similar a SEQ ID NO: 213, e o ácido nucléico que codifica o polipeptídeo CRL.

Os elementos regulatórios adicionais podem incluir melhoradores transcrpcionais, bem como translacionais. Os versados na técnica estão cientes das sequências finalizadoras e melhoradoras que podem ser adequadas para uso na realização da invenção. Uma sequência de intron também pode ser adicionada na região não traduzida 5' (UTR) ou na sequência codificante para aumentar a quantidade da mensagem madura que se acumula no citossol, da maneira descrita na seção de definições. Outras sequências controle (além das sequências promotoras, melhoradoras, silenciadoras, sequências de intron, regiões UTR 3' e/ou UTR 5') podem ser elementos estabilizadores de proteína e/ou RNA. Tais sequências são conhecidas ou podem ser facilmente obtidas pelos versados na técnica.

Os construtos genéticos da invenção podem incluir adicionalmente uma sequência de replicação de origem que é exigida para manutenção e/ou replicação em um tipo de célula específica. Um exemplo é quando um construto genético é exigido para ser mantido em uma célula bacteriana como um elemento genético episomal (por exemplo, molécula de plasmídeo ou cosmídeo). As origens de replicação preferidas incluem, mas sem limitação, a fl-ori e colE1.

Para a deleção da transferência satisfatória das sequências de ácidos nucleicos da maneira usada nos métodos da invenção e/ou seleção de plantas transgênicas compreendendo estes ácidos nucleicos, é vantajoso usar

genes marcadores (ou genes repórteres). Portanto, o construto genético pode compreender opcionalmente um gene de marcador selecionável. Os marcadores selecionáveis são descritos com mais detalhes aqui na seção “definições”. Os genes marcadores podem ser removidos ou excisados da célula transgênica uma vez que eles não são mais necessários. As técnicas para remoção de marcador são conhecidas na tecnologia, as técnicas usadas são descritas anteriormente na seção de definições.

Sabe-se que, mediante a integração estável ou transiente de sequências de ácidos nucleicos nas células de planta, apenas uma minoria das células obtém o DNA estrangeiro e, se desejado, integra-o em seu genoma, dependendo do vetor de expressão usado e o técnica de transfecção usada. Para identificar e selecionar estes integrantes, um gene que codifica um marcador selecionável (tais como aqueles descritos anteriormente) é em geral introduzido nas células hospedeiras junto com o gene de interesse. Estes marcadores, por exemplo, podem ser usados em mutantes nos quais estes genes não são funcionais, por exemplo, por deleção por métodos convencionais. Além disso, as moléculas com sequência de ácidos nucleicos que codifica um marcador selecionável podem ser introduzidas em uma célula hospedeira no mesmo vetor que compreende a sequência que codifica os polipeptídeos da invenção ou usadas nos métodos da invenção, ou ainda em um vetor separado. As células que foram estavelmente transfectadas com a sequência de ácidos nucleicos introduzida podem ser identificadas, por exemplo, por seleção (por exemplo, células que integraram o marcador selecionável sobrevivem, ao passo que as outras células morrem). Os genes marcadores podem ser removidos ou excisados da célula transgênica uma vez que eles não são mais necessários. As técnicas para a remoção de gene marcador são conhecidas na tecnologia, as técnicas usadas são descritas anteriormente na seção de definições.

A invenção também fornece um método para a produção de

plantas transgênicas com melhor tolerância ao estresse abiótico e/ou características relacionadas ao rendimento com relação às plantas controle, compreendendo a introdução e expressão em uma planta de qualquer ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL, da maneira aqui definida anteriormente.

Mais especificamente, a presente invenção fornece um método para a produção de plantas transgênicas com melhor tolerância ao estresse abiótico, particularmente maior tolerância à secura (pouco), cujo método compreende:

(i) introduzir e expressar em uma planta ou célula de planta um ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo do tipo silky-1; e

(ii) cultivar a célula da planta em condições de estresse abiótico.

O ácido nucléico de (i) pode ser qualquer dos ácidos nucleicos capaz de codificar um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo do tipo silky-1, da maneira aqui definida.

Mais especificamente, a presente invenção também fornece um método para a produção de plantas transgênicas com mais características relacionadas ao rendimento com relação às plantas controle, cujo método compreende:

(i) introduzir e expressar em uma planta, parte da planta, ou célula de planta uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL; e

(ii) cultivar a célula da planta, parte da planta ou planta em condições que promovem crescimento e desenvolvimento da planta.

A sequência de ácidos nucleicos de (i) pode ser quaisquer das

sequências de ácidos nucleicos capaz de codificar um polipeptídeo BRXL da maneira aqui definida.

Mais especificamente, a presente invenção também fornece um método para a produção de plantas transgênicas com melhores características relacionadas ao rendimento, particularmente maior rendimento da semente, cujo método compreende:

(i) introduzir e expressar em uma planta ou célula de planta um ácido nucléico que codifica polipeptídeo ARP6; e

(ii) cultivar a célula da planta em condições que promovem crescimento e desenvolvimento da planta.

O ácido nucléico de (i) pode ser qualquer dos ácidos nucleicos capaz de codificar um polipeptídeo ARP6 da maneira aqui definida.

Mais especificamente, a presente invenção também fornece um método para a produção de plantas transgênicas com melhores características relacionadas ao rendimento, em particular maior rendimento cujo método compreende:

(i) introduzir e expressar em uma planta ou célula de planta um ácido nucléico que codifica o polipeptídeo POP; e

(ii) cultivar a célula da planta em condições que promovem crescimento e desenvolvimento da planta.

O ácido nucléico de (i) pode ser qualquer dos ácidos nucleicos capaz de codificar um polipeptídeo de POP da maneira aqui definida.

Mais especificamente, a presente invenção também fornece um método para a produção de plantas transgênicas com melhores características relacionadas ao rendimento, particularmente maior rendimento (semente) e/ou índice da colheita, cujo método compreende:

(i) introduzir e expressar em uma planta ou célula de planta um ácido nucléico que codifica o polipeptídeo CRL e

(ii) cultivar a célula da planta em condições que promovem

crescimento e desenvolvimento da planta.

O ácido nucléico de (i) pode ser qualquer dos ácidos nucleicos capaz de codificar um polipeptídeo CRL da maneira aqui definida.

O ácido nucléico pode ser introduzido diretamente em uma célula de planta 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 nucléico é preferivelmente introduzido em uma planta por transformação. O termo "transformação" é descrito com mais detalhes aqui na seção "definições".

As células de planta geneticamente modificadas podem ser regeneradas por meio de todos os métodos com os quais os versados na técnica são familiares. Os métodos adequados podem ser observados nas publicações anteriormente mencionadas por S. D. Kung e R. Wu, Potrykus ou Hofgen e Willmitzer.

Em geral, após a transformação, as células de planta ou agrupamentos de célula são selecionadas com relação à presença de um ou mais marcadores que são codificados por genes expressíveis em planta co-transferidos com o gene de interesse, após o que o material transformado é regenerado em uma planta completa. Para selecionar plantas transformadas, o material de planta obtido na transformação é, via de regra, submetido a condições seletivas, de maneira tal que as plantas transformadas possam ser distinguidas das plantas não transformadas. Por exemplo, as sementes obtidas da maneira anteriormente descrita podem ser plantadas e, após um período de crescimento inicial, submetidas a uma seleção adequada por aspersão. Uma possibilidade adicional consiste em crescer as sementes, se apropriado após a esterilização, em placas de ágar usando um agente de seleção adequado, de maneira tal que apenas as sementes transformadas possam crescer em plantas. Alternativamente, as plantas transformadas são selecionadas com relação à presença de um marcador selecionável, tais como aqueles descritos

anteriormente.

Após a transferência e regeneração do DNA, as plantas putativamente transformadas também pode ser avaliadas, por exemplo, usando análise Southern, com relação à presença do gene de interesse, número de cópias e/ou organização genômica. Alternativamente, ou
5 adicionalmente, os níveis de expressão do DNA recém introduzido podem ser monitorados usando análise Northern e/ou Western, ambas as técnicas sendo bem conhecidas pelos versados na técnica.

As plantas transformadas geradas podem ser propagadas por
10 uma variedade de meios, tal como por propagação clonal ou técnicas clássicas de cultivo. Por exemplo, uma planta transformada de primeira geração (ou T1) pode ser auto-fecundada e os transformantes homozigotos da segunda geração (ou T2) selecionados, e as plantas T2 podem então ser adicionalmente propagadas por meio de técnicas clássicas de cultivo. Os organismos
15 transformados gerados podem ter uma variedade de formas. Por exemplo, eles podem ser quimeras de células transformadas e células não transformadas; transformantes clonais (por exemplo, todas as células transformadas para conter o cassete de expressão); enxertos de tecidos transformados e não transformados (por exemplo, em plantas, um rizoma transformado enxertado
20 em um descendente não transformado).

A presente invenção estende-se a qualquer célula de planta ou planta produzida por qualquer dos métodos aqui descritos, e a todas as partes da planta e propágulos da mesma. A presente invenção estende-se adicionalmente para incluir a progênie de uma célula primária transformada
25 ou transfectada, tecido, órgão ou a planta completa que foi produzida por qualquer dos métodos anteriormente mencionados, a exigência sendo apenas de que a progênie exiba a(s) mesma(s) característica(s) genotípica(s) e/ou fenotípica(s) daquelas produzidas pelos ancestrais nos métodos de acordo com a invenção.

A invenção também inclui células hospedeiras contendo um ácido nucléico isolado que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL, da maneira aqui definida anteriormente. As células hospedeiras preferidas de acordo com a invenção são células de plantas. As plantas hospedeiras com os ácidos nucleicos ou o vetor usados no método de acordo com a invenção, o cassete de expressão ou construto ou vetor são, a princípio, vantajosamente todas plantas que são capazes de sintetizar os polipeptídeos usados no método inventivo.

Os métodos da invenção são vantajosamente aplicáveis a qualquer planta. As plantas que são particularmente usadas nos métodos da invenção incluem todas as plantas que pertencem à superfamília *Viridiplantae*, em particular plantas monocotiledôneas e dicotiledôneas, incluindo forrageiras de legumes forrageiros, plantas ornamentais, culturas de alimento, árvores ou arbustos. De acordo com uma modalidade preferida da presente invenção, a planta é uma planta de cultivo. Exemplos de plantas de cultivo incluem soja, girassol, canola, alfafa, canola, linhaça, algodão, tomate, batata e tabaco. Preferivelmente de maneira adicional, a planta é uma planta monocotiledônea. Exemplos de plantas monocotiledôneas incluem cana de açúcar. Mais preferivelmente, a planta é um cereal. Exemplos de cereais incluem arroz, milho, trigo, cevada, milheto, centeio, tritcale, sorgo, trigo de emmer, espelta, *Secale*, trigo einkorn, cereal teff, milo e aveias.

A invenção também estende-se a Partes cultiváveis de uma planta tais como, mas sem limitação, sementes, folhas, frutas, flores, caules, raízes, rizomas, tubérculos e bulbões, cujas Partes cultiváveis compreendem um ácido nucléico recombinante que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um

polipeptídeo CRL. A invenção, além disso, diz respeito a produtos derivados, preferivelmente derivados diretamente, de uma parte coletável de uma planta como esta, tais como precipitados secos ou pós, óleo, gordura e ácidos graxos, amido ou proteínas.

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

10 Da maneira anteriormente mencionada, um método preferido para modular a expressão de um ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL, é introduzir e expressar em uma planta um
15 ácido nucléico que codifica um polipeptídeo do tipo alfin, ou um polipeptídeo YRP, ou um polipeptídeo BRXL, ou um polipeptídeo do tipo silky-1, ou um polipeptídeo ARP6, ou um polipeptídeo de POP, ou um polipeptídeo CRL; entretanto os efeitos de realizar o método, isto é, melhorar a tolerância ao estresse abiótico também podem ser atingidos usando outras técnicas bem
20 conhecidas incluindo, mas sem limitação, marcação da ativação do DNA-T, TILLING, recombinação homóloga. Uma descrição da mesmas técnicas é fornecida a seção de definições.

A presente invenção também inclui o uso de ácidos nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou
25 polipeptídeos do tipo silky-1, da maneira aqui descrita e o uso destes polipeptídeos do tipo alfin para melhorar qualquer dos estresses abióticos em plantas anteriormente mencionados.

A presente invenção também inclui o uso de sequências de ácidos nucleicos que codificam polipeptídeos BRXL da maneira aqui descrita

e o uso destes polipeptídeos BRXL para aumentar qualquer das características relacionadas ao rendimento em plantas anteriormente mencionadas, em condições normais de crescimento, em condições de crescimento no estresse abiótico (preferivelmente condições de crescimento em estresse osmótico), e em condições de crescimento com menor disponibilidade de nutriente, preferivelmente em condições com menor disponibilidade de nitrogênio.

A presente invenção também inclui o uso de ácidos nucleicos que codificam polipeptídeos ARP6 da maneira aqui descrita e o uso destes polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, para melhorar qualquer das características relacionadas ao rendimento em plantas anteriormente mencionadas.

Os ácidos nucleicos que codificam polipeptídeo do tipo alfin, ou polipeptídeo YRP, ou polipeptídeo BRXL, ou polipeptídeo do tipo silky-1, ou polipeptídeo ARP6, ou polipeptídeo de POP, ou polipeptídeo CRL aqui descritos, ou os próprios polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, podem ser usados em programas de cultivo nos quais um marcador de DNA é identificado e que pode estar geneticamente ligado ao gene que codifica polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL. Os ácidos nucleicos/genes, ou os próprios polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, podem ser usados para definir um marcador molecular. Este marcador de DNA ou proteína pode então ser usado em programas de cultivo para selecionar plantas com melhor tolerância ao estresse abiótico da maneira aqui definida anteriormente nos métodos da invenção.

Variantes alélicas de um ácido nucléico/gene que codifica

polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, também podem ser usados em programas de cultivo associados a marcador. Tais programas de cultivo exigem algumas vezes a introdução de variação alélica por tratamento mutagênico das plantas usando, por exemplo, mutagênese EMS; alternativamente, o programa pode iniciar com uma coleção de variantes alélicas da assim denominada origem "natural" causada não intencionalmente. A identificação de variantes alélicas então ocorre, por exemplo, por PCR. Isto é seguido por uma etapa para a seleção de variantes alélicas superiores da sequência em questão e que fornecem melhor tolerância ao estresse que pode ser manifestado como o maior rendimento. A seleção é tipicamente realizada monitorando o desempenho do crescimento de plantas contendo diferentes variantes alélicas da sequência em questão. O desempenho do crescimento pode ser monitorado em uma estufa ou no campo. As etapas adicionais incluem cruzar plantas nas quais o variante alélico superior foi identificado com uma outra planta. Isto pode ser usado, por exemplo, para fazer uma combinação de características fenotípicas de interesse.

Os ácidos nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, também pode ser usados como sondas para mapear genética e fisicamente os genes nos quais eles fazem parte, e como marcadores para características ligadas àqueles genes. Tal informação pode ser usada no cultivo de planta a fim de desenvolver linhagens com fenótipos desejados. Tal uso de ácidos nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, exige apenas uma sequência de ácidos nucleicos de pelo menos 15 nucleotídeos de comprimento. Os ácidos

nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, podem ser usados como marcadores de polimorfismo de tamanho de fragmento de restrição (RFLP).

- 5 Southern blots (Sambrook J, Fritsch EF e Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de DNA genômico de planta digerido por restrição podem ser sondados com os ácidos nucleicos que codificam polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky-1, ou polipeptídeos ARP6, ou polipeptídeos POP, ou polipeptídeos CRL. Os padrões de banda resultante podem então ser submetidos a análise genética usando programas de computador tal como marcador Map (Lander et al. (1987) Genomics 1 : 174-181), a fim de
- 10 construir um mapa genético. Além do mais, os ácidos nucleicos podem ser usados para sondar Southern blots contendo DNAs genômicos tratados com endonuclease de restrição de um conjunto de indivíduos que representa
- 15 ancestrais e progênie de um cruzamento genético definido. A segregação dos polimorfismos de DNA é observada e usada para calcular a posição do ácido nucléico que codifica polipeptídeos do tipo alfin, ou polipeptídeos YRP, ou polipeptídeos BRXL, ou polipeptídeos do tipo silky-1, ou polipeptídeos
- 20 ARP6, ou polipeptídeos POP, ou polipeptídeos CRL, no mapa genético previamente obtido usando esta população (Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

- A produção e uso de sondas derivadas de gene de planta para uso no mapeamento genético é descrito em Bernatzky e Tanksley (1986)
- 25 Plant Mol. Biol. Reporter 4: 37-41. Inúmeras publicações descrevem o mapeamento genético de clones de DNA específico usando a metodologia anteriormente representada ou variações da mesmas. Por exemplo, populações de intercruzamento F2, populações de retrocruzamento, populações que foram acasaladas aleatoriamente, linhagens isogênicas próximas, e outros conjuntos

de indivíduos podem ser usados para mapeamento. Tais metodologias são bem conhecidas pelos versados na técnica.

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

Em uma outra modalidade, as sondas de ácido nucléico podem ser usadas no mapeamento de hibridização *in situ* por fluorescência direta (FISH) (Trask (1991) Trends Genet. 7:149-154). Embora os métodos de mapeamento de FISH atuais favoreçam o uso de clones maiores (vários kb a vários mil kb; ver Laan et al. (1995) Genome Res. 5:13-20), melhorias na sensibilidade podem permitir o desempenho do mapeamento de FISH usando sondas menores.

Uma variedade de métodos a base de amplificação de ácido nucléico para mapeamento genético e físico pode ser realizada usando os ácidos nucleicos. Exemplos incluem amplificação específica de alelo (Kazazian (1989) J. Lab. Clin. Med 111:95-96), polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325-332), ligação específica de alelo (Landegren et al. (1988) Science 241:1077-1080), 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 do tipo Happy (Dear e Cook (1989) Nucleic Acid Res. 17:6795-6807). Para estes métodos, a sequência de um ácido nucléico é usada para desenhar e produzir pares de oligonucleotídeo iniciador para uso na reação de amplificação ou nas reações de extensão do oligonucleotídeo iniciador. O desenho de tais oligonucleotídeos iniciadores é bem conhecido pelos versados na técnica. Em métodos que empregam mapeamento genético a base de PCR pode ser necessário para identificar as sequências de DNA as diferenças entre os ancestrais do cruzamento do

mapeamento na região que corresponde à sequência imediata de ácidos nucleicos. Entretanto, em geral, isto não é necessário para os métodos de mapeamento.

Os métodos de acordo com a presente invenção resultam em plantas com melhor tolerância ao estresse abiótico e/ou melhores características relacionadas ao rendimento, da maneira aqui descrita anteriormente. Estas características também podem ser combinadas com outras características economicamente vantajosas, tais como melhores características de tolerância adicional ao estresse abiótico ou biótico, melhores características relacionadas ao rendimento, melhores características relacionadas ao rendimento, tolerância aos herbicidas, insetidas, características que modificam várias características da arquitetura e/ou características bioquímicas e/ou fisiológicas.

Itens

1. Polipectídeos do tipo alfin

1. Método para melhorar a tolerância ao estresse abiótico em plantas modulando a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo do tipo alfin ou um ortólogo ou parólogo deste.

2. Método de acordo com item 1, em que a dita expressão modulada é afetada introduzindo e expressando em uma planta um ácido nucléico que codifica polipeptídeo do tipo alfin.

3. Método de acordo com itens 2 ou 3, em que o dito ácido nucléico que codifica um polipeptídeo do tipo alfin codifica qualquer uma das proteínas listadas na tabela A1, ou é uma porção de um ácido nucléico como este, ou um ácido nucléico capaz de hibridizar em um ácido nucléico como este.

4. Método de acordo com qualquer um dos itens 1 a 4, em que a dita sequência de ácidos nucleicos codifica um ortólogo ou parólogo de quaisquer das proteínas fornecidas na tabela A1.

5. Método de acordo com itens 3 ou 4, em que o dito ácido nucléico é operavelmente ligado a um promotor constitutivo, preferivelmente a um promotor GOS2, acima de tudo preferivelmente a um promotor GOS2 de arroz.

5 6. Método de acordo com qualquer um dos itens 1 a 5, em que o dito ácido nucléico que codifica um polipeptídeo do tipo alfin é da cepa *Allium*.

7. Planta ou parte da mesma, incluindo sementes, obteníveis por um método de acordo com qualquer um dos itens 1 a 6, em que a dita
10 planta ou parte da mesma compreende um ácido nucléico recombinante que codifica um polipeptídeo do tipo alfin.

8. Construto compreendendo:

(i) ácido nucléico que codifica um polipeptídeo do tipo alfin da maneira definida nos itens 1 ou 2;

15 (ii) uma ou mais sequências controle capazes de direcionar a expressão da sequência de ácidos nucleicos de (a); e opcionalmente

(iii) uma sequência de finalização de transcrição.

9. Construto de acordo com item 9, em que uma das ditas sequências controle é um promotor constitutivo, preferivelmente um promotor
20 GOS2, acima de tudo preferivelmente um promotor GOS2 de arroz.

10. Uso de um construto de acordo com item 8 ou 9 em um método para preparar plantas com maior tolerância ao estresse abiótico com relação às plantas controle.

25 11. Planta, parte da planta ou célula de planta transformadas com um construto de acordo com item 8 ou 9.

12. Método para a produção de uma planta transgênica com maior tolerância ao estresse abiótico com relação às plantas controle, compreendendo:

(i) introduzir e expressar em uma planta um ácido nucléico que

codifica um polipeptídeo do tipo alfin; e

(ii) cultivar a célula da planta em condições que promovem o estresse abiótico.

13. Planta transgênica com tolerância ao estresse abiótico, com relação às plantas controle, que resulta na expressão modulada de um ácido nucléico que codifica um polipeptídeo do tipo alfin, ou uma célula transgênica de planta derivada da dita planta transgênica.

14. Planta transgênica de acordo com o item 7, 11 ou 13, ou uma célula transgênica de planta derivada da mesma, em que a dita planta é uma planta de cultivo ou uma monocotiledônea ou um cereal, tais como arroz, milho, trigo, cevada, milheto, centeio, tritcale, sorgo, cana de açúcar, trigo de emmer, espelta, secale, trigo de einkorn, cereal teff, milo e aveias.

15. Partes cultiváveis de uma planta de acordo com o item 14, em que as ditas Partes cultiváveis são preferivelmente biomassa dos ramos e/ou sementes.

16. Produtos derivados de uma planta de acordo com o item 14 e/ou de Partes cultiváveis de uma planta de acordo com o item 15.

17. Uso de um ácido nucléico que codifica um polipeptídeo do tipo alfin para aumentar o rendimento, particularmente para aumentar a tolerância ao estresse abiótico, com relação às plantas controle.

2. Polipeptídeos YRP

1. Método para melhorar a tolerância ao estresse abiótico em plantas modulando a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo YRP ou um ortólogo ou parólogo deste.

2. Método de acordo com o item 1, em que a dita expressão modulada é afetada introduzindo e expressando em uma planta um ácido nucléico que codifica polipeptídeo YRP.

3. Método de acordo com o itens 2 ou 3, em que o dito ácido nucléico que codifica um polipeptídeo YRP codifica qualquer uma das

proteínas listadas na tabela A2 ou é uma porção de um ácido nucléico como este, ou um ácido nucléico capaz de hibridizar com um ácido nucléico como este.

4. Método de acordo com qualquer um dos itens 1 a 4, em que a dita sequência de ácidos nucleicos codifica um ortólogo ou parólogo de quaisquer das proteínas fornecidas na tabela A2.

5. Método de acordo com o itens 3 ou 4, em que o dito ácido nucléico é operavelmente ligado a um promotor constitutivo, preferivelmente a um promotor GOS2, acima de tudo preferivelmente a um promotor GOS2 de arroz.

6. Método de acordo com qualquer um dos itens 1 a 5, em que o dito ácido nucléico que codifica um polipeptídeo YRP é de *Hordeum vulgare*.

7. Planta ou parte da mesma, incluindo sementes, obteníveis por um método de acordo com qualquer um dos itens 1 a 6, em que a dita planta ou parte da mesma compreende um ácido nucléico recombinante que codifica um polipeptídeo YRP.

8. Construto compreendendo:

(i) ácido nucléico que codifica um polipeptídeo YRP da maneira definida nos itens 1 ou 2;

(ii) uma ou mais sequências controle capazes de direcionar a expressão da sequência de ácidos nucleicos de (a); e opcionalmente

(iii) uma sequência de finalização de transcrição.

9. Construto de acordo com o item 9, em que uma das ditas sequências controle é um promotor constitutivo, preferivelmente um promotor GOS2, acima de tudo preferivelmente um promotor GOS2 de arroz.

10. Uso de um construto de acordo com o item 8 ou 9 em um método para preparar plantas com maior tolerância ao estresse abiótico com relação às plantas controle.

11. Planta, parte da planta ou célula de planta transformadas com um construto de acordo com o item 8 ou 9.

12. Método para a produção de uma planta transgênica com maior tolerância ao estresse abiótico com relação às plantas controle, compreendendo:

(i) introduzir e expressar em uma planta um ácido nucléico que codifica um polipeptídeo YRP; e

(ii) cultivar a célula da planta em condições que promovem o estresse abiótico.

13. Planta transgênica com tolerância ao estresse abiótico, com relação às plantas controle, que resulta na expressão modulada de um ácido nucléico que codifica um polipeptídeo YRP, ou uma célula transgênica de planta derivada da dita planta transgênica.

14. Planta transgênica de acordo com o item 7, 11 ou 13, ou uma célula transgênica de planta derivada da mesma, em que a dita planta é uma planta de cultivo ou uma monocotiledônea ou um cereal, tais como arroz, milho, trigo, cevada, milheto, centeio, tritcale, sorgo, cana de açúcar, trigo de emmer, espelta, secale, trigo de einkorn, cereal teff, milo e aveias.

15. Partes cultiváveis de uma planta de acordo com o item 14, em que as ditas Partes cultiváveis são preferivelmente biomassa dos ramos e/ou sementes.

16. Produtos derivados de uma planta de acordo com o item 14 e/ou de Partes cultiváveis de uma planta de acordo com o item 15.

17. Uso de um ácido nucléico que codifica um polipeptídeo YRP para aumentar o rendimento, particularmente para aumentar a tolerância ao estresse abiótico, com relação às plantas controle.

3. Polipeptídeos do tipo de Brevis Radix (BRXL)

1) Um método para aumentar as características relacionadas ao rendimento em plantas com relação às plantas controle, compreendendo

aumentar a expressão em uma planta com uma sequência de ácidos nucleicos que codifica um polipeptídeo do tipo de Brevis Radix (BRXL), cujo polipeptídeo BRXL compreende pelo menos dois domínios BRX com um domínio IPR013591 DZC de entrada InterPro (PFAM entrada PF08381 DZC), e opcionalmente selecionar plantas com mais características relacionadas ao rendimento.

2) Método de acordo com o item 1, em que o dito polipeptídeo BRXL compreende (i) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio BRX da maneira representada por SEQ ID NO: 65; e (ii) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio BRX da maneira representada por SEQ ID NO: 82.

3) Método de acordo com o item 1 ou 2, em que o dito polipeptídeo BRXL compreende (i) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 1 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 83; e (ii) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 2 (compreendendo um domínio BRX) da maneira representada por SEQ ID NO: 84.

4) Método de acordo com o item 3, em que o dito polipeptídeo BRXL compreende (1) em uma ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 3 da maneira representada por SEQ ID NO: 85; e (1) em uma

ordem crescente de preferência pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um domínio conservado 4 da maneira representada por SEQ ID NO: 86.

5 5) Método de acordo com qualquer item precedente, em que o dito polipeptídeo BRXL apresenta em uma ordem crescente de preferência pelo menos 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com um polipeptídeo da maneira representada por SEQ ID NO: 18, ou com
10 quaisquer das sequências de polipeptídeos fornecidas na tabela A aqui.

6) Método de acordo com qualquer item precedente, em que o dito polipeptídeo BRXL interage com com ele mesmo ou com um outro polipeptídeo BRLX em um ensaio de dois híbridos em levedura.

7) Método de acordo com qualquer item precedente, em que a
15 dita sequência de ácidos nucleicos que codifica um polipeptídeo BRXL é representada por qualquer uma das sequências de ácidos nucleicos SEQ ID NOs fornecidas na tabela A3 ou uma porção da mesmas, ou uma sequência capaz de hibridizar em qualquer uma das sequências de ácidos nucleicos SEQ ID NOs fornecidas na tabela A3, ou em um complemento da mesmas.

20 8) Método de acordo com qualquer item precedente, em que a dita sequência de ácidos nucleicos codifica um ortólogo ou parólogo de qualquer sequência de polipeptídeos SEQ ID NOs fornecida na tabela A3.

9) Método de acordo com qualquer item precedente, em que a dita melhor expressão é afetada por qualquer um ou mais de: marcação da
25 ativação do DNA-T, TILLING ou recombinação homóloga.

10) Método de acordo com qualquer item precedente, em que a dita melhor expressão é afetada introduzindo e expressando em uma planta uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL.

11) Método de acordo com qualquer item precedente, em que a

dita melhor característica relacionada ao rendimento é uma ou mais de: maior altura da planta, e aumento no peso de mil grãos (TKW).

12) Método de acordo com qualquer item precedente, em que a dita sequência de ácidos nucleicos é operavelmente ligada a um promotor constitutivo.

13) Método de acordo com o item 12, em que o dito promotor constitutivo é um promotor GOS2, preferivelmente um promotor GOS2 de arroz, acima de tudo preferivelmente uma sequência de GOS2 da maneira representada por SEQ ID NO: 87.

14) Método de acordo com qualquer item precedente, em que a dita sequência de ácidos nucleicos que codifica um polipeptídeo BRXL é de uma planta, preferivelmente de maneira adicional de uma planta dicotiledônea, mais preferivelmente da família *Salicaceae*, acima de tudo preferivelmente a sequência de ácidos nucleicos é de *Populus trichocarpa*.

15) Plantas, parte da mesmas (incluindo sementes), ou células de planta obteníveis por um método de acordo com qualquer item precedente, em que a dita planta, parte ou célula da mesma compreende um transgene isolado de ácido nucléico que codifica um polipeptídeo BRXL.

16) Uma Molécula de ácido nucléico isolado selecionada de:

(i) uma sequência de ácidos nucleicos da maneira representada por qualquer uma de SEQ ID NO: 75, SEQ ID NO: 77, ou SEQ ID NO: 79;

(ii) o complemento de uma sequência de ácidos nucleicos da maneira representada por qualquer uma de SEQ ID NO: 75, SEQ ID NO: 77, ou SEQ ID NO: 79;

(iii) uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL com, em uma ordem crescente de preferência, pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com a sequência de polipeptídeos representada por qualquer uma de SEQ ID NO:

76, SEQ ID NO: 78, ou SEQ ID NO: 80.

17) Um polipeptídeo isolado selecionado de:

(i) uma sequência de polipeptídeos da maneira representada por qualquer uma de SEQ ID NO: 76, SEQ ID NO: 78, ou SEQ ID NO: 80;

5 (ii) uma sequência de polipeptídeos com, em uma ordem crescente de preferência, pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos com uma sequência de polipeptídeos da maneira representada por qualquer uma de SEQ ID NO: 76, SEQ ID NO: 78, ou SEQ
10 ID NO: 80;

(iii) derivados de quaisquer das sequências de polipeptídeos fornecidas em (i) ou (ii) anteriormente.

18) Construto compreendendo:

(a) uma sequência de ácidos nucleicos que codifica um
15 polipeptídeo BRXL da maneira definida em qualquer um dos itens 1 a 8, ou 16;

(b) uma ou mais sequências controle capazes de direcionar a expressão da sequência de ácidos nucleicos de (a); e opcionalmente (c) uma sequência de finalização de transcrição.

20 19) Construto de acordo com o item 18 em que a dita sequência controle é um promotor constitutivo.

20) Construto de acordo com o item 19 em que o dito promotor constitutivo é um promotor GOS2, preferivelmente um promotor GOS2 de arroz, acima de tudo preferivelmente uma sequência de GOS2 da
25 maneira representada por SEQ ID NO: 87.

21) Uso de um construto de acordo com qualquer um dos itens 18 a 20 em um método para preparar plantas com mais características relacionadas ao rendimento com relação às plantas controle, em que melhores características relacionadas ao rendimento são uma ou mais de: maior altura

da planta, maior rendimento da semente por planta, maior número de sementes preenchidas, e aumento no peso de mil grãos (TKW).

22) Planta, parte da planta ou célula de planta transformadas com um construto de acordo com qualquer um dos itens 18 a 20.

5 23) Método para a produção de plantas transgênicas com mais características relacionadas ao rendimento com relação às plantas controle, compreendendo:

10 (i) introduzir e expressar em uma planta, parte da planta, ou célula de planta, uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL da maneira definida em qualquer um dos itens 1 a 8, ou 16; e

(ii) cultivar a célula da planta, parte da planta, ou planta em condições que promovem crescimento e desenvolvimento da planta.

15 24) Planta transgênica com mais características relacionadas ao rendimento com relação às plantas controle, que resultam de melhor expressão de uma sequência isolada de ácidos nucleicos que codifica um polipeptídeo BRXL da maneira definida em qualquer um dos itens 1 a 8, ou 16, ou uma célula transgênica de planta ou parte de planta transgênica derivada da dita planta transgênica.

20 25) Planta transgênica de acordo com o item 14, 22, ou 24, em que a dita planta é uma planta de cultivo ou uma monocotiledônea ou um cereal, tais como arroz, milho, trigo, cevada, milheto, centeio, tritcale, sorgo e aveias, ou uma célula transgênica de planta derivada da dita planta transgênica.

25 26) Partes cultiváveis compreendendo uma sequência isolada de ácidos nucleicos que codifica um polipeptídeo BRXL, de uma planta de acordo com o item 25, em que as ditas Partes cultiváveis são preferivelmente sementes.

27) Produtos derivados de uma planta de acordo com o item 25

e/ou de Partes cultiváveis de uma planta de acordo com o item 26.

28) Uso de uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL da maneira definida em qualquer um dos itens 1 a 8, ou 16, para aumentar as características relacionadas ao rendimento, compreendendo uma ou mais de: maior altura da planta, e aumento no peso de mil grãos (TKW).

4. Polipeptídeos do tipo silky-1

a) Método para melhorar a tolerância ao estresse abiótico em plantas modulando a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo do tipo silky-1 ou um ortólogo ou parólogo deste.

b) Método de acordo com o item 1, em que a dita expressão modulada é afetada introduzindo e expressando em uma planta um ácido nucléico que codifica polipeptídeo do tipo silky-1.

c) Método de acordo com o itens 2 ou 3, em que o dito ácido nucléico que codifica um polipeptídeo do tipo silky-1 codifica qualquer uma das proteínas listadas na tabela A4 ou é uma porção de um ácido nucléico como este, ou um ácido nucléico capaz de hibridizar em um ácido nucléico como este.

d) Método de acordo com qualquer um dos itens 1 a 4, em que a dita sequência de ácidos nucleicos codifica um ortólogo ou parólogo de quaisquer das proteínas fornecidas na tabela A4.

e) Método de acordo com o itens 3 ou 4, em que o dito ácido nucléico é operavelmente ligado a um promotor constitutivo, preferivelmente a um promotor GOS2, acima de tudo preferivelmente a um promotor GOS2 de arroz.

f) Método de acordo com qualquer um dos itens 1 a 5, em que o dito ácido nucléico que codifica um polipeptídeo do tipo silky-1 é de *Populus trichocarpa*.

g) Planta ou parte da mesma, incluindo sementes, obteníveis

por um método de acordo com qualquer um dos itens 1 a 6, em que a dita planta ou parte da mesma compreende um ácido nucléico recombinante que codifica um polipeptídeo do tipo silky-1.

h) Construto compreendendo:

5 (i) ácido nucléico que codifica um polipeptídeo do tipo silky-1 da maneira definida nos itens 1 ou 2;

(ii) uma ou mais sequências controle capazes de direcionar a expressão da sequência de ácidos nucleicos de (a); e opcionalmente

(iii) uma sequência de finalização de transcrição.

10 i) Construto de acordo com o item 9, em que uma das ditas sequências controle é um promotor constitutivo, preferivelmente um promotor GOS2, acima de tudo preferivelmente um promotor GOS2 de arroz.

j) Uso de um construto de acordo com o item 8 ou 9 em um método para preparar plantas com maior tolerância ao estresse abiótico com
15 relação às plantas controle.

k) Planta, parte da planta ou célula de planta transformadas com um construto de acordo com o item 8 ou 9.

l) Método para a produção de uma planta transgênica com maior tolerância ao estresse abiótico com relação às plantas controle,
20 compreendendo:

(i) introduzir e expressar em uma planta um ácido nucléico que codifica um polipeptídeo do tipo silky-1; e

(ii) cultivar a célula da planta em condições que promovem o estresse abiótico.

25 m) Planta transgênica com tolerância ao estresse abiótico, com relação às plantas controle, que resulta na expressão modulada de um ácido nucléico que codifica um polipeptídeo do tipo silky-1, ou uma célula transgênica de planta derivada da dita planta transgênica.

n) Planta transgênica de acordo com o item 7, 11 ou 13, ou

uma célula transgênica de planta derivada da mesma, em que a dita planta é uma planta de cultivo ou uma monocotiledônea ou um cereal, tais como arroz, milho, trigo, cevada, milheto, centeio, triticales, sorgo, cana de açúcar, trigo de emmer, espelta, secale, trigo de einkorn, cereal teff, milo e aveias.

5 o) Partes cultiváveis de uma planta de acordo com o item 14, em que as ditas Partes cultiváveis são preferivelmente biomassa dos ramos e/ou sementes.

p) Produtos derivados de uma planta de acordo com o item 14 e/ou de Partes cultiváveis de uma planta de acordo com o item 15.

10 q) Uso de um ácido nucléico que codifica um polipeptídeo do tipo silky-1 para aumentar o rendimento, particularmente para aumentar a tolerância ao estresse abiótico, com relação às plantas controle.

5. Polipeptídeos ARP6

15 1. Um método para melhorar características relacionadas ao rendimento em plantas com relação às plantas controle, compreendendo modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo ARP6.

2. Método de acordo com o item 1, em que o dito polipeptídeo ARP6 apresenta em uma ordem crescente de preferência pelo menos 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade completa de sequência com o aminoácido representada por qualquer uma de SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, SEQ ID NO: 110, e SEQ ID NO: 112.

3. Método de acordo com o item 1 ou 2, em que a dita

expressão modulada é afetada introduzindo e expressando em uma planta um ácido nucléico que codifica um polipeptídeo ARP6.

4. Método de acordo com qualquer um dos itens 1 a 3, em que o dito ácido nucléico que codifica um polipeptídeo ARP6 codifica qualquer uma das proteínas listadas na tabela A5, ou é uma porção de um ácido nucléico como este, ou um ácido nucléico capaz de hibridizar em um ácido nucléico como este.

5. Método de acordo com qualquer um dos itens 1 a 4, em que a dita sequência de ácidos nucleicos codifica um ortólogo ou parólogo de quaisquer das proteínas fornecidas na tabela A5.

6. Método de acordo com qualquer item precedente, em que as ditas melhores características relacionadas ao rendimento compreendem maior rendimento, preferivelmente maior e biomassa e/ou maior rendimento da semente com relação às plantas controle.

7. Método de acordo com qualquer um dos itens 1 a 6, em que as ditas melhores características relacionadas ao rendimento são obtidas em condições de não estresse.

8. Método de acordo com qualquer um dos itens 1 a 6, em que as ditas melhores características relacionadas ao rendimento são obtidas em condições de estresse por seca, estresse pelo sal ou deficiência de nitrogênio.

9. Método de acordo com qualquer um dos itens 3 a 8, em que o dito ácido nucléico é operavelmente ligado a um promotor constitutivo, preferivelmente a um promotor GOS2, acima de tudo preferivelmente a um promotor GOS2 de arroz.

10. Método de acordo com qualquer um dos itens 1 a 9, em que o dito ácido nucléico que codifica um polipeptídeo ARP6 é de origem vegetal, preferivelmente de uma planta dicotiledônea, preferivelmente de maneira adicional da família *Brassicaceae*, mais preferivelmente do gênero

Arabidopsis, acima de tudo preferivelmente de *Arabidopsis thaliana*.

11. Planta ou parte da mesma, incluindo sementes, obteníveis por um método de acordo com qualquer um dos itens 1 a 10, em que a dita planta ou parte da mesma compreende um ácido nucléico recombinante que codifica um polipeptídeo ARP6.

12. Construto compreendendo:

(ii) ácido nucléico que codifica um polipeptídeo ARP6 da maneira definida nos itens 1 ou 2;

(iii) uma ou mais sequências controle capazes de direcionar a expressão da sequência de ácidos nucleicos de (a); e opcionalmente

(iv) uma sequência de finalização de transcrição.

13. Construto de acordo com o item 12, em que uma das ditas sequências controle é um promotor constitutivo, preferivelmente um promotor GOS2, acima de tudo preferivelmente um GOS2 promotor de arroz.

14. Uso de um construto de acordo com o item 12 ou 13 em um método para preparar plantas com maior rendimento, particularmente maior biomassa e/ou maior rendimento da semente com relação às plantas controle.

15. Planta, parte da planta ou célula de planta transformadas com um construto de acordo com o item 12 ou 13.

16. Método para a produção de uma planta transgênica com maior rendimento, particularmente maior biomassa e/ou maior rendimento da semente com relação às plantas controle, compreendendo:

(i) introduzir e expressar em uma planta um ácido nucléico que codifica um polipeptídeo ARP6 da maneira definida em item 1 ou 2; e

(ii) cultivar a célula da planta em condições que promovem crescimento e desenvolvimento da planta.

17. Planta transgênica com maior rendimento, particularmente maior biomassa e/ou maior rendimento da semente, com relação às plantas

controle, que resulta na expressão modulada de um ácido nucléico que codifica um polipeptídeo ARP6 da maneira definida em item 1 ou 2, ou uma célula transgênica de planta derivada da dita planta transgênica.

18. Planta transgênica de acordo com o item 11, 15 ou 17, ou uma célula transgênica de planta derivada da mesma, em que a dita planta é uma planta de cultivo ou uma monocotiledônea ou um cereal, tais como arroz, milho, trigo, cevada, milheto, centeio, triticale, sorgo trigo de emmer, espelta, secale, trigo de einkorn, cereal teff, milo e aveias.

19. Partes cultiváveis de uma planta de acordo com o item 18, em que as ditas Partes cultiváveis são preferivelmente biomassa dos ramos e/ou sementes.

20. Produtos derivados de uma planta de acordo com o item 18 e/ou de Partes cultiváveis de uma planta de acordo com o item 19.

21. Uso de um ácido nucléico que codifica um polipeptídeo ARP6 para aumentar o rendimento, particularmente para aumentar o rendimento da semente e/ou ramo biomassa em plantas, com relação às plantas controle.

6. Polipeptídeos POP

1. Um método para melhorar características relacionadas ao rendimento em plantas com relação às plantas controle, compreendendo modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo de prolil oligopeptidase (POP), em que o dito polipeptídeo de POP compreende um domínio Peptidase_S9 (entrada Pfam PF00326) e preferivelmente também um domínio DPPIV_N (entrada Pfam PF00930).

2. Método de acordo com o item 1, em que o dito polipeptídeo de POP compreende um ou mais de os motivos 1 a 14 (SEQ ID NO: 118 a SEQ ID NO: 131).

3. Método de acordo com o item 1 ou 2, em que a dita expressão modulada é afetada introduzindo e expressando em uma planta um

ácido nucléico que codifica um polipeptídeo de POP.

4. Método de acordo com qualquer um dos itens 1 a 3, em que o dito ácido nucléico que codifica um polipeptídeo de POP codifica qualquer uma das proteínas listadas na tabela A6 ou é uma porção de um ácido nucléico como este, ou um ácido nucléico capaz de hibridizar em um ácido nucléico como este.

5. Método de acordo com qualquer um dos itens 1 a 4, em que a dita sequência de ácidos nucleicos codifica um ortólogo ou parólogo de quaisquer das proteínas fornecidas na tabela A6.

6. Método de acordo com qualquer item precedente, em que as ditas melhores características relacionadas ao rendimento compreendem maior rendimento, preferivelmente maior biomassa e/ou maior rendimento da semente, e ou período de floração modificada com relação às plantas controle.

7. Método de acordo com qualquer um dos itens 1 a 6, em que as ditas melhores características relacionadas ao rendimento são obtidas em condições de não estresse.

8. Método de acordo com qualquer um dos itens 3 to 7, em que o dito ácido nucléico é operavelmente ligado a um promotor constitutivo, preferivelmente a um promotor GOS2, acima de tudo preferivelmente a um promotor GOS2 de arroz.

9. Método de acordo com qualquer um dos itens 1 a 8, em que o dito ácido nucléico que codifica um polipeptídeo de POP é de origem vegetal, preferivelmente de uma planta dicotiledônea, preferivelmente de maneira adicional da família *Brassicaceae*, mais preferivelmente do gênero *Arabidopsis*, acima de tudo preferivelmente de *Arabidopsis thaliana*.

10. Planta ou parte da mesma, incluindo sementes, obteníveis por um método de acordo com qualquer um dos itens 1 a 9, em que a dita planta ou parte da mesma compreende um ácido nucléico recombinante que codifica um polipeptídeo de POP.

11. Construto compreendendo:

(i) ácido nucléico que codifica um polipeptídeo de POP da maneira definida nos itens 1 ou 2;

5 (ii) uma ou mais sequências controle capazes de direcionar a expressão da sequência de ácidos nucleicos de (a); e opcionalmente

(iii) uma sequência de finalização de transcrição.

12. Construto de acordo com o item 11, em que uma das ditas sequências controle é um promotor constitutivo, preferivelmente um promotor GOS2, acima de tudo preferivelmente um promotor GOS2 de arroz.

10 13. Uso de um construto de acordo com o item 11 ou 12 em um método para preparar plantas com maior rendimento, particularmente maior biomassa e/ou maior rendimento da semente com relação às plantas controle.

15 14. Uso de um construto de acordo com o item 11 ou 12 em um método para preparar plantas com período de floração modificada com relação às plantas controle.

15. Planta, parte da planta ou célula de planta transformadas com um construto de acordo com o item 11 ou 12.

20 16. Método para a produção de uma planta transgênica com maior rendimento e/ou período de floração modificada, particularmente maior biomassa e/ou maior rendimento da semente com relação às plantas controle, compreendendo:

(i) introduzir e expressar em uma planta um ácido nucléico que codifica um polipeptídeo de POP da maneira definida em item 1 ou 2; e

25 (ii) cultivar a célula da planta em condições que promovem crescimento e desenvolvimento da planta.

17. Planta transgênica com maior rendimento, particularmente maior biomassa e/ou maior rendimento da semente, com relação às plantas controle, que resultam da expressão modulada de um ácido nucléico que

codifica um polipeptídeo de POP da maneira definida em item 1 ou 2, ou uma célula transgênica de planta derivada da dita planta transgênica.

18. Planta transgênica de acordo com o item 10, 15 ou 17, ou uma célula transgênica de planta derivada da mesma, em que a dita planta é uma planta de cultivo ou uma monocotiledônea ou um cereal, tais como arroz, milho, trigo, cevada, milheto, centeio, triticale, sorgo trigo de emmer, espelta, secale, trigo de einkorn, cereal teff, milo e aveias.

19. Partes cultiváveis de uma planta de acordo com o item 18, em que as ditas Partes cultiváveis são preferivelmente biomassa dos ramos e/ou sementes.

20. Produtos derivados de uma planta de acordo com o item 18 e/ou de Partes cultiváveis de uma planta de acordo com o item 18.

21. Uso de um ácido nucléico que codifica um polipeptídeo de POP para aumentar o rendimento, particularmente para aumentar o rendimento da semente e/ou ramo biomassa em plantas, com relação às plantas controle.

22. Uso de um ácido nucléico que codifica um polipeptídeo de POP para modificar o período de floração de plantas, com relação às plantas controle.

7. Polipeptídeos CRL

1. Um método para melhorar características relacionadas ao rendimento em plantas com relação às plantas controle, compreendendo modular a expressão em uma planta com um ácido nucléico que codifica um polipeptídeo CRL, em que o dito polipeptídeo CRL compreende um domínio DUF206.

2. Método de acordo com o item 1, em que o dito polipeptídeo CRL compreende um motivo com, em uma ordem crescente de preferência, pelo menos 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 100 %.

72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade de sequência com qualquer um ou mais dos seguintes motivos:

- 5 (i) Motivo 15: EQAFWRxPXKPFRQR (SEQ ID NO: 207);
(ii) Motivo 16: NFCDR (SEQ ID NO: 208);
(iii) Motivo 17: RGKRCLYEGS (SEQ ID NO: 209);
(iv) Motivo 18: QVWGxKXGPYEFK (SEQ ID NO: 210);
em que X representa qualquer aminoácido.
- 10 3. Método de acordo com o item 1 ou 2, em que a dita expressão modulada é afetada introduzindo e expressando em uma planta um ácido nucléico que codifica um polipeptídeo CRL.
4. Método de acordo com qualquer um dos itens 1 a 3, em que o dito ácido nucléico que codifica um polipeptídeo CRL codifica qualquer
- 15 uma das proteínas listadas na tabela A7 ou é uma porção de um ácido nucléico como este, ou um ácido nucléico capaz de hibridizar em um ácido nucléico como este.
5. Método de acordo com qualquer um dos itens 1 a 4, em que a dita sequência de ácidos nucleicos codifica um ortólogo ou parólogo de
- 20 quaisquer das proteínas fornecidas na tabela A7.
6. Método de acordo com qualquer item precedente, em que as ditas melhores características relacionadas ao rendimento compreendem maior rendimento, preferivelmente maior biomassa e/ou maior rendimento da semente com relação às plantas controle.
- 25 7. Método de acordo com qualquer um dos itens 1 a 6, em que as ditas melhores características relacionadas ao rendimento são obtidas em condições de não estresse.
8. Método de acordo com qualquer um dos itens 1 a 6, em que as ditas melhores características relacionadas ao rendimento são obtidas em

condições de estresse por secura, estresse pelo sal ou deficiência de nitrogênio.

5 9. Método de acordo com qualquer um dos itens 3 a 8, em que o dito ácido nucléico é operavelmente ligado a um promotor constitutivo, preferivelmente a um promotor GOS2, acima de tudo preferivelmente a um promotor GOS2 de arroz.

10 10. Método de acordo com qualquer um dos itens 1 a 9, em que o dito ácido nucléico que codifica um polipeptídeo CRL é de origem vegetal, preferivelmente de uma planta dicotiledônea, preferivelmente de maneira adicional da família *Brassicaceae*, mais preferivelmente do gênero *Arabidopsis*, acima de tudo preferivelmente de *Arabidopsis thaliana*.

15 11. Planta ou parte da mesma, incluindo sementes, obteníveis por um método de acordo com qualquer um dos itens 1 a 10, em que a dita planta ou parte da mesma compreende um ácido nucléico recombinante que codifica um polipeptídeo CRL.

12. Construto compreendendo:

(i) ácido nucléico que codifica um polipeptídeo CRL da maneira definida nos itens 1 ou 2;

20 (ii) uma ou mais sequências controle capazes de direcionar a expressão da sequência de ácidos nucleicos de (a); e opcionalmente

(iii) uma sequência de finalização de transcrição.

13. Construto de acordo com o item 12, em que uma das ditas sequências controle é um promotor constitutivo, preferivelmente um promotor GOS2, acima de tudo preferivelmente um promotor GOS2 de arroz.

25 14. Uso de um construto de acordo com o item 12 ou 13 em um método para preparar plantas com maior rendimento, particularmente maior biomassa e/ou maior rendimento da semente com relação às plantas controle.

15. Planta, parte da planta ou célula de planta transformadas

com um construto de acordo com o item 12 ou 13.

16. Método para a produção de uma planta transgênica com maior rendimento, particularmente maior biomassa e/ou maior rendimento da semente com relação às plantas controle, compreendendo:

5 (i) introduzir e expressar em uma planta um ácido nucléico que codifica um polipeptídeo CRL da maneira definida em item 1 ou 2; e

(ii) cultivar a célula da planta em condições que promovem crescimento e desenvolvimento da planta.

10 17. Planta transgênica com maior rendimento, particularmente maior biomassa e/ou maior rendimento da semente, com relação às plantas controle, que resultam na expressão modulada de um ácido nucléico que codifica um polipeptídeo CRL da maneira definida em item 1 ou 2, ou uma célula transgênica de planta derivada da dita planta transgênica.

15 18. Planta transgênica de acordo com o item 11, 15 ou 17, ou uma célula transgênica de planta derivada da mesma, em que a dita planta é uma planta de cultivo ou uma monocotiledônea ou um cereal, tais como arroz, milho, trigo, cevada, milheto, centeio, tritcale, sorgo trigo de emmer, espelta, secale, trigo de einkorn, cereal teff, milo e aveias.

20 19. Partes cultiváveis de uma planta de acordo com o item 18, em que as ditas Partes cultiváveis são preferivelmente biomassa dos ramos e/ou sementes.

20. Produtos derivados de uma planta de acordo com o item 18 e/ou de Partes cultiváveis de uma planta de acordo com o item 19.

25 21. Uso de um ácido nucléico que codifica um polipeptídeo CRL para aumentar o rendimento, particularmente para aumentar o rendimento da semente e/ou ramo biomassa em plantas, com relação às plantas controle.

22. Uma Molécula de ácido nucléico isolado selecionada de:

(i) um ácido nucléico representado por SEQ ID NO: 195;

(ii) o complemento de um ácido nucléico representado por SEQ ID NO: 195;

(iii) um ácido nucléico que codifica o polipeptídeo da maneira representada por SEQ ID NO: 196, preferivelmente em função da degeneração do código genético, o dito ácido nucléico isolado pode ser derivado de uma sequência de polipeptídeos da maneira representada por SEQ ID NO: 196, e preferivelmente de maneira adicional confere melhores características relacionadas ao rendimento com relação às plantas controle;

(iv) um ácido nucléico com, em uma ordem crescente de preferência pelo menos 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade de sequência com quaisquer das sequências de ácidos nucleicos da tabela A7, e preferivelmente de maneira adicional que confere melhores características relacionadas ao rendimento com relação às plantas controle;

(v) uma molécula de ácido nucléico que hibridiza em uma molécula de ácido nucléico de (i) a (iv) em condições de hibridização severa, e preferivelmente confere melhores características relacionadas ao rendimento com relação às plantas controle;

(vi) um ácido nucléico que codifica um polipeptídeo ASPAT com, em uma ordem crescente de preferência, pelo menos 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de

identidade de sequência com a sequência de aminoácidos representada por SEQ ID NO: 196 e quaisquer das outras sequências de aminoácido na tabela A7, e preferivelmente que confere melhores características relacionadas ao rendimento com relação às plantas controle.

5 23. Um polipeptídeo isolado selecionado de:

(i) uma sequência de aminoácidos representada por SEQ ID NO: 196;

(ii) uma sequência de aminoácidos com, em uma ordem crescente de preferência, pelo menos 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 10 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, ou 99 % de identidade de sequência com a sequência de aminoácidos representada por SEQ ID NO: 15 196, e quaisquer das outras sequências de aminoácido na tabela A7, e preferivelmente que confere melhores características relacionadas ao rendimento com relação às plantas controle;

(iii) derivados de quaisquer das sequências de aminoácidos fornecidas em (i) ou (ii) anteriormente.

20 DESCRIÇÃO DAS FIGURAS

A presente invenção será agora descrita com referência às figuras a seguir, nas quais:

A figura 1 representa o vetor binário usado para melhor expressão em *Oryza sativa* de um tipo de ácido nucléico codificante no 25 controle de um promotor GOS2 de arroz(pGOS2).

A figura 2 representa o vetor binário usado para melhor expressão em *Oryza sativa* de um ácido nucléico que codifica YRP no controle de um promotor GOS2 de arroz (pGOS2).

A figura 3 representa como a biossíntese de brassinoesteróide

e sinalização de auxina estão conectadas por meio de um circuito de retroalimentação, que envolve BRX, exigido para crescimento ideal da raiz, de acordo com Mouchel et al. (2006) Nature 443: 458-461.

A figura 4 representa um esquema de um polipeptídeo BRXL da maneira representada por SEQ ID NO: 18, que compreende as seguintes características: (i) um domínio conservado 1 que representa um domínio BRX, que compreende o domínio IPR013591 DZC (entrada PFAM PF08381 DZC); (2) um domínio conservado 2 que representa um domínio BRX, que compreende um domínio C-terminal IPR013591 DZC (entrada PFAM PF08381 DZC); (3) um domínio conservado 3 e (4) um domínio conservado 4, ambos contendo Cys conservada, cujo espaçamento é indicativo de um motivo de ligação de zinco potencial.

A figura 5 mostra um alinhamento de sequência múltiplo ClustalW 1.81 dos polipeptídeos BRXL da tabela A3. As características a seguir são enquadradas de maneira acentuada: (i) um domínio conservado 1 que representa um domínio BRX, que compreende o domínio IPR013591 DZC (entrada PFAM PF08381 DZC; marcado com X); (2) um domínio conservado 2 que representa um domínio BRX, que compreende um domínio C-terminal IPR013591 DZC (entrada PFAM PF08381 DZC; marcado com X); (3) um domínio conservado 3 e (4) um domínio conservado 4, ambos contendo Cys conservada (levemente enquadradas), cujo espaçamento é indicativo de um motivo de ligação de zinco potencial.

A figura 6 mostra o vetor binário para a melhor expressão em plantas *Oryza sativas* de uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL no controle de um promotor constitutivo funcional em plantas.

A figura 7 representa o vetor binário usado para melhor expressão em *Oryza sativa* de um ácido nucléico que codifica um tipo de silky-1 no controle de um promotor GOS2 de arroz (pGOS2).

A figura 8 representa um alinhamento múltiplo de polipeptídeos ARP6.

A figura 9 mostra a árvore filogenética dos polipeptídeos ARP, da maneira descrita por Kandasamy et al. 2004. O grupo de polipeptídeos ARP6 é indicado.

A figura 10 representa o vetor binário usado para melhor expressão em *Oryza sativa* de um ácido nucléico que codifica ARP6 no controle de um promotor GOS2 de arroz (pGOS2).

A figura 11 representa SEQ ID NO: 117 com os domínios conservados DPPIV_N e Peptidase_S9 indicado em itálico, sublinhado e em negrito, respectivamente.

A figura 12 representa um alinhamento múltiplo de várias sequências POP.

A figura 13 mostra uma árvore filogenética de proil peptidases (Tripathi & Sowdhamini, 2006). A ramificação com SEQ ID NO: 117 e seu ortólogo de arroz são indicados pela seta.

A figura 14 representa o vetor binário usado para melhor expressão em *Oryza sativa* de um ácido nucléico que codifica POP no controle de um promotor GOS2 de arroz (pGOS2).

A figura 15 representa um alinhamento múltiplo de proteínas CRL.

A figura 16 mostra uma árvore filogenética de proteínas CRL. Os agrupamentos para dicotiledôneas, monocotiledôneas e outras proteínas CRL de *Viridiplantae* são mostradas.

A figura 17 representa o vetor binário usado para melhor expressão em *Oryza sativa* de um ácido nucléico que codifica CRL no controle de um promotor GOS2 de arroz (pGOS2).

EXEMPLOS

A presente invenção será agora descrita com referência aos

exemplos a seguir, que são apenas a título de ilustração. Não pretende-se que os exemplos a seguir definam completamente ou limitem de outra forma o escopo da invenção.

Manipulação de DNA: a menos que de outra forma declarada, as técnicas de DNA recombinante são realizadas de acordo com protocolos padrões descritos em (Sambrook (2001) Molecular Cloning: um laboratory manual, 3rd Edition Cold Spring Harbor Laboratory Press, CSH, Nova Iorque) ou nos volumes 1 e 2 de Ausubel et al. (1994), Current Protocols in Molecular Biology, Current Protocols. Os materiais e métodos padrões para o trabalho molecular em planta são descritos em Plant Molecular Biology Labfax (1993) por R.D.D. Cray, publicado por BIOS Scientific Publications Ltd (UK) e Blackwell Scientific Publications (UK).

1.1. Polipeptídeos do tipo alfin

As as sequências (DNAC de comprimento total, ESTs ou genômico) relacionadas a SEQ ID NO: 1 e SEQ ID NO: 3 são identificadas entre aquelas mantidas na base de dados de nucleotídeos Entrez no National Center for Biotechnology Information (NCBI) usando ferramentas de pesquisa de sequência em base de dados, tal como o Basic Local Alignment Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa é usado para encontrar regiões de similaridade local entre as sequências comparando as sequências de ácido nucléico ou de polipeptídeos com a base de dados de sequências e calculando a significância estatística das combinações. Por exemplo, o polipeptídeo codificado pelo ácido nucléico de SEQ ID NO: 1 e SEQ ID NO: 3 é usado no algoritmo TBLASTN, com ajustes padrões e o filtro para ignorar sequências de baixa complexidade da mesmacadas. O produto da análise foi visualizado por comparação aos pares, e classificado de acordo com a classificação de probabilidade (valor E), quando a classificação reflete a probabilidade de que um alinhamento particular ocorre ao acaso

(quanto menor o valor E, mais significativo o acerto). Além dos valores E, as comparações também são classificadas por identidade percentual. A identidade percentual refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácido nucléico comparadas (ou polipeptídeos) com relação a um tamanho particular. Em alguns exemplos, os parâmetros padrões são ajustados para modificar a severidade da pesquisa. Por exemplo, o valor E é aumentado para mostrar menos combinações de severidade. Dessa maneira, combinações pequenas quase exatas são identificadas.

10 A tabela A1 fornece uma lista de sequências de ácidos nucleicos do tipo alfin.

Tabela A1 : Exemplos de polipeptídeos do tipo alfin:

Nome	Organismo	SEQ ID NO de ácido nucléico	SEQ ID NO de polipeptídeo
Tipo Ac_ALFIN	<i>Allium cepa</i>	1	2
Tipo Hv_ALFIN	<i>Hordeum vulgare</i>	3	4

Em alguns exemplos, as sequências relacionadas são montadas por tentativas e reveladas publicamente por instituições de pesquisa, tal como The Institute for Genomic Research (TIGR; começando com TA). A base de dados The Eukaryotic Gene Orthologs (EGO) é usada para identificar tais sequências relacionadas, tanto por pesquisa de palavra chave quanto usando o algoritmo BLAST com a sequência de ácidos nucleicos ou sequência de polipeptídeos de interesse. Em outros exemplos, as bases de dados de sequência especial de ácidos nucleicos são criadas para organismos particulares, tal como pelo Joint Genome Institute.

1.2. Polipeptídeos YRP

As sequências (DNAC de comprimento total, ESTs ou genômico) relacionadas a SEQ ID NO: 10 e SEQ ID NO: 12 são identificadas entre aquelas mantidas na base de dados de nucleotídeos Entrez no National Center for Biotechnology Information (NCBI) usando ferramentas de pesquisa de sequência em base de dados, tal como o Basic Local Alignment

Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa é usado para encontrar regiões de similaridade local entre as sequências comparando as sequências de ácido nucléico ou de polipeptídeos com a base de dados de sequências e calculando a significância estatística das combinações. Por exemplo, o polipeptídeo codificado pelo ácido nucléico de SEQ ID NO: 10 e SEQ ID NO: 12 é usado no algoritmo TBLASTN, com ajustes padrões e o filtro para ignorar sequências de baixa complexidade da mesmacadas. O produto da análise foi visualizado por comparação aos pares, e classificado de acordo com a classificação de probabilidade (valor E), quando a classificação reflete a probabilidade de que um alinhamento particular ocorre ao acaso (quanto menor o valor E, mais significativo o acerto). Além dos valores E, as comparações também são classificadas por identidade percentual. A identidade percentual refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácido nucléico comparadas (ou polipeptídeos) com relação a um tamanho particular. Em alguns exemplos, os parâmetros padrões são ajustados para modificar a severidade da pesquisa. Por exemplo, o valor E é aumentado para mostrar menos combinações de severidade. Dessa maneira, combinações pequenas quase exatas são identificadas.

A tabela A2 fornece um lista de sequências de ácidos nucleicos YRP.

Tabela A2: Exemplos polipeptídeos YRP:

Nome	Organismo	SEQ ID NO de ácido nucléico	SEQ ID NO de polipeptídeo
Hv_YRP	<i>Hordeum vulgare</i>	10	11
Hv_YRP	<i>Hordeum vulgare</i>	12	13

Em alguns exemplos, as sequências relacionadas são montadas por tentativas e reveladas publicamente por instituições de pesquisa, tal como The Institute for Genomic Research (TIGR; começando com TA). A base de dados The Eukaryotic Gene Orthologs (EGO) é usada para identificar tais

sequências relacionadas, tanto por pesquisa de palavra chave quanto usando o algoritmo BLAST com a sequência de ácidos nucleicos ou sequência de polipeptídeos de interesse. Em outros exemplos, as bases de dados de sequência especial de ácidos nucleicos são criadas para organismos particulares, tal como pelo Joint Genome Institute.

1.3. Polipeptídeos do tipo de Brevis Radix (BRXL)

As sequências (DNAC de comprimento total, ESTs ou genômico) relacionadas à sequência de ácidos nucleicos usadas nos métodos da presente invenção foram identificadas entre aquelas mantidas na base de dados de nucleotídeos Entrez no National Center for Biotechnology Information (NCBI) usando ferramentas de pesquisa de sequência em base de dados, tal como o Basic Local Alignment Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa é usado para encontrar regiões de similaridade local entre sequências comparando a sequência de ácidos nucleicos ou sequências de polipeptídeos com a base de dados de sequências e calculando a significância estatística das combinações. Por exemplo, o polipeptídeo codificado pela sequência de ácidos nucleicos da presente invenção foi usado pelo algoritmo TBLASTN, com ajustes padrões e o filtro para ignorar sequências de baixa complexidade da mesmacadas. O produto da análise foi visualizado por comparação aos pares, e classificado de acordo com a classificação de probabilidade (valor E), quando a classificação reflete a probabilidade de que um alinhamento particular ocorre ao acaso (quanto menor o valor E, mais significativo 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 sequências comparadas de ácidos nucleicos (ou polipeptídeo) com relação a um tamanho particular. Em alguns exemplos, os parâmetros padrões podem ser ajustados para modificar a severidade da

pesquisa. Por exemplo, o valor E pode ser aumentado para mostrar menos combinações de severidade. Dessa maneira, combinações pequenas quase exatas podem ser identificadas.

5 A tabela A3 fornece uma lista de sequências de ácidos nucleicos relacionadas à sequência de ácidos nucleicos usada nos métodos da presente invenção.

Tabela A3: Exemplos de sequências de polipeptídeos BRXL, e sequência codificantes de ácidos nucleicos

Nome	Número de acesso na base de dados pública	SEQ ID NO de ácido nucléico:	SEQ ID NO de polipeptídeo:
Poptr_BRX	JGI_scaff_III.746#1	17	18
Arath_BRX	NM_102925.2	19	20
Arath_BRXL1	NM_129113.4	21	22
Arath_BRXL2	NM_112254.2	23	24
Arath_BRXL3	NM_104296.2	25	26
Arath_BRXL4	NM_122061.3	27	28
Glyma_BRXL1	JGI_Gm0025x00418#1	29	30
Glyma_BRXL2	JGI_Gm0061x00041#1	31	32
Glyma_BRXL3	JGI_Gm0084x00169#1	33	34
Glyma_BRXL4	JGI_Gm0097x00095#1	35	36
Glyma_BRXL5	JGI_Gm0138x00217#1	37	38
Glyma_BRXL6	TIGR_TA62929_3847	39	40
Medtr_BRXL1	TIGR_TA28312_3880#1	41	42
Nicbe_BRXL1	TIGR_TA8873_4100#1	43	44
Orysa_BRXL1	Os08g36020	45	46
Orysa_BRXL2	Os02g47230	47	48
Orysa_BRXL3	Os04g51170	49	50
Orysa_BRXL4	Os03g63650	51	52
Orysa_BRXL5	CF299403.1,C1189950.1	53	54
Phypa_BRXL1	JGI_P.patens_161871#1	55	56
Picsi_BRXL1	TIGR_TA17584_3332#1	57	58
Poptr_BRXL1	JGI_P.trichocarpa_808986	59	60
Poptr_BRXL2	JGI_scaff_VI.979#1	61	62
Poptr_BRXL3	TIGR_TA12611_3695#1	63	64
Sorbi_BRXL1	Sb07g022540	65	66
Sorbi_BRXL2	Sb02g026020	67	68
Vitvi_BRXL1	GSVIVT00003671001#1	69	70
Vitvi_BRXL2	GGSVIVT00034110001#1	71	72
Zeama_BRXL1	TIGR_TA198521_4577#1	73	74
Zeama_BRXL2	Propriedade_ZM07MC01509_57718871@1504#1	75	76
Zeama_BRXL3	Propriedade_ZM07MC22150_BFb0050E10@22088#	77	78

	1		
Zeama_BRXL4	Propriedade _ZM07MC27026_BFb0199F19@26946# 1	79	80

Em alguns exemplos, as sequências relacionadas foram montadas por tentativa e reveladas publicamente por instituições de pesquisa, tal como The Institute for Genomic Research (TIGR; começando com TA), ou Genoscope (começando com GS). A base de dados The Eukaryotic Gene Orthologs (EGO) pode ser usada para identificar tais sequências relacionadas, tanto por pesquisa de palavra chave quanto usando o algoritmo BLAST com a sequência de ácidos nucleicos ou sequência de polipeptídeos de interesse. Em outros exemplos, as bases de dados de sequência especial de ácidos nucleicos foram criadas para organismos particulares, tal como pelo Joint Genome Institute. Adicionalmente, o acesso às bases de dados prioritárias permitiu a identificação de sequências inéditas de ácido nucléico e de polipeptídeos.

1.4. Polipeptídeos do tipo silky

As sequências (DNAc de comprimento total, ESTs ou genômico) relacionadas a SEQ ID NO: 90, SEQ ID NO: 92 ou SEQ ID NO: 94 são identificadas entre aquelas mantidas na base de dados de nucleotídeos Entrez no National Center for Biotechnology Information (NCBI) usando ferramentas de pesquisa de sequência em base de dados, tal como o Basic Local Alignment Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389- 3402). O programa é usado para encontrar regiões de similaridade local entre as sequências comparando as sequências de ácido nucléico ou de polipeptídeos com a base de dados de sequências e calculando a significância estatística das combinações. Por exemplo, o polipeptídeo codificado pelo ácido nucléico de SEQ ID NO: 90, SEQ ID NO: 92 ou SEQ ID NO: 94 é usado no algoritmo TBLASTN, com ajustes padrões e o filtro para ignorar sequências de baixa complexidade da mesmacadas. O produto da análise foi visualizado por comparação aos pares, e classificado de acordo com a classificação de

probabilidade (valor E), quando a classificação reflete a probabilidade de que um alinhamento particular ocorre ao acaso (quanto menor o valor E, mais significativo o acerto). Além dos valores E, as comparações também são classificadas por identidade percentual. A identidade percentual refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácido nucléico comparadas (ou polipeptídeos) com relação a um tamanho particular. Em alguns exemplos, os parâmetros padrões são ajustados para modificar a severidade da pesquisa. Por exemplo, o valor E é aumentado para mostrar menos combinações de severidade. Dessa maneira, combinações pequenas quase exatas são identificadas.

A tabela A4 fornece uma lista de sequências de ácidos nucleicos do tipo silky-1.

Tabela A4: Exemplos polipeptídeos do tipo silky:

Nome	Organismo	SEQ ID NO de ácido nucléico	SEQ ID NO de polipeptídeo
Homólogo de Pt_silky	<i>Populus trichocarpa</i>	90	91
Homólogo de Le_silky	<i>Solanum lycopersicum</i>	92	93
Ta_silky	<i>Triticum aestivum</i>	94	95

Em alguns exemplos, as sequências relacionadas são montadas por tentativas e reveladas publicamente por instituições de pesquisa, tal como The Institute for Genomic Research (TIGR; começando com TA). A base de dados The Eukaryotic Gene Orthologs (EGO) é usada para identificar tais sequências relacionadas, tanto por pesquisa de palavra chave quanto usando o algoritmo BLAST com a sequência de ácidos nucleicos ou sequência de polipeptídeos de interesse. Em outros exemplos, as bases de dados de sequência especial de ácidos nucleicos são criadas para organismos particulares, tal como pelo Joint Genome Institute.

1.5. Polipeptídeos ARP6

As sequências (DNAC de comprimento total, ESTs ou genômico) relacionadas à sequência de ácidos nucleicos usadas nos métodos da presente invenção foram identificadas entre aquelas mantidas na base de

dados de nucleotídeos Entrez no National Center for Biotechnology Information (NCBI) usando ferramentas de pesquisa de sequência em base de dados, tal como o Basic Local Alignment Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa é usado para encontrar regiões de similaridade local entre as sequências comparando as sequências de ácido nucléico ou de polipeptídeos com a base de dados de sequências e calculando a significância estatística das combinações. Por exemplo, o polipeptídeo codificado pelo ácido nucléico usado na presente invenção foi usado pelo algoritmo TBLASTN, com ajustes padrões e o filtro para ignorar sequências de baixa complexidade da mesmacadas. O produto da análise foi visualizado por comparação aos pares, e classificado de acordo com a classificação de probabilidade (valor E), quando a classificação reflete a probabilidade de que um alinhamento particular ocorre ao acaso (quanto menor o valor E, mais significativo 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 sequências de ácido nucléico comparadas (ou polipeptídeos) com relação a um tamanho particular. Em alguns exemplos, os parâmetros padrões podem ser ajustados para modificar a severidade da pesquisa. Por exemplo, o valor E pode ser aumentado para mostrar menos combinações de severidade. Dessa maneira, combinações pequenas quase exatas podem ser identificadas.

A tabela A5 fornece uma lista de sequências de ácidos nucleicos relacionadas à sequência de ácidos nucleicos usada nos métodos da presente invenção.

Tabela A5: Exemplos de ácidos nucleicos ARP6 e polipeptídeos codificados destes:

ARP6	Organismo fonte	SEQ ID NO de ácido nucléico:	SEQ ID NO de polipeptídeo:
A.thaliana_AT3G33520.1	<i>Arabidopsis thaliana</i>	101	102

O.sativa_LOC_Os01g16414.1	<i>Oryza sativa</i>	103	104
P.patens_126969	<i>Physcomitrella patens</i>	105	106
P.trichocarpa_scaff_XVIII.1149	<i>Populus trichoparca</i>	107	108
Gm0195x00030	<i>Glycine max</i>	109	110
Gm0057x00075	<i>Glycine max</i>	111	112

Em alguns exemplos, as sequências relacionadas foram montadas por tentativa e reveladas publicamente por instituições de pesquisa, tal como The Institute for Genomic Research (TIGR; começando com TA). A base de dados The Eukaryotic Gene Orthologs (EGO) pode ser usada para identificar tais sequências relacionadas, tanto por pesquisa de palavra chave quanto usando o algoritmo BLAST com a sequência de ácidos nucleicos ou sequência de polipeptídeos de interesse. Em outros exemplos, as bases de dados de sequência especial de ácidos nucleicos foram criadas para organismos particulares, tal como pelo Joint Genome Institute. Adicionalmente, o acesso às bases de dados prioritárias permitiu a identificação de sequências inéditas de ácido nucléico e de polipeptídeos.

1.6. Polipeptídeos POP

As sequências (DNAC de comprimento total, ESTs ou genômico) relacionadas à sequência de ácidos nucleicos usadas nos métodos da presente invenção foram identificadas entre aquelas mantidas na base de dados de nucleotídeos Entrez no National Center for Biotechnology Information (NCBI) usando ferramentas de pesquisa de sequência em base de dados, tal como o Basic Local Alignment Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa é usado para encontrar regiões de similaridade local entre as sequências comparando as sequências de ácido nucléico ou de polipeptídeos com a base de dados de sequências e calculando a significância estatística das combinações. Por exemplo, o polipeptídeo codificado pelo ácido nucléico usado na presente invenção foi usado pelo algoritmo TBLASTN, com ajustes padrões e o filtro para ignorar sequências de baixa complexidade da mesmacadas. O produto da análise foi visualizado por

comparação aos pares, e classificado de acordo com a classificação de probabilidade (valor E), quando a classificação reflete a probabilidade de que um alinhamento particular ocorre ao acaso (quanto menor o valor E, mais significativo 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 sequências de ácido nucléico comparadas (ou polipeptídeos) com relação a um tamanho particular. Em alguns exemplos, os parâmetros padrões podem ser ajustados para modificar a severidade da pesquisa. Por exemplo, o valor E pode ser aumentado para mostrar menos combinações de severidade. Dessa maneira, combinações pequenas quase exatas podem ser identificadas.

A tabela A6 fornece uma lista de sequências de ácidos nucleicos relacionadas à sequência de ácidos nucleicos usada nos métodos da presente invenção.

Tabela A6: Exemplos de polipeptídeos POP:

Nome	Fonte vegetal	SEQ ID NO de ácido nucléico	SEQ ID NO de polipeptídeo
	<i>Arabidopsis thaliana</i>	116	117
AC13782818.5#1	<i>Medicago truncatula</i>	135	145
Os02g0290600#1	<i>Oryza sativa</i>	136	146
104540#1	<i>Physcomitrella patens</i>	137	147
105959#1	<i>Physcomitrella patens</i>	138	148
scaff_XII.203#1	<i>Populus trichocarpa</i>	139	149
scaff_XV.146#1	<i>Populus trichocarpa</i>	140	150
AT4G14570	<i>Arabidopsis thaliana</i>	141	151
AT5G36210	<i>Arabidopsis thaliana</i>	142	152
EF085857.1	<i>Picea sitchensis</i>	143	153
AY108871	<i>Zea mays</i>	144	154

Em alguns exemplos, as sequências relacionadas foram montadas por tentativa e reveladas publicamente por instituições de pesquisa, tal como The Institute for Genomic Research (TIGR; começando com TA). A base de dados The Eukaryotic Gene Orthologs (EGO) pode ser usada para identificar tais sequências relacionadas, tanto por pesquisa de palavra chave quanto usando o algoritmo BLAST com a sequência de ácidos nucleicos ou

sequência de polipeptídeos de interesse. Em outros exemplos, as bases de dados de sequência especial de ácidos nucleicos foram criadas para organismos particulares, tal como pelo Joint Genome Institute. Adicionalmente, o acesso às bases de dados prioritárias permitiu a
5 identificação de sequências inéditas de ácido nucléico e de polipeptídeos.

1.7. Polipeptídeos do tipo folha amarrotada (CRL)

As sequências (DNAc de comprimento total, ESTs ou genômico) relacionadas à sequência de ácidos nucleicos usadas nos métodos da presente invenção foram identificadas entre aquelas mantidas na base de dados de nucleotídeos Entrez no National Center for Biotechnology Information (NCBI) usando ferramentas de pesquisa de sequência em base de dados, tal como o Basic Local Alignment Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; e Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). O programa é usado para encontrar regiões de similaridade
10 local entre as sequências comparando as sequências de ácido nucléico ou de polipeptídeos com a base de dados de sequências e calculando a significância estatística das combinações. Por exemplo, o polipeptídeo codificado pelo ácido nucléico usado na presente invenção foi usado pelo algoritmo TBLASTN, com ajustes padrões e o filtro para ignorar sequências de baixa
15 complexidade da mesmacadas. O produto da análise foi visualizado por comparação aos pares, e classificado de acordo com a classificação de probabilidade (valor E), quando a classificação reflete a probabilidade de que um alinhamento particular ocorre ao acaso (quanto menor o valor E, mais significativo o acerto). Além dos valores E, as comparações foram também
20 classificadas pela identidade percentual. A identidade percentual refere-se ao número de nucleotídeos idênticos (ou aminoácidos) entre as duas sequências de ácido nucléico comparadas (ou polipeptídeos) com relação a um tamanho particular. Em alguns exemplos, os parâmetros padrões podem ser ajustados para modificar a severidade da pesquisa. Por exemplo, o valor E pode ser
25

aumentado para mostrar menos combinações de severidade. Dessa maneira, combinações pequenas quase exatas podem ser identificadas.

A tabela A7 fornece uma lista de sequências de ácidos nucleicos relacionadas à sequência de ácidos nucleicos usada nos métodos da presente invenção.

Tabela A7: Exemplos de polipeptídeos CRL:

Nome	SEQ ID NO de ácido nucleico:	SEQ ID NO de proteína:
A.thaliana_AT5G51020.1#1	155	156
B.napus_TA26749_3708#1	157	158
C.clementina_DY292515#1	159	160
C.endivia_TA817_114280#1	161	162
C.sinensis_TA15191_2711#1	163	164
G.max_Gm0065x00465#1	165	166
G.raimondii_TA13121_29730#1	167	168
H.paradoxus_TA3639_73304#1	169	170
I.nil_TA8128_35883#1	171	172
L.saligna_TA1573_75948#1	173	174
M.truncatula_AC139708_31.5#1	175	176
N.tabacum_TA16794_4097#1	177	178
T.officinale_TA2756_50225#1	179	180
V.vinifera_GSVIVT00018055001#1	181	182
P.trichocarpa_834377#1	183	184
S.lycopersicum_TA38444_4081#1	185	186
S.tuberosum_TA25303_4113#1	187	188
S.officinarum_TA38665_4547#1	189	190
T.aestivum_TA92222_4565#1	191	192
S.bicolor_5284384#1	193	194
Z.mays_ZM07MC01438_57666838@1433#1	195	196
O.sativa.indica_BGIOSIBCE034659#1	197	198
O.sativa_Os11g0524300#1	199	200
P.patens_133413#1	201	202
P.taeda_TA18056_3352#1	203	204
S.moellendorffii_170435#1	205	206

Em alguns exemplos, as sequências relacionadas foram montadas por tentativa e reveladas publicamente por instituições de pesquisa, tal como The Institute for Genomic Research (TIGR; começando com TA). A base de dados The Eukaryotic Gene Orthologs (EGO) pode ser usada para identificar tais sequências relacionadas, tanto por pesquisa de palavra chave

quanto usando o algoritmo BLAST com a sequência de ácidos nucleicos ou sequência de polipeptídeos de interesse. Em outros exemplos, as bases de dados de sequência especial de ácidos nucleicos foram criadas para organismos particulares, tal como pelo Joint Genome Institute.

- 5 Adicionalmente, o acesso às bases de dados prioritárias permitiu a identificação de sequências inéditas de ácido nucléico e de polipeptídeos.

Exemplo 2: Alinhamento de sequências relacionadas à sequências de polipeptídeos usadas nos métodos da invenção

2.1. Polipeptídeos do tipo alfin

- 10 O alinhamento de sequências de polipeptídeos é realizado usando o algoritmo ClustalW 2.0 do alinhamento progressivo (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31 :3497-3500) com ajuste padrão (alinhamento lento, matriz de similaridade: Gonnet (ou Blosom 62 (se os polipeptídeos estiverem
15 alinhados), penalidade de abertura de intervalo 10, penalidade de extensão de intervalo: 0,2). A edição do menor manual é realizada para otimizar adicionalmente o alinhamento.

- 20 Uma árvore filogenética de polipeptídeos do tipo alfin é construída usando um algoritmo de agrupamento de união ao redor da maneira fornecida no programa AlignX do vetor NTI (Invitrogen).

- O alinhamento de sequências de polipeptídeos é realizado usando o algoritmo ClustalW 2.0 do alinhamento progressivo (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31 :3497-3500) com ajuste padrão (alinhamento lento, matriz de
25 similaridade: Gonnet, penalidade de abertura de intervalo 10, penalidade de extensão de intervalo: 0,2). A edição do menor manual é realizada para otimizar adicionalmente o alinhamento.

2.2. Polipeptídeos YRP

O alinhamento de sequências de polipeptídeos é realizado

usando o algoritmo ClustalW 2.0 do alinhamento progressivo (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31 :3497-3500) com ajuste padrão (alinhamento lento, matriz de similaridade: Gonnet (ou Blosom 62 (se os polipeptídeos estiverem alinhados), penalidade de abertura de intervalo 10, penalidade de extensão de intervalo: 0,2). A edição do menor manual é realizada para otimizar adicionalmente o alinhamento.

Uma árvore filogenética de polipeptídeos YRP é construída usando um algoritmo de agrupamento de união ao redor da maneira fornecida no programa AlignX do vetor NTI (Invitrogen).

O alinhamento de sequências de polipeptídeos é realizado usando o algoritmo ClustalW 2.0 do alinhamento progressivo (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31 :3497-3500) com ajuste padrão (alinhamento lento, matriz de similaridade: Gonnet, penalidade de abertura de intervalo 10, penalidade de extensão de intervalo: 0,2). A edição do menor manual é realizada para otimizar adicionalmente o alinhamento.

2.3. Polipeptídeos do tipo de Brevis Radix (BRXL)

O alinhamento múltiplo de sequência de todas as sequências de polipeptídeos BRXL na tabela A3 foi realizado usando o algoritmo ClustalW 1.81. Os resultados do alinhamento são mostrados na figura 5 do presente pedido. As características a seguir são enquadradas acentuadamente: (i) um domínio conservado 1 que representa um domínio BRX, que compreende o domínio IPR013591 DZC (entrada PFAM PF08381 DZC; marcado com X); (2) um domínio conservado 2 que representa um domínio BRX, que compreende um domínio C-terminal IPR013591 DZC (entrada PFAM PF08381 DZC; marcado com X); (3) um domínio conservado 3 e (4) um domínio conservado 4, ambos contendo Cys conservada (enquadrada levemente), cujo espaçamento é indicativo de um motivo de ligação de zinco

potencial.

2.4. Polipeptídeos do tipo silky-1

O alinhamento de sequências de polipeptídeos é realizado usando o algoritmo ClustalW 2.0 do alinhamento progressivo (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31 :3497-3500) com ajuste padrão (alinhamento lento, matriz de similaridade: Gonnet (ou Blosum 62 (se os polipeptídeos estiverem alinhados), penalidade de abertura de intervalo 10, penalidade de extensão de intervalo: 0,2). A edição do menor manual é realizada para otimizar adicionalmente o alinhamento.

Uma árvore filogenética de polipeptídeos do tipo silky-1 é construída usando um algoritmo de agrupamento de união ao redor da maneira fornecida no programa AlignX do vetor NTI (Invitrogen).

O alinhamento de sequências de polipeptídeos é realizado usando o algoritmo ClustalW 2.0 do alinhamento progressivo (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31 :3497-3500) com ajuste padrão (alinhamento lento, matriz de similaridade: Gonnet, penalidade de abertura de intervalo 10, penalidade de extensão de intervalo: 0,2). A edição do menor manual é realizada para otimizar adicionalmente o alinhamento.

2.5. Polipeptídeos ARP6

O alinhamento de sequências de polipeptídeos foi realizado usando o algoritmo ClustalW 2.0 do alinhamento progressivo (Thompson et al. (1997) Nucleic Acids Res 25:4876-4882; Chenna et al. (2003). Nucleic Acids Res 31 :3497-3500) com ajuste padrão (alinhamento lento, matriz de similaridade: Gonnet (ou Blosum 62 (se os polipeptídeos estiverem alinhados), penalidade de abertura de intervalo 10, penalidade de extensão de intervalo: 0,2). A edição do menor manual foi realizada para otimizar adicionalmente o alinhamento. Os polipeptídeos ARP6 são alinhados na

figura 8.

O alinhamento de sequências de polipeptídeos foi realizado usando o algoritmo ClustalW 2.0 do alinhamento progressivo (Thompson et al. (1997) *Nucleic Acids Res* 25:4876-4882; Chenna et al. (2003). *Nucleic Acids Res* 31 :3497-3500) com ajuste padrão (alinhamento lento, matriz de similaridade: Gonnet, penalidade de abertura de intervalo 10, penalidade de extensão de intervalo: 0,2). A edição do menor manual foi realizada para otimizar adicionalmente o alinhamento.

2.6. Polipeptídeos POP

O alinhamento de sequências de polipeptídeos foi realizado usando o algoritmo ClustalW 2.0 do alinhamento progressivo (Thompson et al. (1997) *Nucleic Acids Res* 25:4876-4882; Chenna et al. (2003). *Nucleic Acids Res* 31 :3497-3500) com ajustes padrões (alinhamento lento, matriz de similaridade: Gonnet, penalidade de abertura de intervalo 10, penalidade de extensão de intervalo: 0,2). A edição do menor manual foi realizada para otimizar adicionalmente o alinhamento. Os polipeptídeos POP são alinhados na figura 12.

A árvore filogenética de polipeptídeos POP (Figura 3, Tripathi & Sowdhamini, 2006) foi construída da maneira a seguir:

Os alinhamentos múltiplos de sequência dos domínios serina-protease foram construídos usando o programa CLUSTALW. A fim de comparar regiões equivalentes, as regiões de domínio foram recuperadas empregando HMMALIGN, uma sequência para o método de combinar perfil contra a base de dados PfamA. As proteínas que precisam de uma porção significativa do domínio do tipo protease não foram incluídas nos alinhamentos. Uma matriz Blosum 30, uma penalidade de intervalo aberto de 10 e uma penalidade de extensão de 0,05 foram empregadas. Uma árvore filogenética completa foi inferida a partir do alinhamento múltiplo de sequência com PHYLIP (pacote de inferência de fotogenia) 3.65. O

desenvolvimento individual foi realizado 100 vezes usando SEQBOOT para obter valores suportes para cada ramificação interna. As distâncias dos pares foram determinadas com PROTDIST. As árvores filogenéticas de agrupamento de vizinhos foram calculadas com NEIGHBOR usando parâmetros padrões. As árvores consensuais de papel principal de todas as sequências desenvolvidas individualmente foram obtidas com o programa CONSENSE. As representações das árvores calculadas foram construídas usando TreeView. Agrupamentos com valores de desenvolvimento individual maiores que 50 % foram definidos como subgrupos confirmados, e as sequências com valores menores foram adicionadas a estes subgrupos de acordo com sua similaridade de sequência no alinhamento, da maneira julgada pela inspeção visual.

2.7. Polipeptídeos do tipo folha amarrotada (CRL)

O alinhamento de sequências de polipeptídeos foi realizado usando o algoritmo MUSCLE 3.7 (comparação múltipla de comparação de sequência por expectativa Log) (Edgard 2004 de Nucleic Acids Research, 2004, Vol. 32, No. 5 1792-1797) com ajuste padrão (Figura 15).

Uma árvore de agrupamento vizinho foi calculada usando QuickTree (Howe et al. (2002), Bioinformatics 18(11): 1546-7). O suporte da principal ramificação após 100 repetições de desenvolvimento individual é indicado. Um filograma circular foi desenhado usando dendroscópio (Huson et al. (2007), BMC Bioinformatics 8(1):460). Ver figura 16. A proteína CRL parece ser exclusiva na maioria dos organismos.

Exemplo 3: Cálculo de identidade percentual global entre sequências de polipeptídeos usadas na realização dos métodos da invenção

3.1. Polipeptídeos do tipo alfin

Os percentuais globais de similaridade e identidade entre as sequências de comprimento total de polipeptídeos são determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Matriz 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

5 sequências de DNA ou de proteínas sem precisar de pré-alinhamento dos dados. O programa realiza uma série de alinhamentos aos pares usando o algoritmo de alinhamento global de Myers e Miller (com um penalidade de abertura de intervalo de 12, e uma penalidade de extensão de intervalo de 2), calcula a similaridade e identidade usando, por exemplo, Blosun 62 (para

10 polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade de sequência é mostrada na metade inferior da linha de divisão e a identidade de sequência é mostrada na metade superior da linha de divisão diagonal.

Os parâmetros usados na comparação são:

15

Matriz de classificação: Blosun62

Primeiro intervalo: 12

Intervalo de extensão: 2

20

Uma tabela MATGAT para alinhamento local de um domínio específico, ou dados em % de identidade/similaridade entre domínios específicos, também pode ser realizada.

3.2. Polipeptídeos YRP

25

Os percentuais globais de similaridade e identidade entre as sequências de comprimento total de polipeptídeos são determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Matriz 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 sequências de DNA ou de proteínas sem precisar de 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 identidade usando, por exemplo, Blosum 62 (para polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade de sequência é mostrada na metade inferior da linha de divisão e a identidade de sequência é mostrada na metade superior da linha de divisão diagonal.

Os parâmetros usados na comparação são:

Matriz de classificação: Blosum62

Primeiro intervalo: 12

Intervalo de extensão: 2

Uma tabela MATGAT para alinhamento local de um domínio específico, ou dados em % de identidade/similaridade entre domínios específicos, também pode ser realizada.

3.3. Polipeptídeos do tipo de Brevis Radix (BRXL)

Os percentuais globais de similaridade e identidade entre as sequências de comprimento total de polipeptídeos usadas na realização dos métodos da invenção foram determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Matriz 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 sequências de DNA ou de proteínas sem precisar de 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 identidade usando, por exemplo, Blosum 62 (para

polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade de sequência é mostrada na metade inferior da linha de divisão e a identidade de sequência é mostrada na metade superior da linha de divisão diagonal.

5 Os parâmetros usados na comparação foram:

Matriz de classificação: Blosun62

Primeiro intervalo: 12

Intervalo de extensão: 2

Os resultados da análise do software são mostrados na tabela

10 B1 para a similaridade e identidade global com relação ao comprimento total das sequências de polipeptídeos (excluindo as sequências parciais de polipeptídeos).

A identidade percentual entre as sequências de comprimento total de polipeptídeos usadas na realização dos métodos da invenção pode ser
15 tão baixa quanto 43 % de identidade de aminoácido comparada com SEQ ID NO: 18.

Tabela B1: Resultados de MatGAT para a similaridade e identidade globais com relação ao comprimento total das sequências de polipeptídeos da tabela A3.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1. Poptr_BRX		67	65	50	50	46	51	78	52	46	69	48	53	50	44	47	45	45	43	50	50	51	45	45	79	56	45	43	45	47
2. Arath_BRX	79		78	46	48	42	44	70	44	43	64	43	48	45	43	44	42	42	41	43	44	44	42	44	69	49	43	41	42	45
3. Arath_BRXL1	76	86		47	50	42	44	67	44	42	61	42	48	45	42	45	42	42	43	42	44	46	43	42	66	50	41	41	41	43
4. Arath_BRXL2	64	61	59		73	55	61	49	61	54	49	55	70	60	56	39	43	53	43	59	59	68	53	54	52	69	54	51	38	39
5. Arath_BRXL3	66	68	64	82		55	58	49	59	53	51	53	67	60	54	39	40	52	44	57	58	67	50	53	66	52	51	40	40	40
6. Arath_BRXL4	60	57	56	72	69		67	47	68	63	44	62	62	64	54	40	39	53	44	66	67	62	51	53	49	63	53	51	38	41
7. Glyma_BRXL1	66	58	58	77	75	78		50	97	72	48	73	69	72	56	41	40	55	46	75	78	66	52	55	52	72	55	52	40	39
8. Glyma_BRXL2	86	83	79	63	66	60	64		50	46	76	46	52	48	44	46	44	44	45	48	48	50	44	44	77	53	44	42	45	46
9. Glyma_BRXL3	66	59	60	77	75	79	98	65		73	48	74	71	73	56	42	41	55	46	76	80	68	53	56	53	72	56	52	39	38
10. Glyma_BRXL4	63	59	57	72	72	74	82	62	83		45	91	62	64	52	39	41	51	44	66	68	61	50	53	47	63	52	50	39	40
11. Glyma_BRXL5	81	77	74	62	66	60	66	87	66	63		45	50	48	42	43	45	41	42	46	46	47	43	42	71	53	42	41	45	44
12. Glyma_BRXL6	64	60	57	72	70	75	83	62	83	94	63		62	65	53	40	41	53	46	66	67	61	49	54	48	65	54	49	41	40
13. Medtr_BRXL1	70	63	62	83	81	76	82	66	82	75	66	76		67	59	42	44	60	48	65	68	76	57	61	57	80	59	56	38	42
14. Nicbe_BRXL1	66	60	60	75	74	78	83	61	83	78	63	77	78		56	39	42	54	45	67	70	63	53	55	51	70	54	53	39	38
15. Orysa_BRXL1	58	56	55	69	65	67	67	57	67	63	55	64	70	68		37	40	70	40	54	58	58	72	72	48	61	71	69	37	39
16. Orysa_BRXL2	59	56	55	54	52	54	53	57	53	51	56	51	56	53	52		44	41	35	41	41	41	38	39	47	42	39	38	42	72
17. Orysa_BRXL4	61	58	56	59	56	54	56	60	56	52	62	55	61	58	52	57		39	38	40	40	40	38	41	47	44	38	40	76	45
18. Orysa_BRXL5	57	54	53	65	63	64	66	55	66	62	54	63	69	66	78	54	52		42	55	57	59	69	85	48	63	82	68	38	41
19. Pisci_BRXL1	58	59	59	60	66	60	61	59	61	59	58	61	65	59	54	48	55	52		45	47	44	40	41	44	47	42	40	35	36
20. Poptr_BRXL1	64	60	56	75	73	78	86	63	87	79	61	78	78	80	66	51	54	66	60		86	66	53	56	52	70	55	51	37	39

21. Poptr_BRXL2	63	58	57	74	71	81	86	61	87	78	60	77	80	80	70	53	53	70	60	88		67	55	58	53	73	58	54	38	40
22. Poptr_BRXL3	65	59	60	82	78	76	78	63	79	74	63	74	83	75	69	54	55	67	59	79	78		58	60	53	80	59	55	38	40
23. Sorbi_BRXL1	56	55	55	66	62	64	66	56	66	62	56	60	70	66	80	52	49	79	55	65	69	68		70	46	59	70	84	38	37
24. Sorbi_BRXL2	58	54	54	67	64	65	68	57	69	65	55	64	70	65	79	52	53	89	54	67	70	69	82		48	64	92	68	38	40
25. Vitvi_BRXL1	87	79	76	66	69	62	66	84	67	63	79	64	71	66	61	57	63	59	58	65	65	67	58	60		58	47	46	45	47
26. Vitvi_BRXL2	70	63	63	84	81	78	85	66	85	76	69	78	89	83	73	54	60	72	62	82	82	87	71	73	71		61	58	40	43
27. Zeama_BRXL1	58	55	54	66	63	65	66	57	67	64	56	64	69	67	79	52	52	88	53	66	68	70	82	95	60	72		68	39	39
28. Zeama_BRXL2	54	53	53	65	63	64	65	53	65	63	56	62	67	67	78	53	50	78	54	62	67	66	89	79	57	69	79		38	37
29. Zeama_BRXL3	59	59	58	56	55	57	56	60	56	52	59	54	57	55	50	57	86	52	51	53	55	54	53	52	60	58	54	52		42
30. Zeama_BRXL4	61	57	55	54	54	56	53	60	54	53	58	53	56	53	52	82	59	51	51	50	56	53	48	50	61	56	52	49	58	

A identidade de aminoácido percentual pode ser significativamente maior se a maioria da região conservada dos polipeptídeos forem comparadas. Por exemplo, durante a comparação da sequência de aminoácidos de um domínio conservado 1 da maneira representada por SEQ ID NO: 83, ou de um domínio conservado 2 da maneira representada por SEQ ID NO: 84 com os respectivos domínios correspondentes dos polipeptídeos da tabela A3, a identidade de aminoácido percentual aumenta significativamente (em ordem de preferência de pelo menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % ou mais de identidade de sequência de aminoácidos).

3.4. Polipeptídeos do tipo silky-1

Os percentuais globais de similaridade e identidade entre as sequências de comprimento total de polipeptídeos são determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Matriz 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 sequências de DNA ou de proteínas sem precisar de 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 identidade usando, por exemplo, Blosun 62 (para polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade de sequência é mostrada na metade inferior da linha de divisão e a identidade de sequência é mostrada na metade superior da linha de divisão diagonal.

Os parâmetros usados na comparação são:

Matriz de classificação: Blosun62

Primeiro intervalo: 12

Intervalo de extensão: 2

Uma tabela MATGAT para alinhamento local de um domínio específico, ou dados em % de identidade/similaridade entre domínios específicos, também pode ser realizada.

3.5. Polipeptídeos ARP6

Os percentuais globais de similaridade e identidade entre as sequências de comprimento total de polipeptídeos usadas na realização dos métodos da invenção foram determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Matriz 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 sequências de DNA ou de proteínas sem precisar de 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 identidade usando, por exemplo, Blosum 62 (para polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade de sequência é mostrada na metade inferior da linha de divisão e a identidade de sequência é mostrada na metade superior da linha de divisão diagonal.

Os parâmetros usados na comparação foram:

Matriz de classificação: Blosum62

Primeiro intervalo: 12

Intervalo de extensão: 2

Os resultados da análise do software são mostrados na tabela B2 para a similaridade e identidade global com relação ao comprimento total

das sequências de polipeptídeos. A identidade percentual é fornecida acima da diagonal em negrito e similaridade percentual é fornecida abaixo da diagonal (face normal).

A identidade percentual entre as sequências de polipeptídeo

- 5 ARP6 usadas na realização dos métodos da invenção pode ser tão baixa quanto 50,8 % de identidade de aminoácido comparada com SEQ ID NO: 102 (*A.thaliana*_AT3G33520.1).

Tabela B2: Resultados de MatGAT para a global similaridade e identidade globais com relação ao comprimento total das sequências de polipeptídeos.

	1	2	3	4	5	6
1. Gm0057x00075		65,1	53,6	37,0	44,1	50,8
2. Gm0195x00030	72,7		76,7	50,7	64,5	70,1
3. P.trichocarpa_scaff_XVIII.1149	67,5	88,8		52,8	65,4	74,5
4. P.patens_126969	54,9	71,1	69,9		51,8	50,8
5. O.sativa_LOC_Os01g16414.1	62,0	81,9	83,0	70,4		63,1
6. A.thaliana_AT3G33520.1	63,9	83,5	88,9	68,6	79,3	

10 3.6. Polipeptídeos POP

Os percentuais globais de similaridade e identidade entre as sequências de comprimento total de polipeptídeos usadas na realização dos métodos da invenção foram determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Matriz 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 sequências de DNA ou de proteínas sem precisar de 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 identidade usando, por exemplo, Blosum 62 (para polipeptídeos), e então coloca os resultados em uma matriz de distância. A similaridade de sequência é mostrada na metade inferior da linha de divisão e

a identidade de sequência é mostrada na metade superior da linha de divisão diagonal.

Os parâmetros usados na comparação foram:

Matriz de classificação: Blosun62

5

Primeiro intervalo: 12

Intervalo de extensão: 2

Os resultados da análise do software são mostrados na tabela B3 para a similaridade e identidade global com relação ao comprimento total das sequências de polipeptídeos. A identidade percentual é fornecida acima da diagonal em negrito e a similaridade percentual é fornecida abaixo da diagonal (face normal).

A identidade percentual entre as sequências do polipeptídeo de POP usadas na realização dos métodos da invenção pode ser tão baixa quanto 16,8 % de identidade de aminoácido comparada com SEQ ID NO: 1 17 (At5g24260).

15

Tabela B3: Resultados de MatGAT para similaridade e identidade globais com relação ao comprimento total das sequências de polipeptídeos.

	1	2	3	4	5	6	7	8	9	10
1. At5g24260		65,7	51,1	59,3	59,4	65,8	65,5	18,1	16,8	16,9
2. Mt_AC137828	80,4		51,4	58,6	58,6	72,1	72,6	18,5	18,6	18,5
3. Os02g0290600	64,3	64,3		47,8	48,1	53,0	54,0	18,3	18,4	20,7
4. Pp_104540	74,7	74,0	61,6		91,9	58,2	58,5	18,8	17,0	18,2
5. Pp_105959	74,1	73,9	61,3	96,0		57,9	58,7	18,8	16,6	18,0
6. Pt_scaff_XII.203	78,9	85,0	63,5	73,0	73,2		91,1	19,3	17,7	18,8
7. Pt_scaff_XV.146	78,3	85,4	63,9	72,7	73,5	96,0		19,6	17,4	17,6
8. NP_193193.2At	36,0	37,5	32,6	36,3	36,3	37,8	36,2		19,5	18,9
9. NP_198470.3At	36,9	36,6	33,6	33,4	33,9	35,2	35,3	37,6		60,5
10. ABK25153.1Ps	35,9	35,1	35,2	35,2	34,6	34,3	34,2	38,7	74,7	

3.7. Polipeptídeos do tipo folha amarrotada (CRL)

Os percentuais globais de similaridade e identidade entre as sequências de comprimento total de polipeptídeos usadas na realização dos métodos da invenção foram determinados usando um dos métodos disponíveis na técnica, o software MatGAT (Matriz Global Alignment Tool) (BMC Bioinformatics. 2003 4:29. MatGAT: a application that generates

20

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 sequências de DNA ou de proteínas sem precisar de pré-alinhamento dos dados. O programa

5 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 identidade usando, por exemplo, Blosum 62 (para polipeptídeos), e então coloca os resultados em uma matriz de distância. A

10 similaridade de sequência é mostrada na metade inferior da linha de divisão e a identidade de sequência é mostrada na metade superior da linha de divisão diagonal.

Os parâmetros usados na comparação foram:

Matriz de classificação: Blosum62

15 Primeiro intervalo: 12

Intervalo de extensão: 2

Os resultados da análise do software são mostrados na tabela B4 para a similaridade e identidade global com relação ao comprimento total das sequências de polipeptídeos. A identidade percentual é fornecida acima da

20 diagonal em negrito e similaridade percentual é fornecida abaixo da diagonal (face normal).

A identidade percentual entre as sequências de polipeptídeos CRL usadas na realização dos métodos da invenção pode ser tão baixa quanto 48 % de identidade de aminoácido comparada com SEQ ID NO: 156.

de polipeptídeos[illegible]

Exemplo 4: Identificação de domínios compreendidos em sequências de polipeptídeos usadas na realização dos métodos da invenção

4.1. Polipeptídeos do tipo alfin

A base de dados Integrated Resource of Protein Families, Domains and Sites (InterPro) é uma interface integrada para as bases de dados de assinatura comumente usadas nas pesquisas a base de texto e de sequência. A base de dados InterPro combina estas bases de dados, que usam diferentes metodologias e graus variados de informação biológica com relação à proteínas bem caracterizadas para fornecer assinaturas protéicas. As bases de dados de colaboração incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. Pfam é uma ampla coleção dos alinhamentos múltiplos de sequência e modelos ocultos de Markov que cobrem muitos domínios e famílias comuns de proteína. Pfam é hospedado pelo servidor Sanger Institute no Reino Unido, Interpro é hospedado no European Bioinformatics Institute no Reino Unido.

4.2. Polipeptídeos YRP

A base de dados Integrated Resource of Protein Families, Domains and Sites (InterPro) é uma interface integrada para as bases de dados de assinatura comumente usadas nas pesquisas a base de texto e de sequência. A base de dados InterPro combina estas bases de dados, que usam diferentes metodologias e graus variados de informação biológica com relação à proteínas bem caracterizadas para fornecer assinaturas protéicas. As bases de dados de colaboração incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. Pfam é uma ampla coleção dos alinhamentos múltiplos de sequência e modelos ocultos de Markov que cobrem muitos domínios e famílias comuns de proteína. Pfam é hospedado pelo servidor Sanger Institute no Reino Unido, Interpro é hospedado no European Bioinformatics Institute no Reino Unido.

4.3. Polipeptídeos do tipo de Brevis Radix (BRXL)

A base de dados Integrated Resource of Protein Families, Domains and Sites (InterPro) é uma interface integrada para as bases de dados de assinatura comumente usadas nas pesquisas a base de texto e de sequência. A base de dados InterPro combina estas bases de dados, que usam diferentes metodologias e graus variados de informação biológica com relação à proteínas bem caracterizadas para fornecer assinaturas protéicas. As bases de dados de colaboração incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. Interpro é hospedado no European Bioinformatics Institute no Reino Unido.

Os resultados da varredura da InterPro da sequência de polipeptídeos da maneira representada por SEQ ID NO: 18 são apresentados na tabela C1.

Tabela C1: Resultados da varredura de InterPro da sequência de polipeptídeos da maneira representada por SEQ ID NO: 18

Número e nome do acesso InterPro	Nome da base de dados integrada	Número de acesso da base de dados integrada	Nome do acesso da base de dados integrada
IPR13591 Resistência à doença/dedo de zinco/região do tipo de condensação de cromossomo (DZC)	PFAM	PF08381	DZC
No IPR integrado	Panther	PTHR22870	Regulador de condensação de cromossomo
	Panther	PTHR22870_SF25	Regulador de condensação de cromossomo

4.4. Polipeptídeos do tipo silky

A base de dados Integrated Resource of Protein Families, Domains and Sites (InterPro) é uma interface integrada para as bases de dados de assinatura comumente usadas nas pesquisas a base de texto e de sequência. A base de dados InterPro combina estas bases de dados, que usam diferentes metodologias e graus variados de informação biológica com relação à

proteínas bem caracterizadas para fornecer assinaturas protéicas. As bases de dados de colaboração incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. Pfam é uma ampla coleção dos alinhamentos múltiplos de sequência e modelos ocultos de Markov que cobrem muitos domínios e famílias comuns de proteína. Pfam é hospedado pelo servidor Sanger Institute no Reino Unido, Interpro é hospedado no European Bioinformatics Institute no Reino Unido.

4.5. Polipeptídeos ARP6

A base de dados Integrated Resource of Protein Families, Domains and Sites (InterPro) é uma interface integrada para as bases de dados de assinatura comumente usadas nas pesquisas a base de texto e de sequência. A base de dados InterPro combina estas bases de dados, que usam diferentes metodologias e graus variados de informação biológica com relação à proteínas bem caracterizadas para fornecer assinaturas protéicas. As bases de dados de colaboração incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. Pfam é uma ampla coleção dos alinhamentos múltiplos de sequência e modelos ocultos de Markov que cobrem muitos domínios e famílias comuns de proteína. Pfam é hospedado pelo servidor Sanger Institute no Reino Unido, Interpro é hospedado no European Bioinformatics Institute no Reino Unido.

4.6. Polipeptídeos POP

A base de dados Integrated Resource of Protein Families, Domains and Sites (InterPro) é uma interface integrada para as bases de dados de assinatura comumente usadas nas pesquisas a base de texto e de sequência. A base de dados InterPro combina estas bases de dados, que usam diferentes metodologias e graus variados de informação biológica com relação às proteínas bem caracterizadas para fornecer assinaturas protéicas. As bases de dados de colaboração incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. Pfam é uma ampla coleção dos

alinhamentos múltiplos de sequência e modelos ocultos de Markov que cobrem muitos domínios e famílias comuns de proteína. Pfam é hospedado pelo servidor Sanger Institute no Reino Unido, Interpro é hospedado no European Bioinformatics Institute no Reino Unido.

5 Os resultados da varredura da InterPro da sequência de polipeptídeos da maneira representada por SEQ ID NO: 117 são apresentados na tabela C2.

Tabela C2: Resultados da varredura de InterPro (principais números de acesso) da sequência de polipeptídeos da maneira representada por SEQ ID NO: 117.

Base de dados	Número de acesso	Nome de acesso	Coordenadas de aminoácido em SEQ ID NO 117
InterPro	IPR001375	Peptidase S9, região de sítio ativo de prolil oligopeptidase	
HMMPfam	PF00326	Peptidase S9	T[544-746] 1.3e-57
InterPro	IPR002469	Peptidase S9B, terminal N de dipeptidilpeptidase IV	
HMMPfam	PF00930	DPPIV N	T[103-458] 2.1e-129
Gene3D	G3DSA:2.140.10.30	sem descrição	T[77-488] 6.3e-90
Gene3D	G3DSA:3.40.50.1820	sem descrição	T[490-746] 4.7e-73
HMMPanther	PTHR11731	Família protease S9B,C dipeptidil-peptidase IV relacionada	T[116-746] 6.1e-293
HMMPanther	PTHR11731SF12	Dipeptidil peptidase IV	T[116-746] 6.1e-293
Superfamília	SSF53474	alfa/beta-hidrolases	T[490-746] 3.4e-65
Superfamília	SSF82171	Dipeptidil peptidase IV/CD26, domínio N-terminal	T[4-486] 9e-78

4.7. Polipeptídeos do tipo folha amarrotada (CRL)

A base de dados Integrated Resource of Protein Families, Domains and Sites (InterPro) é uma interface integrada para as bases de dados de assinatura comumente usadas nas pesquisas a base de texto e de sequência.

15 A base de dados InterPro combina estas bases de dados, que usam diferentes metodologias e graus variados de informação biológica com relação à proteínas bem caracterizadas para fornecer assinaturas protéicas. As bases de dados de colaboração incluem SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom e Pfam, Smart e TIGRFAMs. Pfam é uma ampla coleção dos

20 alinhamentos múltiplos de sequência e modelos ocultos de Markov que cobrem muitos domínios e famílias comuns de proteína. Pfam é hospedado pelo servidor Sanger Institute no Reino Unido, Interpro é hospedado no

European Bioinformatics Institute no Reino Unido.

Os resultados da varredura da InterPro da sequência de polipeptídeos da maneira representada por SEQ ID NO: 156 são apresentados na tabela C3.

5 Tabela C3: Resultados da varredura de InterPro (principais números de acesso) da sequência de polipeptídeos da maneira representada por SEQ ID NO: 156.

Base de dados	Número de acesso	Descrição	Valor E	Coordenadas de aminoácido [início-fim]
Interpro	IPR010404			
PFAM	PF06206	DUF1001	0,043	[42-236]

Exemplo 5: Previsão de topologia da sequências de polipeptídeos usadas na realização dos métodos da invenção

10 5.1. Polipeptídeos do tipo alfin

TargetP 1.1 prediz a localização subcelular de proteínas eucarióticas. A designação local baseia-se na presença prevista de qualquer das pré-sequências N-terminais: peptídeo de trânsito para cloroplasto (cTP), peptídeo com alvo mitocondrial (mTP) ou peptídeo sinal da via secretória (SP). As classificações nas quais a previsão final baseia-se não são probabilidades reais, e elas não necessariamente são adicionadas. Entretanto, o local com maior classificação é o mais provável de acordo com TargetP, e a relação entre as classificações (a classe de confiança) pode ser uma indicação de como certa previsão ocorre. A classe de confiança (RC) varia de 1 a 5, quando 1 indica a previsão mais forte. TargetP é mantido no servidor da Technical University de Dinamarca. Para as sequências previstas conterem uma pré-sequência N-terminal, um sítio de clivagem potencial também pode ser previsto.

25 Inúmeros parâmetros são selecionados, tais como grupo de organismo (não planta ou planta), conjuntos de acerto (nenhum, conjunto pré-definido de acertos, ou conjunto de acertos especificado pelo usuário), e o cálculo de previsão de sítios clivados (sim ou não).

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

- ChloroP 1.1 hospedado no servidor da Technical University da Dinamarca;

- 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;

- TMHMM, hospedado no servidor da Technical University da Dinamarca

- PSORT (URL: psort.org)

- PLOC (Park e Kanehisa, Bioinformatics, 19, 1656-1663, 2003).

5.2. Polipeptídeos YRP

TargetP 1.1 prediz a localização subcelular de proteínas eucarióticas. A designação local baseia-se na presença prevista de qualquer das pré-sequências N-terminais: peptídeo de trânsito para cloroplasto (cTP), peptídeo com alvo mitocondrial (mTP) ou peptídeo sinal da via secretória (SP). As classificações nas quais a previsão final baseia-se não são probabilidades reais, e elas não necessariamente são adicionadas. Entretanto, o local com maior classificação é o mais provável de acordo com TargetP, e a relação entre as classificações (a classe de confiança) pode ser uma indicação de como certa previsão ocorre. A classe de confiança (RC) varia de 1 a 5, quando 1 indica a previsão mais forte. TargetP é mantido no servidor da Technical University of Dinamarca. Para as sequências previstas conterem uma pré-sequência N-terminal um sítio de clivagem potencial também pode ser previsto.

Inúmeros parâmetros são selecionados, tais como grupo de

organismo (não planta ou planta), conjuntos de acerto (nenhum, conjunto pré-definido de acertos, ou conjunto de acertos especificado pelo usuário), e o cálculo de previsão de sítios clivados (sim ou não).

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

- ChloroP 1.1 hospedado no servidor da Technical University da Dinamarca;
- 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;
- TMHMM, hospedado no servidor da Technical University da Dinamarca
- PSORT (URL: psort.org)
- PLOC (Park e Kanehisa, Bioinformatics, 19, 1656-1663, 2003).

5.3. Polipeptídeos do tipo silky-1

TargetP 1.1 prediz a localização subcelular de proteínas eucarióticas. A designação local baseia-se na presença prevista de qualquer das pré-sequências N-terminais: peptídeo de trânsito para cloroplasto (cTP), peptídeo com alvo mitocondrial (mTP) ou peptídeo sinal da via secretória (SP). As classificações nas quais a previsão final baseia-se não são probabilidades reais, e elas não necessariamente são adicionadas. Entretanto, o local com maior classificação é o mais provável de acordo com TargetP, e a relação entre as classificações (a classe de confiança) pode ser uma indicação de como certa previsão ocorre. A classe de confiança (RC) varia de 1 a 5, quando 1 indica a previsão mais forte. TargetP é mantido no servidor da Technical University of Dinamarca. Para as sequências previstas conterem

uma pré-sequência N-terminal um sítio de clivagem potencial também pode ser previsto.

Inúmeros parâmetros são selecionados, tais como grupo de organismo (não planta ou planta), conjuntos de acerto (nenhum, conjunto pré-definido de acertos, ou conjunto de acertos especificado pelo usuário), e o cálculo de previsão de sítios clivados (sim ou não).

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

- ChloroP 1.1 hospedado no servidor da Technical University da Dinamarca;
- 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;
- TMHMM, hospedado no servidor da Technical University da Dinamarca;
- PSORT (URL: psort.org)
- PLOC (Park e Kanehisa, Bioinformatics, 19, 1656-1663, 2003).

5.4. Polipeptídeos ARP6

TargetP 1.1 prediz a localização subcelular de proteínas eucarióticas. A designação local baseia-se na presença prevista de qualquer das pré-sequências N-terminais: peptídeo de trânsito para cloroplasto (cTP), peptídeo com alvo mitocondrial (mTP) ou peptídeo sinal da via secretória (SP). As classificações nas quais a previsão final baseia-se não são probabilidades reais, e elas não necessariamente são adicionadas. Entretanto, o local com maior classificação é o mais provável de acordo com TargetP, e a relação entre as classificações (a classe de confiança) pode ser uma indicação

de como certa previsão ocorre. A classe de confiança (RC) varia de 1 a 5, quando 1 indica a previsão mais forte. TargetP é mantido no servidor da Technical University of Dinamarca.

Para as sequências previstas conterem uma pré-sequência N-terminal um sítio de clivagem potencial também pode ser previsto.

Inúmeros parâmetros são selecionados, tais como grupo de organismo (não planta ou planta), conjuntos de acerto (nenhum, conjunto pré-definido de acertos, ou conjunto de acertos especificado pelo usuário), e o cálculo de previsão de sítios clivados (sim ou não).

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

- ChloroP 1.1 hospedado no servidor da Technical University da Dinamarca;

- 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 de Alberta, Edmonton, Alberta, Canadá;

- TMHMM, hospedado no servidor da Technical University da

Dinamarca

- PSORT (URL: psort.org)

- PLOC (Park e Kanehisa, Bioinformatics, 19, 1656-1663, 2003).

5.5. Polipeptídeos POP

TargetP 1.1 prediz a localização subcelular de proteínas eucarióticas. A designação local baseia-se na presença prevista de qualquer das pré-sequências N-terminais: peptídeo de trânsito para cloroplasto (cTP), peptídeo com alvo mitocondrial (mTP) ou peptídeo de sinal da via secretória (SP). As classificações nas quais a previsão final baseia-se não são

probabilidades reais, e elas não necessariamente são adicionadas. Entretanto, o local com maior classificação é o mais provável de acordo com TargetP, e a relação entre as classificações (a classe de confiança) pode ser uma indicação de como certa previsão ocorre. A classe de confiança (RC) varia de 1 a 5, quando 1 indica a previsão mais forte. TargetP é mantido no servidor da Technical University da Dinamarca.

Para as sequências previstas conterem uma pré-sequência N-terminal um sítio de clivagem potencial também pode ser previsto.

Inúmeros parâmetros foram selecionados, tais como grupo de organismo (não planta ou planta), conjuntos de acerto (nenhum, conjunto pré-definido de acertos, ou conjunto de acertos especificado pelo usuário), e o cálculo de previsão de sítios clivados (sim ou não).

Os resultados da análise TargetP 1.1 da sequência de polipeptídeos da maneira representada por SEQ ID NO: 117 são apresentados Tabela D1. O grupo de organismo “planta” foi selecionado, nenhum acerto foi definido e o comprimento previsto do peptídeo trânsito solicitado. A localização subcelular da sequência de polipeptídeos da maneira representada por SEQ ID NO: 117 pode estar no citoplasma ou núcleo, nenhum peptídeo trânsito é previsto.

Tabela D1: Análise de TargetP 1.1 da sequência de polipeptídeos da maneira representada por SEQ ID NO: 117. Abreviações: Len, comprimento; cTP, peptídeo de trânsito para cloroplasto; mTP, peptídeo de trânsito mitocondrial, SP, peptídeo sinal da via secretória, outro, outro alvo subcelular, Loc, local previsto; RC, classe de confiança; TPlen, tamanho previsto de peptídeo trânsito.

Nome	Len	cTP	mTP	SP	outro	Loc	RC	TPlen
POP	746	0,083	0,100	0,079	0,914	-	1	
acerto		0,000	0,000	0,000	0,000			

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

- ChloroP 1.1 hospedado no servidor da Technical University

da Dinamarca;

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

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

- TMHMM, hospedado no servidor da Technical University da Dinamarca

- PSORT (URL: psort.org)

10 • PLOC (Park e Kanehisa, Bioinformatics, 19, 1656-1663, 2003).

5.6. Polipeptídeos do tipo folha amarrotada (CRL)

TargetP 1.1 prediz a localização subcelular de proteínas eucarióticas. A designação local baseia-se na presença prevista de qualquer
15 das pré-sequências N-terminais: peptídeo de trânsito para cloroplasto (cTP),
peptídeo com alvo mitocondrial (mTP) ou peptídeo sinal da via secretória
(SP). As classificações nas quais a previsão final baseia-se não são
probabilidades reais, e elas não necessariamente são adicionadas. Entretanto,
o local com maior classificação é o mais provável de acordo com TargetP, e a
20 relação entre as classificações (a classe de confiança) pode ser uma indicação
de como certa previsão ocorre. A classe de confiança (RC) varia de 1 a 5,
quando 1 indica a previsão mais forte. TargetP é mantido no servidor da
Technical University da Dinamarca.

25 Para as sequências previstas conterem uma pré-sequência N-terminal um sítio de clivagem potencial também pode ser previsto.

Inúmeros parâmetros são selecionados, tais como grupo de organismo (não planta ou planta), conjuntos de acerto (nenhum, conjunto pré-definido de acertos, ou conjunto de acertos especificado pelo usuário), e o cálculo de previsão de sítios clivados (sim ou não).

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

- ChloroP 1.1 hospedado no servidor da Technical University da Dinamarca;

- 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;

- TMHMM, hospedado no servidor da Technical University da Dinamarca

- PSORT (URL: psort.org)

- PLOC (Park e Kanehisa, Bioinformatics, 19, 1656-1663, 2003).

O resultado da análise de proteína usando alguns dos algoritmos anteriormente mencionados é mostrado a seguir:

Psort: retículo endoplasmático 0,600

WOLFPsort: cito: 4,0, extr: 3,0, E. R.: 2,5, E.R._plas: 2,5, plas: 1,5, nucl: 1,0, mito: 1,0

TargetP: chloro:0,775 qualidade 4

ChloroP: não associado ao cloroplasto 0,497

MitoProtl: não mitocondrial 0,027

PTS 1 peroxissomo: não alvejado

Fobius: existe um domínio de membrana [19-36].

Exemplo 6: Previsão da localização da subcelular das sequências de polipeptídeos usadas na realização dos métodos da invenção

6.1. Polipeptídeos do tipo de Brevis Radix (BRXL)

Os métodos experimentais para localização de proteína variam de imunolocalização a marcação de proteínas usando proteína verde

fluorescente (GFP) ou beta-glucuronidase (GUS). Tais métodos para identificar compartimentalização subcelular de polipeptídeos BRXL são bem conhecidos na técnica. A previsão computacional de localização de proteína a partir dos dados da sequência foi realizada. Entre os algoritmos bem conhecidos pelos versados na técnica são disponíveis nas ferramentas proteômicas ExPASy hospedadas no Swiss Institute for Bioinformatics, por exemplo, PSort, TargetP, ChloroP, LocTree, Predotar, Lipop, MITOPROT, PATS, PTS1, SignalP, TMHMM, TMpred e outros.

TargetP 1.1 prediz a localização subcelular de proteínas eucarióticas. A designação local baseia-se na presença prevista de qualquer das pré-sequências N-terminais: peptídeo de trânsito para cloroplasto (cTP), peptídeo com alvo mitocondrial (mTP) ou peptídeo sinal da via secretória (SP). As classificações nas quais a previsão final baseia-se não são probabilidades reais, e elas não necessariamente são adicionadas. Entretanto, o local com maior classificação é o mais provável de acordo com TargetP, e a relação entre as classificações (a classe de confiança) pode ser uma indicação de como certa previsão ocorre. A classe de confiança (RC) varia de 1 a 5, quando 1 indica a previsão mais forte. TargetP é mantido no servidor da Technical University da Dinamarca.

Para as sequências previstas conterem uma pré-sequência N-terminal um sítio de clivagem potencial também pode ser previsto.

Inúmeros parâmetros foram selecionados, tais como grupo de organismo (não planta ou planta), conjuntos de acerto (nenhum, conjunto pré-definido de acertos, ou conjunto de acertos especificado pelo usuário), e o cálculo de previsão de sítios clivados (sim ou não).

Os resultados da análise TargetP 1.1 da sequência de polipeptídeos da maneira representada por SEQ ID NO: 1 17 são apresentados na tabela a seguir. O grupo de organismo “planta” foi selecionado, e nenhum acerto foi definido. A localização subcelular prevista da sequência de

polipeptídeos da maneira representada por SEQ ID NO: 117 não é no cloroplasto, não é mitocondrial e não é a via secretória, mas é muito provavelmente no núcleo.

Tabela mostrando a análise TargetP 1.1 da sequência de polipeptídeos da maneira representada por SEQ ID NO: 117

Tamanho (AA)	362
Peptídeo de trânsito de cloroplasto	0,132
Peptídeo de trânsito mitocondrial	0,089
Peptídeo de sinal da via secretória	0,112
Outro alvo celular	0,909
Localização prevista	Outro
Classe de confiança	2

Exemplo 7: Ensaio relacionado à sequências de polipeptídeos usadas na realização dos métodos da invenção

7.1. Polipeptídeos do tipo de Brevis Radix (BRXL)

Os polipeptídeos BRXL usados nos métodos da presente invenção (pelo menos na sua forma nativa) tipicamente, mas não necessariamente, apresentam atividade regulatória transcricional e a capacidade de interagir com outras proteínas. A atividade de ligação de DNA e interações proteína-proteína pode ser determinada *in vitro* ou *in vivo* usando técnicas bem conhecidas na tecnologia (por exemplo, em Current Protocols in Molecular Biology, Volumes 1 e 2, Ausubel et al. (1994), Current Protocols). Os polipeptídeos BRXL compreendem os domínios BRX conservados que mostraram mediar a interação proteína-proteína em experimentos de dois híbridos em levedura. A homodimerização e heterodimerização no BRXL e/ou entre este e também famílias de proteína do tipo PRAF são conhecidas na técnica (Briggs et al. (2006) Plant Physiol 140: 1307-1316; van Leeuwen et al. (2004) Trends Plant Sci 9: 378-384).

7.2. Polipeptídeos POP

A atividade POP pode ser medida da maneira descrita por Bastos et al. (Biochem J. 388: 29-38, 2005): A atividade POP é determinada medindo a fluorescência de AMC (7-amido-4- metilcoumarin) liberada por

hidrólise do substrato enzimático N-Suc-Gly-Pro-Leu-Gly- Pro-AMC, quando Sue corresponde a succinil. O POP é ensaiado no tampão de reação [Hepes 25 mM e DTT 5 mM (ditiotreitól), pH 7,5] contendo substrato 20 μ M em volume final 100 μ L. A fluorescência de AMC liberada pela reação enzimática é registrada da maneira descrita em Grellier et al. (J. Biol. Chem. 276, 47078-47086, 2001). A atividade POP também pode ser ensaiada usando peptídeos diferentes nas mesmas condições experimentais (N-Boc-Val-Leu-Lys-AMC, N-Boc-Leu-Lys-Arg-AMC, N-Cbz-Val-Lys-Met-AMC, N-Boc-Leu-Gly-Arg-AMC, N- Boc-Ile-Gly-Gly-Arg-AMC, N-Suc-Leu-Tyr-AMC, N-Suc-Ala-Ala-Ala-AMC, N-Boc-Val-Pro- Arg-AMC, N-Suc-Gly-Pro-AMC, N-Cbz-Gly-Gly-Arg-AMC, N-Suc-Ala-Ala-Pro-Phe-AMC, N- Cbz-Phe-Arg-AMC, H-Gly-Arg-AMC, H-Gly-Phe-AMC, Ala-Ala-Phe-AMC, L-Arg-AMC e L-Ala-AMC e L-Lys-Ala-AMC, quando Boc e Cbz correspondem a t-butoxicarbonil e benziloxycarbonil respectivamente). Para determinar parâmetros cinéticos, POP recombinante (0,67 ng) ou nativo (0,26 ng) é incubado no tampão de reação com concentrações variáveis de substrato N-Suc-Gly-Pro-Leu-Gly- Pro-AMC (3,12-100 μ M) e o AMC liberado é medido da maneira descrita anteriormente. Os valores K_m e V_{max} são determinados regressão hiperbólica usando o método de Comish-Bowden. O k_{cat} é calculado usando $k_{cat} = V_{max}/[E]_0$ quando $[E]_0$ representa a concentração enzimática ativa. A quantificação de POP ativo é realizada por incubação da proteína com concentrações séricas do inibidor clorometano POP Tc80 irreversível, da maneira descrita em Grellier et al. (2001).

Exemplo 8: Clonagem da sequência de ácidos nucleicos usada nos métodos da invenção

8.1. Polipeptídeos do tipo alfin

A sequência de ácidos nucleicos é amplificada por PCR usando como molde de uma biblioteca de DNAc (em pCMV Sport 6.0; Invitrogen, Paisley, UK). A PCR é realizada usando Hifi Taq DNA

polimerase em condições padrões, usando 200 ng de molde em 50 µL de uma mistura de PCR. Os oligonucleotídeos iniciadores incluem os sítios AttB para recombinação do tipo Gateway. O fragmento de PCR amplificado também é purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, é então realizada, durante a qual o fragmento de PCR recombina *in vivo* com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada". O plasmídeo pDONR201 é obtido da Invitrogen, como parte da tecnologia Gateway®.

O clone de entrada compreendendo SEQ ID NO: 1 ou SEQ ID NO: 3 é então usado em uma reação LR com um vetor de destino usado para a transformação de *Oryza sativa*. Este vetor contém como elementos funcionais nas bordas do DNA-T: um marcador selecionável de planta; um cassete de expressão de marcador de classificação; e um cassete do tipo Gateway pretendido para recombinação LR *in vivo* com a sequência de ácidos nucleicos de interesse já clonada no clone de entrada. Um promotor GOS2 de arroz (SEQ ID NO: 7) para expressão constitutiva está localizada à montante deste cassete Gateway.

Após a etapa de recombinação de LR, o vetor resultante de expressão pGOS2:: tipo alfin (Figura 1) é transformado na cepa *Agrobacterium* LBA4044 de acordo com métodos bem conhecidos na técnica.

8.2. Polipeptídeos YRP

A sequência de ácidos nucleicos é amplificada por PCR usando como molde de uma biblioteca de DNAC (em pCMV Sport 6.0; Invitrogen, Paisley, UK). A PCR é realizada usando Hifi Taq DNA polimerase em condições padrões, usando 200 ng de molde em 50 µL de uma mistura de PCR. Os oligonucleotídeos iniciadores incluem os sítios AttB para recombinação do tipo Gateway. O fragmento de PCR amplificado também é purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, é então realizada, durante a qual o fragmento de PCR

recombina *in vivo* com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada". O plasmídeo pDONR201 é obtido da Invitrogen, como parte da tecnologia Gateway®.

O clone de entrada compreendendo SEQ ID NO: 10 ou SEQ ID NO: 12 é então usado em uma reação LR com um vetor de destino usado para a transformação de *Oryza sativa*. Este vetor contém como elementos funcionais nas bordas do DNA-T: um marcador selecionável de planta; um cassete de expressão de marcador de classificação; e um cassete do tipo Gateway pretendido para recombinação LR *in vivo* com a sequência de ácidos nucleicos de interesse já clonada no clone de entrada. Um promotor GOS2 de arroz (SEQ ID NO: 16) para expressão constitutiva está localizada à montante deste cassete Gateway.

Após a etapa de recombinação de LR, o vetor resultante de expressão pGOS2::YRP (Figura 2) é transformado na cepa *Agrobacterium* LBA4044 de acordo com métodos bem conhecidos na técnica.

8.3. Polipeptídeos do tipo Brevis Radix (BRXL)

A menos que de outra forma declarada, as técnicas de DNA recombinante são realizadas de acordo com protocolos padrões descritos em (Sambrook (2001) Molecular Cloning: um laboratory manual, 3rd Edition Cold Spring Harbor Laboratory Press, CSH, Nova Iorque) ou em Volumes 1 e 2 de Ausubel et al. (1994), Current Protocols em Biologia molecular, Current Protocols. Os materiais e métodos padrões para o trabalho molecular em planta são descritos em Plant Molecular Biology Labfax (1993) por R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (UK) e Blackwell Scientific Publications (UK).

A sequência de ácidos nucleicos de álamo que codifica uma sequência de polipeptídeo BRXL da maneira representada por SEQ ID NO: 18 foi amplificada por PCR usando como molde um banco de DNAC construído usando RNA de plantas de álamo em diferentes estágios de

desenvolvimento. Os oligonucleotídeos iniciadores a seguir, que incluem os sítios AttB para recombinação do tipo Gateway, foram usados para amplificação por PCR:

1) prm 11475 (SEQ ID NO: 88, sentido):

5'-ggggacaagtttgtagcaaaaaagcaggcttaacaatgtttacgtgcatagc-3'

2) prm1 1476 (SEQ ID NO: 89, anti-sentido, complementar):

5'-ggggaccactttgtacaagaaagctgggtgcaatttaggtcatgggaaat-3'

A PCR foi realizado usando Hifi Taq DNA polimerase em condições padrões. Um fragmento de PCR do tamanho esperado (incluindo os sítios AttB) foi amplificado e purificado também usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante a qual o fragmento de PCR recombinau *in vivo* com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada". O plasmídeo pDONR201 foi obtido da Invitrogen, como parte da tecnologia Gateway®.

O clone de entrada compreendendo SEQ ID NO: 17 foi subsequentemente usado em uma reação LR com um vetor de destino usado para a transformação de *Oryza sativa*. Este vetor continha como elementos funcionais nas bordas do DNA-T: um marcador selecionável de planta; um cassete de expressão de marcador de classificação; e um cassete do tipo Gateway pretendido para recombinação LR *in vivo* com a sequência de ácidos nucleicos de interesse já clonada no clone de entrada. Um promotor GOS2 de arroz (SEQ ID NO: 87) para expressão constitutiva estava localizado à montante deste cassete Gateway.

Após a etapa de recombinação de LR, o vetor resultante de expressão pGOS2::BRXL (Figura 6) para expressão constitutiva foi transformado na cepa de *Agrobacterium* LBA4044 de acordo com métodos bem conhecidos na técnica.

8.4. Polipeptídeos do tipo silky

A sequência de ácidos nucleicos é amplificada por PCR usando como molde de uma biblioteca de DNAC (em pCMV Sport 6.0; Invitrogen, Paisley, UK). A PCR é realizada usando Hifi Taq DNA polimerase em condições padrões, usando 200 ng de molde em 50 µL de uma
 5 mistura de PCR. Os oligonucleotídeos iniciadores incluem os sítios AttB para recombinação do tipo Gateway. O fragmento de PCR amplificado também é purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, é então realizada, durante a qual o fragmento de PCR recombina *in vivo* com o plasmídeo pDONR201 para produzir, de acordo com
 10 a terminologia Gateway, um "clone de entrada". O plasmídeo pDONR201 é obtido da Invitrogen, como parte da tecnologia Gateway®.

O clone de entrada compreendendo SEQ ID NO: 90, SEQ ID NO: 92 ou SEQ ID NO: 94 é então usado em uma reação LR com um vetor de destino usado para a transformação de *Oryza sativa*. Este vetor contém
 15 como elementos funcionais nas bordas do DNA-T: um marcador selecionável de planta; um cassete de expressão de marcador de classificação; e um cassete do tipo Gateway pretendido para recombinação LR *in vivo* com a sequência de ácidos nucleicos de interesse já clonada no clone de entrada. Um promotor GOS2 de arroz (SEQ ID NO: 96) para expressão constitutiva está localizada à
 20 montante deste cassete Gateway.

Após a etapa de recombinação de LR, o vetor resultante de expressão pGOS2:: tipo silky-1 (Figura 7) é transformado na cepa de *Agrobacterium* LBA4044 de acordo com métodos bem conhecidos na técnica.

8.5. polipeptídeos ARP6

25 A sequência de ácidos nucleicos usada nos métodos da invenção foi amplificada por PCR usando como molde uma biblioteca de DNAC de mudas de *Arabidopsis thaliana* feita sob medida (em pCMV Sport 6.0; Invitrogen, Paisley, UK). A PCR foi realizado usando Hifi Taq DNA polimerase em condições padrões, usando 200 ng de molde em 50 µL de uma

mistura de PCR. Os oligonucleotídeos iniciadores usados foram: (SEQ ID NO: 113; sentido, códon inicial em negrito): 5'-ggggacaagtttgtacaaaaagcaggcttaacaatgtc aaacatcggtgttcta-3' e (SEQ ID NO: 114; anti-sentido, complementar): 5'-ggggaccact
 5 ttgtacaagaaagctgggttcattgtgatattgttttggtt-3', que incluem os sítios AttB para recombinação do tipo Gateway. O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante a qual o fragmento de PCR recombinau *in vivo* com o plasmídeo pDONR201 para produzir, de acordo
 10 com a terminologia Gateway, um "clone de entrada", pARP6. Plasmídeo pDONR201 foi obtido da Invitrogen, como parte da tecnologia Gateway®.

O clone de entrada compreendendo SEQ ID NO: 101 foi então usado em uma reação LR com um vetor de destino usado para a transformação de *Oryza sativa*. Este vetor continha como elementos
 15 funcionais nas bordas do DNA-T: um marcador selecionável de planta; um cassete de expressão de marcador de classificação; e um cassete do tipo Gateway pretendido para recombinação LR *in vivo* com a sequência de ácidos nucleicos de interesse já clonada no clone de entrada. Um promotor GOS2 de arroz (SEQ ID NO: 115) para a expressão constitutiva específica estava
 20 localizado à montante deste cassete Gateway.

Após a etapa de recombinação de LR, o vetor resultante de expressão pGOS2::ARP6 (Figura 10) foi transformado na cepa *Agrobacterium* LBA4044 de acordo com métodos bem conhecidos na técnica.

8.6. Polipeptídeos POP

25 A sequência de ácidos nucleicos usada nos métodos da invenção foi amplificada por PCR usando como molde uma biblioteca de DNAC de mudas de *Arabidopsis thaliana* feita sob medida (em pCMV Sport 6.0; Invitrogen, Paisley, UK). A PCR foi realizado usando Hifi Taq DNA polimerase em condições padrões, usando 200 ng de molde em 50 µL de uma

mistura de PCR. Os oligonucleotídeos iniciadores usados foram prmO561 1 (SEQ ID NO: 132; sentido, códon inicial em **negrito**): 5'-ggggacaagtttgtacaaaaaagcaggc **taaacaatggcggataaggacgtt**-3' e prmO5612 (SEQ ID NO: 133; anti-sentido, complementar): 5'-
 5 ggggaccactttgtacaagaaagctgggtaagcaacaacaggttctg tga- 3', que incluem os
 sítios AttB para recombinação do tipo Gateway. O fragmento de PCR
 amplificado também foi purificado usando métodos padrões. A primeira etapa
 do procedimento Gateway, a reação BP, foi então realizada, durante a qual o
 fragmento de PCR recombinau *in vivo* com o plasmídeo pDONR201 para
 10 produzir, de acordo com a terminologia Gateway, um "clone de entrada",
 pPOP. Plasmídeo pDONR201 foi obtido da Invitrogen, como parte da
 tecnologia Gateway®.

O clone de entrada compreendendo SEQ ID NO: 116 foi então
 usado em uma reação LR com um vetor de destino usado para a
 15 transformação de *Oryza sativa*. Este vetor continha como elementos
 funcionais nas bordas do DNA-T: um marcador selecionável de planta; um
 cassete de expressão de marcador de classificação; e um cassete do tipo
 Gateway pretendido para recombinação LR *in vivo* com a sequência de ácidos
 nucleicos de interesse já clonada no clone de entrada. Um promotor GOS2 de
 20 arroz (SEQ ID NO: 131) para a expressão constitutiva específica estava
 localizado à montante deste cassete Gateway.

Após a etapa de recombinação de LR, o vetor resultante de
 expressão pGOS2::POP (Figura 13) foi transformado na cepa de
Agrobacterium LBA4044 de acordo com métodos bem conhecidos na técnica.

25 8.7. Polipeptídeos do tipo folha amarrotada (CRL)

A sequência de ácidos nucleicos usada nos métodos da
 invenção foi amplificada por PCR usando como molde uma biblioteca de
 DNAC de mudas de *Arabidopsis thaliana* feita sob medida (em pCMV Sport
 6.0; Invitrogen, Paisley, UK). A PCR foi realizado usando Hifi Taq DNA

polimerase em condições padrões, usando 200 ng de molde em 50 µL de uma mistura de PCR. Os oligonucleotídeos iniciadores usados foram um oligonucleotídeo iniciador da maneira representada por (SEQ ID NO: 211; sentido, códon inicial em negrito): 5'-ggggacaagtttgta caaaaaagcaggcttaacaatgggtaccgagtcggg-3¹ e um oligonucleotídeo iniciador da maneira representada por SEQ ID NO: 212; anti-sentido, complementar): 5'-ggggaccactttgtacaagaaagctgggt tcagacaatagaaaaggggt- 3', que incluem os sítios AttB para recombinação do tipo Gateway. O fragmento de PCR amplificado também foi purificado usando métodos padrões. A primeira etapa do procedimento Gateway, a reação BP, foi então realizada, durante a qual o fragmento de PCR recombinau *in vivo* com o plasmídeo pDONR201 para produzir, de acordo com a terminologia Gateway, um "clone de entrada", pCRL. O plasmídeo pDONR201 foi obtido da Invitrogen, como parte da tecnologia Gateway®.

O clone de entrada compreendendo SEQ ID NO; 155 foi então usado em uma reação LR com um vetor de destino usado para a transformação de *Oryza sativa*. Este vetor continha como elementos funcionais nas bordas do DNA-T: um marcador selecionável de planta; um cassete de expressão de marcador de classificação; e um cassete do tipo Gateway pretendido para recombinação LR *in vivo* com a sequência de ácidos nucleicos de interesse já clonada no clone de entrada. Um promotor GOS2 de arroz (SEQ ID NO: 213) para a expressão constitutiva específica estava localizado à montante deste cassete Gateway.

Após a etapa de recombinação de LR, o vetor resultante de expressão pGOS2::CRL (Figura 3) foi transformado na cepa de *Agrobacterium* LBA4044 de acordo com métodos bem conhecidos na técnica.

Exemplo 9: Transformação de planta

Transformação de arroz

O *Agrobacterium* contendo a vetor de expressão é usado para

transformar plantas de *Oryza sativa*. As sementes secas maduras do cultivar de arroz japonica Nipponbare são descascadas. A esterilização é realizada incubando por um minuto em etanol a 70 %, seguido por 30 minutos em 0,2 % de HgCb, seguido por uma lavagem de 6 vezes por 15 minutos com água destilada estéril. As sementes estéreis são então germinadas em um meio contendo 2,4-D (meios de indução de calo). Após a incubação no escuro por quatro semanas, os calos embriogênicos derivados de escutelo são excisados e propagados no mesmo meio. Após duas semanas, os calos são multiplicados ou propagados por subcultura no mesmo meio por mais 2 semanas. Pedacos de calo embriogênico são sub-cultivados em meio fresco 3 dias antes da co-cultivação (para estimular a atividade de divisão celular).

A cepa de *Agrobacterium* LBA4404 contendo a vetor de expressão é usada para co-cultivo. O *Agrobacterium* é inoculado no meio AB com os antibióticos apropriados e cultivados por 3 dias a 28 °C. As bactérias são então coletadas e suspensas em meio líquido de co-cultivo em uma densidade (OD₆₀₀) de cerca de 1. A suspensão é então transferida para uma placa de Petri e o calo é imerso na suspensão por 15 minutos. O tecido caloso é então absorvido em um papel de filtro e transferido para o meio de co-cultivo solidificado e incubado por 3 dias no escuro a 25 °C. Os calos co-cultivados são crescidos em meio contendo 2,4-D por 4 semanas no escuro a 28 °C, na presença de um agente de seleção. Durante este período, ilhas de calos que crescem rapidamente se desenvolveram. Após a transferência deste material para um meio de regeneração e incubação na luz, o potencial embriogênico é liberado e os ramos desenvolvidos nas próximas quatro a cinco semanas. Os ramos são excisados a partir dos calos e incubados por 2 a 3 semanas em um meio contendo auxina, a partir do qual eles são transferidos para o solo. Os ramos endurecidos são crescidos em alta umidade e dias curtos em uma estufa.

Aproximadamente 35 transformantes de arroz TO

independentes são geradas para um construto. Os transformantes primários são transferidos de uma câmara de cultura de tecido para uma estufa. Após uma análise de PCR quantitativa para verificar o número de cópias do DNA-T inserido, apenas cópias simples de plantas transgênicas que exibem tolerância
5 ao agente de seleção são mantidas para coleta de semente T1. As sementes são então coletadas três a cinco meses após o transplante. Ver Aldemita e Hodges1996, Chan et al. 1993, Hiei et al. 1994.

Transformação de milho

A transformação de milho (*Zea mays*) é realizada com uma
10 modificação do método descrito por Ishida et al. (1996) Nature Biotech 14(6): 745-50. A transformação é dependente do genótipo em milho e apenas genótipos específicos são amenos á transformação e regeneração. A linhagem natural A188 (University de Minnesota) ou híbridos com A188 como um ancestral são boas fontes de material doador para transformação, mas outros
15 genótipos podem ser usados também satisfatoriamente. As espigas são coletadas de planta de milho aproximadamente 11 dias após a polinização (DAP), quando o comprimento do embrião imaturo é cerca de 1 a 1,2 mm. Os embriões imaturos são co-cultivados com *Agrobacterium tumefaciens* contendo a vetor de expressão, e as plantas transgênicas são recuperadas por
20 meio da organogênese. Os embriões excisados são crescidos em meio de indução de calo, a seguir em meio de regeneração de milho contendo o agente de seleção (por exemplo, imidazolinona, mas vários marcadores de seleção podem ser usados). As placas de Petri são incubadas na luz a 25 °C por 2-3 semanas, ou até que os ramos se desenvolvam. Os ramos verdes são
25 transferidos de cada embrião para o meio de enraizamento de milho e incubados a 25 °C por 2-3 semanas, até que as raízes se desenvolvam. Os ramos enraizados são transplantados para o solo na estufa. As sementes T1 são produzidas de plantas que exibem a tolerância para o agente de seleção e que contêm uma única cópia do DNA-T inserido.

Transformação do trigo

A transformação de trigo é realizada de acordo com o método descrito por Ishida et al. (1996) Nature Biotech 14(6): 745-50. O cultivar Bobwhite (disponível pela CIMMYT, México) é comumente usado na transformação. Os embriões imaturos são co-cultivados com *Agrobacterium tumefaciens* contendo o vetor de expressão, e plantas transgênicas são recuperadas por meio de organogênese. Após a incubação com *Agrobacterium*, os embriões são crescidos *in vitro* no meio de indução de calo, e então em meio de regeneração contendo o agente de seleção (por exemplo, imidazolinona, mas vários marcadores de seleção podem ser usados). As placas de Petri são incubadas na luz a 25 °C por 2-3 semanas, ou até que os ramos se desenvolvam. Os ramos verdes são transferidos de cada embrião para o meio de enraizamento e incubados a 25 °C por 2-3 semanas, até que as raízes se desenvolvam. Os ramos enraizados são transplantados para o solo na estufa. As sementes T1 são produzidas a partir das plantas que exibem tolerância ao agente de seleção e que contêm uma única cópia do DNA-T inserido.

Transformação da soja

A soja é transformada de acordo com uma modificação do método descrito na patentes de Texas A&M U.S. 5.164.310. Diversas variedades de soja comercial são propensas à transformação por este método. O cultivar Jack (disponível pela Illinois Seed foundation) é comumente usado para a transformação. As sementes de soja são esterilizadas para o semeio *in vitro*. O hipocótilo, a radícula e um cotilédone são excisados de mudas de sete dias. O epicótilo e o cotilédone remanescente são crescidos adicionalmente para desenvolver nódulos auxiliares. Estes nódulos auxiliares são excisados e incubados com *Agrobacterium tumefaciens* contendo o vetor de expressão. Após o tratamento de co-cultivo, os explantes são lavados e transferidos para os meios de seleção. Os ramos regenerados são excisados e colocados em um

meio de alongamento de ramo. Ramos de não mais que 1 cm são colocados em meio de enraizamento até que as raízes se desenvolvam. Os ramos enraizados são transplantados para o solo na estufa. As sementes T1 são produzidas a partir de plantas que exibem tolerância ao agente de seleção e

5 que contêm uma única cópia do DNA-T inserido.

Transformação de colza/canola

Os pecíolos cotiledonários e hipocótilos de mudas de 5-6 dias são usados como explantes para cultura de tecido e transformados de acordo com Babic et al. (1998, Plant Cell Rep 17: 183- 188). O cultivar Westar

10 comercial (Agricultura do Canadá) é a variedade padrão usada para a transformação, mas outras variedades também podem ser usadas. As sementes de canola são esterilizadas na superfície para semeio *in vitro*. Os explantes de pecíolo de cotilédone o cotilédone anexado são excisados das mudas *in vitro* e inoculados com *Agrobacterium* (contendo o vetor de expressão) imergindo a

15 extremidade do explante de pecíolo na suspensão bacteriana. Os explantes são então cultivados por 2 dias em meio MSBAP-3 contendo 3 mg/L de BAP, sacarose 3 %, Phytagar 0,7 % a 23 °C, 16 horas na luz. Após dois dias de co-cultivo com *Agrobacterium*, os explantes de pecíolo são transferidos para o meio MSBAP-3 contendo 3 mg/L de BAP, cefotaxima, carbenicilina ou

20 timentina (300 mg/L) por 7 dias, e então cultivados em meio MSBAP-3 com cefotaxima, carbenicilina, ou timentina e agente de seleção, até a regeneração do ramo. Quando os ramos apresentam 5 - 10 mm de comprimento, eles são cortados e transferidos para o meio de alongamento de ramo (MSBAP-0,5, contendo 0,5 mg/L de BAP). Os ramos de cerca de 2 cm de comprimento são

25 transferidos para o meio de enraizamento (MSO) para indução de raiz. Os ramos enraizados são transplantados para o solo na estufa. As sementes T1 são produzidas a partir de plantas que exibem tolerância ao agente de seleção e que contêm uma única cópia do DNA-T inserido.

Transformação de alfafa

Um clone de regeneração de alfafa (*Medicago sativa*) é transformado usando o método de (McKersie et al., 1999 Plant Physiol 119: 839-847). A regeneração e transformação de alfafa são dependentes do genótipo e, portanto, uma planta de regeneração é exigida. Os métodos para obter as plantas de regeneração foram descritos. Por exemplo, estes podem ser selecionados do cultivar Rangelander (Agricultura do Canadá) ou qualquer outra variedade comercial de alfafa da maneira descrita por Brown DCW e A. Atanassov (1985. Plant Cell Tissue Organ Culture 4: 111-112). Alternativamente, a variedade RA3 (University of Wisconsin) foi selecionada para uso em cultura de tecido (Walker et al., 1978 Am J Bot 65:654-659). Os explantes de pecíolo são co-cultivados com uma cultura de *Agrobacterium tumefaciens* C58C1 pMP90 crescida por toda a noite (McKersie et al., 1999 Plant Physiol 119: 839-847) ou LBA4404 contendo o vetor de expressão. Os explantes são co-cultivados por 3 dias no escuro em meio de indução SH contendo 288 mg/L de Pro, 53 mg/L de tioprolina, 4,35 g/L de K₂SO₄, e 100 µm de acetossiringinona. Os explantes são lavados em meio de Murashige-Skoog com metade da concentração (Murashige e Skoog, 1962) e colocados no mesmo meio de indução SH sem acetossiringinona, mas com um agente de seleção adequado e antibiótico adequado para inibir o crescimento de *Agrobacterium*. Após várias semanas, os embriões somáticos são transferidos para o meio de desenvolvimento BOi2Y sem reguladores de crescimento, sem antibióticos, e 50 g/L de sacarose. Os embriões somáticos são subsequentemente germinados em meio de Murashige-Skoog com a metade da concentração. As mudas enraizadas foram transplantadas em potes e crescidas em uma estufa. As sementes T1 são produzidas a partir das plantas que exibem tolerância ao agente de seleção e que contêm uma única cópia do DNA-T inserido.

Transformação de algodão

O algodão é transformado usando *Agrobacterium tumefaciens*

de acordo com o método descrito em U.S. 5.159.135. As sementes de algodão são esterilizadas na superfície em solução de hipoclorito de sódio a 3 % durante 20 minutos e lavadas em água destilada com 500 µg/mL de cefotaxima. As sementes são então transferidas para meio SH com 50µg/mL de benomil para a germinação. Os hipocótolos de mudas de 4 a 6 dias são removidos, cortados em pedaços de 0,5 cm e são colocados em 0,8 % de agar. Uma suspensão de *Agrobacterium* (aproximadamente 10⁸ células por mL, diluída de uma cultura crescida por toda a noite transformada com o gene de interesse e marcadores de seleção adequados) é usada para a inoculação dos explantes de hipocótilo. Após 3 dias em temperatura ambiente e na luz, os tecidos são transferidos para um meio sólido (1,6 g/L de Gelrite) com sais de Murashige e Skoog com vitamina B5 (Gamborg et al., Exp. Cell Res. 50:151-158 (1968)), 0,1 mg/L de 2,4-D, 0,1 mg/L de 6- furfurilaminopurina e 750 µg/mL de MgCL₂, e com 50 a 100 µg/mL de cefotaxima e 400-500 µg/mL de carbenicilina para matar as bactérias residuais. As linhagens celulares individuais são isoladas após dois a três meses (com subculturas a cada quatro a seis semanas) e são cultivadas adicionalmente em meio seletivo para a amplificação do tecido (30 °C, 16 horas em fotoperíodo). Os tecidos transformados são subsequentemente cultivados de maneira adicional em meio não seletivo durante 2 a 3 meses para fornecer a origem de embriões somáticos. Os embriões que parecem saudáveis de pelo menos 4 mm de tamanho são transferidos para tubos com meio SH em vermiculita fina, suplementado com 0,1 mg/L de ácido indol acético, 6 furfurilaminopurina e ácido giberélico. Os embriões são cultivados a 30 °C com um fotoperíodo de 16 horas, e plântulas no estágio de 2 a 3 folhas são transferidas para potes com vermiculita e nutrientes. As plantas são endurecidas e subsequentemente movidas para a estufa para cultivo adicional.

Exemplo 10: Procedimento de avaliação fenotípica

10.1 Estrutura de avaliação

Aproximadamente 35 transformantes de arroz TO independentes são gerados. Os transformantes primários são transferidos de uma câmara de cultura de tecido para uma estufa para crescimento e coleta de semente T1. Seis eventos, dos quais a progênie T1 segregada 3:1 quanto a presença/ausência do transgene, são mantidos. Para cada um destes eventos, aproximadamente 10 mudas T1 contendo o transgene (hetero e homozigotos) e aproximadamente 10 mudas T1 que precisam do transgene (nulizigotos) são selecionadas monitorando a expressão de marcador visual. As plantas transgênicas e os nulizigotos correspondentes são crescidos lado a lado em posições aleatórias. As condições da estufa são para dias curtos (12 horas na luz), 28 °C na luz e 22 °C no escuro, e uma umidade relativa de 70 %.

Quatro eventos T1 foram avaliados adicionalmente na geração T2 após o mesmo procedimento de avaliação como para a geração T1, mas com mais indivíduos por evento. A partir do estágio de semeio até o estágio de maturidade, as plantas passam várias vezes por uma cabine de imageamento digital. Em cada ponto de tempo, imagens digitais (2048x1536 pixels, 16 milhões de cores) são obtidas de cada planta a partir de pelo menos 6 ângulos diferentes.

Seleção por secura

As plantas com sementes T2 são crescidas em potes com solo em condições normais, até que elas alcancem o estágio principal. Elas são então transferidas para uma seção "seca" quando a irrigação é retirada. As sondas de umidade são inseridas em potes escolhidos aleatoriamente para monitorar o teor de água no solo (SWC). Quando SWC pertence a certos limites, as plantas são automaticamente re-regadas continuamente até que um nível normal seja atingido novamente. As plantas são então re-transferidas novamente para as condições normais. O restante do cultivo (maturação da planta, coleta da semente) é o mesmo das plantas não crescidas em condições de estresse abiótico. Os parâmetros e rendimentos de crescimento são

registrados da maneira detalhada para o crescimento em condições normais.

Seleção pela eficiência do uso de nitrogênio

As plantas de arroz de sementes T2 são crescidas em potes com solo em condições normais, exceto com relação à solução nutriente. Os
5 potes são aguados a partir do transplante até a maturação com uma solução de nutriente específica contendo menos teor de nitrogênio (N), em geral entre 7 a 8 vezes menos. O restante do cultivo (maturação da planta, coleta da semente) é o mesmo das plantas não crescidas em estresse abiótico. Os parâmetros e rendimentos de crescimento são registrados da maneira detalhada para o
10 crescimento em condições normais.

Seleção pelo estresse pelo sal

As plantas são crescidas em um substrato feito de fibras de coco e argex (razão 3 a 1). Uma solução de nutriente normal é usada durante as primeiras duas semanas após o transplante das plântulas na estufa. Após as
15 duas primeiras semanas, sal 25 mM (NaCl) é adicionado à solução de nutriente, até que as plantas sejam coletadas. Os parâmetros relacionados às sementes são então medidos.

10.2 Análise estatística: teste F

Um ANOVA de dois fatores (análise de variantes) foi usado
20 como um modelo estatístico para a avaliação completa das características fenotípicas da planta. Um teste F foi realizado em todos os parâmetros medidos de todas as plantas de todos os eventos transformadas com o gene da presente invenção. O teste F foi realizado para conferir um efeito do gene com relação a todos os eventos de transformação e para verificar um efeito total do
25 gene, também conhecido como um efeito global do gene. O limite para a significância para um efeito global verdadeiro de gene foi ajustado em uma probabilidade de 5 % para o teste F. Um valor de teste F significativo aponta para um efeito de gene, o que significa que não é apenas a mera presença ou posição do gene que é o que causa as diferenças no fenótipo.

Em decorrência de dois experimentos com eventos que se sobrepõem serem realizados, uma análise combinada foi realizada. Isto é usado é para conferir a consistência dos efeitos com relação a dois experimentos, e se este for o caso, para acumular a evidência de ambos os experimentos a fim de aumentar a confiança na conclusão. O método usado foi uma abordagem de modelo misto que leva em consideração a estrutura em multinível dos dados (isto é, experimento - evento - segregantes). Os valores P foram obtidos comparando o teste da razão de probabilidade com as distribuições do quiquadrado.

10.3 Parâmetros medidos

Medição de parâmetro relacionado à biomassa

A partir do estágio de semeio até o estágio de maturidade as plantas passaram várias vezes por uma cabine de imageamento digital. Em cada ponto do tempo as imagens (2048x1536 pixels, 16 milhões de cores) foram obtidas de cada planta de pelo menos 6 ângulos diferentes.

A área aérea da planta (ou biomassa folhosa) foi determinada contando o número total de pixels nas imagens digitais da parte aérea das plantas discriminadas nos fundamentos da invenção. Este valor foi o valor médio das fotos obtidas ao mesmo tempo dos diferentes ângulos e foi convertido em um valor de superfície física expresso em mm quadrado por calibração. Os experimentos mostram que a área aérea da planta medida da mesma maneira corresponde à biomassa de parte aérea das plantas. A área aérea é a área medida no momento em que a planta atingiu sua biomassa folhosa máxima. O vigor inicial é a área aérea da planta (muda) em três semanas após a germinação. O aumento na biomassa da raiz é expresso como um aumento na biomassa total da raiz (medido como biomassa de raízes máxima observada durante a vida útil de uma planta); ou como um aumento no índice de raiz/ramo (medido como a razão entre massa da raiz e massa do ramo no período de crescimento ativo de raiz e ramo).

O vigor inicial foi determinado contando o número total de pixels da parte aérea das plantas discriminadas nos fundamentos da invenção. Este valor obtido pela média das figuras obtidas ao mesmo tempo de diferentes ângulos e foi convertido em um valor de superfície física expresso em mm quadrado por calibração. Os resultados descritos a seguir são para plantas três semanas após a germinação.

A determinação do início da floração foi da maneira descrita em WO 2007/093444

Medições de parâmetro relacionado à semente

As panículas maduras primárias foram coletadas, contadas, embaladas, marcadas com código de barra e então secas por três dias em um forno a 37 °C. As panículas foram então debulhadas e todas as sementes foram coletadas e contadas. As cascas preenchidas foram separadas das vazias usando um dispositivo de soprar. As cascas vazias foram descartadas e a fração remanescente foi contada novamente. As cascas preenchidas foram pesadas em uma balança analítica. O número de sementes preenchidas foi determinado contando o número de cascas preenchidas que permaneceu após a etapa de separação. O peso total da semente por planta foi medido pesando todas as cascas preenchidas coletadas de uma planta. O número total de semente por planta foi medido contando o número de cascas coletadas de uma planta. O peso de mil grãos (TKW) é extrapolado do número de sementes preenchidas contado e seu total peso. O índice da colheita (HI) na presente invenção é definido como a razão entre o peso total da semente por planta e a área aérea (mm^2), multiplicado por um fator 10^6 . O número total de flores por panícula da maneira definida na presente invenção é a razão entre o número total de sementes e o número de panículas primárias maduras. A taxa de preenchimento da semente da maneira definida na presente invenção é a proporção (expressa com um %) do número de sementes preenchidas com relação ao número total de sementes (ou flores pequenas).

Exemplos 11: Resultados da avaliação fenotípica das plantas transgênicas

11.1. Polipeptídeos do tipo de Brevis Radix (BRXL)

Os resultados da avaliação de plantas de arroz transgênicas da geração T1 que expressam a sequência de ácidos nucleicos que codifica um polipeptídeo BRXL da maneira representada por SEQ ID NO: 18, no controle de um promotor constitutivo, são apresentados na tabela D a seguir.

Houve um aumento significativo em altura da planta, e no peso de mil grãos (TKW), das plantas transgênicas com relação aos controles.

Tabela D1: Resultados da avaliação de plantas de arroz transgênicas da geração T1 que expressam a sequência de ácidos nucleicos que codifica um polipeptídeo BRXL da maneira representada por SEQ ID NO: 18, no controle de um promotor para expressão constitutiva.

Característica	Aumento % na média total em 6 eventos na geração T1
Altura da planta	4%
Peso por mil grãos	7%

11.2. Polipeptídeos ARP6

Os resultados da avaliação de plantas transgênicas de arroz na geração T2 e que expressam um ácido nucléico compreendendo o quadro de leitura aberto mais longo em SEQ ID NO: 101 em condições de não estresse são apresentados a seguir. Ver os exemplos prévios para detalhes nas gerações das plantas transgênicas.

Os resultados da avaliação das plantas transgênicas de arroz em condições de não estresse são apresentados a seguir. Um aumento de pelo menos 5 % foi observado para o rendimento total da semente (peso total da semente), número de sementes preenchidas (número de sementes preenchidas), taxa (taxa de preenchimento) e índice da colheita (índice da colheita).

Tabela D2:

Característica relacionada ao rendimento	Aumento % em planta transgênica comparada à planta controle
Peso total da semente	12,0

Taxa de preenchimento	13,2
Índice da colheita	13,0
Número de semente preenchida	9,8

11.3. Polipeptídeos POP

Os resultados da avaliação das plantas transgênicas de arroz na geração T2 e que expressam um ácido nucléico que codifica o polipeptídeo de POP de SEQ ID NO: 117 em condições de não estresse são apresentados a seguir na tabela D3. Quando crescidas em condições de não estresse, um aumento de pelo menos 5 % foi observado para a biomassa aérea (área máxima), biomassa da raiz (raiz máxima e espessura máxima da raiz) e para o rendimento da semente (peso total de sementes, número de sementes preenchidas, taxa de preenchimento, índice da colheita). Além do mais, as plantas que expressam um ácido nucléico POP mostraram uma taxa de crescimento mais rápida (um período mais curto (em dias) necessário entre o semeio e o dia que a planta atinge 90 % de sua biomassa final (área do ciclo) e um início precoce de floração (Tempo de floração: tempo (em dias) entre o semeio e a emergência da primeira panícula).

Tabela D3: Sumário de dados para plantas transgênicas de arroz; para cada parâmetro, o aumento percentual completo é mostrado para a confirmação (geração T2), para cada parâmetro o valor p é <0,05.

Parâmetro	Aumento total
Área máxima	11,0
Tempo para florar	7,6
Raiz máxima	7,9
Peso total da semente	35,0
Número de semente preenchida	30,4
Taxa de preenchimento	31,3
Índice de coleta	21,0
Área do ciclo	5,1
Espessura máxima da raiz	12,6

11.4. Polipeptídeos do tipo folha amarrotada (CRL)

Os resultados da avaliação de plantas transgênicas de arroz na geração T2 e que expressam um ácido nucléico compreendendo o quadro de

leitura aberto mais longo em SEQ ID NO: 155 avaliado em condições de não estresse são apresentados a seguir. Ver os exemplos prévios para detalhes nas gerações das plantas transgênicas.

Os resultados da avaliação de plantas transgênicas de arroz em condições de não estresse são apresentados a seguir. Um aumento de pelo menos 5 % foi observado para o rendimento total da semente (peso total da semente), número de sementes preenchidas (número de sementes preenchidas), taxa de preenchimento (taxa de preenchimento) e índice da colheita (índice da colheita).

10 Tabela D4.

Parâmetro	Aumento % em plantas transgênicas comparado às controle
Peso total da semente	17,4
Número de semente preenchida	15,5
Taxa de preenchimento	17,7
Índice de colheita	15,9

REIVINDICAÇÕES

1. Método para aumentar tolerância a estresse abiótico e/ou para aumentar uma característica relacionada ao rendimento em uma planta em relação a uma planta controle correspondente, caracterizado pelo fato de que compreende modular a expressão em uma célula de planta, planta ou

5

parte da mesma de um ácido nucleico selecionado do grupo consistindo em:

10

a) uma sequência de ácidos nucleicos que codifica um polipeptídeo do tipo Brevis Radix (BRXL), em que dito polipeptídeo BRXL compreende pelo menos dois domínios BRX com um domínio DZC de entrada InterPro IPR013591 (DZC de entrada Pfam PF08381);

b) um ácido nucleico que codifica um polipeptídeo prolil oligopeptidase (POP), em que dito polipeptídeo POP compreende um domínio Peptidade_S9 (entrada PFAM PF00326) e domínio DPPIV_N (entrada Pfam PF00930);

15

c) um ácido nucleico que codifica um polipeptídeo ARP6;

d) um ácido nucleico que codifica um polipeptídeo CRL, em que dito polipeptídeo CRL compreende um domínio DUF206;

e) um ácido nucleico que codifica um polipeptídeo tipo alfin ou um ortólogo ou parálogo do mesmo;

20

f) um ácido nucleico que codifica um polipeptídeo YRP ou um ortólogo ou parálogo do mesmo;

g) um ácido nucleico que codifica um polipeptídeo tipo silky-1 ou um ortólogo ou parálogo do mesmo;

e opcionalmente selecionar uma célula de planta, planta ou parte da mesma tendo tolerância a estresse abiótico aumentado ou uma característica relacionada ao rendimento aumentada.

25

2. Método de acordo com a reivindicação 1, caracterizado pelo fato de que dita expressão modulada é efetuada pela introdução e expressão do ácido nucleico em uma planta ou célula de planta.

3. Método de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que:

a) o ácido nucleico que codifica um polipeptídeo BRXL compreende

5 (i) um polinucleotídeo codificando uma sequência de aminoácidos tendo pelo menos 50% de identidade de sequência com a sequência de aminoácidos do domínio BRX de SEQ ID NO: 81 e uma sequência de aminoácidos tendo pelo menos 50% de identidade de sequência com a sequência de aminoácidos do domínio BRX de SEQ ID NO:82;

10 (ii) um polinucleotídeo codificando uma sequência de aminoácidos tendo pelo menos 50% de identidade de sequência com a sequência de aminoácidos do Domínio Conservado 1 de SEQ ID NO: 83 e uma sequência de aminoácidos tendo pelo menos 50% de identidade com a sequência de aminoácidos do Domínio Conservado 2 de SEQ ID NO: 84;

15 (iii) um polinucleotídeo codificando uma sequência de aminoácidos tendo pelo menos 50% de identidade de sequência com a sequência de aminoácidos do Domínio Conservado 3 de SEQ ID NO: 85 e uma sequência de aminoácidos tendo pelo menos 50% de identidade com a sequência de aminoácidos do Domínio Conservado 4 de SEQ ID NO: 86;

20 (iv) um polinucleotídeo codificando uma sequência de aminoácidos tendo pelo menos 40% de identidade de sequência com o polipeptídeo de SEQ ID NO: 18 ou com qualquer uma das sequências de polipeptídeos listadas na Tabela A;

25 (v) um polinucleotídeo codificando uma sequência de aminoácidos que interage com ela mesma ou com qualquer outro polipeptídeo BRXL em um experimento de duplo híbrido em levedura;

(vi) qualquer uma das sequências de ácido nucleico listadas na Tabela A3 ou uma porção das mesmas;

(vii) uma sequência de ácidos nucleicos capaz de hibridizar

com qualquer uma das sequências de ácidos nucleicos listadas na Tabela A3 ou com um complemento das mesmas; ou

(viii) uma sequência de ácidos nucleicos codificando um ortólogo ou parálogo de qualquer uma das sequências de polipeptídeos listadas na Tabela A3.

b) o ácido nucleico que codifica um polipeptídeo POP compreende

(i) um polinucleotídeo codificando um polipeptídeo compreendendo um ou mais dos motivos 1 a 14 (SEQ ID NO: 118 a SEQ 10 131); ou

(ii) um polinucleotídeo codificando um polipeptídeo listado na Tabela A6 ou um ortólogo ou parálogo do mesmo, uma porção do dito nucleotídeo ou um ácido nucleico capaz de hibridizar com dito polinucleotídeo;

c) o ácido nucleico que codifica um polipeptídeo ARP6 compreende

(i) um polinucleotídeo que codifica um polipeptídeo tendo pelo menos 30% de identidade de sequência com a sequência de aminoácidos de SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106; SEQ ID NO: 108, 20 SEQ ID NO: 110, SEQ ID NO: 112; ou

(ii) um polinucleotídeo codificando um polipeptídeo listado na Tabela A5 ou um ortólogo ou parálogo do mesmo, uma porção do dito polinucleotídeo ou um ácido nucleico capaz de hibridizar com dito polinucleotídeo;

d) o ácido nucleico codificando um polipeptídeo CRL compreende

(i) um polinucleotídeo codificando um polipeptídeo compreendendo um motivo tendo pelo menos 50% de identidade de sequência com qualquer um dos seguintes motivos:

(1) Motivo 15: EQAFWRXPXKPFRQR (SEQ ID NO: 207);

(2) Motivo 16: NFCDR (SEQ ID NO: 208);

(3) Motivo 17: RGKRCLYEGS (SEQ ID NO: 209);

(4) Motivo 18: QVWGXXKXGPYEFK (SEQ ID NO: 210);

5 em que X representa qualquer aminoácido; ou

(ii) um polinucleotídeo codificando um polipeptídeo listado na Tabela A7 ou um ortólogo ou parálogo do mesmo, uma porção do dito polinucleotídeo ou um ácido nucleico capaz de hibridizar com dito polinucleotídeo;

10 e) o ácido nucleico que codifica um polipeptídeo tipo alfin ou um ortólogo ou parálogo do mesmo compreende um polinucleotídeo codificando um polipeptídeo listado na Tabela A1 ou um ortólogo ou parálogo do mesmo, uma porção do dito polinucleotídeo ou um ácido nucleico capaz de hibridizar com dito polinucleotídeo;

15 (f) o ácido nucleico que codifica um polipeptídeo YRP ou um ortólogo ou parálogo do mesmo compreende um polinucleotídeo codificando um polipeptídeo listado na Tabela A2 ou um ortólogo ou parálogo do mesmo, uma porção do dito polinucleotídeo ou um ácido nucleico capaz de hibridizar com dito polinucleotídeo; ou

20 (g) o ácido nucleico que codifica um polipeptídeo tipo silky-1 ou um ortólogo ou parálogo do mesmo compreende um polinucleotídeo codificando um polipeptídeo listado na Tabela A4 ou um ortólogo ou parálogo do mesmo, uma porção do dito polinucleotídeo ou um ácido nucleico capaz de hibridizar com dito polinucleotídeo.

25 4. Método de acordo com qualquer uma das reivindicações 1 a 3, caracterizado pelo fato de que dito ácido nucleico é operavelmente ligado a um promotor constitutivo, um promotor GOS2 ou um promotor GOS2 de arroz.

5. Método de acordo com qualquer uma das reivindicações 1 a

4, caracterizado pelo fato de que:

a) o ácido nucleico codificando um polipeptídeo tipo alfin ou um ortólogo ou parálogo do mesmo é de uma planta, de uma planta monocotiledônea ou da cepa *Allium*;

5 b) o ácido nucleico codificando um polipeptídeo YRP ou um ortólogo ou parálogo do mesmo é de uma planta, de uma planta monocotiledônea ou de *Hordeum Vulgare*;

c) o ácido nucleico codificando um polipeptídeo BRXL é de uma planta, de uma planta dicotiledônea ou da família *Salicaceae* ou *Populus trichocarpa*;

d) o ácido nucleico codificando um polipeptídeo tipo silky-1 ou um ortólogo ou parálogo do mesmo é de uma planta, de uma planta dicotiledônea ou da família *Salicaceae* ou *Populus trichocarpa*;

e) o ácido nucleico codificando um polipeptídeo ARP6 é de uma planta, uma planta dicotiledônea ou da família *Brassicaceae*, do gênero *Arabidopsis* ou *Arabidopsis thaliana*;

f) o ácido nucleico codificando um polipeptídeo POP é de uma planta, uma planta dicotiledônea ou da família *Brassicaceae*, do gênero *Arabidopsis* ou *Arabidopsis thaliana*;

20 g) o ácido nucleico codificando um polipeptídeo CRL é de uma planta, uma planta dicotiledônea ou da família *Brassicaceae*, do gênero *Arabidopsis* ou *Arabidopsis thaliana*.

6. Método de acordo com qualquer uma das reivindicações 1 a 5, caracterizado pelo fato de que dita característica relacionada ao rendimento aumentada é um ou mais de aumento da altura da planta, aumento no peso de mil grãos (TKW), maior rendimento da semente por planta, maior número de sementes preenchidas, maior biomassa ou modificação do tempo de floração em relação a plantas controle.

7. Método de acordo com a reivindicação 6, caracterizado pelo

fato de que ditas características relacionadas ao rendimento são obtidas sob condições sem estresse.

8. Método de acordo com a reivindicação 6, caracterizado pelo fato de que ditas características relacionadas ao rendimento são obtidas sob condições de estresse por seca, estresse salino ou deficiência de nitrogênio.

9. Método de acordo com qualquer uma das reivindicações 1 a 8, caracterizado pelo fato de que dita expressão modulada é efetuada por um ou mais de: marcação da ativação do DNA-T, TILLING ou recombinação homóloga.

10. Construto, caracterizado pelo fato de que compreende:
(i) um ácido nucleico como definido na reivindicação 1 ou 3;
(ii) uma ou mais sequências controle capazes de dirigir a expressão da sequência de ácidos nucleicos de (i); e opcionalmente
(iii) uma sequência de terminação de transcrição.

11. Construto de acordo com a reivindicação 10, caracterizado pelo fato de que ditas uma ou mais sequências controle é um promotor constitutivo, um promotor GOS2, um promotor GOS2 de arroz ou um promotor GOS2 de SEQ ID NO: 87.

12. Planta transgênica tendo tolerância a estresse abiótico e/ou características relacionadas ao rendimento aumentadas em relação a plantas controle correspondentes, ou uma célula de planta transgênica derivada da dita planta transgênica; célula de planta, planta ou parte da mesma, incluindo sementes obtíveis pelo método como definido em qualquer uma das reivindicações 1 a 9; partes cultiváveis da dita planta; produtos derivados de dita planta e/ou de ditas partes cultiváveis, caracterizadas pelo fato de que dita planta transgênica é resultante da expressão modulada de um ácido nucleico como definido na reivindicação 1 ou 3; e/ou pelo fato de que dita planta ou parte da mesma obtível pelo método de qualquer uma das reivindicações 1 a 9 compreende um ácido nucleico como definido na reivindicação 1 ou 3;

e/ou pelo fato de que dita planta, parte de planta ou célula de planta é transformada com um construto como definido na reivindicação 10 ou 11, em que dita planta ou parte da mesma é uma planta de cultivo, uma monocotiledônea, um cereal, tais como arroz, milho, trigo, cevada, milheto, centeio, tritcale, sorgo e aveias, ou uma célula de planta transgênica derivada da dita planta transgênica; e/ou pelo fato de que ditas partes cultiváveis da dita planta compreendem um ácido nucleico como definido na reivindicação 1 ou 3 e em que ditas partes cultiváveis são biomassa dos ramos e/ou sementes; e/ou pelo fato de que ditos produtos derivados de dita planta e/ou de ditas partes cultiváveis compreendem o constructo como definido na reivindicação 10 ou 11.

13. Método para a produção de plantas transgênicas com tolerância a estresse abiótico aumentada e/ou características relacionadas ao rendimento aumentadas com relação a plantas controle, caracterizado pelo fato de que compreende:

- (i) introduzir e expressar em uma célula de planta o construto como definido na reivindicação 10 ou 11; e
 - (ii) cultivar a célula da planta sob condições que promovem estresse abiótico ou
- cultivar a célula da planta sob condições que promovem crescimento e desenvolvimento da planta.

14. Uso de um construto como definido na reivindicação 10 ou 11 ou uso de um ácido nucleico como definido na reivindicação 1 ou 3, caracterizado pelo fato de ser em um método para produzir plantas com tolerância a estresse abiótico aumentada e/ou características relacionadas ao rendimento aumentadas com relação a plantas controle.

15. Molécula de ácido nucleico isolada, caracterizada pelo fato de que compreende:

- (i) a sequência de ácidos nucleicos de SEQ ID NO: 75, SEQ

ID NO: 77, SEQ ID NO: 79 ou SEQ ID NO: 195, ou o complemento da mesma;

(ii) uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL tendo pelo menos 50% de identidade de sequência com a sequência de polipeptídeos de SEQ ID NO: 76, SEQ ID NO: 78, ou SEQ ID NO: 80;

(iii) um ácido nucleico que codifica o polipeptídeo de SEQ ID NO: 196;

(iv) um ácido nucleico tendo pelo menos 30% de identidade de sequência com qualquer uma das sequências de aminoácidos da Tabela A7;

(v) um ácido nucleico que codifica um polipeptídeo CRL tendo pelo menos 50% de identidade de sequência com a sequência de aminoácidos de SEQ ID NO: 196 ou com qualquer uma das outras sequências de aminoácidos da Tabela A7; ou

(vi) uma molécula de ácido nucleico que hibridiza com um ácido nucleico de (i) a (v) sob severas condições de hibridização.

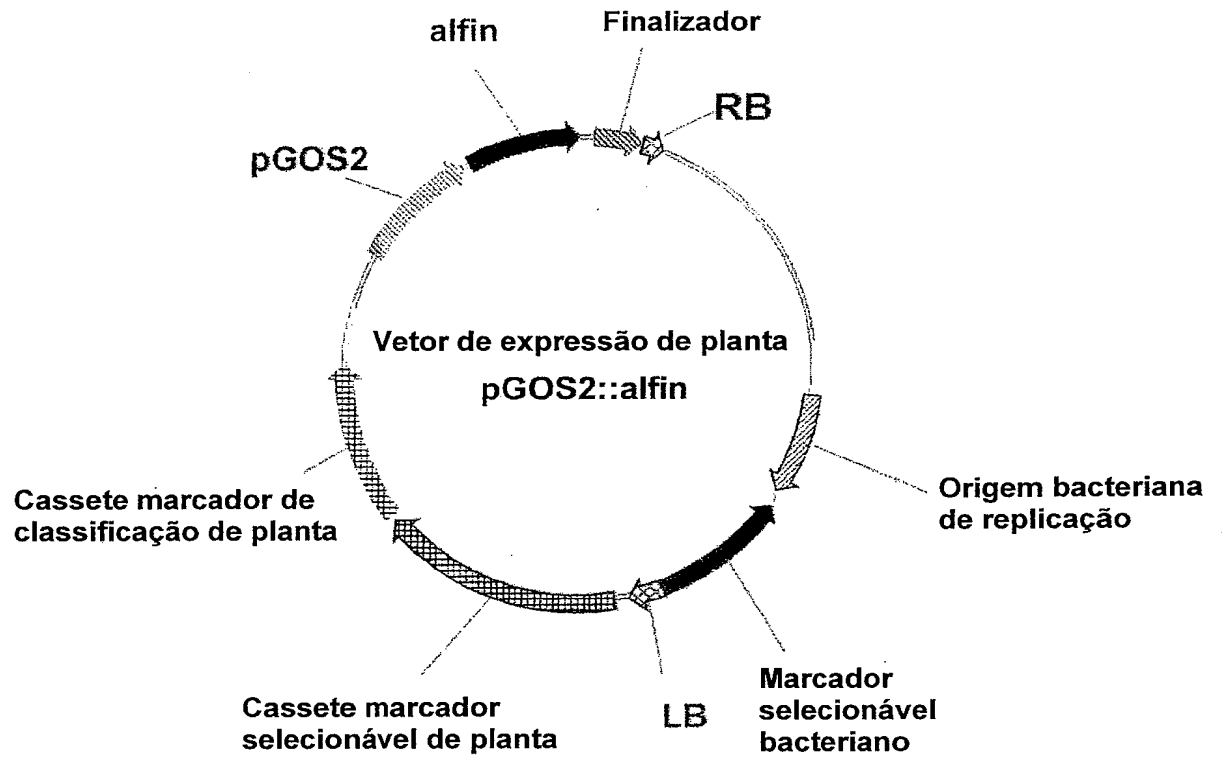
16. Polipeptídeo isolado, caracterizado pelo fato de que compreende:

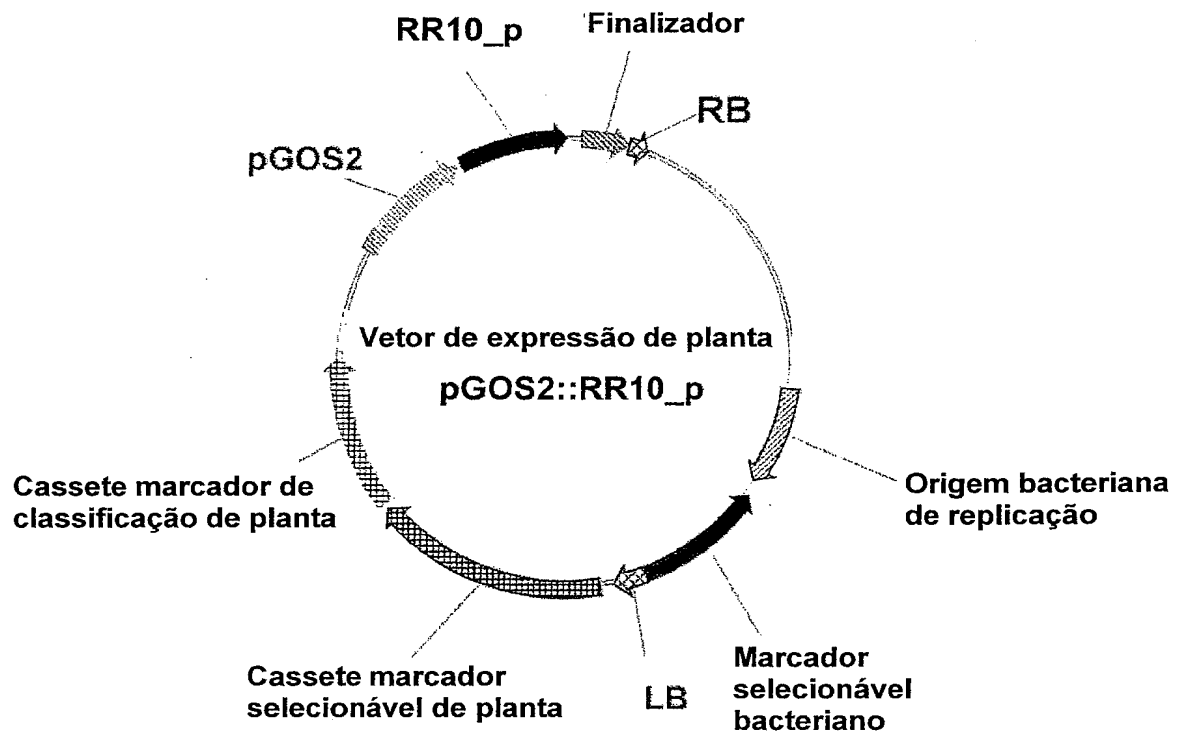
(i) a sequência de aminoácidos de SEQ ID NO: 76, SEQ ID NO: 78, SEQ ID NO: 80 ou SEQ ID NO: 196;

(ii) uma sequência de aminoácidos tendo pelo menos 50% de identidade de sequência com a sequência de aminoácidos de SEQ ID NO: 76, SEQ ID NO: 78 ou SEQ ID NO: 80;

(iii) uma sequência de aminoácidos tendo pelo menos 50% de identidade de sequência com a sequência de aminoácidos de SEQ ID NO: 196, ou com qualquer outra sequência de aminoácidos da Tabela A7; ou

(iv) derivados de qualquer uma das sequências de aminoácidos dadas em (i), (ii) ou (iii).

**Figura 1**

**Figura 2**

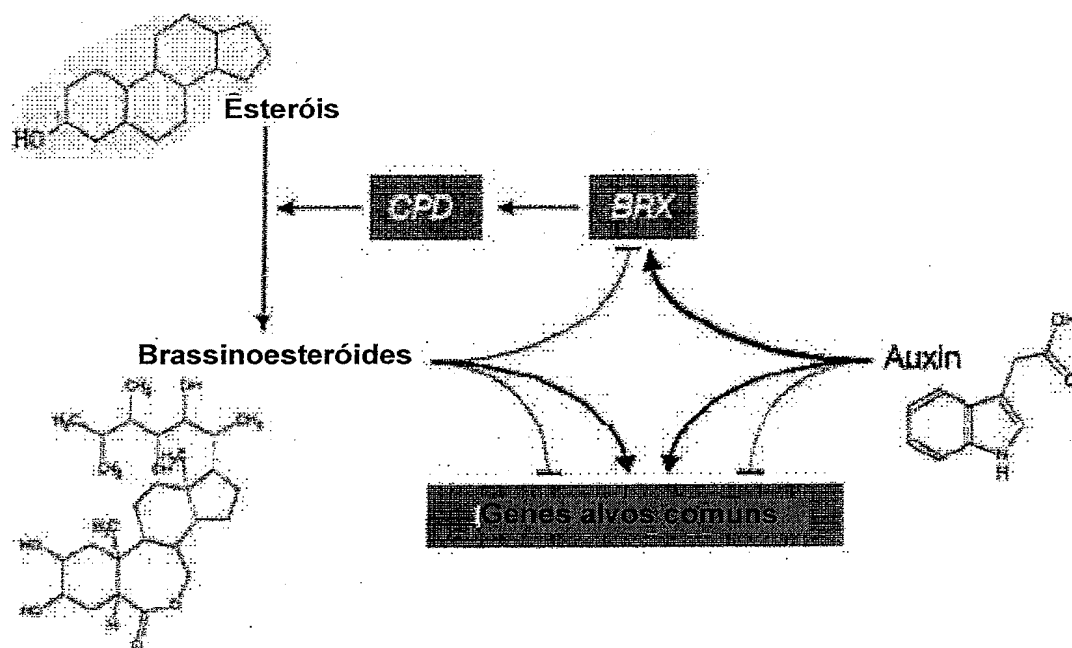


Figura 3

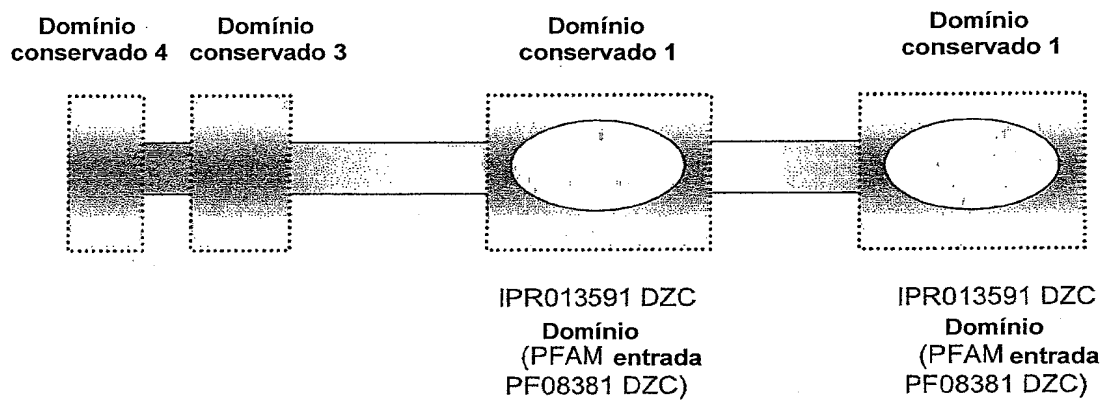


Figura 4

CLUSTAL W (1.81) Alinhamento múltiplo de sequência

Poptr_BRX	MFTCI	ACTKE	VAED	GRGEE	GGARGS	-----	GTPST	KEAVK	SLTSQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Glyma_BRXL2	MFTCI	ACTKI	DDKD	---EE	GGSR	ES-----	GTPST	KEAVK	SLTTQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Vitvi_BRXL1	MLTCI	ISCE	KQTE	EDGR	EEGSGGAR	-----	GTPST	KEAVK	SLTAQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Glyma_BRXL5	MFTCI	ACTKQ	AAEE	KEEEE	EGESGT	-----	TSTN	KEAVK	SLSAQL																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Arath_BRX	MFSCI	ACTKAD	--GGEE	VEHG	GARGG	-----	TPNT	KEAVK	SLTIQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Arath_BRXL1	MFTCI	INCTK	MADR	GEEDE	EDARGS	-----	TPNT	KEAVK	SLTTQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Poptr_BRXL3	MLTCI	ACSKRL	NNRCS	--PPR	DREE	-----DVDVAA	-----	FETL	RTHAMKALTAQM																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Vitvi_BRXL2	MLTCI	ACSKQ	VSGRS	----L	HEQE	-----EGE	GEA	-----	VATPSTKHAIKALTAQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Medtr_BRXL1	MLTCI	ACSKQ	LNNGS	-----L	HQ	-----QE	DEA	-----	VQTPSTKQAIKALTAQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Arath_BRXL2	MLTCI	ACTKQ	LNTN	NGGSK	QEE	-----EEE	EDR	-----	VIETPRSKQIKSLTSQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Arath_BRXL3	MLTCI	ACTKQ	LNTN	NGGST	REED	-----EEH	GV	-----	IGTPRIKQAIKSLTSQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Glyma_BRXL1	MLTCI	ARPKP	DES	DPDN	NATS	-----	AAK	SQAIK	SLTSQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Glyma_BRXL3	MLTCI	ARPKP	DES	DPNN	NATS	-----	AAK	SQAVK	SLTSQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Poptr_BRXL1	MLTCI	AREKQ	P	GDD	-----	-----	-----	SLT	IKSLTSQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Poptr_BRXL2	MLTCI	AREKQ	P	GDD	SLSQ	ADDDSA	-----ATTAN	HHPSA	AQKQKQKQKQAIKSLTSQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Arath_BRXL4	MLTCI	AREK	RAG	DESS	GQPD	-----	-----	-----	PDSKNAKSLTSQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Glyma_BRXL4	MLTCI	ARPKP	G	GDS	ASDDP	-----	SSR	SQGVK	SLTCQL																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Glyma_BRXL6	MLTCI	ARPKL	V	GDS	AASE	-----	DP	SR	GVKSLTGQL																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Nicbe_BRXL1	MLTCI	AREKQ	P	S	DDSR	DHQPEKFN	-----PNS	-----	APANKQAVKLTSLTSQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Sorbi_BRXL2	MLTCI	ACSKQ	QFA	AAGG	GPPLHEP	-----PEDD	DVVD	GGGA	IGGAGTPSTRHAIKALTAQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Zeama_BRXL1	MLTCI	ACSKQ	QF	S	GGGP	PLHEPP	-----EDD	DVVG	GGGAGTSAGAATPSTRHAIKALTAQI																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
Orysa_BRXL5	MLTCI	ACSKQ	P	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G

Domínio conservado 4

Domínio conservado 3

Figura 5

Poptr_BRX	KDMALKMSGAYKQCKPCTSPSS--YKKGQRPYPDFDAASEGVP-----
Glyma_BRXL2	KDMALKFSGAYKQCKPCTGSSS--YKKGHRPYPDFDTISEGVP-----
Vitvi_BRXL1	KDMALKFSGAYKQCKPCTGSSS--YKKGHRPYPDFDTISEGVP-----
Glyma_BRXL5	KDMALKFSGAYKQCKPCTGSSS--YKNGQRSYPDFDTISEGVP-----
Arath_BRX	KDMALKFSGAYKQCKPCTGSSSSPLKKGHRSPDYDNASEGVP-----
Arath_BRXL1	KDMASKFSGSHKQSKPTPGSSSSNLRK----FPDFDTASESVP-----
Poptr_BRXL3	KDMAVKASGAYRNCKPCSGSSSNNNNRNYAESDAASDSARFH-----
Vitvi_BRXL2	KDMALKASGAYRNCKPCSGSSSQNQDRNYADSESASDSARFH-----
Medtr_BRXL1	KDMAVKASGAYKNCKPCSGSSNGNKNKKYADSDMGSDSARFN-----
Arath_BRXL2	KDMAVKASGAYKSCKPCSGSSNQKNRSYADSDVASNSGRFR-----
Arath_BRXL3	KDMAVKASGAYKNCKPCSGTTNRNQNRNYADSDAASDSGRFH-----
Glyma_BRXL1	RDMALKASGAYKHCACTGPAT----QGRVRSNATELDADS----DRFR-----
Glyma_BRXL3	RDMALKASGAYKHCACTGPAT----QGRFRSNTTESDADS----DRFR-----
Poptr_BRXL1	KDMALKASGAYRHCNPCTVPTTTTTTQSRLRSNWTASDAES----ERFR-----
Poptr_BRXL2	KDMALKASGAYRHCNPCTAPN--TTTQSRLRSNSTESDAES----EKFR-----
Arath_BRXL4	KDMALKASGAYRHCTPCTAAQGGQGGQGPICKNNPSSSSSVKS----DFESDQRFKMLYGRS
Glyma_BRXL4	KEMALKASGAYKQCGPCA-TAPSRPSRSG-----
Glyma_BRXL6	KEMALKASGAYKQCGPCA-TAPSRPSRSG-----
Nicbe_BRXL1	KDMALKASGAYKQCSPT-TQPTTLRKNGSQNESSDSASLDKFRW-----
Sorbi_BRXL2	KDMALKASGAYRHCKPCAG----SSAAASRRHHPYHHRGGSGVF-----GGSDAGSA
Zeama_BRXL1	KDMALKASGAYRQCKPCAG----SSAAASRRHHPYHHRGGSAFG-----GSDADSGA
Orysa_BRXL5	KDIALKASGAYRHCKPCAGSS--SAAGASRRHHPYHHRGGGGGF-----GDPDAASG
Orysa_BRXL1	KDMALKASGAYRHCKPCAGSSSSSPAAGARRHHPYHAYADSG-----
Sorbi_BRXL1	KDMALKASGAYRHCKPCAGSS--SPAASRRQPPYYHG-----AYAESR
Zeama_BRXL2	KDMALKASGAYRHCKPCAGSS--SPAASRRQPPYYHG-----AYAESG
Picsi_BRXL1	KDMALKLSGAHRHCRPFVAVNLSREGQLQRCTASEVGSSENGTP-----
Orysa_BRXL2	KDMVLKFSGSGRHQYKSGGSPSLRTSRFHRSSRLAAYPGIIDE-----
Zeama_BRXL4	KDMVLKFSGSGRQYKAAASPSFRGNRFHRNSRLAAYTGVIDC-----
Orysa_BRXL4	KDMVMKLSGTSRHHGQQR-----RGGSPPPRGRTTTSVYR-----
Zeama_BRXL3	KDMVLKLSGTHHRQHGAQ-----HRRGGSPPPRGRATSLYR-----
	::: . * ** : :

Domínio conservado 3
(continua)

Figura 5 (cont.)

Poptr_BRX	-----YPYFGGGSSSSTPAWDFTTP-----	-----KHNRGTRADSREFT
Glyma_BRXL2	-----YPYIGGASSSSTPAWDFTT-----	-----SHYPGGRSDPRFAG
Vitvi_BRXL1	-----YPYLRPGSSSSTPAWDFTTS-----	-----SHNPGAGSDSRFTG
Glyma_BRXL5	-----YPYIGGASSTSTPAWDFTS-----	-----SNFPGGRSDQRFMG
Arath_BRX	-----YPFMGG-SAGSTPAWDFTNS-----	-----SHHPAGRLESKFTS
Arath_BRXL1	-----YPYPGG-STSSSTPAWDLPRS-----	-----SYHQSGRPDSRFTS
Poptr_BRXL3	-----CLYRRAGSSNSTPRKWGKES-----	-----EARLKGILSS-
Vitvi_BRXL2	-----CSYRRTGSSSSTPRLLGKEM-----	-----EARSKRLSS-
Medtr_BRXL1	-----WAYRRTGSASSTPRMWCKEM-----	-----EARLKGISS-
Arath_BRXL2	-----YAYKRAGSGSSTPKILGKEM-----	-----ESRLKGFLS-
Arath_BRXL3	-----YSYQRAGTATSTPKIWGNEM-----	-----ESRLKGISS-
Glyma_BRXL1	-----WSYRRTGSSSSTTTTRTWGKEM-----	-----EARLKGISS-
Glyma_BRXL3	-----WSYRRTGSSSSTTTTRTWGKEM-----	-----EARLKGISS-
Poptr_BRXL1	-----WPLQRTGSSSSTTPRTWGKEM-----	-----EARLKGISSS
Poptr_BRXL2	-----WSLRRTGSSSSTTPRTWGKEM-----	-----EARLKGISSS
Arath_BRXL4	NSSITATAA VAATQQQQPRVWGKEM-----	-----EARLKGISS-
Glyma_BRXL4	-----TESDSESSSSRRRWGKEL-----	-----EARLKGISS-
Glyma_BRXL6	-----DSESSSSSSRRRWGKEL-----	-----EARLKGISS-
Nicbe_BRXL1	-----SYKRTGSSSSTTAGRKEL-----	-----EARLKGISS-
Sorbi_BRXL2	SDRFHYAYRRAGSSAASTPRLRSGG--AALSS-----	-----GDATPSMSVRT--DF
Zeama_BRXL1	SDRFHYAYRRAGSSAASTPRLRSGG--AALSS-----	-----GDATPSISVRTGTFD
Orysa_BRXL5	SDRFHYAYRRATSSAASTPRFRGGGGGALSS-----	-----GDATPSMSARSDFPI
Orysa_BRXL1	SDRFHYAYRRAGS-----G-----	-----GDATPSVSARTDFLA
Sorbi_BRXL1	SDRFHYAYQCAGSSAASTPRLRTGG---AMSS-----	-----GDVTSPVSARTDFLA
Zeama_BRXL2	SDRFHCAYQRAGSSAASTPGLRTGG---AMSS-----	-----GDITPSVSARTDFLA
Picsi_BRXL1	-----RGSSSSTPAWSIVSS-----	-----SSKGHVLGDRLCAN
Orysa_BRXL2	-----SGFTSDGAGEAYTYMRTTTASAGARAAPSTWDLPPKVNHRSFQPRVIRSPSASGV	
Zeama_BRXL4	-----SGFTSDGATEGYGYMRTTTTHATGATAG-----	-----TKVGRGFPHQVRSASWI
Orysa_BRXL4	-----SGYYRPGMVQDDMAVPPATYLGGGGTS-----	-----MSSASST
Zeama_BRXL3	-----SGYYRPGVVQDDMAVPPATYLGGGGAG-----	-----ASSASST

Figura 5 (cont.)

Poptr_BRX	LYGGDRTPGGAES----	CDVVLEDEDEPKWMAQVEPGVHITFVSLPN-GGNDLKIRIFS
Glyma_BRXL2	AYGGDRTPGRDSSSV-	CDVVLEDEDEPKWMAQVEPGVHITFVSLPN-GGNDLKIRIFS
Vitvi_BRXL1	VLRGDQTPGGVSISAQSCDVVLEDEDEPKWMAQVEPGVHITFVSLPH-GGNDLKIRIFS	
Glyma_BRXL5	RFSGDRTPRGPQSAPASDVVVVEDEDEPKWMAQVEPGVHITFVSLPN-GGNDLKIRIFS	
Arath_BRX	IYGNDRE-----	SISAQSCDVVLEDDGPKWMAQVEPGVHITFASLPT-GGNDLKIRIFS
Arath_BRXL1	MYGGERE-----	SISAQSCDVVLEDDGPKWMAQVEPGVHITFVSLPS-GGNDLKIRIFS
Poptr_BRXL3	----	GEGTPASVSGRTESVVFMEH-DEPKWVAQVEPGVLITFVSLPD-GGNDLKIRIFS
Vitvi_BRXL2	----	GEGTPASVSGRAESVVFMEH-DEPKEWIAQVEPGVLITFVSLMPQ-GGNDLKIRIFS
Medtr_BRXL1	----	GEGTPSVSGRTESVVFMEHDEDEPKWVAQVEPGVLITFVSLPE-GGNDLKIRIFS
Arath_BRXL2	----	GEGTPESMSGRTESVVFMEHDEDELKEWVAQVEPGVLITFVSLPE-GGNDLKIRIFS
Arath_BRXL3	----	EGGTPSVSGRTESVVFMEH-DEVKEWVAQVEPGVLITFVSLPQ-GGNDLKIRIFS
Glyma_BRXL1	-GEGTPN--	SASGRRAEPVVLFEHENEPEKWEVAQVEPGVLITFVSLPR-GGNDLKIRIFS
Glyma_BRXL3	-GEGTPN--	SASGRRAEPVVLFEHENEPEKWEVAQVEPGVLITFVSLPR-GGNDLKIRIFS
Poptr_BRXL1	SGEGTPNSVNSSGRRVDPIAFVHEKEPEKWEVAQVEPGVLITLVS LPR-GGNDLKIRIFS	
Poptr_BRXL2	SGEGTPNSVNGSGRRVDPIVFEHKEPEKWEVAQVEPGVLITFVSLPR-GGNDLKIRIFS	
Arath_BRXL4	---	GEATPKSASGRNRVDPIVFEHKEPEKWEVAQVEPGVLITFVSLPG-GGNDLKIRIFS
Glyma_BRXL4	--	GEGTPSSS-----GRRVLMLEDEEPEKWEVAQVEPGVLISFVSLPR-GGNHLKIRIFS
Glyma_BRXL6	--	GEGTPSSS-----GRRVLMLEDEEPEKWEVAQVEPGVLITFVSLPR-GGNHLKIRIFS
Nicbe_BRXL1	--	GEVTPVSASASGRRVEPVVFEHSEPEKWEVAQVEPGVLITFVSLPR-GGNDLKIRIFS
Sorbi_BRXL2	PAGDDEEDDEMASEAAGCGGGGKDDDAREWVAQVEPGVLITFVSLAQ-GGNDLKIRIFS	
Zeama_BRXL1	PAGDDD--	DEMTPEATGGCGG--KDDAKEWVAQVEPGVLITFVSLAE-GGNDLKIRIFS
Orysa_BRXL5	GDEEDEEEDDDDEMSTGGGGGKEEDAKEWVAQVEPGVLITFVSLPQ-GGNDLKIRIFS	
Orysa_BRXL1	GDEEEEEEEEE--	EGTTADGSEDEAKEWVAQVEPGVLITFVSLPE-GGNDLKIRIFS
Sorbi_BRXL1	GDE-----	DEETAAGSSEDEAKEWVAQVEPGVLITFVSLPR-GGNGLKIRIFS
Zeama_BRXL2	DDEEGD-----	DEEGTATGSSEDEAKEWVAQVEPGVLITFVSLPR-GGNGLKIRIFS
Picsi_BRXL1	TSRIGTPMLHTSSGPVETTMEEVEEESKEWISQVEPGVLITLVS VKG-GGNELKIRIFS	
Orysa_BRXL2	PSIGEEDYDDDDDDDEETVLEEDRVPREWTAQVEPGVQITFVSLPQGGAGNDLKIRIFS	
Zeama_BRXL4	PSIGE-----	DDEEEDDEEVVVEEDRVPREWTAQVEPGVQITFVSLTAGGAGNDLKIRIFS
Orysa_BRXL4	P-----	AWDFARPAEG--EAREWVAQVEPGVQITFVSLAGGGGNDLKIRIFS
Zeama_BRXL3	P-----	AWDLPAARADGEACREWVAQVEPGVQITFVSLPQGGAGNDLKIRIFS

** :***** *: : * * ** :*****
 XXXXXXXXXXXXXXX

PF08381

Domínio conservado 1

Figura 5 (cont.)

Poptr_BRX	REMFNKGQAQRWWGENYDRITELYNVQRFNR----	QALHTPPRCEDEQ-RDSSYSRLESA
Glyma_BRXL2	REMFNKGQAQRWWGENYDRIMELYNVQRFNR----	QALNTPPRSEDEQ-RDSSYSRLTSA
Vitvi_BRXL1	REMFNKGQAQRWWGENYDRIMELYNVQRFNR----	QALHTPPRSEDEVNRDSSYSRMGSA
Glyma_BRXL5	REIFDKWQAQKWWGENYDRIMELYNVQRFNR----	QALNTPSRSEDER-RDSSYSRMTSG
Arath_BRX	REMFDKWQAQRWWGENYDKIVELYNVQRFNR----	QALQTPARSDQSQRDSTYSKMDSA
Arath_BRXL1	REVFDKWQAQRWWGENYDRIVELYNVQRFNR----	QALQTPGRSEDQSQRDSTYTRIDSA
Poptr_BRXL3	REMFNKGQAQSWWAENYDKVMELYNVQRFNR----	QSVPLPPPPRSEDESSKPES-AKDSPT
Vitvi_BRXL2	REIFNKGQAQRWWAENYDKVMELYNVQRFNR----	QAVPLPTPPRSEDESSRMES-IQNSPV
Medtr_BRXL1	REMFNKGQAQRWWAENYDKVMELYNVQRFNR----	EAGPLPTPPRSEDESSKIES-ARDSPV
Arath_BRXL2	REMFDKWQAQKWWAENFDKVMELYNVQRFNR----	QSVPLPTPPRSEDGSSRIQS-TKNGPA
Arath_BRXL3	REMFNKGQAQKWWVENFEKVMELYNVQRFNR----	QSVPLQTPPVSEDESSQIQS-VKDSPV
Glyma_BRXL1	REMFNKGQAQRWWAENYDKVMELYNVQRFNR----	QAFPLPTPLRSEDESSKLES-VEESPV
Glyma_BRXL3	REMFNKGQAQRWWAENYDKVMELYNVQRFNR----	QAFPLPTPPRSEDESSKLES-AEESPV
Poptr_BRXL1	RDMFNKGQAQRWWAENYDRIMELYNVQRFNR----	QAFPLPPPPRSEDESSKMES-AEDIPV
Poptr_BRXL2	RDMFNKGQAQRWWAENYDRIMELYNVQRFNR----	QAFPLPTPPRSEDESSKMES-AEDSPV
Arath_BRXL4	RDMFNKLQAQRWWADNYDKVMELYNVQKLSR----	QAFPLPTPPRSEDENAKVEYHPEDTPA
Glyma_BRXL4	REIFNKGQAQRWWAENYDKVMELYNVQRLDR----	QAFPLPTPPRSEDESSKRES-IEDFPV
Glyma_BRXL6	REIFNKGQAQRWWAENYDKVMELYNVQRLNR----	QAFPLPTPPRSEDESSKRES-IEDFPV
Nicbe_BRXL1	REIFNKYQAQRWWAENCKVMELYNVQRLNR----	QAFPLPTPPRSEDGSSKIES-IEDSPV
Sorbi_BRXL2	REMFNKGQAQRWWAENYDKVMELYNVQRFNR----	QTVPLPTTPKSEDESSKEDS-----PV
Zeama_BRXL1	REMFNKGQAQRWWAENYDKVMELYNVQKFN----	QTVPLPATPKSEDESSKEDS-----PV
Orysa_BRXL5	REMFNKGQAQRWWAENYDKVMELYNVQRFNR----	QAVPLPATPKSEDEVSKEDS-----PV
Orysa_BRXL1	REIFNKGQAQRWWAENYEKVMELYNVQRFNR----	QQTPLPTTPKSEDESLKEDI-----PA
Sorbi_BRXL1	REMFNKGQAQRWWTENYEKVMELYNVQKFDN----	QAASLPSIPRSENESSKDDN-----SA
Zeama_BRXL2	REMFNKGQAQRWWTENYEKVMELYNVQKSDN----	QVDPLPSIPRSDSEISKDDT-----PA
Picsi_BRXL1	RELFNKGQAQRWWAENYDKVMELYNVHAHAKDDSAVAVPTPPRSEDERDSKMQ-----	
Orysa_BRXL2	REMFNKGQAQRWWGENYDRVVELYNVQTFNR----	QQGFSTPTSSVDEAMQORDS-FYSRV
Zeama_BRXL4	RDMFNKGQAQRWWGENYDRVVELYNVQTFNR----	QQGVSTPTSSIDDATQDSSFYSSRA
Orysa_BRXL4	REMYDKWQAQKWWGENNERIMELYNVRRFSR----	QVLPTPPRSDDGGERESFYSQVGSTR
Zeama_BRXL3	REMYDKWQAQKWWGDNNERIMELYNVRRFSR----	QVLPTPPRSDDAERESFYSQS--QV
	::: ** * :* ::* *****:	
PF08381	XXXXXXXXXXXXXXXXXXXXXXXXXXXX	

Domínio conservado 1 (continua)

Figura 5 (cont.)

Poptr_BRX	RESPMAPS----	FTPRNY-----	-----YKPAGSKGYFPSDTMDQG
Glyma_BRXL2	RESPMASNK--	DWTPRSH-----	-----YKPSGSRGYYPSEPLDHG
Vitvi_BRXL1	RESPMTPSLNKEWI	PRNY-----	-----YKPSGSK-----
Glyma_BRXL5	QDSPMHMSLSLKGW	TPRNNH-----	-----YKPSGNNPSEAMEQSGG
Arath_BRX	RESKDWT	TPRNNH--	FRPPG-----SVP-----
Arath_BRXL1	RESRDWTQ	RDNNFRPPG-----	-----SVP-----
Poptr_BRXL3	TPPLGKERPSN--	FHHPTG-----	-----MGYSSSDSLDHHPMQSHQYYESAGL
Vitvi_BRXL2	TPPLSKERLPRNFH	HRPTG-----	-----MGYSSSDSLDHHPLQSRHYIDSAGL
Medtr_BRXL1	TPPLTKERLPR--	HLHHPMT-----	-----MGYSSSDSLDHHMQPQPCYETSGL
Arath_BRXL2	TPPLNKE-----	CSRGKG-----	-----YASSGSLAHQPTTQTQSRHHDSSGL
Arath_BRXL3	TPPLERE-----	RPHRN-----	-----IPGSSGF
Glyma_BRXL1	TPPLNSERLPRNMY	RPTG-----	-----MGMGYSSSDSFDHQSMQS--RHFYDSNGM
Glyma_BRXL3	TPPLNRERLPRNMY	RPTG-----	-----MGMGYSSSDSFDHQSMQS--RHFYDSNGM
Poptr_BRXL1	TPPLNRERLPRNLY	RPTG-----	-----TGMGYSSSDSLDHHPIQA--RHYCDSTGL
Poptr_BRXL2	TPPLNRERLPRNLY	RPTG-----	-----MGMGYSSSDSLDHHPLQA--RHYCDSIGF
Arath_BRXL4	TPPLNKERLPRTI	HRPPG-----	-----LAA--YSSSDSLDHNMQS--QQFYDSGLL
Glyma_BRXL4	TPPLSRERPPCNLY	RAGGRGGG-----	-----MGMGYSSSDSFDHTSMQSSRHYIDPNGV
Glyma_BRXL6	TPPLSRERPPCNLF	RAGG--	GG-----MGMGYSSSDSFDHHSMQSSRHYIDPNDV
Nicbe_BRXL1	TPPLTKERLPHTLY	RPMYS-----	-----SSDSLEQHSSTHFRYNYDSCGV
Sorbi_BRXL2	TPPLDKERLPRTF	HR--	QGGGAMGYSSSDSLEHHSNHYCTG---HHHHHG--HQCCDSMGL
Zeama_BRXL1	TPPLDKERLPRTF	HR--	QGGGAMGSSSDSLEHHSNHYCTGRHHHHHHHG--HQCCDSMGL
Orysa_BRXL5	TPPLGKERLPRS	FHRPLSGGAVGSSSDSLEHHSNHYCNG---GHHHHG--HQCYDSVGL	
Orysa_BRXL1	TPPLNSERLPHTL	HRSLTGGRTTGYGQPD	SLGHQHNLGNHGRHQHHHCYTGHQCYGSVGL
Sorbi_BRXL1	IAPLNKGQLD	TLHRPLKVSGAIGYSSSDCLQHQP	NHLGNIYR---QDRYLGHQCCDSVGL
Zeama_BRXL2	TAPLNKGQLL	RTSPRPLKGSEAI	GYSSSDCPQHQS
Picsi_BRXL1	-----	-----	-----ESG
Orysa_BRXL2	GSTRESPAMMMPP	PPPLPSSGAGREHPISRTASSKAQLSSSSSVAAARPPFP	PSTAVPDP
Zeama_BRXL4	GSTRDSPVILPP	-----	TAVGREQPIARATSCRATAAAASTARAACNPSS--SAVPDP
Orysa_BRXL4	GSPAATPSPAPLT	PDVTS-----	WSAFVRPPSASRQQQHSFRPLSPPPP---SSSNP
Zeama_BRXL3	GSPSATPSPAPLT	PERISWGAFARQVAAPPASSGATGAARQHSFRPYSPPPP---SSSNP	

Figura 5 (cont.)

```

Poptr_BRX      GSHHYHAGSSSYGMGGPRFEA---SSLEASRTTTSSRDEPSI-----SVSNASDLET-
Glyma_BRXL2    GSGSQYHAG-----PSMEPARDTTASRDEP-----SISNASEMET-
Vitvi_BRXL1    -GHQYNAGSSAYGTGGPRGET---SSMDASRTTTSSRDEASV-----SISNASDMET-
Glyma_BRXL5    GQDFHAAAS-----SVEASRTTTSSRDER-----SMSNASDLET-
Arath_BRX      ---HHFYGGSSNYGPGSYHGG---PPMDAARTTTSSRDDPP-----SMSNASEMQA-
Arath_BRXL1    ---HHFYG-----PPMDAARITTTSSRDEPP-----SMSNASEMQG-
Poptr_BRXL3    ASTPKLSSIAGAKTETSS-----IDGSVRTSMSRESDRS---EELSISNASDMET-
Vitvi_BRXL2    ASTPKLSSISGAKTETSS-----IDASVRTSSSREADRS---GELSISNASDMET-
Medtr_BRXL1    PSKSNLSNIGVPKTERSS-----IDASVRTSSSEEDHS---GELSISNASDMET-
Arath_BRXL2    ATTPKLSSISGKTETSS-----VDESARSSFSREEEEADHSGEELSVSNASDIET-
Arath_BRXL3    ASTPKLSSISGKTETSS-----IDGSARSSSVDRSE-----EVSVSNASDMES-
Glyma_BRXL1    NSTPKVSTISAAKTEISS-----MEASIRSSSSSREADRS---GDFSISNASELET-
Glyma_BRXL3    NSTPKVSTISAAKTEISS-----MDASIRSSSSSREADRS---GDFSISNASDLET-
Poptr_BRXL1    TSTPKLSSISGAKTETSS-----MDASIRSSSSSREADCS---GELSISNASDMET-
Poptr_BRXL2    TSTPKLSSISGAKTETSS-----MDASIRSSSSSREADRS---GELSISNASDMET-
Arath_BRXL4    NSTPKVSSISVAKTETSS-----IDASIRSSSSSRDADRS---EEMSVSNASDVDN-
Glyma_BRXL4    NSTPKAS-TISAAAKTDISSI---MDADASIRSSSSSREADRS---GDLSISNASDFDN-
Glyma_BRXL6    NSTPKASSTISAAAKTDISS---MDVDASIRSSSSSREADRS---GDLSISNASDLDT-
Niche_BRXL1    ASTPKLSSISGAKTEISS-----MDASIRTSTSRDADRS---GELSISNASDLET-
Sorbi_BRXL2    ASTPKLSSISGAKTETSS-----MDASMRTSSSPPEEVDRSGE-LSVSVSNASDQER-
Zeama_BRXL1    ASTPKLSSISGAKTETSS-----MDASMRTSSSPPEIDRSGE-LSVSVSNASDQER-
Orysa_BRXL5    VSTPKLSSISGAKTETSS-----MDASMRTSSSPPEEVDRSGE-LSVSVSNASDQER-
Orysa_BRXL1    ASTPKLSSISGAKTETSS-----MDASMRSSSSPPEEVDRSRE-LSVSVSNASDQER-
Sorbi_BRXL1    ASTPKLSSISGAKTETS-----IDASVRTSSSPPEEVDRSGE-LSASVSNASDQER-
Zeama_BRXL2    ASTPELSSISGAKTETS-----IGASVRTSSSPPEEVDGSGE-LSASVSNASDEER-
Picsi_BRXL1    ADSPVTPPIQNMSLLPRG-----LYASRTSSSHYADRSDDFQSSGNSSEQEQQEQEQ-
Orysa_BRXL2    SDHVWAHHFNLLNSAAAG-----PAAP---YDPSRGTTSSRDEASVSI SNASDLEAT
Zeama_BRXL4    SDHVWAHHFNLLNSAPAP-----APAAPHLDPSTRATTSSLDEASVSVSNASDLEAT
Orysa_BRXL4    SERAWQQQQQPQRAG-----KSPAAASDAMDARTTSCSSRDEVSISNASELEV
Zeama_BRXL3    SERAWQQQQRHIGGGGGGAAGGKSPAASEAAAAATEARTTTSSRDDVSVSNASELEVS

```

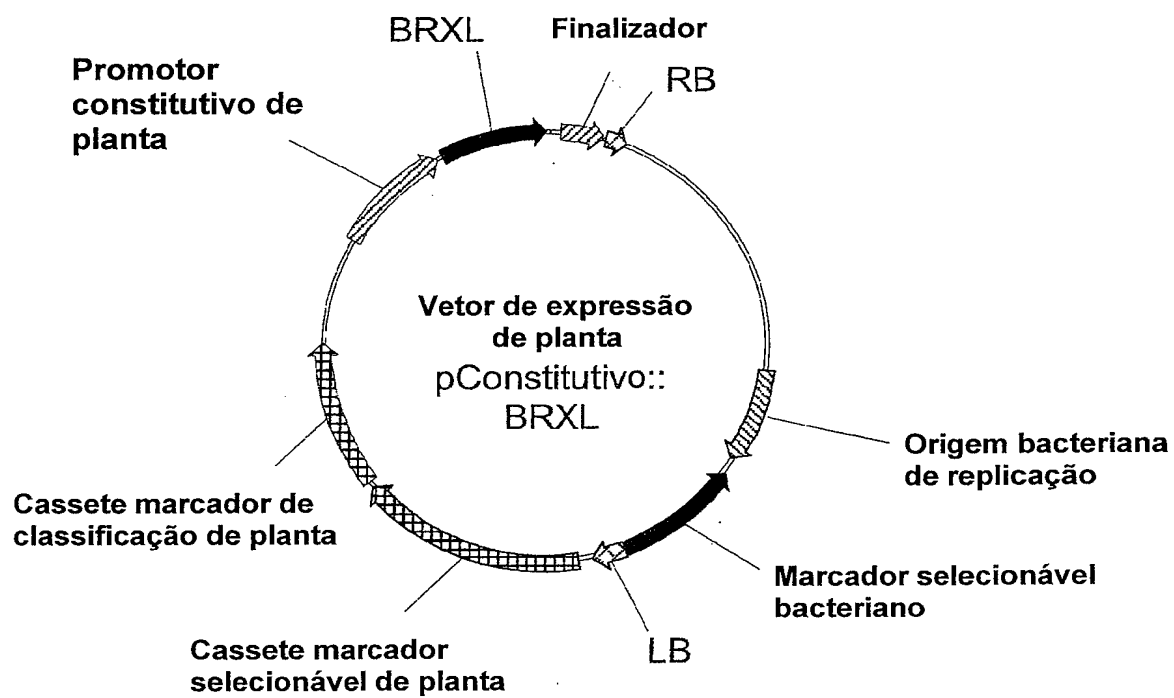
* . : . : .

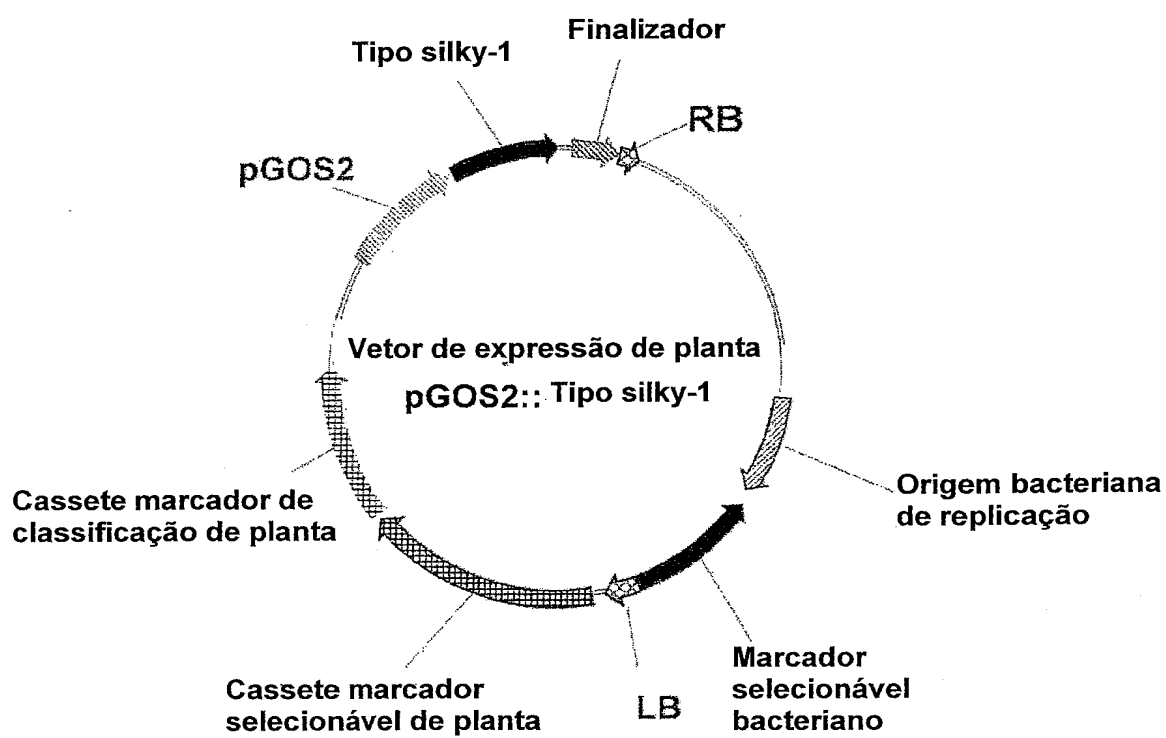
Figura 5 (cont.)

Poptr_BRXL	EWVEQDEPGVYITIRQLADGTRELRRVRFSSREQFGEVHAKTWEEQNRRERIQAOYL
Glyma_BRXL2	EWVEQDEPGVYITIRQLADGTRELRRVRFSSRERFGEVNAKTWEEENRRIQAQYI-----
Vitvi_BRXL1	EWVEQDEPGVYITIRQLADGTRELRRVRFSSRERFGEVHAKTWEEENRRIQAQYI-----
Glyma_BRXL5	EWIEQDEPGVYITIRQLADGTEKELRRVRFSSRERFGEGHAKKWEDNRERIQAQYI-----
Arath_BRX	EWIEEDEPGVYITIRQLSDGTRELRRVRFSSRERFGEVHAKTWEEQNRRERIQTQYI-----
Arath_BRXL1	EWVEEDEPGVYITIRQLPDGTRELRRVRFSSRERFGEVHAKTWEEQNDRDIQTQYI-----
Poptr_BRXL3	EWVEQDEPGVYITIRALPGGTRELRRVRFSSRETFTGETRARLWEEENRDVVHEQYLYR--
Vitvi_BRXL2	EWVEQDEPGVYITIRALPDGTRELRRVRFSSRERFGEMHARLWEEENRARIQEQQYI-----
Medtr_BRXL1	EWVEQDEPGVYITIRALPGGTRELRRVRFSSREKFGEHARLWEEENRARIQEQQYI-----
Arath_BRXL2	EWVEQDEAGVYITIRALPDGTRELRRVRFSSREKFGETNARLWEEQNRRARIQQQYI-----
Arath_BRXL3	EWVEQDEPGIYITIRALPDGNRELRRVRFSSDKDFGETHARLWEEQNRRARIQQQYI-----
Glyma_BRXL1	EWVEQDEPGVYITIRALPGGARELKRVRFSSREKFGEHARLWEEENRARIHEQYI-----
Glyma_BRXL3	EWVEQDEPGVYITIRALPGGARELKRVRFSSREKFGEHARLWEEENRARIHEQYI-----
Poptr_BRXL1	EWVEQDEQGVIYITIRALPGGKREIRRVRFSRERFGETHAKVWEEENRARIHQQYM-----
Poptr_BRXL2	EWVEQDEPGVYITIRALPGGKRELRRVRFSSRERFGEMHARVWEEENRARIHEQYI-----
Arath_BRXL4	EWVEQDEPGVYITIKVLPGGKRELRRVRFSSRERFGEMHARLWEEENRARIHEQYI-----
Glyma_BRXL4	EWVEQDEPGVYITIRALLGGKKELRRVRFSSREKFGEHARLWEEENRARIHEQYI-----
Glyma_BRXL6	EWVEQDEPGVYITIRALPGGKKELRRVRFSSREKFGEHARLWEEENRARIHEQYI-----
Nicbe_BRXL1	EWVEQDEPGVYITIQUALPNGRRELKRVRFSSREKFGEVHARLWEEENRARIHKQYIG----
Sorbi_BRXL2	EWVEEDEPGVYITIRALPGGIRELRRVRFSSREKFSEM HARLWEEENRARIHEQYI-----
Zeama_BRXL1	EWVEEDEPGVYITIRALPGGIRELRRVRFSSREKFSEM HARLWEEENRARIHEQYI-----
Orysa_BRXL5	EWVEEDEPGVYITIRALPGGIRELRRVRFSSRERFSEM HARLWEEENRARIHEQYI-----
Orysa_BRXL1	EWVEEDEPGVYITIRALPGGIRELRRVRFSSREKFSEM HARLWEEENRARIHDQYI-----
Sorbi_BRXL1	EWVEEDEPGVYITIRALPGGIRELRRVRFSSRERFNEM HARLWEEENRARIHDQYI-----
Zeama_BRXL2	EWVEEDEPGVYITIRALPGSIRELRRVRFSSRERFSEM HARLWEEENQARIYEQYI-----
Pisci_BRXL1	EWVEEDEPGVVTIRCSPAGSREIRRVRFSREKFSQM ARLWEEENRLARIHEQYISGRSL
Orysa_BRXL2	EWVEQDEPGVSITIREFGDGTRELRRVRFSSRERFGEERAKVWVEQNDRRIHAQYI-----
Zeama_BRXL4	EWVEQDEPGVSITIREFGDGTRELRRVRFSSRERFGEERAKVWWDQNNRRIHAQYI-----
Orysa_BRXL4	EWVIQDEPGVYITVRELADGTRELRRVRFSSRERFAELNAKLWEEENKERIQAQYI-----
Zeama_BRXL3	EWIIQDPQGVYITVRELADGSRELRRVRFSSRERFAELNAKLWEEENKERIQAQYI-----
	** : * : * : * : * : * : * : * : * : * : * : * : *

Domínio conservado 2

Figura 5 (cont.)

**Figura 6**

**Figura 7**

A.thaliana_AT3G33520.1 P.trichocarpa_scaff_XVIII.1149 Gm0057x00075 Gm0195x00030 P.patens_126969 O.sativa_LOC_Os01g16414.1 Consenso	(1)	1	-----MSNIWVLDNGGGLIKAGQGGERDPTTVPNClyKPLSSKKFI	50
			-----MSSVVLDNGGGLIKAGYGGERDPSTIIPNClyRPLSSKKFL	
			-----MSTSTNLVLDNGGGLIKAGIDGEHDP-----FA	
			-----MSTSTNVVLDNGGGLIKAGIGGERDPSAIVPNClyRPPS---SK	
			MKRTGGVTSSVVLDNGAGYCKAGMAGQSEPTAVVPCMARPRSACKWL	
A.thaliana_AT3G33520.1 P.trichocarpa_scaff_XVIII.1149 Gm0057x00075 Gm0195x00030 P.patens_126969 O.sativa_LOC_Os01g16414.1 Consenso	(1)	51	--MTGG---SGVVLDNGGGLIKAGFGGDMNPTAVVPCMAKPPGSKKWL	100
			SNVVLDNGGGLIKAG GGERDPTAVVPClyRP SSKKFL	
			-----LTSAAVRRRPIDRGYLINSDLQREIWSHLFTSLHIAF	
			-----DLTSTAMRYPVDRGYLINPDQLQHKIWSHLFSSVLHTNP	
			-----DLTSAAVRRPMDRGYLINPDQLQREIWSHLFSSVLHINP	
A.thaliana_AT3G33520.1 P.trichocarpa_scaff_XVIII.1149 Gm0057x00075 Gm0195x00030 P.patens_126969 O.sativa_LOC_Os01g16414.1 Consenso	(1)	101	-----DIQSIARKRPIDRGYMI SPEIEREIWDRVFKVHLKVNP	150
			VADQLQAQDV--DVTGMTLRRPIDRGYLINQEVQREVWERVIRNLLQVDP	
			L D DLTSAAVRRRPIDRGYLINPDQLQREIWSHLFSSVLHINP	
			-----LTSAAVRRRPIDRGYLINSDLQREIWSHLFTSLHIAF	
			-----DLTSTAMRYPVDRGYLINPDQLQHKIWSHLFSSVLHTNP	
A.thaliana_AT3G33520.1 P.trichocarpa_scaff_XVIII.1149 Gm0057x00075 Gm0195x00030 P.patens_126969 O.sativa_LOC_Os01g16414.1 Consenso	(1)	150	-----DLTSAAVRRPMDRGYLINPDQLQREIWSHLFSSVLHINP	200
			-----DLTSAAVRRRPIDRGYLINSDLQREIWSHLFTSLHIAF	
			-----DLTSTAMRYPVDRGYLINPDQLQHKIWSHLFSSVLHTNP	
			-----DLTSAAVRRPMDRGYLINPDQLQREIWSHLFSSVLHINP	
			-----DLTSAAVRRRPIDRGYLINSDLQREIWSHLFTSLHIAF	

Figura 8

A.thaliana_AT3G33520.1	(143)	Q PDS-ILSKTQCSLVVDCGFSFTHAVPVLHNFNLHAIKRIDLGGKAFN	200
P.trichocarpa_scaff_XVIII.1149	(139)	RPYG-LVSKAQCSLVVDCGFSFTHAAPVFQNFNTLNYGVKRIDLGGKALT	
Gm0057x00075	(128)	NNANGILSKAQCSLVMDLGFSTHTSPVFHNFALNYAIRRIDLSGKALT	
Gm0195x00030	(141)	NNANGLLSKAQCSLVLDAGFSFTHASPVFHNFALNYAVKRIDLGGKALT	
P.patens_126969	(148)	APTS-ILARGRTSLVVDGFSFTHHTVPVFQNKAVTTAAKRINLGKALT	
O.sativa_LOC_Os01g16414.1	(144)	QPS---LFRAQCSLVVDCGFSFTHASPVLQNFNTLNYAVRRMDLGKALT	
Consenso	(151)	P ILSKAQCSLVVDCGFSFTHASPVF NF LNYAVKRIDLGGKALT	
A.thaliana_AT3G33520.1	201	Y LKELVSYRSINVMDETFVLVDDAKEKLCFVSLDLLRLDLARN--GNTLI	250
P.trichocarpa_scaff_XVIII.1149	(188)	FLKELVSYRSVSVMDSEFTMDDVKEKLCFVSLDVARDLKIARRRGNDNFF	
Gm0057x00075	(178)	Y LKDLVSFRSVNIMEETFIIN-----	
Gm0195x00030	(191)	Y LKELVSFRSVNVMEETFIIDDVKEKLCFVSLDVNRDLTIARKSGKENLF	
P.patens_126969	(197)	Y LKELVSYRAWNVMDETYTMEDVKEKLCFVSLDIERDLNIARVKGKSNAL	
O.sativa_LOC_Os01g16414.1	(191)	Y LKELISYRSLNVMDETLLIDDVKEKLCFVSLDVPGDLRLARLSSNDNPF	
Consenso	(201)	Y LKELVSYRSVNVMDETFIIDDVKEKLCFVSLDV RDL IAR G N F	
A.thaliana_AT3G33520.1	251	K STYVLPDGVTHTKGYVKDPQAARFSLSEKES-----VVVMDKVG E	300
P.trichocarpa_scaff_XVIII.1149	(238)	R CTYVLPDGVTHTKGFVKDPDEAKKYLTVGDGAYLETR--KMDRTEIMD	
Gm0057x00075	(199)	-----DEFVKYPDQAQHYALRECGLPSSPSVDAPGDVKCLE	
Gm0195x00030	(241)	R CTYVLPDGVTYTKGFVKYPDQAQRYALREGGLHSSSPVQAQEDVNFE	
P.patens_126969	(247)	R REYVLPDGVKHKRGFVKLEPLVNKHLKSSK-----IRYLLPPCAE	
O.sativa_LOC_Os01g16414.1	(241)	R CSYILPDGITYKKGFVKDLDEACRYSSLPANGESVRKDSSDSDRSKFED	
Consenso	(251)	R CTYVLPDGVTHTKGFVKDPD A RYL L E G K E	

Figura 8 (cont.)

A.thaliana_AT3G33520.1 P.trichocarpa_scaff_XVIII.1149 Gm0057x00075 Gm0195x00030 P.patens_126969 O.sativa_LOC_Os01g16414.1	(283) (286) (236) (291) (291) (291) (301)	301	RKKADMNKNEIDLTNER-----FLVPETLFQPADLGMNQAGLAECIVR	350
			RKKVDLTKNFEFDTNER-----FLVPFMI FHPADLGMNQAGLAECIVR	
			IAKQPEDRKIVDLTKN-----FLVPKMI FHPADLGLSIQYICIFLN-	
			IAEHPENRKRVDLTKNFEFDTNERFLVPFMI FHPADLGMNQAGLAECIVR	
			KLREFFVNQEI TLTNER-----FLVPFMI FHPADLGMNEAGLAECIVR	
Consenso	(301)	301	KKKPELSQNEFVLTNER-----FLVPFMI FHPIDLGMNQAGLAECIVR	350
			KK NE DLTNER FLVPEM FHPADLGMNQAGLAECIVR	
A.thaliana_AT3G33520.1 P.trichocarpa_scaff_XVIII.1149 Gm0057x00075 Gm0195x00030 P.patens_126969 O.sativa_LOC_Os01g16414.1	(326) (329) (277) (341) (334) (334) (351)	351	AINSCHYLQPVLYQSIILTGSTLFPQLKERLEGEELRPLVPDHFVVKIT	400
			AVNSCHPLLHPLLYQSIILTGSTLFPFAERLEMELELRPLVPDDYQVKIT	
			-----LKHSFLNVVVCVTLFPQFVERLEKELELRPLVPNDYRVKIA	
			AVNACHPLRPVLYESIILTGSTLFPQFAERLEKELELRPLVPDDYRVKIT	
			AINACEPELHALLYQSVLLTGNTLLPGFKNRLEHELELRPLVTDDFNINIE	
Consenso	(351)	351	AIQACHPHLQPVLFERIILTGSTLFPRTERLEKELELRPLVPDDYQVKII	400
			AINACHP L PVLYQSIILTGSTLFPQF ERLEKELELRPLVPDDY VKIT	
A.thaliana_AT3G33520.1 P.trichocarpa_scaff_XVIII.1149 Gm0057x00075 Gm0195x00030 P.patens_126969 O.sativa_LOC_Os01g16414.1	(376) (379) (315) (391) (384) (384) (401)	401	TQEDPILGVWRGGSLIASSPDFESMCVTKAEYEEELGSARCRRRFFH----	450
			TQEDPILGVWRGGSLIASSPDFEAMCVTKAEYEEELGSARCRRRFFH----	
			TQEDPLLGVWRGGSLIASSPDFEAMCVTKSEYEEELGSARCRKRFFH----	
			TQEDPILGVWRGGSLIASSPDFEAMCVTKSEYEEELGSARCRKRFFH----	
			IVDCPILAAWKASLMAARPEFHTWAVTKAEYEEEDGTLRCRQRFVSADWS	
Consenso	(401)	401	AQEDPILGAWRGGSLLAHRPDFESMCITKSEYEEEMGSMRCRRRRFFH----	450
			TQEDPILGVWRGGSLIASSPDFEAMCVTK EYEEELGSARCRRRRRFFH	
A.thaliana_AT3G33520.1 P.trichocarpa_scaff_XVIII.1149 Gm0057x00075 Gm0195x00030 P.patens_126969 O.sativa_LOC_Os01g16414.1	(422) (425) (361) (437) (434) (430) (451)	451	-----	

			LFKQSP	
Consenso	(451)	451	-----	

Figura 8 (cont.)

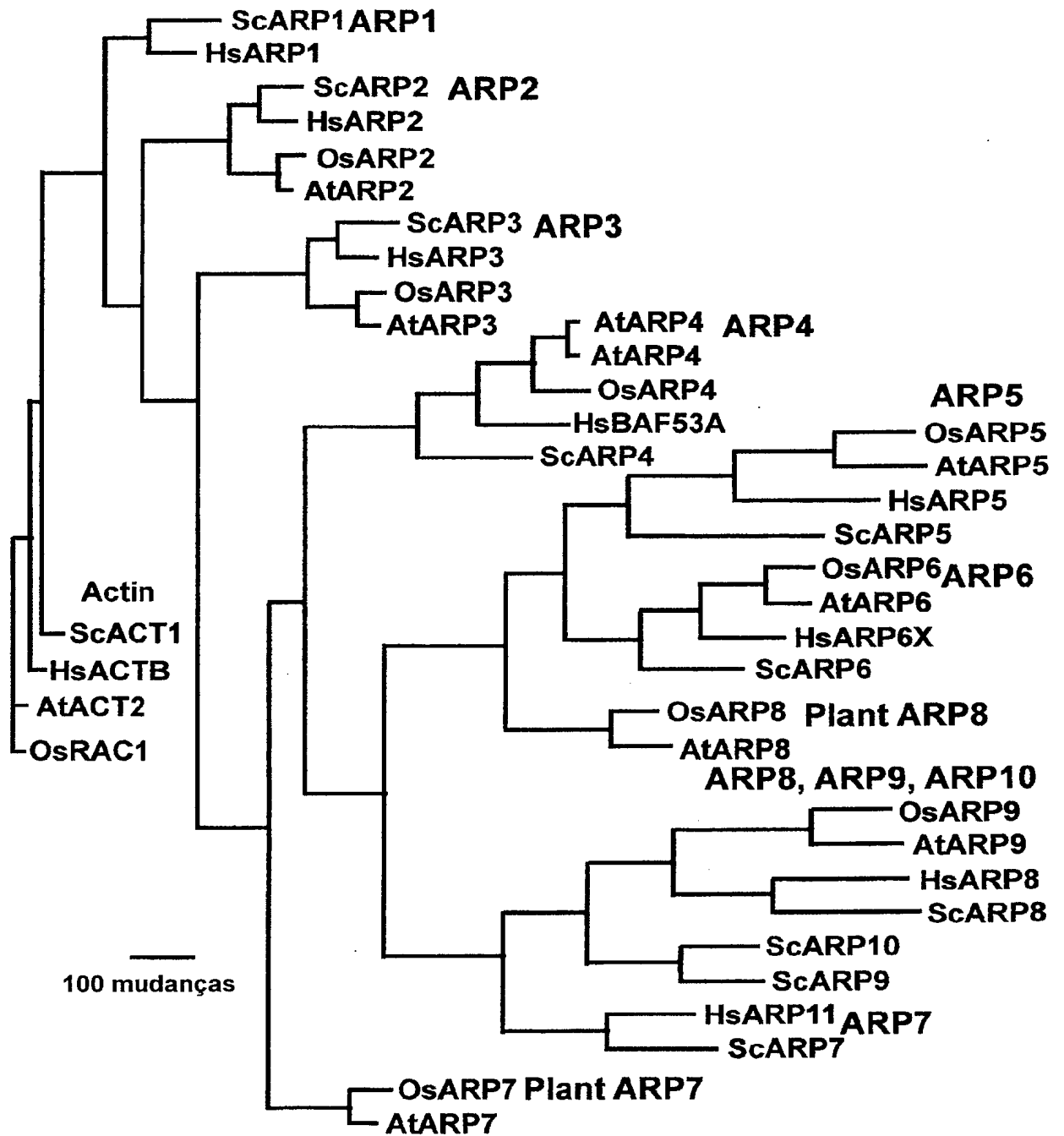
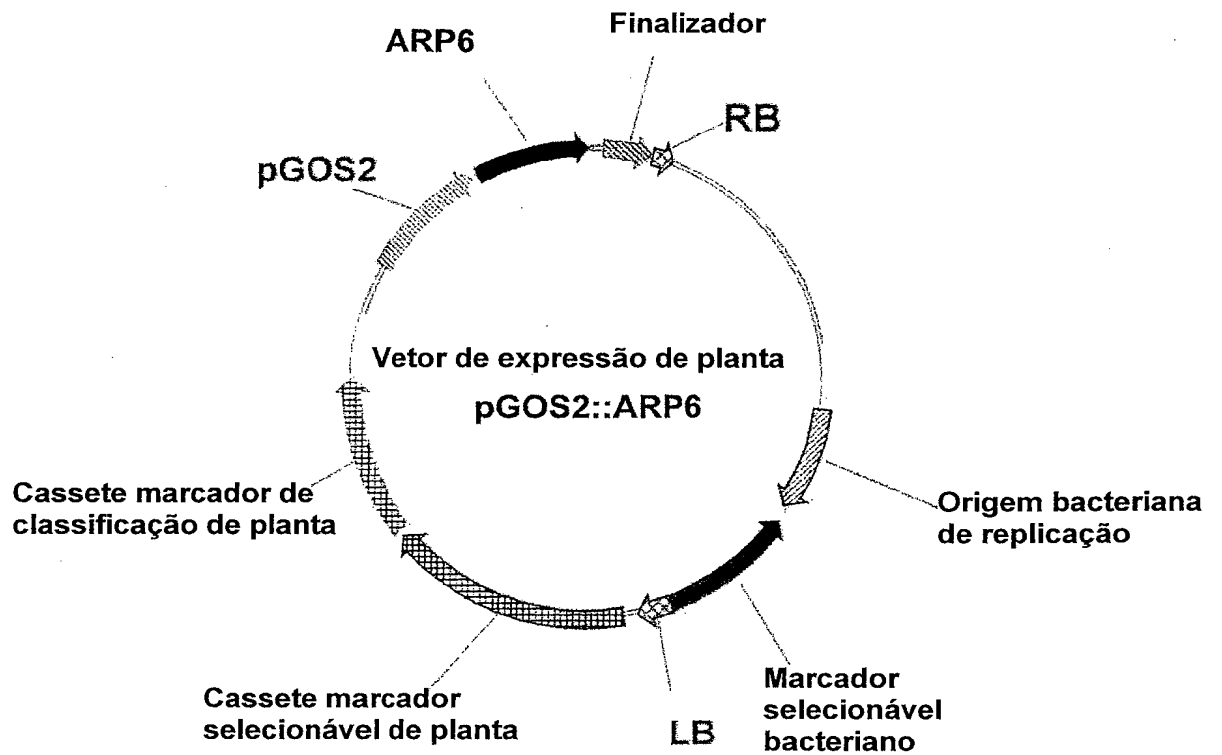


Figura 9

**Figura 10**

MADKDVPFVGVEDIVQTPLPGYVAPTAVSFSPDDSLITYLESPEK
 NLKRRVYAFDVNKGESNLVFSPPDGGVDESNI SPEEKLRRERLR
 ERGLGVTRYEWVKTNSKMRFIVVPLPAGVYMKDLSSSPNPELIV
PSSPTSPIIDPRLSPNGLFLAYVRESELHVLNLLKNQTQQLTSG

PF00930 - DPPIV_N

ANGSTLSHGLAEYIAQEEMDRRNGYWWSLDSKFIAYTEVDSSQI
 PLFRIMHQGKR SVGSEAQEDHAYPFAGALNSTLR LGIVSSAGGG
 KTTWMNLVCGGRGNTED EYLGRVNWLPGNVLIVQVLNRSQSKLK
 IISFDINTGQGNVLLTEESDTWVTLHDCFTPLEKVPSSRGSGGF
 IWASERTGFRHLYLYESDGTCLGAITSGEWMVEQIAGVNEPMSL
 VYFTGTLDGPLETNLYCAKLEAGNTSRCQPMRLTHGKGKHIVVL
 DHQMKNFVDIHDSVDSPPRVSLCSLSDGTVLKIPIYEQTSP IQIL
 KSLKLEPPEFVQIQANDGKTTLYGAVYKPDSSSKFGPPPYKTMIN
 VYGGPSVQLVYDSWINTVDMRTQYLR SRGILVWKLDNRGTARRG
 LKFESWMKHNCGYVDAEDQVTGAKW LIEQGLAKPDHIGVYGWSY

PF00326 - Peptidase_S9

GGYLSATLLTRYPEIFNCAVSGAPVTSWDGYDSFYTEKYMGLPT
 EEERYLKSSVMHHVGNLTDKQKLMLVHGMIDENVHFRHTARLVN
 ALVEAGKRYELLIFPDERHMPRRKKDRIYMEQRIWEFIEKNL

Figura 11

CLUSTAL 2.0.8 Alinhamento múltiplo de sequência

```

P.patens_104540#1 -----MPTTD--SQSLGDEYTLFPVEDIVQNP 25
P.patens_105959#1 -----MPTTD--SQALGDEYTLFPVEDIVQYP 25
P.trichocarpa_scaff_XII.203#1 MQSVDENESQNKKLRLRLSLNNDMELTDNTIPQNVEDSILFPPIEEIVQSP 50
P.trichocarpa_scaff_XV.146#1 MQSVDENESQNKKLRLRLSLNNDMELTDNTIPQNVEDSILFPPIEEIVQSP 50
M.truncatula_AC137828.18.5#1 -----MRGIRSVKNHFLSKCLFVTFDFNDVQNLDDGILFPVEEIAQYP 43
A.thaliana_AT5G24260.1#1 -----MADKDVFFGVEDIVCTP 17
O.sativa_Os02g0290600#1 -----
gi|145358557|ref|NP_198470.3| -----MALLLTSLNHLVSFSLTRLPPSS 24
gi|116789192|gb|ABK25153.1| -----MASFSFS---CAY 10
gi|42566792|ref|NP_193193.2| -----MDSSGTDSAKELHVLDPTEEEYA 25

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828.18.5#1
A.thaliana_AT5G24260.1#1
C.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

LPGYVAPGSVQFSPDDKLITYLYSPDSTLSRKIYAFDVVAKQQRLLVNPP 75
LPGYVAPASVQFSPDDKLITYLYSPDSTLSRKIYAFDVVAKQQRLLVNPP 75
LPGYEAPTSIGFSADDSLLTYLFSPDHTLSRKVFAPFDLKSQKQELFFGPP 100
LPGYVAPTSIGFSADDSLVTCFLSPDHTLSRKVFAPFDLKSQKQELFFGPP 100
LPGYVSPTSISFSPDDSLISYLFSPDNTLNKRIFFDLKTNKQELFFSFP 93
LPGYVAPTAVSFSPDDSLITYLFSPEKNLKRKYAFDVNKGESNLVFSFP 67
-----
AHNLFSLRSFSSSIR-RFNRFSLKPLRSFASMSSSSPD--AAQTPLTTAP 71
LRPLFYCRSITSRRG-IYTVNCLRQSSADHKTTDMYAEIRIEVHKEKMSAP 59
TQSKLLQEFINIPSIDKAWIFNSDSGSQAMFALSQANLLANKKKKFMLSG 75

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828.18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

GGGVDEGNLSTVEKLRRLRLRERGLG-VTRYEWAKGVPHR--LMVPLPSG 122
GGGVDEGNLSTVEKLRRLRLRERGLG-VTRYEWGKGVPHR--LMVPLPSG 122
DGGDLDESNISSPEEKLRLRLRERGLG-VTCYEWVKTGSKKKAIMVPLPAG 149
DGGDLDESNISSAEKLRRLRLRERGLG-VTRYEWVKTGLKKKAIMVPLPAG 149
DGGDLDESNISSPEEKLRLRLRERGLG-VTRYEWVKTNSKKKAVLVLPLPAG 142
DGGVDENISNIPPEEKLRLRLRERGLG-VTRYEWVKTNSKMRFTVPLPAG 116
-----
YGSWKSPITADIVSGASKRLGGTAVDSHGRVLLESRPNESGRGVVLVQ 121
YGSWKSPITADIVSGADKRLGGFALDGEGRVILWLEGRPTEAGRSVLVREA 109
-----HISNESNQSVNFHWAPFFIEMTGASAFVPSPSGLKLLVIRNPEN 119

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828.18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

IY-VQG-VGSELRLRVASTPSNPILDPQLSPDASSIAYVRDDELFFVPLA 170
IY-VQEGVRAELQRRVASTSWSPILDPQLSPDACSIAAYVRDDEIFVVPVA 171
----ELHSS-KPELKLPSALSPIIDPHVSPDGTMLAYIRDSELHVLNLL 194
----ELYSP-KPELKLPSSSLSPIIDPHISPDGTMLAYVRDSELHVLNLL 194
IYVHDISHS-KTELKLPSIPASPIIDPHLSPDGSMLAYVRDCELHVMNLL 191
VYMKDLSSSPNELIVPSSPTSPIIDPRLSPNGLFLAYVRESELHVLNLL 166
-----MIAYVRDDELETVGFS 16
-----ETSIDITPKDFAVRTLQYEGGGAQFISSDDTLVFSNY 159
AD-----EEGTAEDITPAGFNVRTLVHEYGGGAFTVSG-DVVVFSNY 150
ESPTKFEIWNSSQLEKEFHIPQKVHGSVYVDGWFEGISWDSDETHVAYVA 169

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828.18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

FGEPIQITYGARGTGKTHGLAEYIAQEEMDRRNGFWWSPDSRYIAFAEVD 220
FGDPIQITSGARGTGKTHGLAEYIAQEEMDRRNGFWWSPDSRYIAFAEVD 221
YNESKQLTHGAQGNVTTHGLAEYIAQEEMDRKNGYWWSLDSKFIAFTQVD 244
FNESKQLTHGAQGNVTTHGLAEYIAQEEMDRKNGYWWSLDSKFIAFTQVD 244
SDESKQLTHGAKENGFTTHGLAEYIAQEEMDRKNGYWWSLDSKFIAFTQVD 241
KNQTQQLTSGANGSTLSHGLAEYIAQEEMDRRNGYWWSLDSKFIAFTQVD 216
DGQTQQLTSGASESGKIHGLAEYIAQEEMERKMGFWWSPDSKHLAFTEVD 66
KDQRLYKQDITDKDSSPKFITPDYGTFAVYADGVFDSRFNRYVTVREDG 209
KDQRLYKQSIKG-GHAPIALTDPYGA PVVRYADGVMDLHLGCVTVVREDH 199
EESRPRKPTFDHLGYKKENSLDKIGISWKGEEDWEEWGEAYAGKRQPA 219

```

Figura 12

```

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828.18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

SSRIPTFTIVHQQKVTGGEAEVHAYPFAGHSNVSVRLGVVPSNG-DSV 269
STRIPFFRIMHQQKASTGGEAEETHAYPFAGHSNVSVRLGVVPSNG-GSV 270
SSEIPLFRIMHQQKSSVSGSEAQEDHYPFAGASNVKVRGLGVVSVHG-DSI 293
SSEIPLFRIMHQQKSSVSGSEAQEDHYPFAGASNVKVRGLGVVSVHG-GSV 293
YSEIPLFRIMHQQGRSSVGTDAQEDHYPFAGASNAKIRLGVVSVAG-GST 290
SSQIPLFRIMHQQGRSVSGSEAQEDHAYPFAGALNSTLRLGVVSSAGGGKT 266
SSEIPLYRIMHQQKSSVGPDAQEDHAYPFAGASNVKVRGLGVVSSHG-GEI 115
RQDRSNPITTIIVEVNLSGETLEBPVKLVSGNDFYAFPRLPDPK----CERL 255
RESDTNPTTTIVSVELNGAGTTEPHVLVSGSDFYAFPRLPDPD----GGKM 245
LFVINVDSEGEVEPIKGI PRSISVGQVWVSPNSNGSAQYLVFAGWLGDKRK 269

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828.18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

VWMDLEVGLCGEGHEDEYLARVTWLPNGYLTAQVQNRQTQSKLKLRLRFD 319
VWMDLEVGLCGEGRDDEYLARVTWLPNGSLTAQVQNRQTQSKLKLRLRFD 320
TWMDLLCGGTKEPDNEDEYLARVNMHGNVLIQVNLNRSHSKLKLRLRFDI 343
TWLDLLCGGTKEPDNEDEYLARINWMHGNILIAQVNLNRSHSKLKLRLRFDI 343
TWMDLVCGGKELDNEDEYLARVNMHGNILTAQIINRHQTKIKIVKFDI 340
TWMNLVCGGRG--NTEDEYLGRVNWLPNGVLIQVNLNRSHSKLKLRLRFDI 314
TWMDLLCGEPNSIHGDEEYLARVNMHNSAIAVQVNLNRTHSKLKLRLRFDI 165
AWIEWSH-----PNMPWDKAEWLVGYISEGGNIDKRVCVAGC 292
AWIEWNH-----PNMPWDKSELVWGYMSAEGKVEKRICIAGN 282
FGIKYCNRPCTAIYAIKFTSDEPKDDANEFFIHNLTSSSGFCPRFSK 319

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828.18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

LTGKRLVLLTEESEVWVNLHDSFTPLRKY-HGPLSGGFIWASEKTGFRHL 368
LTGKRTVLLFTEESDVWVNLHDCFTSLRKC-HGLLSGGFIWASEKTGFRHL 369
KTKKEVLYAEQPLPWNLHDCFTPLDKG-ITKYSGGFIWASEKSGFRHL 392
KAGRKVEIYVEEQFPWNLHDCFTPLDKG-ITKYSGGFIWASEKTGFRHL 392
RTGQRRLVLEENKTWINIHDCFTPLDKG-VTKFSGGFIWASEKSGFRHL 389
NTGQGNVLLTEESDTWVTLHDCFTPLKVPSSRGSGGFIWASEKTGFRHL 364
ASGKREVLLEEHDWTWTLHDCFTPLDKGVNSKHPGGFIWASEKTGFRHL 215
DP---KYVESPTPEKWSRGELFFVTDRK--NGCWNHKKWIESTNEVSV 337
DP---NMIESPTPEKWSRGELFFVTDRK--SGFWNLKXWVESTNEVKAL 327
DG---KFLVFVSATAVDSGARWATESLH-RIDWPSDGKLPESTNIVDVI 365

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828.18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

YMHD-EFGNCIRALTQGSWMVEQVAGVDEAGLVYFTGAFDSPLEVHLYS 417
YVHD-EFGNCIRALTQGSWMVEQVAGVDEEAGLVYFTGTHDSPLEMHLYS 418
CVHD-ANGACLGPIITEGWMVEQIAGVNEAAGIYFTATLDGPLESILYR 441
YLHD-ANGTCLGPITEGDMVVEQIAGVNEAAGMIYFTATRDGPLESILYR 441
YLHD-ANGICLGPITEGWMVEQIAGVNEATGLVYFTGTLDSPLESNLYC 438
YLIE-SDGTCLGAITSGEWMVEQIAGVNEPMSLVYFTGTLDGPLETNLYC 413
YLHD-KNGVCLGPLTQGDWMIDQIAGVNESSGVYFTGTLDGPLETNLYS 264
YPLDGEFAKPLWIFGTNSYELIECSEENLIACSYRQKKG-SYLGIVDDS 386
YPLDAEFTRPSSVVFGNSSYAFIEQKQGNKNIACYRQKGM-SYLGILDHV 376
QVNCPCPKDCGCFGLYVTGLLSDPWLSDGHSIMLSTYWRSCRVLNLLS 415

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828.18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

ASLDCT-EPNLAEPKRLTQAGRHIVVLDHEMKKFVDIHDSDLTPPRLLL 466
VSLDCT-QPNLVEPKRLTQAGRHIVVLDHQMKNFVDIEDSLDTPPRLLL 467
AKLYPIENNPLQAPVRLTNGKSKHSVLDHHLQNFVDIEDSLDSPRVSL 491
AKLFPDEKNALQAPVRLTNGKSKHSVLDHHLQNFVDIEDSLDCPPRVLL 491
AKLFVDGTQPLQAPTRLTHSKGKHIVVLDHMMRTFVDIHDSDLSCPPRVLL 488
AKLEAG-NTSRCQPMRLTHGKSKHIVVLDHQMKNFVDIEDSVDSPPRVSL 462
TNLFPDWSLPQVPPKRLTHGTGRHSVILDHQLLRFIDVDSIKSPFVILL 314
QGSCLLDIPLTDFDSITLGNOCLYVEGASAVLPPSVARVTLDOHKTAL 436
LGSFSLVDLPFTDIYNITSIGSHLYLEGASPLHPLSIVKVSYEENLIAVR 426
GEVSRASPSDSYWNALALDGDSTIVAVSSSPVSPVEIKYKKGKGLDSACK 465

```

Figura 12 (cont.)

```

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828_18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

RSLDTGKLLISIIYEQPAPTYHARRLQLTPPEFKSLTA----SDG-TVLHG 511
RSLDTGKLLVTIYEQPAPTYHARRLQLTPPEFKSLTA----SDG-TVLHG 512
CSLFDGREIMPLFEQSFTIPRYKRLELEPPKIVQIQQA----NDG-TILYG 536
CSLIDGREIMPLFEQAFTIPRFKRLELEPPKIVQIQQA----NDG-TILYG 536
CSLEDGTIIMPLYEQIPIPKSKKLQLEPPEIVEIQS----DDG-TILYG 533
CSLSDGTIVLKILYEQTSPIQILKSLKLEPPEFVQIQQA----NDGKTTLYG 508
CSLLDGSVIMPLYEQPLTVQPLKKFQQLSPEIVQIEG----KDG-TALYG 359
S-----SEIVWSSS-PDVLKYKAYFSVPELIEFPT----EVPQONAYA 474
G-----ISIIWSSSSLNISEYSAFISSPEIVEFST----KVPQQTAF 465
PSWLWSNIQSPRIYSEKVMAGLSLQFKILKVPISDVSEGLAEGAKNP 515
: . :.

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828_18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

LIYIPDSKQFGRP---PYRTVSVYGGPNVQIVCNSWMNTVDMRAQYLR 558
LVYIPDSKQFGRP---PYRTVSVYGGPNVQIVCNSWMNTVDMRAQYLR 559
ALYDPDPTRFGRPP---PYKTVISVYGGPGVQYVCDSWIGTADMRAQYLR 583
ALYEPDPTRFGRPP---PYKTLSVYGGPSVQYVCDSWISTVDMRAQYLR 583
ALYKPPDSRFGRPP---PYKTMINVYGGPSVQLVSNLWNTVDLRAQYLR 580
AVYKPPDSKFGPP---PYKTMINVYGGPSVQLVSDSWINTVDMRTQYLR 555
TLYLPDEKKYGP---PYKTLNVYGGPSVQLVSDSWISTVDMRAQFLR 406
YFYPPPTNPLYNASMEEEKPPLLKSHGGPTAESR-----SLNLIQYWT 519
YLYLPSNINYEAPEGEKPPLLVKS HGGPTSESHS-----ALDLIQYWT 510
AIYVSSSSKSKENG--KCDPLIAVLHGGPHSVSPCS-----FSRTMAYLES 558
.* . : *** : .

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828_18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

RGILVWKLNDNRGSARRGLDFEGAIXHNMGMIDVDDQQTGVQWLVSQGLAM 608
RGILVWKLNDNRGSARRGLDFEGAIXHNMGMIDVEDQQTGVQWLVSQGLAM 609
QGILVWKLNDNRGSARRGLKFEGLKGNPGRFDAQEDQLTGAEWLIKQGLAK 633
KGILVWKLNDNRGSARRGLKFEGLKGNPGRFDAQEDQLTGAEWLIKQGLAK 633
KGILVWKLNDNRGTSRRGLKFEGLKQKLGQIDADDQFTGAEWLVKNGLA 630
RGILVWKLNDNRGTARRGLKFESWMKHNCGYVDAEDQVTGAKWLIEQGLAK 605
KGILVWKMNDNRGTARRGLQFEGQLKYNIGRVAEDQLAGAELIKKGLAK 456
RGWAFVDVNYGGSTGYGREYRERLLRQWGVVDVDDCCGCAKYLVS S GKAD 569
RGWAFADVNYGGSTGYGREYRERLNGSWGIVDVNDCCSCAEFLVTTGRVD 560
IGYSQLIINYRGSGLGYGDALQSLPGKVGSDQDVKDCLLAVDHA IEMGIAD 608
* : : * . : . * * . . : * .

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828_18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

PNRIGIYGWSYGGYMASMALAR-----CPETFSCAVAGAPVT----- 645
SNRIGIYGWSYGGYMASMALAR-----CPQTFSCAVAGAPVT----- 646
AGHIGLCGWSYGGYMSAVILAR-----FPDVFCVAVSGAPVT----- 670
AGHIGLYGWSYGGYMSAMILAR-----FPDVFCVAVSGAPVT----- 670
FGHIGLYGWSYGGYLSAMTLR-----YPDFKCAVAGAPVT----- 667
PDHIGVYGWSYGGYLSATLLR-----YPEIFNCVAVSGAPVT----- 642
PGHIGLYGWSYGGFLSAMCLAR-----FPDTFSCAVSGAPVT----- 493
VKRLCISGGSGAGGYTTLASLAFRDVFKAGASLYGVADLKMLKE----- 612
GERLCITGRSAGGYTTLAALVFRETFFKAGASLFGVADVSLKLA----- 603
PSRITVLGSGHGGFLTTHLIG-----APDKFVAAAARNPVCNMASMGV 652
: : * * * : : . : *

P.patens_104540#1
P.patens_105959#1
P.trichocarpa_scaff_XII.203#1
P.trichocarpa_scaff_XV.146#1
M.truncatula_AC137828_18.5#1
A.thaliana_AT5G24260.1#1
O.sativa_Os02g0290600#1
gi|145358557|ref|NP_198470.3|
gi|116789192|gb|ABK25153.1|
gi|42566792|ref|NP_193193.2|

-----SWDGYDTYYTEKYMGRP--ASNAGYQYSSVMHHVYQITG--KLL 686
-----SWDGYDTHYTEKYMGRP--ASNAGYQYSSVMHHVYQITG--KLL 687
-----SWDGYDTFYTEKYMGRP--SDNPKGYEYGSVMHHVHKLKG--RLL 711
-----SWDGYDTFYTEKYMGRP--YENPTGYEYGSVMHHVHKLKG--RLL 711
-----SWDGYDTFYTEKYMGRP--SEYKSGYARASVMHHVHKMRG--RLL 708
-----SWDGYDSFYTEKYMGRP--TEER--YKSSVMHHVGNLTDKQKLM 684
-----AWDGYDTFYTEKYMGRP--SEQRDAYRYGSIIMHHVKNLRG--RLL 534
-----EGHKFESRYIDNLVG-----DEKDFYERSPINFVDKFSCT--II 649
-----DTHKFESYTDLSLVGK-----DESLLYERSPINFVDRLSCP--MI 641
ITDIPJCWCFEAYGDSHYTEAPSAEDLSRFHQMSPIHSISKVKTPT--TL 700
: : : . * : : . :

```

Figura 12 (cont.)

P.patens_104540#1	LVHGMIDENVHFRHTARLINTLTAASKDYELLVFPDERHMPRGLDRMYM	736
P.patens_105959#1	LVHGMIDENVHFRHTARLINSLTAAASKDYELLVFPDERHMPRGLDRMYM	737
P.trichocarpa_scaff_XII.203#1	LVHGMIDENVHFRHTARLVNALVAAGKPYELLIFPDERHMPRRHNDRIYM	761
P.trichocarpa_scaff_XV.146#1	LVHGMIDENVHFRHTARLVNALVAAGKPYELLIFPDERHMPRRHTDRIYM	761
M.truncatula_AC137828.18.5#1	IVHGMIDENVHFRHTARLINALVAAGKTYELLIFPDERHMPRRYSDRVYM	758
A.thaliana_ATSG24260.1#1	LVHGMIDENVHFRHTARLVNALVEAGKRYELLIFPDERHMPRKKKDRIYM	734
O.sativa_Os02g0290600#1	LIHGMIDENVHFRHTARLINSLEAEGKPYDILLFPDERHMPRRLGDRYM	584
gi 145358557 ref NP_198470.3	LFQGLEDKVVPDQSRKIYEALKKKGKLPVALVEYEGEGHQHGFKAENIKYT	699
gi 16789192 gb ABK25153.1	LFQGLEDKVVPPEQARKIYAAVKEKGLPVALVEYEGEGHQHGFKAENIKYT	691
gi 42566792 ref NP_193193.2	FLLGTKDLRVPSINGFQYVRALKEKGVEVKVLVFPDNDHFLDRPQTDYES	750
	:. * * * : : : : . : : . : *	
P.patens_104540#1	EERICDFLERNI-----	748
P.patens_105959#1	EERICDFLERNI-----	749
P.trichocarpa_scaff_XII.203#1	EERIWEFFQSL-----	773
P.trichocarpa_scaff_XV.146#1	EERIWEFFERNL-----	773
M.truncatula_AC137828.18.5#1	EERMWEFIDRNL-----	770
A.thaliana_ATSG24260.1#1	EQRIWEFIEKNL-----	746
O.sativa_Os02g0290600#1	EERIWDVFERNL-----	596
gi 145358557 ref NP_198470.3	LEQQMVFVFARVVGGFKVADDITPLKIDNFD	730
gi 16789192 gb ABK25153.1	LEQQMVFVFARLIGNFKVADDIIPVHIENFD	721
gi 42566792 ref NP_193193.2	FLNIADVFNKYCKL-----	764
	.. :	

Figura 12 (cont.)

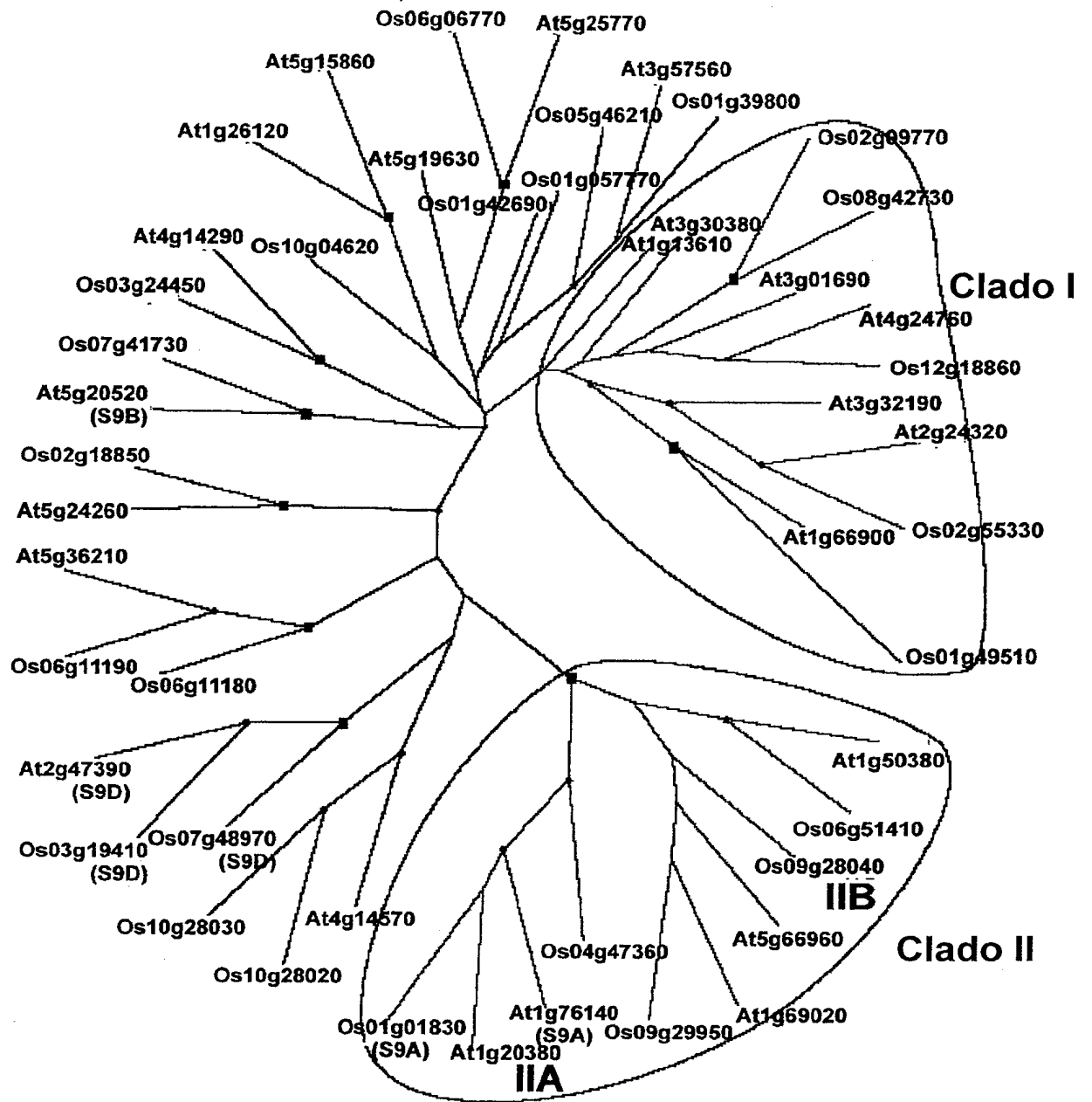
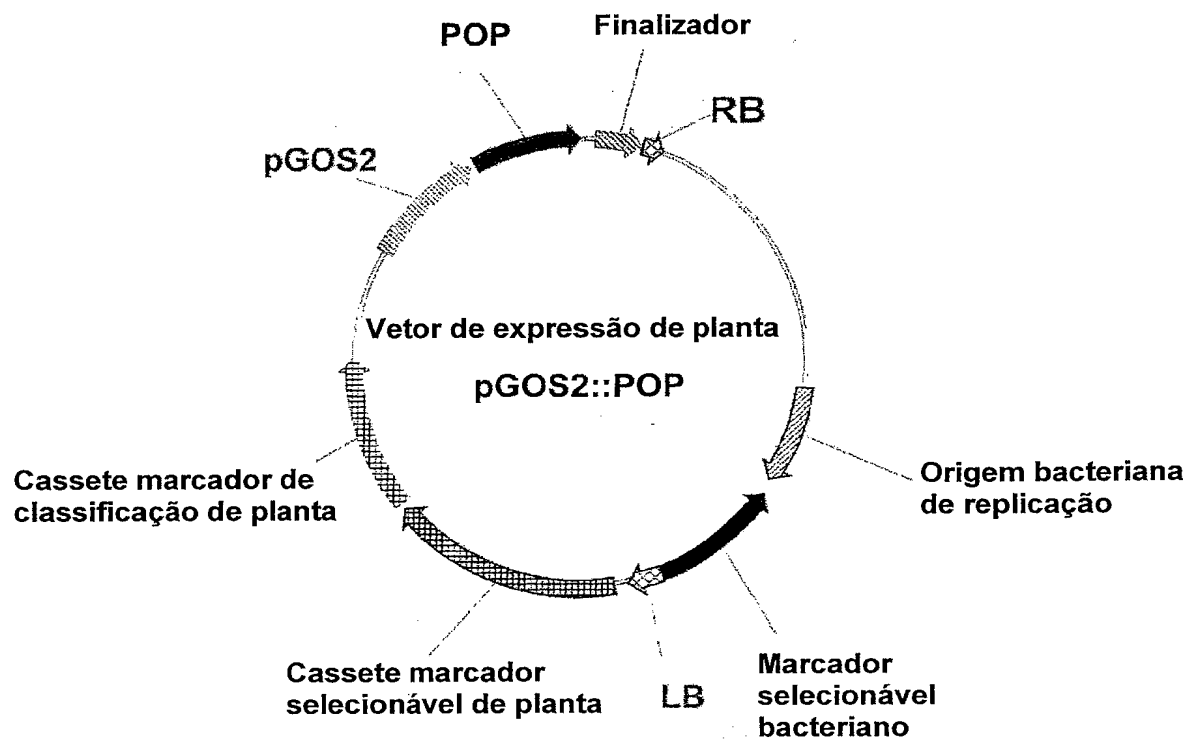


Figura 13

**Figura 14**

P.taeda	TA18056	3352#1	---	---	MVDGICLREETSRCGAPAYTDDNNKAGSGSGSRRVINAVILKALCV
S.bicolor	5284384#1		---	---	MSGG
S.officinatum	TA38665	4547#1	---	---	MSGG
Z.mays	ZM07MC01438	576683801433	---	---	MSGG
O.sativa.indica	BGIOSIBCE034659#		---	---	MSGG
O.sativa	Os1lg0524300#1		---	---	MSGG
T.aestivum	TA92222	4565#1	---	---	MSGG
A.thaliana	AT5G51020.1#1		---	---	MSGG
B.napus	TA26749	3708#1	---	---	MSGG
H.paradoxus	TA3639	73304#1	---	---	MSGG
T.officinale	TA2756	50225#1	---	---	MSGG
C.undiviva	TA817	114280#1	---	---	MSGG
L.saligna	TA1573	75948#1	---	---	MSGG
I.nil	TA8128	35883#1	---	---	MSGG
N.tabacum	TA16794	4097#1	---	---	MSGG
S.lycopersicum	TA38444	4081#1	---	---	MSGG
S.tuberosum	TA25303	4113#1	---	---	MSGG
V.vinifera	GSVIVT00018055001#1		---	---	MSGG
G.raimondii	TA13121	29730#1	---	---	MSGG
C.clementina	DY292515#1		---	---	MSGG
C.sinensis	TA15191	2711#1	---	---	MSGG
P.trichocarpa	834377#1		---	---	MSGG
G.max	Gm0065x00465#1		---	---	MSGG
M.truncatula	AC139708	31.5#1	---	---	MSGG
S.moellendorffii	170435#1		---	---	MSGG
P.patens	133413#1		---	---	MSGG

[illegible]

P.taeda_TA18056_3352#1
 S.bicolor_5284384#1
 S.officinatum_TA38665_4547#1
 Z.mays_ZM07MC01438_57666838@1433
 O.sativa.indica_BGIOSIBCE034659#
 O.sativa_Os1lg0524300#1
 T.aestivum_TA92222_4565#1
 A.thaliana_AT5G51020.1#1
 B.napus_TA26749_3708#1
 H.paradoxus_TA3639_73304#1
 T.officinale_TA2756_50225#1
 C.undivia_TA817_114280#1
 L.saligna_TA1573_75948#1
 I.nil_TA8128_35883#1
 N.tabacum_TA16794_4097#1
 S.lycopersicum_TA38444_4081#1
 S.tuberosum_TA25303_4113#1
 V.vinifera_GSVIVT00018055001#1
 G.raimondii_TA13121_29730#1
 C.clementina_DY292515#1
 C.sinensis_TA15191_2711#1
 P.trichocarpa_834377#1
 G.max_Gm0065x00465#1
 M.truncatula_AC139708_31.5#1
 S.moellendorffii_170435#1
 P.patens_133413#1

Figura 15 (cont.)

P.taeda_TAI18056_3352#1
S.bicolor_5284384#1
S.officinatum_TA38665_4547#1
Z.mays_ZM07MC01438_5766683801433
O.sativa_indica_BGIOSIBCE034659#
O.sativa_Os11g0524300#1
T.aestivum_TA92222_4565#1
A.thaliana_AT5G51020.1#1
B.napus_TA26749_3708#1
H.paradoxus_TA3639_73304#1
T.officinale_TA2756_50225#1
C.undivisa_TA817_114280#1
L.saligna_TA1573_75948#1
I.nil_TA8128_35883#1
N.tabacum_TA16794_4097#1
S.lycopersicum_TA38444_4081#1
S.tuberosum_TA25303_4113#1
V.vinifera_GSVIVT00018055001#1
G.raimondii_TA13121_29730#1
C.clementina_DY292515#1
C.sinensis_TA15191_2711#1
P.trichocarpa_834377#1
G.max_Gm0065x00465#1
M.truncatula_AC139708_31.5#1
S.moellendorffii_170435#1
P.patens_133413#1

Figura 15 (cont.)

P.taeda_TA18056_3352#1	-----
S.bicolor_5284384#1	K-LDKAYVIDDQ-----
S.officinatum_TA38665_4547#1	K-LDKAYVIDDQ-----
Z.mays_ZM07MC01438_5766683801433	K-LDKAYVIDDQ-----
O.sativa_indica_BGIOSIBCE034659#	K-IESSFAIDDQ-----
O.sativa_Os1lg0524300#1	K-IESSFAIDDQ-----
T.aestivum_TA92222_4565#1	--MERSYAIDDQ-----
A.thaliana_AT5G51020.1#1	P-IQGSFILQD-----
B.napus_TA26749_3708#1	P-IKGSFILQE-----
H.paradoxus_TA3639_73304#1	R-IEGSFVLQE-----
T.officinale_TA2756_50225#1	R-IEGSFWPQE-----
C.endivia_TA817_114280#1	R-IEGSFVLQE-----
L.saligna_TA1573_75948#1	R-IEGSFVLQE-----
I.nil_TA8128_35883#1	R-IEGSFVIQD-----
N.tabacum_TA16794_4097#1	R-IEGSFVLQE-----
S.lycopersicum_TA38444_4081#1	R-IEGSFVLQE-----
S.tuberosum_TA25303_4113#1	R-IEGSFVLQE-----
V.vinifera_GSVIVT00018055001#1	R-IDGSFVLQE-----
G.raimondii_TA13121_29730#1	R-IEGSFVLQE-----
C.clementina_DY292515#1	RDLDPPFVLPKMITSVHGN
C.sinensis_TA15191_2711#1	R-IEGSFVLQE-----
P.trichocarpa_834377#1	R-IEGSFVLQV-----
G.max_Gm0065x00465#1	R-IEGSFVLQE-----
M.truncatula_AC139708_31.5#1	R-IEGSFVLQE-----
S.moellendorffii_170435#1	-----
P.patens_133413#1	-----

Figura 15 (cont.)

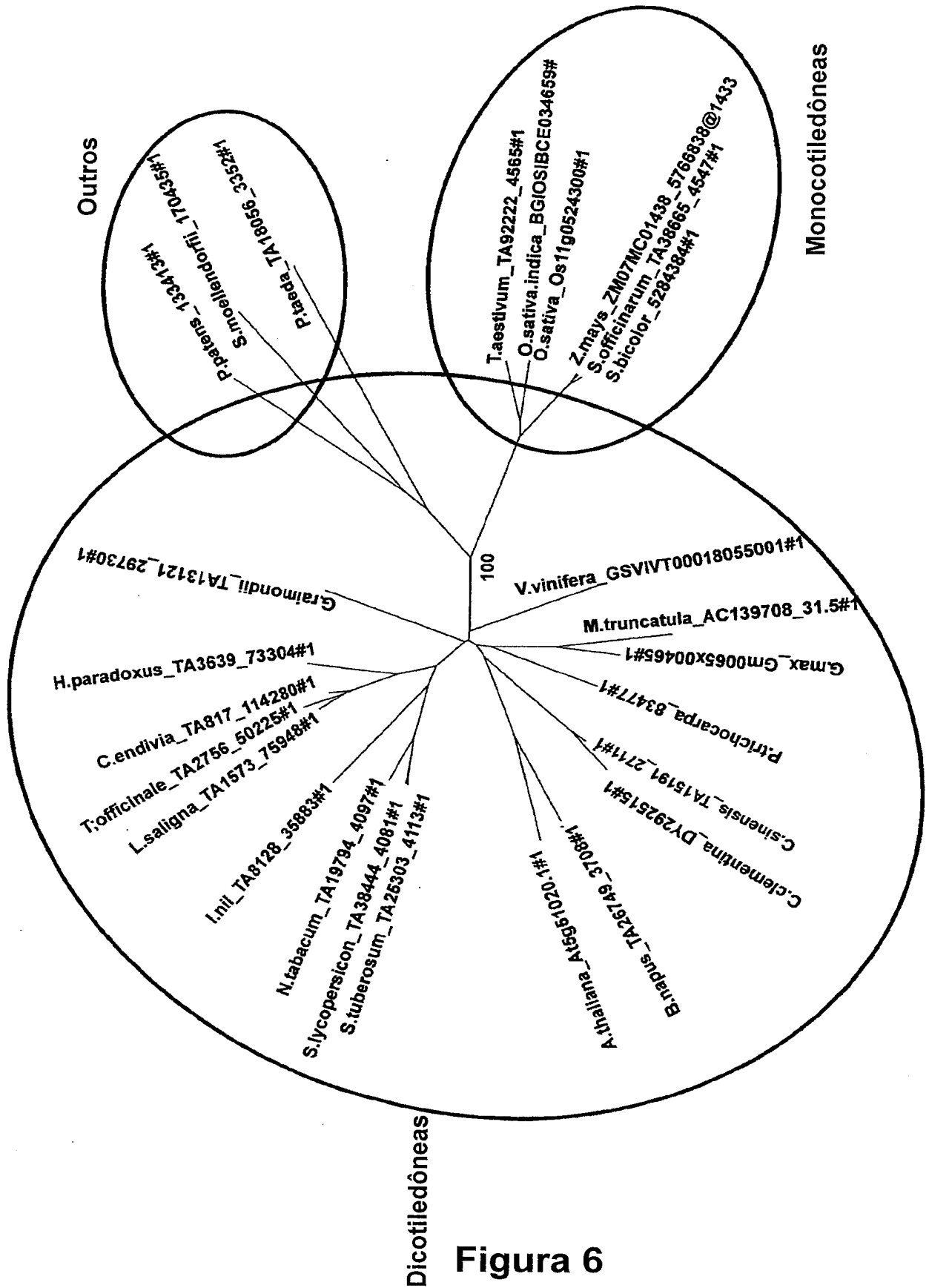
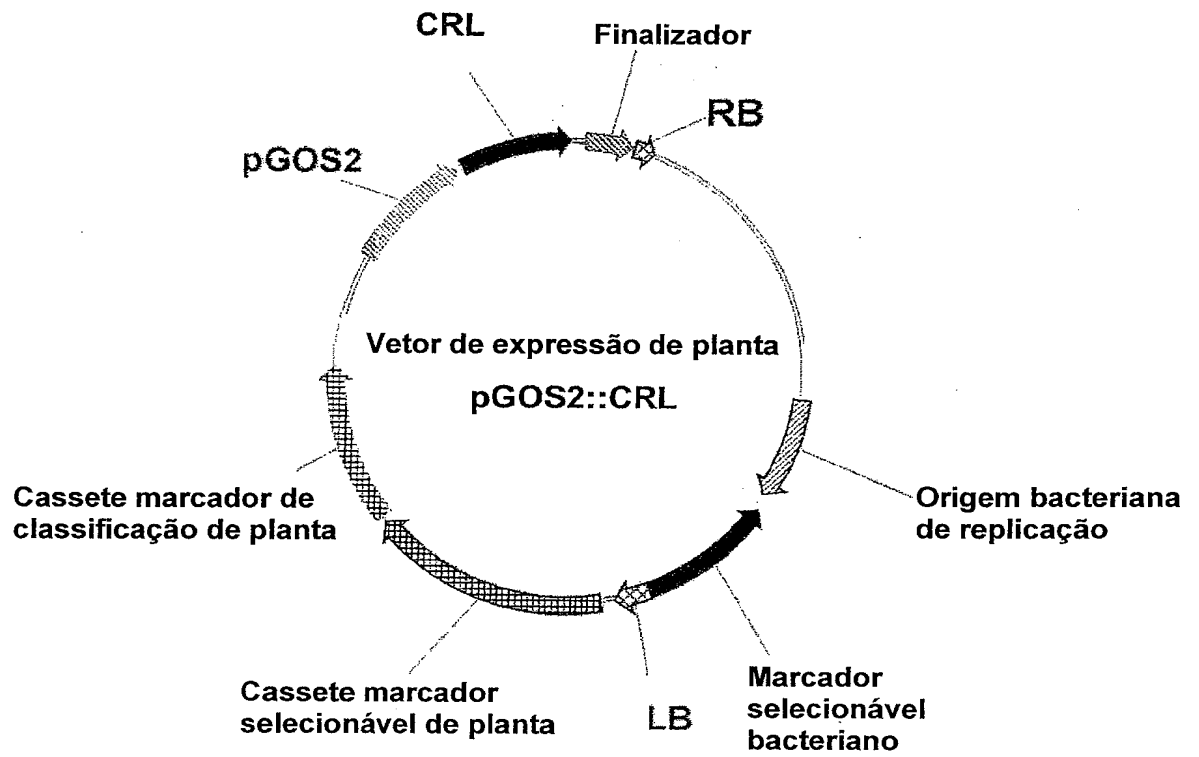


Figura 6

**Figura 17**

RESUMO

“MÉTODO PARA AUMENTAR A TOLERÂNCIA A ESTRESSE ABIÓTICO E/OU PARA AUMENTAR UMA CARACTERÍSTICA RELACIONADA AO RENDIMENTO EM UMA PLANTA, CONSTRUTO, PLANTA, CÉLULA DE PLANTA OU PARTE DA MESMA, INCLUINDO SEMENTES, PARTES CULTIVÁVEIS DA DITA PLANTA, PRODUTOS DERIVADOS DE DITA PLANTA E/OU DE DITAS PARTES CULTIVÁVEIS, MÉTODO PARA A PRODUÇÃO DE PLANTAS TRANSGÊNICAS, USO DE UM CONSTRUTO OU DE UM ÁCIDO NUCLEICO, MOLÉCULA DE ÁCIDO NUCLEICO ISOLADA, E, POLIPEPTÍDEO ISOLADO”

A presente invenção diz respeito em geral ao campo da biologia molecular e está relacionada a um método para aumentar várias características da planta relacionadas ao rendimento aumentando a expressão em uma planta com uma sequência de ácidos nucleicos que codifica um polipeptídeo do tipo de Brevis Radix (BRXL). A presente invenção também está relacionada a plantas com melhor expressão de uma sequência de ácidos nucleicos que codifica um polipeptídeo BRXL, cujas plantas apresentam melhores características relacionadas ao rendimento com relação às plantas controle. A invenção diz respeito adicionalmente a sequências de ácidos nucleicos, construtos de ácido nucléico, vetores e plantas contendo as ditas sequências de ácidos nucleicos.