



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

(21) PI 0720759-0 A2



(22) Data de Depósito: 13/12/2007
(43) Data da Publicação: 21/01/2014
(RPI 2246)

(51) Int.Cl.:
A01H 1/00
C12N 15/82

(54) Título: PRODUÇÃO DE PLANTAS COM
CONTEÚDO ALTERADO DE ÓLEOS, PROTEÍNAS OU
FIBRAS

(57) Resumo:

(30) Prioridade Unionista: 15/12/2006 US 60/870,345

(73) Titular(es): Agrinomics LLC

(72) Inventor(es): Hein Tsoeng (Medard) NG, John P. Davies

(74) Procurador(es): Dannemann, Siemsen, Bigler &
Ipanema Moreira

(86) Pedido Internacional: PCT US2007087437 de
13/12/2007

(87) Publicação Internacional: WO 2008/076831 de
26/06/2008

Relatório Descritivo da Patente de Invenção para **"PRODUÇÃO DE PLANTAS COM CONTEÚDO ALTERADO DE ÓLEOS, PROTEÍNAS OU FIBRAS"**.

REFERÊNCIA CRUZADA A PEDIDO(S) DE PATENTE(S) RELACIONA-
DO(S)

Este pedido de patente reivindica o benefício do Pedido de Patente U.S. Provisório Nº 60/870.345, depositado em 15 de dezembro de 2006, cuja totalidade é incorporada aqui como referência.

CAMPO DA DESCRIÇÃO

A presente invenção refere-se a plantas modificadas com conteúdo alterado de óleos, proteínas e/ou fibras, assim como a métodos de produção de plantas modificadas que possuem conteúdo modificado de óleos, proteínas e/ou fibras partindo de tais plantas.

ANTECEDENTES

A capacidade de manipular a composição de sementes de plantas de cultivo, particularmente o conteúdo e a composição de óleo e de proteína das sementes, assim como a energia metabolizável disponível ("AME") na farinha de sementes no gado, possui aplicações importantes nas indústrias agrícolas, em relação tanto a óleos para alimentos processados quanto a produtos alimentícios para animais. As sementes de plantas de cultivo agrícolas contêm uma variedade de constituintes valiosos, incluindo óleo, proteína e amido. O processamento industrial pode separar parte ou todos estes constituintes para venda individual em aplicações específicas. Por exemplo, quase 60% da plantação de soja nos Estados Unidos da América são triturados pela indústria de processamento da soja. O processamento da soja produz óleo purificado, que é vendido com alto valor, enquanto que o restante da farinha das sementes é vendido para a alimentação de gado (U.S. Soybean Board, 2001 Soy Stats). A semente de canola também é triturada para produzir óleo e o co-produto farinha de canola (Canola Council of Canada). A farinha de canola contém uma alta porcentagem de proteína e um bom equilíbrio de aminoácidos, mas devido ao fato de que possui um alto teor de fibras e fitato, não é facilmente digerida pelo gado (Slominski,

B.A. e outros, 1999 Proceedings of the 10th International Rapeseed Congress, Canberra, Austrália) e possui um valor inferior ao da farinha de soja.

Mais de 55% do milho produzido nos EUA são utilizados como produto alimentício para animais (Iowa milho Growers Association). O valor
5 do milho está diretamente relacionado com sua capacidade de ser digerido pelo gado. Assim, é desejável maximizar tanto o conteúdo de óleo das sementes quanto a AME da farinha. Para as sementes oleosas processadas tais como soja e canola, o aumento do conteúdo absoluto de óleo da semente aumentará o valor de tais grãos, enquanto que o aumento da AME da fa-
10 rinha aumentará seu valor. Para o milho processado, um aumento ou um decréscimo no conteúdo de óleo pode ser desejado, dependendo de como os outros constituintes principais serão utilizados. A diminuição de óleo pode aumentar a qualidade do amido isolado através da redução de sabores inde-
15 sejados associados com a oxidação do óleo. Alternativamente, quando o amido é utilizado para a produção de etanol, quando o sabor não é importante, o aumento do conteúdo de óleo pode aumentar o valor geral.

Em muitos grãos para alimentação, tais como milho e trigo, é desejável aumentar o conteúdo de óleo das sementes, porque o óleo possui maior conteúdo de energia que os outros constituintes da semente tal como
20 o carboidrato. O processamento de sementes oleosas, como a maior parte dos negócios de processamento de grãos, é um negócio de capital intensivo; assim pequenas alterações na distribuição dos produtos dos componentes de menor valor para o componente de óleo de alto valor podem ter impactos econômicos substanciais para os processadores de grãos. Em adição, o
25 aumento da AME da farinha através do ajuste do conteúdo e da composição de proteínas e de fibras das sementes, sem diminuir o conteúdo de óleo das sementes, pode aumentar o valor do produto alimentício para animais.

A manipulação biotecnológica de óleos tem mostrado fornecer alteração na composição e aumento do rendimento dos óleos. As alterações
30 na composição incluem óleo de soja e de milho com alto conteúdo de ácido oleico (Patentes U.S. Nºs 6.229.033 e 6.248.939) e sementes contendo laurato (Patente U.S. Nº 5.639.790), entre outros. O trabalho na alteração da

composição se concentrou predominantemente em sementes oleosas processadas, mas pode ser facilmente estendido a plantas de cultivo sem ser semente oleosa, incluindo milho. Embora haja interesse considerável no aumento do conteúdo de óleo, a única biotecnologia praticada atualmente
5 nesta área é a tecnologia High-Oil Corn (HOC) (Dupont, Patente U.S. Nº 5.704.160). HOC emprega polinizadores de alto conteúdo de óleo desenvolvidos através do cruzamento por seleção clássico com fêmeas híbridas de elite (estéreis em relação ao sexo masculino) em um sistema de produção referido como TopCross. O sistema TopCross High Oil aumenta o conteúdo
10 de óleo no grão colhido no milho de aproximadamente 3,5% até aproximadamente 7%, aumentando o conteúdo de energia do grão.

Embora este tenha sido produtivo, o sistema de produção HOC possui limitações inerentes. Em primeiro lugar, o sistema de ter uma baixa porcentagem de polinizadores responsáveis por um conjunto inteiro de se-
15 mentes do campo contém riscos inerentes, particularmente em anos de seca. Em segundo lugar, o conteúdo de óleo nos campos de HOC atuais não foi capaz de atingir um conteúdo de óleo de semente acima de 9%. Finalmente, o milho com alto conteúdo de óleo não é primariamente uma alteração bioquímica, mas ao invés disso um mutante anatômico (tamanho de
20 embrião maior) que possui o resultado indireto de aumentar o conteúdo de óleo. Por estas razões, uma estratégia de alto conteúdo de óleo alternativa, particularmente uma que deriva de um produto bioquímico alterado, seria especialmente valiosa.

A manipulação da composição das sementes identificou vários
25 componentes que aumentam a qualidade nutritiva, a digestibilidade e a AME na farinha de semente. O aumento do conteúdo de lisina na canola e na soja (Falco e outros, 1995 *Bio/Technology* 13:577-582) aumenta a disponibilidade deste aminoácido essencial e diminui a necessidade de suplementos nutricionais. Foi mostrado que variedades de soja com maior proteína na semente
30 contêm consideravelmente mais energia metabolizável que as variedades convencionais (Edwards e outros, 1999, *Poultry Sci.* 79:525-527). Foi mostrado que a diminuição do conteúdo de fitato da semente de milho aumenta

a disponibilidade biológica de aminoácidos nos produtos alimentícios para animais (Douglas e outros, 2000, *Poultry Sci.* 79:1586-1591) e a diminuição do conteúdo de oligossacarídeos na farinha de soja aumenta a energia metabolizável na farinha (Parsons e outros, 2000, *Poultry Sci.* 79:1127-1131).

5 A soja e a canola são as plantas de cultivo alvos mais óbvias para os mercados de óleo processado e de farinha de semente uma vez que ambas as plantas de cultivo são trituradas para óleo e a farinha remanescente é vendida para a alimentação de animais. Um grande grupo de trabalho comercial (por exemplo, Patente U.S. Nº 5.952.544; Pedido de Patente PCT
10 Nº WO9411516) demonstra que *Arabidopsis* é um modelo excelente para o metabolismo de óleo nestas plantas de cultivo. Verificações bioquímicas da composição do óleo das sementes identificaram genes de *Arabidopsis* para muitas enzimas biossintéticas críticas e levaram à identificação de ortólogos de genes importantes para a agronomia. Por exemplo, verificações utilizando
15 populações mutagenizadas quimicamente identificaram mutantes lipídicos cujas sementes exibem composição alterada de ácidos graxos (Lemieux e outros, 1990, *Theor. Appl. Genet.* 80: 234-240; James e Dooner, 1990, *Theor. Appl. Genet.* 80: 241-245). Verificações de mutagênese por T-DNA (Feldmann e outros, 1989, *Science* 243: 1351-1354) que detectaram compo-
20 sição alterada de ácidos graxos identificaram os genes da ômega 3 dessaturase (*FAD3*) e da delta-12 dessaturase (*FAD2*) (Patente U.S. Nº 5952544; Yadav e outros, 1993, *Plant Physiol.* 103: 467-476; Okuley e outros, 1994, *Plant Cell* 6(1):147-158). Uma verificação que se concentrou no conteúdo de óleo ao invés de na qualidade do óleo, analisou mutantes induzidos quimi-
25 camente em relação a sementes enrugadas ou à densidade alterada de sementes, cujo conteúdo alterado de óleo da semente foi inferido (Focks e Benning, 1998, *Plant Physiol.* 118:91-101).

Uma outra verificação, planejada para identificar enzimas envolvidas na produção de ácidos graxos de cadeia muito longa, identificou uma
30 mutação no gene que codifica uma diacilglicerol aciltransferase (DGAT) como sendo responsável pelo acúmulo reduzido de triacil glicerol nas sementes (Katavic V e outros, 1995, *Plant Physiol.* 108(1):399-409). Foi também

mostrado que a superexpressão específica à semente do cDNA de DGAT estava associada com o maior conteúdo de óleo na semente (Jako e outros, 2001, *Plant Physiol.* 126(2):861-74). *Arabidopsis* é também um modelo para o entendimento do acúmulo dos componentes da semente que afetam a qualidade da farinha. Por exemplo, *Arabidopsis* contém as proteínas de armazenamento de semente albumina e globulina encontradas em muitas plantas dicotiledôneas incluindo canola e soja (Shewry 1995, *Plant Cell* 7:945-956). As vias bioquímicas para a síntese de componentes de fibra, tais como celulose e lignina, são conservadas dentro das plantas vasculares e mutantes de *Arabidopsis* que afetam estes componentes foram isolados (revisado em Chapel e Carpita 1998, *Current Opinion in Plant Biology* 1:179-185).

A marcação por ativação em plantas refere-se a um método de produção de mutações aleatórias através da inserção de um constructo de ácido nucleico heterólogo que compreende sequências reguladoras (por exemplo, um intensificador) no genoma de uma planta. As sequências reguladoras podem atuar para aumentar a transcrição de um ou mais genes vegetais nativos; consequentemente, a marcação por ativação é um método produtivo para a produção de mutantes de ganho de função, geralmente dominantes (ver, por exemplo, Hayashi e outros, 1992, *Science* 258: 1350-1353; Weigel D e outros, 2000, *Plant Physiology*, 122:1003-1013). O constructo inserido fornece uma marcação molecular para a identificação rápida da planta nativa cuja expressão errada causa o fenótipo mutante. A marcação por ativação também pode causar fenótipos de perda de função. A inserção pode resultar na interrupção de um gene vegetal nativo, em cujo caso o fenótipo é geralmente recessivo.

A marcação por ativação foi utilizada em várias espécies, incluindo tabaco e *Arabidopsis*, para a identificação de muitos tipos diferentes de fenótipos mutantes e os genes associados com estes fenótipos (Wilson e outros, 1996, *Plant Cell* 8: 659-671; Schaffer e outros, 1998, *Cell* 93: 1219-1229; Fridborg e outros, 1999, *Plant Cell* 11: 1019-1032; Kardailsky e outros, 1999, *Science* 286: 1962-1965; e Christensen S e outros, 1998, 9th Internati-

onal Conference on Arabidopsis Research, Univ. of Wisconsin-Madison, 24-28 de junho, Resumo 165).

SUMÁRIO

São fornecidas aqui plantas modificadas que possuem um fenótipo alterado. As plantas modificadas com um fenótipo alterado podem incluir um fenótipo de quantidade maior de óleo e/ou de melhor qualidade de farinha. O fenótipo alterado em uma planta modificada pode também incluir conteúdo de óleos, proteínas e/ou fibras alterado em qualquer parte da planta modificada, por exemplo, nas sementes. Em algumas modalidades de uma planta modificada, o fenótipo alterado é um aumento no conteúdo de óleo da semente (um fenótipo de alto conteúdo de óleo). Em outras modalidades, o fenótipo alterado pode ser um aumento no conteúdo de proteínas na semente e/ou um decréscimo no conteúdo de fibras da semente. É também fornecida uma farinha de semente derivada das sementes de plantas modificadas, em que as sementes possuem conteúdo alterado de proteínas e/ou conteúdo alterado de fibras. É também fornecido um óleo derivado das sementes de plantas modificadas, em que as sementes possuem conteúdo de óleo alterado. Qualquer uma destas alterações pode levar a um aumento na AME da semente ou da farinha da semente das plantas modificadas, em relação às plantas de controle ou do tipo selvagem. É também fornecida aqui uma farinha, um produto alimentício ou um alimento produzido partindo de qualquer parte da planta modificada com um fenótipo alterado.

Em certas modalidades, as plantas modificadas descritas incluem plantas transgênicas que possuem um vetor de transformação que compreende uma sequência de nucleotídeos de HIO (ou pseudônimo do gene HIO) que codifica ou é complementar a uma sequência que codifica um polipeptídeo "HIO". Em modalidades particulares, a expressão de um polipeptídeo de HIO em uma planta transgênica causa um conteúdo alterado de óleo, um conteúdo alterado de proteínas e/ou um conteúdo alterado de fibras na planta transgênica. Em modalidades preferidas, a planta transgênica é selecionada do grupo que consiste em plantas das espécies de *Brassica*, incluindo canola e semente de colza, soja, milho, girassol, algodão, cacau, açá-

froa, dendezeiro, coqueiro, linho, mamona, amendoim, trigo, aveia e arroz. É também fornecido um método de produção de óleo ou de farinha de semente, que compreende o cultivo da planta transgênica e a recuperação do óleo e/ou da farinha da semente da dita planta. A descrição fornece ainda produto alimentício, farinha, grão ou semente que compreende uma sequência de ácido nucleico que codifica um polipeptídeo de HIO. A descrição fornece ainda produto alimentício, farinha, grão ou semente que compreende o polipeptídeo de HIO ou um ortólogo ou um parálogo do mesmo.

Os exemplos da planta transgênica descrita são produzidos através de um método que compreende a introdução nas células progenitoras da planta de um vetor de transformação de planta que compreende uma sequência de nucleotídeos de HIO que codifica, ou é complementar a uma sequência que codifica, um polipeptídeo de HIO e o crescimento das células progenitoras transformadas para produzir uma planta transgênica, em que a sequência de polinucleotídeo de HIO é expressa, causando um fenótipo alterado na planta transgênica. Em alguns exemplos não limitantes específicos, o método produz plantas transgênicas em que a expressão do polipeptídeo de HIO causa um fenótipo de alto conteúdo (maior) de óleos, de alto conteúdo (maior) de proteínas e/ou de baixo conteúdo (menor) de fibras na planta transgênica, em relação às plantas não transgênicas ou do tipo selvagem de controle.

São descritos aqui métodos adicionais de produção de uma planta que possui um fenótipo alterado, em que é identificado que uma planta possui uma mutação ou um alelo em sua sequência de ácido nucleico de HIO que resulta em um fenótipo alterado, comparada a plantas que não possuem a mutação ou o alelo. A planta mutada pode ser gerada utilizando um ou mais agentes mutagênicos, por exemplo, um agente mutagênico químico, radiação ou luz ultravioleta. Em algumas modalidades do método, a planta é cruzada para gerar progênie que herda o alelo e expressa o fenótipo alterado. Em modalidades particulares do método, o método emprega a metodologia de gene/QTL candidato ou a metodologia de TILLING.

É também fornecida aqui uma célula vegetal modificada que

possui um fenótipo alterado. Em algumas modalidades, a célula vegetal modificada inclui um vetor de transformação que compreende uma sequência de nucleotídeos de HIO que codifica ou é complementar a uma sequência que codifica um polipeptídeo de HIO. Em modalidades preferidas, a célula vegetal transgênica é selecionada do grupo que consiste em plantas das espécies de *Brassica*, incluindo canola e semente de colza, soja, milho, girassol, algodão, cacau, açafrão, dendezeiro, coqueiro, linho, mamona, amendoim, trigo, aveia e arroz. Em outras modalidades, a célula vegetal é uma semente, um pólen, um propágulo ou uma célula embrionária. A descrição fornece ainda células vegetais provenientes de uma planta que é a progênie direta ou a progênie indireta de uma planta crescida partindo das ditas células progenitoras.

DESCRIÇÃO DETALHADA

Termos

A não ser que seja indicado de outra maneira, todos os termos técnicos e científicos utilizados aqui possuem o mesmo significado que teriam para um versado na técnica da presente descrição. Os profissionais se direcionam particularmente a Sambrook e outros (*Molecular Cloning: A Laboratory Manual* (Segunda Edição), Cold Spring Harbor Press, Plainview, N.Y., 1989) e Ausubel FM e outros (*Current Protocols in Molecular Biology*, John Wiley & Sons, Nova York, N.Y., 1993) para definições e termos da técnica. Deve ser entendido que esta descrição não está limitada à metodologia, aos protocolos e aos reagentes particulares descritos, uma vez que estes podem variar.

Como utilizado aqui, o termo "fenótipo alterado" refere-se a plantas ou a qualquer parte de uma planta (por exemplo, sementes ou farinha produzida partindo de sementes), com um conteúdo de óleos, proteínas e/ou fibras alterado (fenótipo). Como fornecido aqui, conteúdo de óleos, de proteínas (por exemplo, proteína que pode ser digerida) e/ou de fibras alterado inclui um nível maior ou menor de conteúdo de óleos, de proteínas (por exemplo, proteína que pode ser digerida) e/ou de fibras nas plantas, nas sementes ou na farinha de sementes. Qualquer combinação destas alterações

pode levar a um fenótipo alterado. Por exemplo, em um exemplo não-limitante específico, um fenótipo alterado pode se referir ao conteúdo maior de óleos e menor de fibras. Em um outro exemplo não limitante específico, um fenótipo alterado pode se referir ao conteúdo inalterado de proteínas e menor de fibras. Em um outro exemplo não limitante específico, um fenótipo alterado pode se referir ao conteúdo maior de óleos e de proteínas e menor de fibras. Ainda em outros exemplos não limitantes, um fenótipo alterado pode se referir ao conteúdo maior de óleos e de proteínas e inalterado de fibras; ao conteúdo inalterado de óleos, maior de proteínas e menor de fibras; ou ao conteúdo maior de óleos, maior de proteínas e menor de fibras. É também fornecido que qualquer combinação destas alterações pode levar a um aumento na AME (energia metabolizável disponível) da semente ou farinha gerada partindo da semente. Um fenótipo alterado também inclui um fenótipo de maior qualidade de semente (ISQ) ou um fenótipo de maior qualidade de farinha da semente.

Como utilizado aqui, o termo "energia metabolizável disponível" (AME) refere-se à quantidade de energia no produto alimentício que é capaz de ser extraída através da digestão em um animal e está correlacionada com a quantidade de proteínas que podem ser digeridas e de óleo na farinha para animais. A AME é determinada através da estimativa da quantidade de energia no produto alimentício antes da alimentação e da medida da quantidade de energia nos excrementos do animal após o consumo do produto alimentício. Em um exemplo não limitante específico, uma planta modificada com um aumento na AME inclui plantas modificadas com conteúdo alterado de óleo da semente, de proteínas que podem ser digeridas, de proteína total e/ou de fibras, resultando em um aumento no valor do produto alimentício para animais derivado da semente.

Como utilizado aqui, o termo "conteúdo" refere-se ao tipo e à quantidade relativa de, por exemplo, um componente da semente ou de farinha da semente.

Como utilizado aqui, o termo "óleo da semente" refere-se à quantidade total de óleo dentro da semente.

Como utilizado aqui, o termo "fibra da semente" refere-se aos componentes que não podem ser digeridos da semente da planta incluindo componentes celulares tais como celulose, hemicelulose, pectina, lignina e compostos fenólicos.

5 Como utilizado aqui, o termo "farinha" refere-se aos componentes da semente remanescentes após a extração de óleo da semente. Os exemplos de componentes de farinha incluem proteína e fibra.

 Como utilizado aqui, o termo "proteína total da semente" refere-se à quantidade total de proteína dentro da semente.

10 Como utilizado aqui, o termo "proteína que pode ser digerida da semente" refere-se à proteína da semente que é capaz de ser digerida por enzimas no trato digestivo de um animal. É um subconjunto do conteúdo de proteína total.

 Como utilizado aqui, o termo "vetor" refere-se a um constructo
15 de ácido nucleico planejado para a transferência entre células hospedeiras diferentes. Um "vetor de expressão" refere-se a um vetor que possui a capacidade de incorporar e expressar fragmentos de DNA heterólogos em uma célula estranha. Muitos vetores de expressão procarióticos e eucarióticos estão disponíveis comercialmente. A seleção de vetores de expressão apropriados está dentro do conhecimento dos versados na técnica.
20

 Um constructo ou uma sequência de ácido nucleico "heteróloga" possui uma parte da sequência que não é nativa à célula vegetal em que é expressa. Heteróloga, em relação a uma sequência de controle refere-se a uma sequência de controle (isto é, promotor ou intensificador) que não funciona na natureza para regular o mesmo gene cuja expressão é, no momento, de regulação. Geralmente, sequências de ácidos nucleicos heterólogas não são endógenas à célula ou à parte do genoma em que estão presentes e foram adicionadas na célula por infecção, transfecção, microinjeção, eletroporação ou similar. Um constructo de ácido nucleico "heterólogo" pode
25 conter uma combinação de sequência de controle/sequência codificadora de DNA que é a mesma que ou diferente de, uma combinação de sequência de controle/sequência codificadora de DNA encontrada na planta nativa. Os
30

exemplos não limitantes específicos de uma sequência de ácido nucleico heteróloga incluem uma sequência de ácido nucleico de HIO ou um fragmento, um derivado (variação) ou um ortólogo ou um parálogo da mesma.

5 Como utilizado aqui, o termo "gene" significa o segmento de DNA envolvido na produção de uma cadeia de polipeptídeo, que pode ou não incluir regiões que precedem e que sucedem a região codificadora, por exemplo, sequências não- traduzidas a 5' (5' UTR) ou "líderes" e sequências 3' UTR ou "trailers", assim como sequências que se estendem (introns) entre segmentos codificadores individuais (éxons) e sequências reguladoras não
10 transcritas.

O termo "homólogo" refere-se a qualquer gene que está relacionado a um gene de referência por herança de uma sequência de DNA ancestral comum. O termo "ortólogo" refere-se a homólogos em espécies diferentes que evoluíram partindo de um gene ancestral comum através da es-
15 peciação. Tipicamente, os ortólogos mantêm a mesma ou uma função similar apesar das diferenças em sua estrutura primária (mutações). O termo "parálogo" refere-se a homólogos na mesma espécie que evoluíram através da duplicação genética de um gene ancestral comum. Em muitos casos, os parálogos exibem funções relacionadas (mas nem sempre idênticas). Como
20 utilizado aqui, o termo homólogo abrange tanto ortólogos quanto parálogos. Até a extensão que uma espécie particular desenvolveu vários genes relacionados partindo de uma sequência de DNA ancestral compartilhada com uma outra espécie, o termo ortólogo pode abranger o termo parálogo.

Como utilizado aqui, "recombinante" inclui referência a uma célu-
25 la ou um vetor, que foi modificado através da introdução de uma sequência de ácido nucleico heteróloga ou que a célula é derivada de uma célula modificada dessa maneira. Assim, por exemplo, as células recombinantes expressam genes que não são encontrados na forma idêntica dentro da forma nativa (não recombinante) da célula ou expressam genes nativos que são de
30 outra maneira expressos de forma anormal, subexpressos ou não expressos de forma alguma como um resultado da intervenção humana deliberada.

Como utilizado aqui, o termo "expressão gênica" refere-se ao

processo através do qual um polipeptídeo é produzido com base na sequência de ácido nucleico de um gene. O processo inclui tanto transcrição quanto tradução; conseqüentemente, "expressão" pode se referir a uma sequência de polinucleotídeos ou de polipeptídeos ou ambas. Algumas vezes, a expressão de uma sequência de polinucleotídeos não levará à tradução de proteínas. "Superexpressão" refere-se à maior expressão de uma sequência de polinucleotídeos e/ou de polipeptídeos em relação a sua expressão em uma planta do tipo selvagem (ou outra de referência [por exemplo, não-transgênica]) e pode se referir a uma sequência que ocorre naturalmente ou que não ocorre naturalmente. "Expressão ectópica" refere-se à expressão em um momento, local e/ou maior nível que não ocorre naturalmente na planta não modificada ou do tipo selvagem. "Subexpressão" refere-se à menor expressão de uma sequência de polinucleotídeos e/ou de polipeptídeos, geralmente de um gene endógeno, em relação a sua expressão em uma planta do tipo selvagem. Os termos "expressão errada" e "expressão alterada" abrangem superexpressão, subexpressão e expressão ectópica.

O termo "introduzida" no contexto de inserir uma sequência de ácido nucleico em uma célula, inclui "transfecção", "transformação" e "transdução" e inclui referência à incorporação de uma sequência de ácido nucleico em uma célula eucariótica ou procariótica em que a sequência de ácido nucleico pode ser incorporada no genoma da célula (por exemplo, DNA cromossômico, plasmídeo, de plastídeo ou mitocondrial), convertida em um replicon autônomo ou expressa de forma transitória (por exemplo, mRNA transfectado).

Como utilizado aqui, uma "célula vegetal" refere-se a qualquer célula derivada de uma planta, incluindo células de tecido não diferenciado (por exemplo, calo), assim como de sementes, pólen, propágulos e embriões de plantas.

Como utilizado aqui, os termos "nativa" e "do tipo selvagem" relativos a uma certa característica ou fenótipo da planta referem-se à forma em que tal característica ou fenótipo é encontrado na mesma variedade de planta na natureza. Em uma modalidade, uma planta do tipo selvagem ou

nativa é também uma planta de controle. Em uma outra modalidade, uma planta do tipo selvagem ou nativa é uma planta não-transgênica ou não-mutada. Ainda em uma outra modalidade, uma planta do tipo selvagem ou nativa é uma planta não modificada.

5 Como utilizado aqui, o termo "modificada" em relação a uma planta, refere-se a uma planta com um fenótipo alterado (por exemplo, uma planta gerada através de métodos de engenharia genética, de mutagênese ou de cruzamento). Uma planta engenheirada geneticamente também pode ser uma planta transgênica. Em modalidades particulares, as plantas modifi-
10 cadas geradas através de métodos de cruzamento são primeiramente mutagenizadas utilizando qualquer um de uma variedade de agentes mutagênicos, tal como um agente mutagênico químico, radiação ou luz ultravioleta. As plantas modificadas podem ter qualquer combinação de um conteúdo alterado de óleos, um conteúdo alterado de proteínas e/ou um conteúdo alte-
15 rado de fibras em qualquer parte da planta transgênica, por exemplo, nas sementes, em relação a uma planta não modificada similar.

 Como utilizado aqui, o termo "alterada" refere-se a uma modificação (um aumento ou um decréscimo) de uma característica ou de um fenótipo da planta (por exemplo, conteúdo de óleos, conteúdo de proteínas
20 e/ou conteúdo de fibras) em uma planta modificada, em relação uma planta não modificada similar. Em um exemplo não limitante específico, uma planta modificada com uma característica alterada inclui uma planta com um maior conteúdo de óleos, maior conteúdo de proteínas e/ou menor conteúdo de fibras em relação a uma planta não modificada similar. Em um outro exem-
25 plo não limitante específico, uma planta modificada com uma característica alterada inclui conteúdo de óleos inalterado, maior conteúdo de proteínas e/ou menor conteúdo de fibras em relação a uma planta não modificada similar. Ainda em um outro exemplo não limitante específico, uma planta modificada com uma característica alterada inclui um maior conteúdo de óleos,
30 um maior conteúdo de proteínas e/ou um conteúdo inalterado de fibras em relação a uma planta não modificada similar.

 Um "fenótipo interessante (característica)" com referência a uma

planta modificada refere-se a um fenótipo que pode ser observado ou medido demonstrado por uma T1 e/ou uma planta de geração subsequente, que não é exibido pela planta não modificada correspondente (isto é, uma planta similar genotipicamente que foi crescida ou analisada sob condições similares). Um fenótipo interessante pode representar uma melhoria na planta (por exemplo, maior conteúdo de óleos, maior conteúdo de proteínas e/ou menor conteúdo de fibras nas sementes da planta) ou pode fornecer um meio para produzir melhorias em outras plantas. Uma "melhoria" é uma característica que pode aumentar a utilidade de uma espécie ou variedade de planta através do fornecimento à planta de um fenótipo ou uma qualidade exclusiva e/ou nova. Tais plantas modificadas podem ter um fenótipo melhorado, tal como um fenótipo de óleo, proteína e/ou fibra alterado. A farinha gerada partindo de sementes de uma planta modificada com um fenótipo melhorado pode ter melhor (maior) qualidade de farinha. Em um exemplo não limitante específico de farinha com um fenótipo de melhor (maior) qualidade, a farinha é gerada partindo de uma semente de uma planta modificada, em que a semente possui maior conteúdo de proteínas e/ou menor conteúdo de fibras, em relação a uma planta não modificada similar.

A expressão "fenótipo de conteúdo alterado de óleos" refere-se a um fenótipo que pode ser medido de uma planta modificada, em que a planta exibe um aumento ou um decréscimo estatisticamente significativo no conteúdo total de óleos (isto é, a porcentagem de massa da semente que é óleo), quando comparada à planta similar, mas não modificada (por exemplo, uma não transgênica ou uma não mutada). Um fenótipo de alto conteúdo de óleos refere-se a um aumento no conteúdo total de óleos. Um aumento no conteúdo de óleos inclui, em várias modalidades, um aumento de aproximadamente 1,0%, 1,5%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5%, 5,0%, 7,5%, 10% ou mais no conteúdo de óleos. Similarmente, um decréscimo no conteúdo de óleos inclui um decréscimo de aproximadamente 1,0%, 1,5%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5%, 5,0%, 7,5%, 10% ou mais no conteúdo de óleos, em várias modalidades.

A expressão "fenótipo de conteúdo alterado de proteínas" refere-

se ao fenótipo que pode ser medido de uma planta modificada, em que a planta exibe um aumento ou um decréscimo estatisticamente significativo no conteúdo total de proteínas (isto é, a porcentagem de massa da semente que é proteína), quando comparada à planta similar, mas não modificada (por exemplo, não transgênica ou não mutada). Um fenótipo de alto conteúdo de proteínas refere-se a um aumento no conteúdo total de proteínas. Um aumento no conteúdo de proteínas inclui, em várias modalidades, um aumento de aproximadamente 1,0%, 1,5%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5%, 5,0%, 7,5%, 10% ou mais no conteúdo total de proteínas. Similarmente, um aumento no conteúdo de proteínas que podem ser digeridas inclui, em várias modalidades, um aumento de aproximadamente 1,0%, 1,5%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5%, 5,0%, 7,5%, 10% ou mais no conteúdo de proteínas que podem ser digeridas. Um decréscimo no conteúdo de proteínas inclui um decréscimo de aproximadamente 1,0%, 1,5%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5%, 5,0%, 7,5%, 10% ou mais no conteúdo total de proteínas, em várias modalidades. Similarmente, um decréscimo no conteúdo de proteínas que podem ser digeridas inclui, em várias modalidades, um decréscimo de aproximadamente 1,0%, 1,5%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5%, 5,0%, 7,5%, 10% ou mais no conteúdo de proteínas que podem ser digeridas. A expressão "fenótipo de conteúdo alterado de fibras" refere-se ao fenótipo que pode ser medido de uma planta modificada, em que a planta exibe um aumento ou um decréscimo estatisticamente significativo no conteúdo total de fibras (isto é, a porcentagem de massa da semente que é fibra), quando comparada à planta similar, mas não modificada (por exemplo, não transgênica ou não mutada). Um fenótipo de baixo conteúdo de fibras refere-se à diminuição no conteúdo total de fibras. Um aumento no conteúdo de fibras inclui um aumento de aproximadamente 1,0%, 1,5%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5%, 5,0%, 7,5%, 10% ou mais no conteúdo de fibras. Similarmente, um decréscimo no conteúdo de fibras inclui um decréscimo de aproximadamente 1,0%, 1,5%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5%, 5,0%, 7,5%, 10% ou mais no conteúdo de fibras.

Como utilizado aqui, uma sequência de polinucleotídeos ou um

gene "mutante" ou "mutado" difere da sequência de polinucleotídeos ou do gene do tipo selvagem correspondente em termos de sequência ou de expressão, em que a diferença contribui para um fenótipo ou uma característica alterada da planta. Em relação a uma planta ou uma linhagem de planta, o termo "mutante" ou "mutada" refere-se a uma planta ou uma linhagem de planta que possui um fenótipo ou uma característica alterada da planta, em que o fenótipo ou a característica alterada está associada com a expressão alterada de uma sequência de polinucleotídeos ou um gene do tipo selvagem. A sequência de polinucleotídeos ou o gene mutado pode ser gerado através de métodos de engenharia genética (tal como marcação por ativação ou transformação), através da utilização de um ou mais agentes mutagênicos (por exemplo, agentes mutagênicos químicos, radiação ou luz ultravioleta) ou através da utilização de métodos para alterar uma sequência de DNA (por exemplo, PCR propensa a erros, cruzamento molecular com embaralhamento de DNA, mutagênese direcionada ao sítio ou introdução do gene em um organismo que induz mutagênese tal *E. coli* ou cepas de leveduras que são deficientes em relação à atividade de reparo do DNA).

Como utilizado aqui, o termo "T1" refere-se à geração de plantas partindo da semente de plantas T0. A geração T1 é o primeiro conjunto de plantas modificadas que pode ser selecionado através da aplicação de um agente de seleção, por exemplo, um antibiótico ou um herbicida, para o qual a planta modificada contém o gene de resistência correspondente. O termo "T2" refere-se à geração de plantas através de autofertilização de flores de plantas T1, previamente selecionadas como sendo modificadas. As plantas T3 são geradas partindo de plantas T2 etc. Como utilizado aqui, a "progênie direta" de uma certa planta deriva da semente (ou, algumas vezes, outro tecido) de tal planta e está na geração imediatamente subsequente; por exemplo, para uma certa linhagem, uma planta T2 é a progênie direta de uma planta T1. A "progênie indireta" de uma certa planta deriva da semente (ou outro tecido) da progênie direta de tal planta ou da semente (ou outro tecido) de gerações subsequentes em tal linhagem; por exemplo, uma planta T3 é a progênie indireta de uma planta T1.

Como utilizado aqui, o termo "parte da planta" inclui qualquer órgão ou tecido da planta, incluindo, sem limitação, sementes, embriões, regiões meristemáticas, tecido de calo, folhas, raízes, brotos, gametófitos, esporófitos, pólen e micrósporos. As células vegetais podem ser obtidas partindo de qualquer órgão ou tecido da planta e culturas preparadas partindo do mesmo. É fornecida aqui uma célula vegetal modificada que possui um fenótipo alterado. Em modalidades particulares, a célula vegetal modificada é uma célula vegetal transgênica. A célula vegetal transgênica inclui um vetor de transformação que compreende uma sequência de nucleotídeos de HIO que codifica ou é complementar a uma sequência que codifica um polipeptídeo de HIO. Em modalidades preferidas, a célula vegetal transgênica é selecionada do grupo que consiste em plantas das espécies de *Brassica*, incluindo canola e semente de colza, soja, milho, girassol, algodão, cacau, açafrão, dendezeiro, coqueiro, linho, mamona, amendoim, trigo, aveia e arroz. Em outras modalidades, a célula vegetal é uma semente, um pólen, um propágulo ou uma célula embrionária. A descrição fornece ainda células vegetais de uma planta que é a progênie direta ou a progênie indireta de uma planta crescida partindo das ditas células progenitoras. A classe de plantas que pode ser utilizada nos métodos da presente invenção é geralmente tão ampla quanto a classe de plantas superiores sensíveis às técnicas de transformação, incluindo tanto plantas monocotiledôneas quanto dicotiledôneas.

Como utilizado aqui, "planta transgênica" inclui uma planta que compreende dentro de seu genoma um polinucleotídeo heterólogo. O polinucleotídeo heterólogo pode estar integrado de forma estável no genoma ou pode ser extracromossômico. Preferencialmente, o polinucleotídeo da presente invenção está integrado de forma estável no genoma de forma que o polinucleotídeo é passado a diante para gerações sucessivas. Uma célula, um tecido, um órgão vegetal ou uma planta na qual os polinucleotídeos heterólogos foram introduzidos é considerada "transformada", "transfectada" ou "transgênica". As progênies diretas e indiretas de plantas ou células vegetais transformadas que também contêm o polinucleotídeo heterólogo também são consideradas transgênicas.

São descritas aqui plantas modificadas que possuem um fenótipo alterado. As plantas modificadas com um fenótipo alterado podem incluir uma quantidade melhorada (maior) de óleos e/ou uma qualidade melhor (maior) de farinha, quando comparadas à planta similar, mas não modificada (por exemplo, não transgênica ou não mutada). As plantas modificadas com um fenótipo alterado podem incluir conteúdo alterado de óleos, proteínas e/ou fibras em qualquer parte da planta modificada, por exemplo, nas sementes, quando comparadas à planta similar, mas não modificada (por exemplo, não transgênica ou não mutada). Em algumas modalidades de uma planta modificada, por exemplo, em plantas com um fenótipo de conteúdo melhorado ou maior de óleos, o fenótipo alterado inclui um aumento no conteúdo de óleos da semente (um fenótipo de alto conteúdo de óleos) da planta, quando comparado ao da planta similar, mas não modificada (não transgênica ou não mutada). Um aumento no conteúdo de óleos inclui, em várias modalidades, um aumento de aproximadamente 1,0%, 1,5%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5%, 5,0%, 7,5%, 10% ou mais no conteúdo de óleos. O fenótipo alterado pode ser um aumento em um ou mais ácidos graxos, tal como ácido oleico, com um decréscimo concomitante em outros ácidos graxos tais como os ácidos linoleicos ou linolínicos. Uma alteração no conteúdo de ácidos graxos inclui um aumento de aproximadamente 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% ou mais em um ácido graxo específico. Em outras modalidades de uma planta modificada, por exemplo, em plantas com um fenótipo de qualidade melhorada ou maior de farinha, o fenótipo alterado pode ser um aumento no conteúdo de proteínas na semente e/ou um decréscimo no conteúdo de fibras da semente. Um aumento no conteúdo de proteínas inclui um aumento de aproximadamente 1,0%, 1,5%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5%, 5,0%, 7,5%, 10% ou mais no conteúdo de proteínas, por exemplo, no conteúdo total de proteínas ou no conteúdo de proteínas que podem ser digeridas. Esta alteração no conteúdo de proteínas da semente pode ser o resultado de quantidades alteradas de proteínas de armazenamento da semente tais como albuminas, globulinas prolaminas e glutelinas. Um decréscimo no conteúdo de fibras inclui um decréscimo de

aproximadamente 1,0%, 1,5%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5%, 5,0%, 7,5%, 10% ou mais no conteúdo de fibras. Esta alteração no conteúdo de fibras pode ser o resultado de quantidades alteradas de componentes fibrosos tais como celulose, hemicelulose, lignina e pectinas.

5 É também fornecida farinha de semente derivada das sementes de plantas modificadas, em que as sementes possuem conteúdo alterado (por exemplo, maior) de proteínas (por exemplo, que pode ser digerida) e/ou conteúdo alterado (por exemplo, menor) de fibras. É adicionalmente fornecido óleo derivado das sementes de plantas modificadas, em que as sementes
10 possuem conteúdo alterado de óleos. Qualquer uma destas alterações pode levar a um aumento na AME da semente ou da farinha da semente de plantas modificadas, em relação a plantas de controle, não transgênicas ou do tipo selvagem. Um aumento na AME inclui um aumento de aproximadamente 1,0%, 1,5%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5%, 5,0%, 7,5%, 10% ou
15 mais na AME na semente ou na farinha da semente, em várias modalidades. É também fornecida aqui farinha, produto alimentício ou alimento produzido partindo de qualquer parte da planta modificada com um fenótipo alterado.

Em certas modalidades, as plantas transgênicas descritas compreendem um vetor de transformação que compreende uma sequência de
20 nucleotídeos de HIO que codifica ou é complementar a uma sequência que codifica um polipeptídeo de "HIO". Em modalidades particulares, a expressão de um polipeptídeo de HIO em uma planta transgênica causa um conteúdo alterado de óleos, um conteúdo alterado de proteínas e/ou um conteúdo alterado de fibras na planta transgênica. Em modalidades preferidas, a planta transgênica é selecionada do grupo que consiste em plantas das espécies
25 de *Brassica*, incluindo canola e semente de colza, soja, milho, girassol, algodão, cacau, açafrão, dendezeiro, coqueiro, linho, mamona, amendoim, trigo, aveia e arroz. É também fornecido um método de produção de óleo ou farinha da semente, que compreende o cultivo da planta transgênica e a recuperação do óleo e/ou da farinha da semente da dita planta. A descrição fornece ainda produto alimentício, farinha, grão ou semente que compreende
30 uma sequência de ácido nucleico que codifica um polipeptídeo de HIO. A

descrição fornece ainda produto alimentício, farinha, grão ou semente que compreende o polipeptídeo de HIO ou um ortólogo ou um parálogo do mesmo.

Vários métodos para a introdução de uma sequência de polinucleotídeos desejada que codifica a proteína desejada nas células vegetais estão disponíveis e são conhecidos pelos versados na técnica e incluem, mas não estão limitados a: (1) métodos físicos tais como microinjeção, eletroporação e fornecimento mediado por microprojéteis (biolística ou tecnologia de pistola gênica); (2) métodos de fornecimento mediado por vírus; e (3) métodos de transformação mediada por *Agrobacterium*.

Os métodos mais comumente utilizados para a transformação de células vegetais são o processo de transferência de DNA mediada por *Agrobacterium* e o processo de biolística ou mediado por bombardeio de microprojéteis (isto é, a pistola gênica). Tipicamente, a transformação nuclear é desejada, mas quando é desejável transformar especificamente plastídeos, tais como cloroplastos ou amiloplastos, os plastídeos vegetais podem ser transformados utilizando um fornecimento mediado por microprojéteis do polinucleotídeo desejado.

A transformação mediada por *Agrobacterium* é conseguida através do uso de uma bactéria do solo engenheirada geneticamente que pertence ao gênero *Agrobacterium*. Um número de cepas do tipo selvagem e desarmadas de *Agrobacterium tumefaciens* e *Agrobacterium rhizogenes* portando os plasmídeos Ti ou Ri pode ser utilizado para a transferência de genes para as plantas. A transferência de genes é feita através da transferência de um DNA específico conhecido como "T-DNA" que pode ser engenheirado geneticamente para carregar qualquer pedaço desejado de DNA para dentro de muitas espécies de plantas.

A transformação genética de plantas mediada por *Agrobacterium* envolve várias etapas. A primeira etapa, na qual o *Agrobacterium* virulento e as células vegetais são primeiramente colocados em contato um com o outro, é geralmente chamada de "inoculação". Após a inoculação, é permitido que o *Agrobacterium* e as células/tecidos vegetais sejam cultivados juntos

durante um período de várias horas até vários dias ou mais sob condições adequadas para crescimento e para transferência do T-DNA. Esta etapa é denominada de "co-cultura". Após a co-cultura e o fornecimento do T-DNA, as células vegetais são tratadas com agentes bactericidas ou bacteriostáticos para matar ou limitar o crescimento do *Agrobacterium* remanescente em contato com o explante e/ou no recipiente contendo o explante. Se isto for realizado na ausência de quaisquer agentes seletivos para promover o crescimento preferencial de células vegetais transgênicas versus não transgênicas, então esta é tipicamente referida como a etapa de "retardo". Se for realizada na presença de pressão seletiva favorecendo as células vegetais transgênicas, então é referida como uma etapa de "seleção". Quando um "retardo" é utilizado, este é tipicamente seguido por uma ou mais etapas de "seleção".

Em relação ao bombardeio de microprojéteis (Patente U.S. Nº 5.550.318; Patente U.S. Nº 5.538.880. Patente U.S. Nº 5.610.042; e Publicação PCT WO 95/06128; das quais cada uma é especificamente incorporada aqui como referência em sua totalidade), as partículas são revestidas com ácidos nucleicos e fornecidas para dentro das células através de uma força de propulsão. Os exemplos de partículas incluem as compreendidas de tungstênio, platina e preferencialmente, ouro.

Uma modalidade ilustrativa de um método para o fornecimento de DNA para dentro das células vegetais através da aceleração é o Biolistics Particle Delivery System (BioRad, Hercules, CA), que pode ser utilizado para empurrar partículas revestidas com DNA ou células através de uma tela, tal como uma tela de aço inoxidável ou uma tela Nytex, sobre uma superfície de filtro coberta com células de plantas monocotiledôneas cultivadas em suspensão.

As técnicas de bombardeio de microprojéteis podem ser amplamente aplicadas e podem ser utilizadas para transformar virtualmente qualquer espécie de planta. Os exemplos de espécies que foram transformadas através do bombardeio de microprojéteis incluem espécies de monocotiledôneas tais como milho (Publicação PCT Nº WO 95/06128), cevada, trigo (Pa-

tente U.S. Nº 5.563.055, incorporada aqui como referência em sua totalidade), arroz, aveia, centeio, cana-de-açúcar e sorgo, assim como um número de dicotiledôneas incluindo tabaco, soja (Patente U.S. Nº 5.322.783, incorporada aqui como referência em sua totalidade), girassol, amendoim, algodão, tomate e verduras em geral (Patente U.S. Nº 5.563.055, incorporada aqui como referência em sua totalidade).

Para selecionar ou classificar as células vegetais independentemente da metodologia de transformação, o DNA introduzido na célula contém um gene que funciona em um tecido vegetal que pode ser regenerado para produzir um composto que confere ao tecido vegetal resistência a um composto que seria de outra maneira tóxico. Os genes de interesse para uso como um marcador que pode ser selecionado, verificado ou classificado incluiriam, mas não estão limitados a GUS, proteína fluorescente verde (GFP), luciferase (LUX), genes de tolerância a antibióticos ou herbicidas. Os exemplos de genes de resistência a antibióticos incluem penicilinas, canamicina, neomicina, G418, bleomicina, metotrexato (e trimetoprim), cloranfenicol e tetraciclina. As moléculas de polinucleotídeos que codificam as proteínas envolvidas na tolerância a herbicidas são conhecidas na técnica e incluem, mas não estão limitadas a uma molécula de polinucleotídeo que codifica a 5-enolpiruvilshikimato-3-fosfato sintase (EPSPS) descrita na Patente U.S. Nº 5.627.061, na Patente U.S. Nº 5.633.435 e na Patente U.S. Nº 6.040.497 e aroA descrita na Patente U.S. Nº 5.094.945 para tolerância ao glifosato; uma molécula de polinucleotídeo que codifica bromoxinil nitrilase (Bxn) descrita na Patente U.S. Nº 4.810.648 para tolerância ao Bromoxinil; uma molécula de polinucleotídeo que codifica fitoeno dessaturase (crtI) descrita em Misawa e outros, (*Plant J.* 4:833-840, 1993) e Misawa e outros, (*Plant J.* 6:481-489, 1994) para tolerância ao norflurazon; uma molécula de polinucleotídeo que codifica acetohidroxiácido sintase (AHAS, também conhecida como ALS) descrita em Sathasiivan e outros (*Nucl. Acids Res.* 18:2188-2193, 1990) para tolerância a herbicidas de sulfoniluréia; e o gene bar descrito em DeBlock e outros, (*EMBO J.* 6:2513-2519, 1987) para tolerância ao glufosinato e ao bialaphos.

A regeneração, o desenvolvimento e o cultivo de plantas partindo de vários explantes são bem documentados na técnica. Este processo de regeneração e crescimento inclui tipicamente as etapas de seleção de células transformadas e de cultivo de células individualizadas através dos estágios usuais de desenvolvimento embrionário até o estágio de planta jovem enraizada. Os embriões e as sementes transgênicos são similarmente regenerados. Os brotos enraizados transgênicos resultantes são depois disso plantados em um meio de cultivo de planta apropriado tal como solo. As células que sobrevivem à exposição ao agente seletivo ou as células que tiveram classificação positiva em um ensaio de verificação, podem ser cultivadas em meio que sustenta a regeneração de plantas. As plantas jovens em desenvolvimento são transferidas para o solo menos a mistura de crescimento de planta e estabilizadas, antes da transferência para uma estufa ou câmara de crescimento para maturação.

A presente invenção pode ser utilizada com qualquer célula ou tecido que pode ser transformado. Como utilizado aqui, o que pode ser transformado é entendido como uma célula ou um tecido capaz de se propagar adicionalmente para dar origem a uma planta. Os versados na técnica reconhecem que pode ser transformado um número de células ou tecidos vegetais em que após a inserção do DNA exógeno e de condições de cultura apropriadas as células ou os tecidos vegetais podem ser formar em uma planta diferenciada. O tecido adequado para estas finalidades pode incluir, mas não está limitado a embriões imaturos, tecido escutelar, culturas de células em suspensão, inflorescência imatura, meristema de broto, explantes nodais, tecido de calo, tecido hipocotiledôneo, cotilédones, raízes e folhas.

Qualquer meio de cultura de planta adequado pode ser utilizado. Os exemplos de meios adequados incluiriam, mas não estão limitados a meios à base de MS (Murashige e Skoog, *Physiol. Plant*, 15:473-497, 1962) ou a meios à base de N6 (Chu e outros, *Scientia Sinica* 18:659, 1975) suplementados com reguladores de crescimento vegetal adicionais incluindo, mas não limitados a auxinas, citocininas, ABA e giberelinas. Os versados na técnica estão familiarizados com a variedade de meios de cultura de tecidos,

que quando suplementados apropriadamente, sustentam o crescimento e o desenvolvimento do tecido vegetal e são adequados para a transformação e para a regeneração de plantas. Estes meios de cultura de tecidos podem ser obtidos na forma de uma preparação comercial ou preparados e modificados pelo usuário. Os versados na técnica estão a par dos meios e dos suplementos de meios tais como nutrientes e reguladores do crescimento para uso na transformação e na regeneração e outras condições de cultura tais como intensidade da luz durante a incubação, pH e temperaturas de incubação que podem ser otimizadas para a variedade particular de interesse.

Um versado na técnica considerará que, após um cassete de expressão ser incorporado de forma estável em plantas transgênicas e confirmado como sendo operacional, este pode ser introduzido em outras plantas através de cruzamento sexual. Qualquer uma de um número de técnicas de cruzamento padronizadas pode ser utilizada, dependendo da espécie que será cruzada.

Identificação de Plantas com um Fenótipo Alterado

Uma varredura de marcação por ativação com *Arabidopsis* (ACTTAG) foi utilizada para identificar a associação entre 1) linhagens de plantas ACTTAG com um conteúdo alterado de óleos, proteínas e/ou fibras (ver as colunas 4, 5 e 6 respectivamente, da Tabela 1, a seguir) e 2) as sequências de ácidos nucleicos identificadas na coluna 3 das Tabelas 2 e 3, em que cada sequência de ácido nucleico é fornecida dentro de um pseudônimo de gene ou uma denominação de HIO (HIO#; ver a coluna 1 nas Tabelas 1, 2 e 3). A denominação de HIO é arbitrária e não se refere necessariamente a uma planta que possui um fenótipo de alto conteúdo de óleos (HIO).

Sucintamente e como descrito adicionalmente nos Exemplos, um grande número de plantas de *Arabidopsis* foi mutado com o vetor pS-KI015, que compreende um T-DNA do plasmídeo Ti de *Agrobacterium tumefaciens*, um elemento intensificador viral e um gene marcador que pode ser selecionado (Weigel e outros, 2000, *Plant Physiology*, 122:1003-1013). Quando os insertos de T-DNA no genoma das plantas transformadas, o elemento intensificador pode causar a regulação para mais dos genes na vizi-

nhança, geralmente dentro de aproximadamente nove kilobases (kb) dos intensificadores. As plantas T1 foram expostas ao agente seletivo com a finalidade de recuperar especificamente as plantas transformadas que expressavam o marcador que pode ser selecionado e, portanto, carregavam as inserções do T-DNA. Foi permitido que as plantas T1 crescessem até a maturidade, se autofertilizassem e produzissem semente. A semente T2 foi colhida, rotulada e armazenada. Para a amplificação dos estoques de sementes, aproximadamente dezoito T2 foram plantadas no solo e, após a germinação, expostas ao agente seletivo para recuperar as plantas T2 transformadas. A semente T3 destas plantas foi colhida e agrupada. O conteúdo de óleos, proteínas e fibras da semente foi estimado utilizando Near Infrared Spectroscopy (NIR) como descrito nos Exemplos.

A associação de uma sequência de ácido nucleico de HIO com um fenótipo alterado foi descoberta através da análise da sequência de DNA genômico flanqueando a inserção do T-DNA na linhagem ACTTAG identificada na coluna 3 da Tabela 1. Uma linhagem ACTTAG é uma família de plantas derivada de uma única planta que foi transformada com um elemento de T-DNA contendo quatro cópias em tandem dos intensificadores do 35S do CaMV. Consequentemente, as sequências de ácidos nucleicos e/ou de polipeptídeos de HIO descritas podem ser empregadas no desenvolvimento de plantas transgênicas que possuem um fenótipo alterado, por exemplo, alto conteúdo de óleos. As sequências de ácidos nucleicos de HIO podem ser utilizadas na produção de plantas transgênicas, tais como plantas de cultivo de sementes oleosas, que fornecem maior rendimento de óleo partindo do processamento das sementes oleosas e resultam em um aumento na quantidade de óleo recuperado das sementes da planta transgênica. As sequências de ácidos nucleicos de HIO também podem ser utilizadas na produção de plantas transgênicas, tais como plantas de cultivo de grãos para produto alimentício, que fornecem um fenótipo alterado resultando em maior energia para a alimentação de animais, por exemplo, sementes ou farinha da semente com um conteúdo alterado de proteínas e/ou de fibras, resultando em um aumento na AME. As sequências de ácidos nucleicos de HIO po-

dem ser adicionalmente utilizadas para aumentar o conteúdo de óleos de plantas de cultivo oleosas especiais, com a finalidade de aumentar o rendimento e/ou a recuperação de ácidos graxos não usuais desejados. Os exemplos não limitantes específicos de ácidos graxos não usuais são ácido ricinoleico, ácido vernólico e os ácidos graxos poli-insaturados de cadeia muito longa ácido docosahexaenoico (DHA) e ácido eicosapentaenoico (E-PA). As plantas transgênicas que foram modificadas geneticamente para expressarem polipeptídeos de HIO podem ser utilizadas na produção de sementes, em que as plantas transgênicas são crescidas e o óleo e a farinha da semente são obtidos de partes da planta (por exemplo, semente) utilizando métodos padronizados.

Ácidos Nucleicos e Polipeptídeos de HIO

A denominação de HIO para cada uma das sequências de ácidos nucleicos de HIO descobertas na verificação com a marcação por ativação descrita aqui estão listadas na coluna 1 das Tabelas 1-3, a seguir. Os polipeptídeos de HIO descritos estão listados na coluna 4 das Tabelas 2 e 3, a seguir. A denominação de HIO é arbitrária e não se refere necessariamente a uma planta que possui um fenótipo de alto conteúdo de óleos (HIO). Como utilizado aqui, o pseudônimo de gene ou a denominação de HIO refere-se a qualquer sequência de polipeptídeo (ou à sequência de ácido nucleico que codifica a mesma) que quando expressa em uma planta causa um fenótipo alterado em qualquer parte da planta, por exemplo, nas sementes. Em uma modalidade, um polipeptídeo de HIO refere-se a uma proteína HIO de comprimento completo ou a um fragmento, um derivado (variação) ou um ortólogo ou um parálogo da mesma que é "funcionalmente ativo", de forma que o fragmento, o derivado ou o ortólogo ou o parálogo da proteína exibe uma ou mais ou as atividades funcionais associadas com um ou mais dos polipeptídeos de HIO de comprimento completo descritos, por exemplo, as sequências de aminoácidos fornecidas na entrada do GenBank referida na coluna 4 das Tabelas 2 e 3 que correspondem às sequências de aminoácidos apresentadas como a SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID

NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32 ou SEQ ID NO: 34 ou um ortólogo ou um parálogo da mesma. Em uma modalidade preferida, um polipeptídeo de HIO funcionalmente ativo causa um fenótipo alterado em uma planta transgênica. Em uma outra modalidade, um polipeptídeo de HIO funcionalmente ativo causa um fenótipo de conteúdo alterado de óleos, proteínas e/ou fibras (por exemplo, um fenótipo de conteúdo alterado da farinha da semente) quando expresso de forma errada em uma planta. Em outras modalidades preferidas, a expressão errada do polipeptídeo de HIO causa um fenótipo de alto conteúdo de óleos (tal como, mais óleo), de alto conteúdo de proteínas (tal como, maior conteúdo de proteína total ou de proteínas que podem ser digeridas) e/ou de baixo conteúdo de fibras (tal como, menos fibra) em uma planta. Ainda em outras modalidades preferidas, a expressão errada do polipeptídeo de HIO causa fenótipo inalterado de óleos, de alto conteúdo de proteínas (tal como, maior conteúdo de proteína total ou de proteínas que podem ser digeridas) e/ou de baixo conteúdo de fibras (tal como, menos fibra) em uma planta. Em uma outra modalidade, a expressão errada do polipeptídeo de HIO causa uma maior AME da farinha. Ainda em uma outra modalidade, um polipeptídeo de HIO funcionalmente ativo pode recuperar a atividade defeituosa do polipeptídeo de HIO endógeno (incluindo deficiente) quando expresso em uma planta ou em células vegetais; o polipeptídeo de recuperação pode ser proveniente da mesma ou de uma espécie diferente da espécie com a atividade defeituosa do polipeptídeo. A descrição fornece ainda produto alimentício, farinha, grão, alimento ou semente que compreende o polipeptídeo de HIO ou um fragmento, um derivado (variação) ou um ortólogo ou um parálogo do mesmo.

Em uma outra modalidade, um fragmento funcionalmente ativo de um polipeptídeo de HIO de comprimento completo (por exemplo, um fragmento funcionalmente ativo de um polipeptídeo nativo que possui a sequência de aminoácidos apresentada como a SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ

ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32 ou SEQ ID NO: 34 ou um ortólogo ou um parálogo que ocorre naturalmente da mesma) mantém uma ou mais das propriedades biológicas associadas com o polipeptídeo de HIO de comprimento completo, tal como atividade de

5 sinalização, atividade de ligação, atividade catalítica ou atividade de localização celular ou extracelular. Um fragmento de HIO preferencialmente compreende um domínio de HIO, tal como um domínio C- ou N-terminal ou catalítico, entre outros e compreende preferencialmente pelo menos 10, preferencialmente pelo menos 20, mais preferencialmente pelo menos 25 e mais

10 preferencialmente pelo menos 50 aminoácidos contíguos de uma proteína HIO. Os domínios funcionais dos genes HIO estão listados na coluna 6 da Tabela 2 e podem ser identificados utilizando o programa PFAM (Bateman A e outros, 1999, *Nucleic Acids Res.* 27:260-262) ou o programa INTERPRO (Mulder e outros, 2003, *Nucleic Acids Res.* 31, 315-318). As variações fun-

15 cionalmente ativas de polipeptídeos de HIO de comprimento completo ou fragmentos dos mesmos, incluem polipeptídeos com inserções, deleções ou substituições de aminoácidos que mantêm uma ou mais das propriedades biológicas associadas com o polipeptídeo de HIO de comprimento completo. Em alguns casos, são geradas variações que alteram o processamento pós-

20 tradução de um polipeptídeo de HIO. Por exemplo, as variações podem ter características alteradas de transporte de proteínas ou de localização de proteínas ou meia-vida alterada de proteínas, comparadas com o polipeptídeo nativo.

Como utilizado aqui, o termo "ácido nucleico de HIO" refere-se a

25 qualquer polinucleotídeo que quando expresso em uma planta causa um fenótipo alterado em qualquer parte da planta, por exemplo, nas sementes. Em uma modalidade, um polinucleotídeo de HIO abrange ácidos nucleicos com a sequência fornecida na ou complementar à entrada do GenBank referida na coluna 3 das Tabelas 2 e 3, que correspondem às sequências de

30 ácidos nucleicos apresentadas como SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO:

23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31 ou SEQ ID NO: 33, assim como fragmentos, derivados ou ortólogos ou parálogos funcionalmente ativos das mesmas. Um ácido nucleico de HIO desta descrição pode ser DNA, derivado de DNA genômico ou cDNA ou RNA.

5 Em uma modalidade, um ácido nucleico de HIO funcionalmente ativo codifica ou é complementar a um ácido nucleico que codifica um polipeptídeo de HIO funcionalmente ativo. Um ácido nucleico de HIO funcionalmente ativo também inclui DNA genômico que serve como um molde para um produto da transcrição primária de RNA (isto é, um precursor de mRNA)

10 que requer processamento, tal como "splicing", antes de codificar o polipeptídeo de HIO funcionalmente ativo. Um ácido nucleico de HIO pode incluir outras sequências não codificadoras, que podem ou não ser transcritas; tais sequências incluem 5' e 3' UTRs, sinais de poliadenilação e sequências reguladoras que controlam a expressão gênica, entre outros, que são conhecidos na técnica. Alguns polipeptídeos requerem eventos de processamento, tais como clivagem proteolítica, modificação covalente etc., com a finalidade de se tornarem completamente ativos. Consequentemente, os ácidos nucleicos funcionalmente ativos podem codificar o polipeptídeo de HIO maduro ou pré-processado ou uma forma intermediária. Um polinucleotídeo de HIO

20 também pode incluir sequências codificadoras heterólogas, por exemplo, sequências que codificam um marcador incluído para facilitar a purificação do polipeptídeo fundido ou um marcador de transformação. Em uma outra modalidade, um ácido nucleico de HIO funcionalmente ativo é capaz de ser utilizado na produção de fenótipos de HIO de perda de função, por exemplo,

25 através da supressão antissenso, da co-supressão, etc. A descrição fornece ainda produto alimentício, farinha, grão, alimento ou semente que compreende uma sequência de ácido nucleico que codifica um polipeptídeo de HIO.

 Em uma modalidade preferida, um ácido nucleico de HIO utilizado nos métodos descritos compreende uma sequência de ácido nucleico que

30 codifica ou é complementar a uma sequência que codifica, um polipeptídeo de HIO que possui pelo menos 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% ou 99% de identidade de sequência em relação a uma sequência

de polipeptídeo de HIO descrita, por exemplo, a sequência de aminoácidos apresentada como a SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32 ou SEQ ID NO: 34.

Em uma outra modalidade, um polipeptídeo de HIO compreende uma sequência de polipeptídeo com pelo menos 50% ou 60% de identidade em relação a uma sequência de polipeptídeo de HIO descrita (por exemplo, a sequência de aminoácidos apresentada como a SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32 ou SEQ ID NO: 34) e pode ter pelo menos 70%, 80%, 85%, 90%, 95%, 97%, 98% ou 99% de identidade de sequência em relação a uma sequência de polipeptídeo de HIO descrita. Em uma modalidade adicional, um polipeptídeo de HIO compreende 50%, 60%, 70%, 80%, 85%, 90%, 95%, 97%, 98% ou 99% de identidade de sequência em relação a uma sequência de polipeptídeo de HIO descrita e pode incluir um domínio de proteína conservado do polipeptídeo de HIO (tal como o(s) domínio(s) listado(s) na coluna 6 da Tabela 2). Em uma outra modalidade, um polipeptídeo de HIO compreende uma sequência de polipeptídeo com pelo menos 50%, 60%, 70%, 80%, 85%, 90%, 95%, 97%, 98% ou 99% de identidade de sequência em relação a um fragmento funcionalmente ativo do polipeptídeo referido na coluna 4 da Tabela 2. Ainda em uma outra modalidade, um polipeptídeo de HIO compreende uma sequência de polipeptídeo com pelo menos 50%, 60%, 70%, 80%, 90%, 95%, 97%, 98% ou 99% de identidade em relação à sequência de polipeptídeo da entrada do GenBank referida na coluna 4 da Tabela 2 ao longo de todo seu comprimento e compreende um(ou mais) domínio(s) de proteína conservado(s) listado(s) na coluna 6 da Tabela 2.

Em um outro aspecto, uma sequência de polinucleotídeo de HIO é pelo menos 50% até 60% idêntica ao longo de seu comprimento inteiro a uma sequência de ácido nucleico de HIO descrita, tal como a sequência de

ácido nucleico apresentada como SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31 ou

5 SEQ ID NO: 33 ou as sequências de ácidos nucleicos que são complementares a tal sequência de HIO e pode compreender pelo menos 70%, 80%, 85%, 90%, 95%, 97%, 98% ou 99% de identidade de sequência em relação à sequência de HIO descrita ou a um fragmento funcionalmente ativo da mesma ou sequências complementares. Em uma outra modalidade, um áci-

10 do nucleico de HIO descrito compreende uma sequência de ácido nucleico que é mostrada na SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31 ou SEQ ID NO: 33

15 ou sequências de ácidos nucleicos que são complementares a tal sequência de HIO e sequências de ácidos nucleicos que possuem homologia de sequência substancial a tais sequências de HIO. Como utilizado aqui, a expressão "homologia de sequência substancial" refere-se às sequências de ácidos nucleicos que possuem variações de sequência pequenas ou sem

20 importância partindo de tais sequências de HIO, isto é, as sequências funcionam substancialmente da mesma maneira e codificam um polipeptídeo de HIO.

Como utilizado aqui, "porcentagem (%)" de identidade de sequência em relação a uma sequência de objetivo especificada ou a uma

25 parte especificada da mesma, é definida como a porcentagem de nucleotídeos ou de aminoácidos em uma sequência identificada idêntica aos nucleotídeos ou aos aminoácidos na sequência de objetivo (ou na parte especificada da mesma), após o alinhamento das sequências e da introdução de espaços, se necessário, para atingir uma porcentagem de identidade de sequência máxima, que é gerada através do programa WU-BLAST-2,0 a 19

30 (Altschul e outros, *J. Mol. Biol.*, 1990, 215:403-410) com parâmetros de busca ajustados nos valores predeterminados. Os parâmetros HSP S e HSP S2

são valores dinâmicos e são estabelecidos pelo próprio programa dependendo da composição da sequência particular e da composição do banco de dados particular contra o qual a sequência de interesse será pesquisada. Um "valor de identidade percentual (%)" é determinado pelo número de

5 combinações de nucleotídeos ou aminoácidos idênticos dividido pelo comprimento da sequência para a qual a porcentagem de identidade está sendo relatada. A "porcentagem (%) de similaridade de sequência de aminoácidos" é determinada realizando o mesmo cálculo que para a determinação da %

10 de identidade de sequência de aminoácidos, mas incluindo substituições conservativas de aminoácidos em adição a aminoácidos idênticos no cálculo. Uma substituição conservativa de aminoácido é uma em que um aminoácido é substituído por um outro aminoácido que possui propriedades similares de forma que o dobramento ou a atividade da proteína não seja significativamente afetada. Os aminoácidos aromáticos que podem ser substituídos

15 por cada outro são fenilalanina, triptofano e tirosina; os aminoácidos hidrofóbicos intercambiáveis são leucina, isoleucina, metionina e valina; os aminoácidos polares intercambiáveis são glutamina e asparagina; e os aminoácidos básicos intercambiáveis são arginina, lisina e histidina; os aminoácidos ácidos intercambiáveis são ácido aspártico e ácido glutâmico; e os aminoácidos

20 pequenos intercambiáveis são alanina, serina, treonina, cisteína e glicina.

Moléculas de ácidos nucleicos derivadas das moléculas de ácidos nucleicos de objetivo incluem sequências que se hibridizam seletivamente às sequências de ácidos nucleicos de HIO descritas (por exemplo, a sequência de ácido nucleico apresentada como a SEQ ID NO: 1, SEQ ID

25 NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31 ou SEQ ID NO: 33). A estringência da hibridização pode ser controlada pela temperatura, pela força iônica, pelo pH e pela presença de

30 agentes desnaturantes tais como formamida durante a hibridização e a lavagem. As condições utilizadas rotineiramente são bem-conhecidas (ver, por exemplo, *Current Protocol in Molecular Biology*, Vol. 1, Cap. 2,10, John Wi-

ley & Sons, Publishers (1994); Sambrook e outros, 1989, *Molecular Cloning: A Laboratory Manual* (Segunda Edição), Cold Spring Harbor Press, Plainville, N.Y.).

Em algumas modalidades, uma molécula de ácido nucleico da

5 descrição é capaz de se hibridizar com uma molécula de ácido nucleico que contém a sequência de nucleotídeos descrita sob condições de hibridização estringentes que são: pré-hibridização de filtros contendo ácido nucleico durante 8 horas até durante a noite a 65°C em uma solução que compreende 6X citrato de concentração simples (SSC) (1X SSC é 0,15 M de NaCl, 0,015

10 M de de citrato de Na; pH 7,0), 5X solução de Denhardt, 0,05% de pirofosfato de sódio e 100 µg/mL de DNA de esperma de arenque; hibridização durante 18-20 horas a 65° C em uma solução contendo 6X SSC, 1X solução de Denhardt, 100 µg/mL de tRNA de levedura e 0,05% de pirofosfato de sódio; e lavagem of filtros a 65° C durante 1 h em uma solução contendo 0,1X SSC

15 e 0,1% de SDS (dodecil sulfato de sódio). Em outras modalidades, são utilizadas condições de hibridização de estringência moderada que são: pré-tratamento dos filtros contendo ácido nucleico durante 6 h a 40° C em uma solução contendo 35% de formamida, 5X SSC, 50 mM de Tris-HCl (pH 7,5), 5 mM de EDTA, 0,1% de PVP, 0,1% de Ficoll, 1% de BSA e 500 µg/mL de

20 DNA de esperma de salmão desnaturado; hibridização durante 18-20 h a 40° C em uma solução contendo 35% de formamida, 5X SSC, 50 mM de Tris-HCl (pH 7,5), 5 mM de EDTA, 0,02% de PVP, 0,02% de Ficoll, 0,2% de BSA, 100 µg/mL de DNA de esperma de salmão e 10% (p/vol) de sulfato de dextran; seguida pela lavagem duas vezes durante 1 hora a 55° C em uma

25 solução contendo 2X SSC e 0,1% de SDS. Alternativamente, podem ser utilizadas condições de baixa estringência que compreendem: incubação durante 8 horas até durante a noite a 37° C em uma solução que compreende 20% de formamida, 5 x SSC, 50 mM de fosfato de sódio (pH 7,6), 5X solução de Denhardt, 10% de sulfato de dextran e 20 µg/mL de DNA de esperma

30 de salmão cisalhado desnaturado; hibridização no mesmo tampão durante 18 até 20 horas; e lavagem dos filtros em 1 x SSC a aproximadamente 37° C durante 1 hora.

Como um resultado da degeneração do código genético, pode ser produzido um número de sequências de polinucleotídeos que codificam um polipeptídeo de HIO. Por exemplo, os códons podem ser selecionados para aumentar a taxa em que a expressão do polipeptídeo ocorre em uma espécie hospedeira particular, de acordo com o uso ótimo de códons determinado pelo organismo hospedeiro particular (ver, por exemplo, Nakamura e outros, 1999, *Nucleic Acids Res.* 27:292). Tais variações de sequências podem ser utilizadas nos métodos descritos aqui.

Os métodos descritos podem utilizar ortólogos (e/ou parálogos) de uma sequência de ácido nucleico de HIO de *Arabidopsis* descrita. Os supostos ortólogos (e/ou parálogos) representantes de cada um dos genes HIO de *Arabidopsis* descritos são identificados na coluna 5 da Tabela 3, a seguir. Os métodos de identificação dos ortólogos em outras espécies de plantas são conhecidos na técnica. Em geral, os ortólogos em espécies diferentes mantêm a mesma função, devido à presença de um ou mais motivos de proteína e/ou estruturas tridimensionais. Na evolução, quando um evento de duplicação gênica sucede a especiação, um único gene em uma espécie, tal como *Arabidopsis*, pode corresponder a vários genes (parálogos) em uma outra. Quando os dados das sequências estão disponíveis para uma espécie de planta particular, os ortólogos são geralmente identificados através da análise de homologia de sequência, tal como a análise de BLAST, geralmente utilizando sequências de isca de proteínas. As sequências são designadas como um ortólogo potencial se a melhor coincidência de sequência partindo do resultado de BLAST para frente recuperar a sequência de busca original no BLAST inverso (Huynen MA e Bork P, 1998, *Proc. Natl. Acad. Sci.*, 95:5849-5856; Huynen MA e outros, 2000, *Genome Research*, 10:1204-1210).

Programas para alinhamento de várias sequências, tal como CLUSTAL (Thompson JD e outros, 1994, *Nucleic Acids Res.* 22:4673-4680) podem ser utilizados para destacar regiões e/ou resíduos conservados de proteínas homólogas (ortólogas e/ou parálogas) e para gerar árvores filogenéticas. Em uma árvore filogenética que representa várias sequências ho-

mólogas provenientes de espécies distintas (por exemplo, recuperadas através da análise de BLAST), as sequências ortólogas de duas espécies geralmente aparecem mais próximas na árvore em relação a todas as outras sequências destas duas espécies. O "rosqueamento" estrutural ou outras

5 análises de dobramento das proteínas (por exemplo, utilizando o software da ProCeryon, Biosciences, Salzburg, Áustria) podem também identificar ortólogos potenciais. Os métodos de hibridização de ácidos nucleicos também podem ser utilizados para encontrar genes ortólogos e são preferidos quando os dados das sequências não estão disponíveis. A PCR degenerada e a

10 verificação de bibliotecas de cDNA ou de DNA genômico são métodos comuns para encontrar sequências gênicas relacionadas e são bem-conhecidas na técnica (ver, por exemplo, Sambrook, 1989, *Molecular Cloning: A Laboratory Manual* (Segunda Edição), Cold Spring Harbor Press, Plainview, N.Y.; Dieffenbach e Dveksler, 1995, *PCR Primer: A Laboratory*

15 *Manual*, Cold Spring Harbor Laboratory Press, NY). Por exemplo, os métodos para gerar uma biblioteca de cDNA partindo da espécie de planta de interesse e submetendo a biblioteca a sondas com sondas de genes parcialmente homólogos são descritos em Sambrook e outros. Uma parte altamente conservada da sequência codificadora de HIO de *Arabidopsis* pode

20 ser utilizada como uma sonda. Os ácidos nucleicos ortólogos de HIO podem se hibridizar ao ácido nucleico da SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31 ou SEQ ID

25 NO: 33, sob condições de estringência alta, moderada ou baixa. Após a amplificação ou o isolamento de um segmento de um suposto ortólogo, tal segmento pode ser clonado e sequenciado através de técnicas padronizadas e utilizado como uma sonda para isolar um clone completo de cDNA ou de DNA genômico.

30 Alternativamente, é possível iniciar um projeto de EST para gerar um banco de dados de informação de sequência para a espécie de planta de interesse. Em uma outra abordagem, os anticorpos que se ligam espe-

cificamente a polipeptídeos de HIO conhecidos são utilizados para o isolamento de ortólogos (e/ou parálogos) (ver, por exemplo, Harlow e Lane, 1988, 1999, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Nova York). A análise de Western blot pode determinar se um ortólogo de HIO (isto é, um ortólogo de proteína para um polipeptídeo de HIO descrito) está presente em um extrato bruto de uma espécie de planta particular. Quando a reatividade é observada, a sequência que codifica o ortólogo candidato pode ser isolada através da verificação de bibliotecas de expressão que representam a espécie de planta particular. As bibliotecas de expressão podem ser construídas em uma variedade de vetores disponíveis comercialmente, incluindo lambda gt11, como descrito em Sambrook e outros, 1989. Uma vez que o(s) ortólogo(s) candidato(s) é(são) identificado(s) através de qualquer um destes meios, as sequências ortólogas candidatas são utilizadas como isca (a "busca") para o BLAST inverso contra sequências de *Ara-*
15 *bidopsis* ou de outras espécies em que as sequências de ácidos nucleicos de HIO e/ou de polipeptídeos foram identificadas.

Os ácidos nucleicos e os polipeptídeos de HIO podem ser obtidos através da utilização de qualquer método disponível. Por exemplo, as técnicas para o isolamento de sequências de cDNA ou de DNA genômico de interesse através da verificação de bibliotecas de DNA ou através da utilização da reação em cadeia da polimerase (PCR), como descrito anteriormente, são bem-conhecidas na técnica. Alternativamente, a sequência de ácido nucleico pode ser sintetizada. Qualquer método conhecido, tal como a mutagênese direcionada ao sítio (Kunkel TA e outros, 1991, *Methods Enzymol.*
20 204:125-39), pode ser utilizado para introduzir alterações desejadas em um ácido nucleico clonado.

Em geral, os métodos descritos aqui envolvem a incorporação da forma desejada do ácido nucleico de HIO em um vetor de expressão de planta para a transformação de células vegetais e o polipeptídeo de HIO é
30 expresso na planta hospedeira. As plantas e as células vegetais transformadas que expressam um polipeptídeo de HIO expressam um fenótipo alterado e, em um exemplo não limitante específico, podem ter conteúdo alto (maior)

de óleos, alto (maior) de proteínas e/ou baixo (menor) de fibras.

Uma molécula de ácido nucleico de HIO "isolada" é outra que na forma ou no ambiente em que é encontrada na natureza e é identificada e separada de pelo menos uma molécula de ácido nucleico contaminante com a qual está comumente associada na fonte natural do ácido nucleico de HIO. Entretanto, uma molécula de ácido nucleico de HIO isolada inclui moléculas de ácidos nucleicos de HIO contidas nas células que expressam comumente o polipeptídeo de HIO quando, por exemplo, a molécula de ácido nucleico está em uma localização cromossômica diferente daquela das células naturais.

Produção de Plantas Modificadas Geneticamente com um Fenótipo Alterado

Os ácidos nucleicos e os polipeptídeos de HIO descritos podem ser utilizados na produção de plantas transgênicas que possuem um fenótipo modificado ou alterado, por exemplo, um fenótipo de conteúdo alterado de óleos, proteínas e/ou fibras. Como utilizado aqui, um "conteúdo (fenótipo) alterado de óleos" pode se referir ao conteúdo alterado de óleos em qualquer parte da planta. Em uma modalidade preferida, a expressão alterada do gene HIO em uma planta é utilizada para gerar plantas com um conteúdo (fenótipo) alto de óleos. Como utilizado aqui, um "conteúdo (fenótipo) alterado de proteínas totais" ou um "conteúdo (fenótipo) alterado de proteínas que podem ser digeridas" pode se referir ao conteúdo alterado de proteínas (totais ou que podem ser digeridas) em qualquer parte da planta. Em uma modalidade preferida, a expressão alterada do gene HIO em uma planta é utilizada para gerar plantas com um conteúdo (fenótipo) alto (ou maior) de proteínas totais ou que podem ser digeridas. Como utilizado aqui, um "conteúdo (fenótipo) alterado de fibras" pode se referir ao conteúdo alterado de fibras em qualquer parte da planta. Em uma modalidade preferida, a expressão alterada do gene HIO em uma planta é utilizada para gerar plantas com um conteúdo (fenótipo) baixo (ou menor) de fibras. O conteúdo alterado de óleos, proteínas e/ou fibras é frequentemente observado nas sementes. Os exemplos de uma planta transgênica incluem plantas que compreendem um vetor de transformação de planta com uma sequência de nucleotídeos que

codifica ou é complementar a uma sequência que codifica um polipeptídeo de HIO que possui a sequência de aminoácidos que é apresentada na SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32 ou SEQ ID NO: 34 ou um ortólogo ou um parálogo da mesma.

As plantas transgênicas, tais como milho, soja e canola que contêm as sequências de ácidos nucleicos descritas, podem ser utilizadas na produção de óleo e farinha vegetal. O óleo vegetal é utilizado em uma variedade de produtos alimentícios, enquanto que a farinha da semente é utilizada como um produto alimentício para animais. Após a colheita da semente das plantas transgênicas, a semente é limpa para remover os talos da planta e outro material e então fendidas em lascas em um moinhos de rolo para quebrar as cascas. A semente triturada é aquecida a 75-100° C para desnaturar as enzimas hidrolíticas, lisar as células não rompidas que contêm óleo e permitir que as pequenas gotículas de óleo se unam. A maior parte do óleo é então removida (e pode ser recuperada) através da pressão do material da semente em uma prensa de rosca. O restante do óleo é removido da torta da prensa através da extração com solventes orgânicos, tal como hexano. O solvente é removido da farinha através do aquecimento da mesma a aproximadamente 100° C. Após a secagem, a farinha é então granulada em uma forma consistente. A farinha, contendo a proteína, o carboidrato que pode ser digerido e a fibra da semente, pode ser misturada com outros materiais antes de ser utilizada como um produto alimentício para animais.

Os métodos descritos aqui para a produção de plantas transgênicas podem ser geralmente aplicados a todas as plantas. Embora a marcação por ativação e a identificação gênica sejam realizadas em *Arabidopsis*, a sequência de ácido nucleico de HIO (ou um ortólogo, um parálogo, uma variação ou um fragmento da mesma) pode ser expressa em qualquer tipo de planta. Em uma modalidade preferida, as plantas que produzem óleo produzem e armazenam triacilglicerol em órgãos específicos, primariamente nas

sementes. Tais espécies incluem soja (*Glicine max*), semente de colza e canola (incluindo *Brassica napus*, *B. campestris*), girassol (*Helianthus annuus*), algodão (*Gossypium hirsutum*), milho (*Zea mays*), cacau (*Theobroma cacao*), açafrão (*Carthamus tinctorius*), dendezeiro (*Elaeis guineensis*), coqueiro (*Cocos nucifera*), linho (*Linum usitatissimum*), mamona (*Ricinus communis*) e amendoim (*Arachis hypogaea*), assim como trigo, arroz e aveia. As plantas que carregam frutos e verduras, as plantas que produzem grãos, as plantas que produzem nozes, as espécies de *Brassica* de ciclo rápido, alfalfa (*Medicago sativa*), tabaco (*Nicotiana*), turfa (família Poaceae), outras plantas de cultivo de forragem e espécies selvagens também podem ser uma fonte de ácidos graxos exclusivos. Em outras modalidades, qualquer planta que expressa a sequência de ácido nucleico de HIO também pode expressar conteúdo maior de proteínas e/ou menor de fibras em uma parte ou um órgão específico da planta, tal como nas sementes.

O versado na técnica reconhecerá que existe uma ampla variedade de técnicas de transformação na técnica e novas técnicas estão continuamente se tornando disponíveis. Qualquer técnica que seja adequada para a planta hospedeira alvo pode ser empregada dentro do âmbito da presente invenção. Por exemplo, os constructos podem ser introduzidos em uma variedade de formas incluindo, mas não limitadas a, na forma de um filamento de DNA, em um plasmídeo ou em um cromossomo artificial. A introdução dos constructos nas células vegetais alvos pode ser realizada através de uma variedade de técnicas, incluindo, mas não limitadas a transformação mediada por *Agrobacterium*, eletroporação, microinjeção, bombardeio de microprojéteis, co-precipitação de DNA com fosfato de cálcio ou transformação mediada por lipossomos de um ácido nucleico heterólogo. A transformação da planta é preferencialmente permanente, isto é, através da integração dos constructos de expressão introduzidas no genoma da planta hospedeira, de forma que os constructos introduzidos são passados para gerações de plantas sucessivas. Dependendo do uso pretendido, um constructo de ácido nucleico heterólogo que compreende um polinucleotídeo de HIO pode codificar a proteína inteira ou uma parte biologicamente ativa da

mesma.

Em uma modalidade, sistemas de vetores binários baseados em Ti podem ser utilizados para transferir polinucleotídeos. Vetores binários padronizados de *Agrobacterium* são conhecidos pelos versados na técnica e muitos estão disponíveis comercialmente (por exemplo, pBI121 Clontech Laboratories, Palo Alto, CA). Um constructo ou um vetor pode incluir um promotor de planta para expressar a molécula de ácido nucleico de escolha. Em uma modalidade preferida, o promotor é um promotor de planta.

O procedimento ótimo para a transformação de plantas com vetores de *Agrobacterium* variará com o tipo de planta que será transformado. Os exemplos de métodos para a transformação mediada por *Agrobacterium* incluem a transformação de explantes hipocotiledôneos, de extremidade do broto, de tecido do caule ou foliar, derivado de mudas e/ou plantas jovens esterilizadas. Tais plantas transformadas podem ser reproduzidas sexualmente ou através da cultura de células ou de tecido. A transformação com *Agrobacterium* foi descrita anteriormente para um grande número de tipos diferentes de plantas e os métodos para tal transformação podem ser encontrados na literatura científica. São de relevância particular os métodos para a transformação de plantas de cultivo importantes comercialmente, tais como plantas das espécies de *Brassica*, incluindo canola e semente de colza, (De Block e outros, 1989, *Plant Physiol.*, 91:694-701), milho (Ishida e outros, 1996 *Nature Biotechnol.* 14:745-750, Zhang e outros, 2002 *Plante Cell Rep.* 21:263-270), girassol (Everett e outros, 1987, *Bio/Technology*, 5:1201), soja (Christou e outros, 1989, *Proc. Natl. Acad. Sci USA*, 86:7500-7504; Kline e outros, 1987, *Nature*, 327:70), trigo, arroz e aveia.

A expressão (incluindo a transcrição e a tradução) de uma sequência de ácido nucleico de HIO pode ser regulada em relação ao nível de expressão, ao(s) tipo(s) de tecido(s) onde a expressão ocorre e/ou ao estágio do desenvolvimento da expressão. Está disponível um número de sequências reguladoras heterólogas (por exemplo, promotores e intensificadores) para o controle da expressão de um ácido nucleico de HIO. Estas incluem promotores constitutivos, que podem ser induzidos e que podem ser re-

regulados, assim como promotores e intensificadores que controlam a expressão de uma maneira específica ao tecido ou ao tempo. Os exemplos de promotores constitutivos incluem o promotor E4 da framboesa (Patentes U.S. N^{os} 5.783.393 e 5.783.394), o promotor da nopalina sintase (NOS) (E-
 5 bert e outros, *Proc. Natl. Acad. Sci. (U.S.A.)* 84:5745-5749, 1987), o promotor da octopina sintase (OCS) (que é carregado nos plasmídeos indutores de tumor de *Agrobacterium tumefaciens*), os promotores de caulimovírus tais como o promotor 19S do vírus mosaico da couve-flor (CaMV) (Lawton e outros, *Plant Mol. Biol.* 9:315-324, 1987) e o promotor 35S do CaMV (Odell e
 10 outros, *Nature* 313:810-812, 1985 e Jones JD e outros, 1992, *Transgenic Res.*, 1:285-297), o promotor 35S do vírus mosaico da escrofulária (Patente U.S. N^o 5.378.619), o promotor que pode ser induzido pela luz da subunidade pequena da ribulose-1,5-bis-fosfato carboxilase (ssRUBISCO), o promotor Adh (Walker e outros, *Proc. Natl. Acad. Sci. (U.S.A.)* 84:6624-6628,
 15 1987), o promotor da sacarose (Yang e outros, *Proc. Natl. Acad. Sci. (U.S.A.)* 87:4144-4148, 1990), o promotor do complexo do gene R (Chandler e outros, *The Plant Cell* 1:1175-1183, 1989), o promotor do gene da proteína de ligação à clorofila a/b, o promotor do CsVMV (Verdaguer B e outros, 1998, *Plant Mol Biol.*, 37:1055-1067) e o promotor da actina do melão (Pedi-
 20 do de Patente PCT publicado WO0056863). Os exemplos de promotores específicos ao tecido incluem os promotores E4 e E8 do tomate (Patente U.S. N^o 5.859.330) e o promotor do gene 2All do tomate (Van Haaren MJJ e outros, 1993, *Plant Mol Bio.*, 21:625-640).

Em uma modalidade preferida, a expressão da sequência de
 25 ácido nucleico de HIO está sob o controle de sequências reguladoras dos genes cuja expressão está associada com o desenvolvimento inicial da semente e/ou do embrião. Na verdade, em uma modalidade preferida, o promotor utilizado é um promotor intensificado na semente. Os exemplos de tais promotores incluem as regiões reguladoras a 5' de genes tais como da napina (Kridl e outros, *Seed Sci. Res.* 1:209:219, 1991), da globulina (Belanger e
 30 Kriz, *Genet.*, 129: 863-872, 1991, N^o de Acesso de GenBank L22295), da gama zeína Z 27 (Lopes e outros, *Mol Gen Genet.*, 247:603-613, 1995), do

promotor da L3 oleosina (Patente U.S. Nº 6,433,252), da faseolina (Bustos e outros, *Plant Cell*, 1(9):839-853, 1989), da arcelina5 (Pedido de Patente U.S. Nº 2003/0046727), de um promotor 7S da soja, de um promotor 7S α (Pedido de Patente U.S. Nº 2003/0093828), do promotor da 7S α ' beta conglucina da soja, de um promotor 7S α ' (Beachy e outros, *EMBO J.*, 4:3047, 1985; Schuller e outros, *Nucleic Acid Res.*, 10(24):8225-8244, 1982), inibidor da tripsina da soja (Riggs e outros, *Plant Cell* 1(6):609-621, 1989), ACP (Baerson e outros, *Plant Mol. Biol.*, 22(2):255-267, 1993), da esteroil-ACP dessaturase (Slocum e outros, *Plant Physiol.* 104(4):167-176, 1994), da subunidade α ' da β -conglucina da soja (Chen e outros, *Proc. Natl. Acad. Sci.* 83:8560-8564, 1986), de USP de *Vicia faba* (P-Vf.Usp, SEQ ID NO: 1, 2 e 3 no (Pedido de Patente U.S. Nº 2003/229918) e do promotor da L3 oleosina de *Zea mays* (Hong e outros, *Plant Mol. Biol.*, 34(3):549-555, 1997). São também incluídas as zeínas, que são um grupo de proteínas de armazenamento encontradas no endosperma do milho. Clones genômicos para os genes da zeína foram isolados (Pedersen e outros, *Cell*, 29:1015-1026, 1982; e Russell e outros, *Transgenic Res.* 6(2):157-168) e os promotores destes clones, incluindo 15 kD, 16 kD, 19 kD, 22 kD, 27 kD e os genes, podem ser utilizados. Outros promotores conhecidos como sendo funcionais, por exemplo, no milho incluem os promotores para os genes a seguir: *waxy*, *Brittle*, *Shrunken 2*, Enzimas de ramificação I e II, amido sintases, enzimas de desramificação, oleosinas, glutelinas e sacarose sintases. Os genes de verduras cujos promotores estão associados com o desenvolvimento inicial da semente e do embrião incluem *legumin* de *V. faba* (Baumlein e outros, 1991, *Mol. Gen. Genet.* 225:121-8; Baumlein e outros, 1992, *Plant J.* 2:233-9), *usp* de *V. faba* (Fiedler e outros, 1993, *Plant Mol. Biol.* 22:669-79), *convicilin* da ervilha (Bown e outros, 1988, *Biochem. J.* 251:717-26), *lectin* da ervilha (dePater e outros, 1993, *Plant Cell* 5:877-86), *beta phaseolin* de *P. vulgaris* (Bustos e outros, 1991, *EMBO J.* 10:1469-79), *DLEC2* e *PHS* de *P. vulgaris* [beta] (Bobb e outros, 1997, *Nucleic Acids Res.* 25:641-7) e *beta-Conglicinin* da soja, proteína 7S de armazenamento (Chamberland e outros, 1992, *Plant Mol. Biol.* 19:937-49).

Os genes de cereais cujos promotores estão associados com o desenvolvimento inicial da semente e do embrião incluem *glutelin* do arroz ("GluA-3", Yoshihara e Takaiwa, 1996, *Plant Cell Physiol.* 37:107-11; "GluB-1", Takaiwa e outros, 1996, *Plant Mol. Biol.* 30:1207-21; Washida e outros, 1999, *Plant Mol. Biol.* 40:1-12; "Gt3", Leisy e outros, 1990, *Plant Mol. Biol.* 14:41-50), *prolamin* do arroz (Zhou & Fan, 1993, *Transgenic Res.* 2:141-6), *prolamin* do trigo (Hammond-Kosack e outros, 1993, *EMBO J.* 12:545-54), *zein* do milho (Z4, Matzke e outros, 1990, *Plant Mol. Biol.* 14:323-32) e *B-hordeins* da cevada (Entwistle e outros, 1991, *Plant Mol. Biol.* 17:1217-31).

Outros genes cujos promotores estão associados com o desenvolvimento inicial da semente e do embrião incluem GLO7A do dendezeiro (7S globulina, Morcillo e outros, 2001, *Physiol. Plant* 112:233-243), *napin* de *Brassica napus*, proteína de armazenamento 2S e gene napA (Josefsson e outros, 1987, *J. Biol. Chem.* 262:12196-201; Stalberg e outros, 1993, *Plant Mol. Biol.* 1993 23:671-83; Ellerstrom e outros, 1996, *Plant Mol. Biol.* 32:1019-27), *oleosin* de *Brassica napus* (Keddie e outros, 1994, *Plant Mol. Biol.* 24:327-40), *oleosin* de *Arabidopsis* (Plant e outros, 1994, *Plant Mol. Biol.* 25:193-205), FAE1 de *Arabidopsis* (Rossak e outros, 2001, *Plant Mol. Biol.* 46:717-25), conA de *Canavalia gladiata* (Yamamoto e outros, 1995, *Plant Mol. Biol.* 27:729-41) e *estrictosidina* sintase de *Catharanthus roseus* (Str, Ouwerkerk e Memelink, 1999, *Mol. Gen. Genet.* 261:635-43). Em uma outra modalidade preferida, são utilizadas as sequências reguladoras de genes expressos durante a biossíntese de óleo (ver, por exemplo, a Patente U.S. Nº 5.952.544). Os promotores alternativos são de genes de proteínas de armazenamento de plantas (Bevan e outros, 1993, *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 342:209-15). Os promotores adicionais que podem ser utilizados são descritos, por exemplo, nas Patentes U.S. Nºs 5.378.619; 5.391.725; 5.428.147; 5.447.858; 5.608.144; 5.608.144; 5.614.399; 5.633.441; 5.633.435; e 4.633.436.

Em uma outra modalidade, o gene HIO endógeno pode ser colocado sob o controle de um fator de transcrição transgênico ou utilizado para planejar sítios de ligação que modulam sua expressão. Uma tal classe de

fatores de transcrição é constituída das proteínas de dedo de zinco Cys₂-His₂ (ZFPs). As ZFPs são proteínas de ligação ao DNA comuns e podem ser planejadas para se ligarem especificamente a sequências de DNA específicas (Beerli & Barbas, *Nat Biotechnol.*, 2002, 20:135-141,; Gommans e outros, *J Mol Biol.*, 2005, 354:507-519). Os domínios de dedo de zinco individuais são compostos de aproximadamente 30 aminoácidos, são estruturalmente conservados e podem interagir com 3-4 pb de DNA. Pode ser sintetizado um polipeptídeo que contém vários dedos de zinco planejados para se ligarem a uma sequência de DNA específica no promotor de um gene HIO.

Os princípios para o planejamento dos domínios de dedo de zinco para a interação com sequências de DNA específicas foram descritos em Segal e outros, (Segal e outros, *Proc Natl Acad Sci U S A.*, 1999, 96:2758-2763), Dreier e outros (Dreier e outros, *J Mol Biol.*, 2000, 303:489-502) e Beerli e Barbas (Beerli & Barbas, *Nat Biotechnol.*, 2002, 20:135-141). Estes domínios de ligação ao DNA podem ser fundidos a domínios efetores para formar uma ZFP sintética que pode regular a transcrição de genes aos quais se liga. Os domínios efetores que podem ativar a transcrição incluem, mas não estão limitados à parte ácida da proteína VP16 do vírus herpes simplex (Sadowski e outros, *Nature.*, 1988, 335:563-564) e da VP64 (Beerli e outros, *Proc Natl Acad Sci U S A.*, 1998, 95:14628-14633) e o domínio p65 do fator de transcrição do NF- κ B (Bae e outros, *Nat Biotechnol.*, 2003, 21:275-280, Liu e outros, *J Biol Chem.*, 2001, 276:11323-11334). Os domínios efetores que podem reprimir a transcrição incluem, mas não estão limitados a mSIN3 e KRAB (Ayer e outros, *Mol Cell Biol.*, 1996, 16:5772-5781, Beerli & Barbas, *Nat Biotechnol.*, 2002, 20:135-141, Beerli e outros, *Proc Natl Acad Sci U S A.*, 1998, 95:14628-14633, Margolin e outros, *Proc Natl Acad Sci U S A.*, 1994, 91:4509-4513). Foi mostrado que estas abordagens funcionam em plantas (Guan e outros, *Proc Natl Acad Sci U S A.*, 2002, 99:13296-13301, Stege e outros, *Plant J.*, 2002, 32:1077-1086, Van Eenennaam e outros, *Metab Eng.*, 2004, 6:101-108).

Ainda em um outro aspecto, em alguns casos pode ser desejável inibir a expressão da sequência de ácido nucleico de HIO endógena em

uma célula hospedeira. Os exemplos de métodos para a prática deste aspecto da invenção incluem, mas não estão limitados à supressão antissenso (Smith e outros, 1988, *Nature*, 334:724-726; van der Krol e outros, 1988, *BioTechniques*, 6:958-976); co-supressão (Napoli e outros, 1990, *Plant Cell*, 2:279-289); ribozimas (Publicação PCT WO 97/10328); e combinações de 5 senso e antissenso (Waterhouse e outros, 1998, *Proc. Natl. Acad. Sci. USA*, 95:13959-13964). Os métodos para a supressão de sequências endógenas em uma célula hospedeira empregam tipicamente a transcrição ou a trans-
 10 crição e a tradução de pelo menos uma parte da sequência que deve ser suprimida. Tais sequências podem ser homólogas às regiões codificadoras assim como não codificadoras da sequência endógena. A inibição antissenso pode utilizar a sequência de cDNA inteira (Sheehy e outros, 1988, *Proc. Natl. Acad. Sci. USA*, 85:8805-8809), uma sequência de cDNA parcial inclu-
 15 indo fragmentos da sequência codificadora a 5' (Cannon e outros, 1990, *Plant Mol. Biol.*, 15:39-47) ou sequências não codificadoras a 3' (Ch'ng e outros, 1989, *Proc. Natl. Acad. Sci. USA*, 86:10006-10010). As técnicas de co-supressão podem utilizar a sequência de cDNA inteira (Napoli e outros, 1990, *Plant Cell*, 2:279-289; van der Krol e outros, 1990, *Plant Cell*, 2:291-299) ou uma sequência de cDNA parcial (Smith e outros, 1990, *Mol. Gen.*
 20 *Genetics*, 224:477-481).

Os testes moleculares e genéticos padronizados podem ser realizados para analisar adicionalmente a associação entre uma sequência de ácido nucleico e um fenótipo observado. Os exemplos de técnicas são descritos a seguir.

25 1. Análise de DNA/RNA

Os padrões de expressão gênica específica ao estágio e ao tecido em linhagens mutantes versus do tipo selvagem podem ser determina-
 dos, por exemplo, através da hibridização *in situ*. A análise da condição de metilação do gene, especialmente das regiões reguladoras flanqueadoras,
 30 pode ser realizada. Outras técnicas adequadas incluem a superexpressão, a expressão ectópica, a expressão em outras espécies de plantas e o nocaute gênico (genética inversa, nocaute direcionado, silenciamento gênico induzi-

do por vírus (VIGS; ver, Baulcombe D, 1999, *Arch. Virol. Suppl.* 15:189-201).

Em uma aplicação preferida, a obtenção de um perfil de expressão, geralmente através da análise de microarranjos, é utilizada para medir simultaneamente diferenças ou alterações induzidas na expressão de vários genes diferentes. As técnicas para a análise de microarranjo são bem-conhecidas na técnica (Schena M e outros, *Science* 1995 270:467-470; Baldwin D e outros, 1999, *Cur. Opin. Plant Biol.* 2(2):96-103; Dangond F, *Physiol Genomics* (2000) 2:53-58; van Hal NL e outros, *J Biotechnol.* (2000) 78:271-280; Richmond T e Somerville S, *Curr. Opin. Plant Biol.* 2000 3:108-116). Pode ser realizada a obtenção do perfil de expressão de linhagens marcadas individuais. Tal análise pode identificar outros genes que são coordenadamente regulados como uma consequência da superexpressão do gene de interesse, que pode ajudar a colocar um gene desconhecido em uma via particular.

15 2. Análise dos Produtos Gênicos

A análise dos produtos gênicos pode incluir a expressão de proteínas recombinantes, a produção de antissoro, a imunolocalização, os ensaios bioquímicos para atividade catalítica ou outra, a análise da condição de fosforilação e a análise da interação com outras proteínas através de ensaios de duplo híbrido de levedura.

20 3. Análise das Vias

A análise das vias pode incluir a colocação de um gene ou de um produto gênico dentro de uma via bioquímica, metabólica ou de sinalização particular com base em seu fenótipo de expressão errada fenótipo ou através da homologia de sequência com genes relacionados. Alternativamente, a análise pode compreender cruzamentos genéticos com linhagens do tipo selvagem e outras linhagens mutantes (criando duplos mutantes) para ordenar o gene em uma via ou determinar o efeito de uma mutação sobre a expressão de genes "repórteres" a jusante em uma via.

30 Produção de Plantas Mutadas com um Fenótipo Alterado

São descritos aqui métodos adicionais de produção de uma planta que possui um fenótipo alterado, em que é identificada uma planta

que possui uma mutação ou um alelo em sua sequência de ácido nucleico de HIO que resulta em um fenótipo alterado, comparada a plantas que possuem a mutação ou o alelo. A planta mutada pode ser gerada utilizando um ou mais agentes mutagênicos, por exemplo, um agente mutagênico químico (tal como sulfonato de etilmetano, sulfonato de metil metano, dietilsulfato e nitrosoguanidina ou 5-bromo-deoxiuridina), radiação ou luz ultravioleta. Em algumas modalidades do método, a planta mutada pode ser cruzada para gerar progênie, que herda a mutação ou o alelo e possui um fenótipo alterado. Por exemplo, é fornecido aqui um método de identificação de plantas que possui uma ou mais mutações na sequência de ácido nucleico de HIO endógena que confere um fenótipo alterado e de produção de progênie destas plantas mutadas que possuem tal fenótipo que não é transgênica. As plantas mutadas com um fenótipo alterado podem ter um conteúdo alterado de óleos, proteínas e/ou fibras ou um conteúdo alterado de farinha da semente.

Em uma modalidade específica do método, chamado de "TILLING" (para **targeting induced local lesions in genomes** – direcionamento de lesões locais induzidas nos genomas), são induzidas mutações na semente de uma planta de interesse, por exemplo, utilizando tratamento com EMS (sulfonato de etilmetano) tratamento. As plantas resultantes são crescidas e autofertilizadas e a progênie é utilizada para preparar amostras de DNA. A amplificação através da PCR e o sequenciamento da sequência de ácido nucleico de HIO são utilizados para identificar se uma planta mutada possui uma mutação na sequência de ácido nucleico de HIO. As plantas que possuem mutações HIO podem ser então testadas em relação ao conteúdo alterado de óleos, proteínas e/ou fibras. Para confirmar se a mutação HIO causa o fenótipo modificado, podem ser realizados experimentos que correlacionam a presença do gene modificado e o fenótipo modificado através de cruzamentos genéticos. O TILLING pode identificar mutações que alteram a expressão de genes específicos ou a atividade de proteínas codificadas por estes genes (ver Colbert e outros, 2001, *Plant Physiol.* 126:480-484; McCallum e outros, 2000, *Nature Biotechnology* 18:455-457).

Em uma outra modalidade específica do método, uma abordagem de gene/Locus de Característica Quantitativa (QTLs) candidato pode ser utilizada em um programa de cruzamento auxiliado por marcadores para identificar alelos de ou mutações na sequência de ácido nucleico de HIO ou ortólogos (e/ou parálogos) da sequência de ácido nucleico de HIO que podem conferir conteúdo alterado de óleos, proteínas e/ou fibras (ver Bert e outros, *Theor Appl Genet.*, 2003 Jun;107(1):181-9; e Lionneton e outros, *Genome*, 2002 Dez;45(6):1203-15). Assim, em um aspecto adicional da descrição, um ácido nucleico de HIO é utilizado para identificar se uma planta que possui conteúdo alterado de óleos, proteínas e/ou fibras possui uma mutação em uma sequência de ácido nucleico de HIO endógena ou possui um alelo particular que causa conteúdo alterado de óleos, proteínas e/ou fibras na planta.

Embora a descrição tenha sido descrita com referência a métodos e modalidades específicos, será considerado que várias modificações e alterações podem ser feitas sem divergir da descrição. Todas as publicações citadas aqui são expressamente incorporadas aqui como referência com a finalidade de descrever e divulgar composições e metodologias que poderiam ser utilizadas em associação com a descrição. Todas as patentes, os pedidos de patentes e a informação de sequências citados nos bancos de dados públicos referidos também são incorporados aqui como referência.

EXEMPLOS

EXEMPLO 1

Produção de Plantas com um Fenótipo HIO através da Transformação com um Constructo de Marcação por Ativação

Este Exemplo descreve a produção de plantas transgênicas com conteúdo alterado de óleos, proteínas e/ou fibras.

Os mutantes foram gerados utilizando o vetor "ACTTAG" de marcação por ativação, pSKI015 (GI#6537289; Weigel D e outros, 2000, *Plant Physiology*, 122:1003-1013). Métodos padronizados foram utilizados para a produção de plantas transgênicas de *Arabidopsis* e eram essencialmente como descritos no pedido de patente publicado PCT WO0183697.

Sucintamente, plantas de *Arabidopsis* T0 (Col-0) foram transformadas com *Agrobacterium* carregando o vetor pSKI015, que compreende o T-DNA derivado do plasmídeo Ti de *Agrobacterium*, um gene marcador que pode ser selecionado de resistência a herbicida e o elemento intensificador 4X CaMV 35S. As plantas transgênicas foram selecionadas na geração T1 com base na resistência ao herbicida. A semente T2 (proveniente das plantas T1) foi colhida e plantada no solo. As plantas T2 foram expostas ao herbicida para matar as plantas que não possuem o vetor ACTTAG. As plantas T2 foram crescidas até a maturidade, permitidas que fossem autofertilizadas e produzissem semente. A semente T3 (proveniente das plantas T2) foi colhida em grandes quantidades para cada linhagem.

A semente T3 foi analisada através de Near Infrared Spectroscopy (NIR) no momento da colheita. Os espectros de NIR foram capturados utilizando um espectrômetro perto do infravermelho Bruker 22. O Software Bruker foi utilizado para estimar o conteúdo total de óleos da semente, o conteúdo total de proteínas da semente e o conteúdo total de fibras da semente utilizando os dados provenientes da análise de NIR e métodos de referência de acordo com as instruções do fabricante. As calibrações que prevêm o conteúdo de óleos foram desenvolvidas seguindo o método geral do AOCS Procedure Am1-92, Official Methods and Recommended Practices of the American Oil Chemists Society, 5ª Ed., AOCS, Champaign, Ill. Uma calibração que prevê o conteúdo total de proteínas por NIR foi desenvolvida utilizando dados do conteúdo de nitrogênio total de amostras de sementes seguindo o método geral do Dumas Procedure AOAC 968.06 (Official Methods of Analysis of AOAC International 17ª Edição AOAC, Gaithersburg, MD). Uma calibração que prevê o conteúdo de fibras por NIR foi desenvolvida utilizando os dados de conteúdo de fibras brutas de amostras de sementes seguindo o método geral do AOAC Official Method 962.09 (Official Methods of Analysis of AOAC International 17ª Edição AOAC, Gaithersburg, MD). Uma calibração que prevê o conteúdo de ácido oleico por NIR foi desenvolvida utilizando os dados do conteúdo de ácido oleico de amostras de semente determinados seguindo o método de Browse e outros (1986 Anal. Bio-

chem. 152:141-145). Foi desenvolvida uma curva de calibração de NIR para prever o conteúdo de proteínas que podem ser digeridas através da medida do conteúdo de proteínas que podem ser digeridas em um conjunto de amostras de sementes. O conteúdo total de proteínas em uma massa conhecida de semente foi determinado através da medida do conteúdo de nitrogênio total da semente utilizando o método de Dumas (AOAC Official Method 968.06). A fibra da semente é extraída de uma amostra de semente separada utilizando o método de Honig e Rackis, (1979, *J. Agri. Food Chem.*, 27: 1262-1266). A proteína não digerida remanescente associada com a fibra é medida através do método de Dumas (AOAC Official Method 968.06). O conteúdo de proteínas que podem ser digeridas é determinado através da subtração da quantidade de proteína não digerida associada com a fibra da quantidade total de proteína na semente.

As previsões de óleos, proteínas e fibras partindo dos espectros de NIR foram comparadas para 82.274 linhagens ACTTAG individuais. Subsequente à análise da composição das sementes, a posição do elemento ACTTAG no genoma em cada linhagem foi determinada através da PCR inversa e do sequenciamento. 37.995 linhagens com sequências flanqueadoras recuperadas foram consideradas nesta análise.

Os valores de óleos e de proteínas das sementes em 82.274 linhagens foram determinados através da espectroscopia por NIR e normalizados para permitir a comparação dos valores dos componentes das sementes nas plantas crescidas em tempos diferentes. Os valores de óleos, proteínas e fibras foram normalizados através do cálculo dos valores médios de óleos, proteínas e fibras na semente de todas as plantas plantadas no mesmo dia (incluindo um grande número de outras plantas ACTTAG, incluindo plantas de controle, do tipo selvagem ou não transgênicas). Os componentes das sementes para cada linhagem foram expressos na forma de um "valor relativo percentual" que foi calculado através da divisão do valor do componente para cada linhagem pelo valor médio do componente para todas as linhagens plantadas no mesmo dia (que deve se aproximar ao valor nas plantas de controle, do tipo selvagem ou não transgênicas). A "porcentagem

relativa de proteínas" e a "porcentagem relativa de fibras" foram calculadas
similarmente.

A PCR inversa foi utilizada para recuperar o DNA genômico que
flanqueia a inserção do T-DNA. O produto da PCR foi submetido à análise
5 de sequência e posicionado no genoma utilizando uma busca de BLASTN
básica e/ou uma busca do banco de dados Information Resource (TAIR) de
Arabidopsis (disponível no website disponível ao público). Geralmente, pro-
motores dentro de 9 kb dos intensificadores no elemento ACTTAG são con-
siderados como estando dentro do "espaço de ativação". Os genes com in-
10 sertos de T-DNA dentro das sequências codificadoras não foram considera-
dos como estando dentro do "espaço de ativação". As linhagens ACTTAG
identificadas estão listadas na coluna 3 da Tabela 1. Em alguns casos, mais
de uma linhagem ACTTAG está associada com um gene. Os valores relati-
vos de óleos, proteínas, fibras e ácido oleico nas colunas 4, 5, 6 e 7, respec-
15 tivamente, são determinados através da comparação do componente da
semente na planta identificado na coluna 3 em relação a outras plantas
crescidas ao mesmo tempo e que não exibem a característica.

TABELA 1

1. Pseudônimo	2. TAIR ID	3. ID da Planta	4. (%) Relativa de Óleo	5. (%) Relativa de Proteína	6. (%) Relativa de Fibra	7. Ácido Oleico Relativo
HIO2011 B	At4g10540	IN081274	121,52	85,77	112,19	170,91
HIO2022 A	At1g21160	IN070576	117,87	86,36	105,35	131,71
HIO2022 A	At1g21160	IN081625	108,6	94,24	102,16	127,55
HIO2024 A	At5g43820	IN078194	117,12	84,22	98,87	93,6
HIO2024 A	At5g43820	IN077533	110,52	97,43	97,63	101,01
HIO2029 A	At4g21880	IN040221	112,12	91,32	105,08	98,4
HIO2036 A	At1g35610	IN003383	110,63	95,26	102,61	107,39
HIO2044 A	At1g07200	IN085633	136,02	215,34	33,73	
HIO2044 A	At1g07200	IN065584	110,22	94,36	107,98	109,02
HIO2062 B	At3g21530	IN075911	112,25	97,43	128,26	121,82
HIO2062 B	At3g21530	IN010283	105,43	87,91	103,95	102,14
HIO2070 B	At1g03230	IN051870	110,64	97,08	92,8	99,68
HIO2070 B	At1g03230	IN027452	106,6	95,11	104,3	103,25
HIO2079 B	At1g71420	IN083604	131,01	83,85	99,94	139,75
HIO2079 B	At1g71420	IN090912	111,54	94,45	102,28	121,4

TABELA 1 - continuação

1. Pseudônimo	2. TAIR ID	3. ID da Planta	4. (%) Relativa de Óleo	5. (%) Relativa de Proteína	6. (%) Relativa de Fibra	7. Ácido Oleico Relativo
HIO2084 A	At3g12020	IN003802	115,88	86,87	92,23	35,51
HIO2084 A	At3g12020	IN090757	114,25	89,69	104,06	190,46
HIO2091 B	At1g09850	IN086447	139,97	124,86	65,1	
HIO2091 B	At1g09850	IN056625	104,65	106,59	100,56	103,87
HIO2093 A	At5g61370	IN088828	137,19	136,19	56,9	
HIO2093 A	At5g61370	IN013278	107,74	98,35	105,92	114,61
HIO2093 C	At5g61390	IN088828	137,19	136,19	56,9	
HIO2093 C	At5g61390	IN000614	109,24	92,2	104,2	106,81
HIO2094 C	At3g13224	IN090531	128,14	75,88	117,6	
HIO2094 C	At3g13224	IN041823	108,28	98,85	97,35	102,03
HIO2094 D	At3g13227	IN090531	128,14	75,88	117,6	
HIO2094 D	At3g13227	IN041823	108,28	98,85	97,35	102,03
HIO2104 B	At4g33900	IN067514	117,59	88,13	96,25	121,81

TABELA 2

1. Locus	2. Tair ID	3. Seq. de ácido nucleico GI#	4. Seq. de Polipeptídeo GI#	5. Suposta função bioquímica/nome da proteína	6. Domínio conservado da proteína
HIO2022 A	At1g21160	gi 42562216	gi 15218159	proteína da família do fator de início da tradução eucariótica 2/proteína da família eIF-2	IPR000795 Fator de síntese de proteína, ligação a GTP
					IPR002917 proteína de ligação a GTP, relacionada a HSR1
					IPR004161 Fator de alongamento Tu, domain 2
					IPR010528 ToIA
					IPR004130 Proteína de ligação a ATP hipotética conservada
HIO2024 A	At5g43820	gi 18422375	gi 15239991	proteína contendo repetição de pentatricopeptídeo (PPR)	IPR002885 Repetição de pentatricopeptídeo
HIO2029 A	At4g21880	gi 30685605	gi 30685606	proteína contendo repetição de pentatricopeptídeo (PPR)	IPR002885 Repetição de pentatricopeptídeo
HIO2036 A	At1g35610	gi 18399770	gi 15219660	Proteína contendo domínio DC1	IPR004146 DC1
					IPR011424 similar a C1

TABELA 2 - continuação

1. Locus	2. Tair ID	3. Seq. de ácido nucleico GI#	4. Seq. de Polipeptídeo GI#	5. Suposta função bioquímica/home da proteína	6. Domínio conservado da proteína
					IPR001965 similar a dedo de Zn, dedo de PHD
HIO2044 A	At1g07200	gi 30680018	gi 22329388	Relacionada à proteína ClpB Clp protease dependente de ATP	IPR001270 Chaperonina clpA/B
HIO2062 B	At3g21530	GI:42565085	gi 42565086	proteína da família de endonucleases/exonuclease/fosfatase	IPR005135 Endonuclease/ exonuclease/fosfatase
					IPR004808 Exodeoxirribonuclease III xth
					IPR000097 AP endonuclease, família 1
HIO2070 B	At1g03230	gi 42561644	gi 18379072	suposta glicoproteína dérmica extracelular/suposta EDGP	
HIO2079 B	At1g71420	gi 18409850	gi 15217470	proteína contendo repetição de pentatricopeptídeo (PPR)	IPR02885 Repetição de pentatricopeptídeo
HIO2084 A	At3g12020	gi 22331005	gi 22331006	relacionada à proteína motora quinolina	IPR01752 Quinesina, região motora
HIO2091 B	At1g09850	gi 30681401	gi 18391078	cisteína protease, similar à papaína (XBCP3)	IPR00169 Peptidase, sítio ativo da cisteína peptidase eucariótica

TABELA 2 - continuação

1. Locus	2. Tair ID	3. Seq. de ácido nucleicoGI#	4. Seq. de Polipeptídeo GI#	5. Suposta função bioquímica/nome da proteína	6. Domínio conservado da proteína
					IPR000668 Peptidase C1A, papain
					IPR000118 Granulina
HIO2093 A	At5g61370	gi 30697519	gi 15240232	proteína contendo repetição de pentatricopeptídeo (PPR)	IPR002885 Repetição de pentatricopeptídeo
HIO2093 C	At5g61390	gi 42568688	gi 42568689	proteína da família de exonuclease	IPR006055 Exonuclease
					IPR006054 DNA polimerase III, subunidade épsilon
HIO2094 C	At3g13224	GI:30682549	GI:30682550	Proteína contendo o motivo de reconhecimento de RNA (RRM)	IPR000504 Região RNP-1 de ligação ao RNA (motivo de reconhecimento de RNA)
					IPR009909 Nmi/IFP 35
					IPR002343 Antígeno de encefalomielite paraneoplásica
HIO2094 C	At3g13224	GI:30682552	GI:30682553	Proteína contendo o motivo de reconhecimento de RNA (RRM)	IPR000504 Região RNP-1 de ligação ao RNA (motivo de reconhecimento de RNA)

TABELA 2 - continuação

1. Locus	2. Tair ID	3. Seq. de ácido nucleico GI#	4. Seq. de Polipeptídeo GI#	5. Suposta função bioquímica/nome da proteína	6. Domínio conservado da proteína
					IPR009909 Nmi/IFP 35
					IPR002343 Antígeno de encefalomielite paraneoplásica
HIO2094 D	At3g13227	GI:30682556	gi 22331053	relacionada à proteína rica em serina	
HIO2011 B	At4g10540	gi 18413350	gi 18413351	proteína da família subtilase	IPR010259 Inibidor de proteinase 19, pró-peptídeo de subtilisina
					IPR003137 Associado à protease PA
					IPR000209 Peptidase S8 e S53, subtilisina, kexina, sedolisina
HIO2104 B	At4g33900	gi 18418316	gi 18418317	proteína da família F-box contendo repetições kelch	IPR006652 Repetição kelch
					IPR011498 Kelch
					IPR001810 F-box similar à ciclina

TABELA 3

1. Locus	2. Tair ID	3. Seq. de ácido nucleico GI#	4. Seq. de Polipeptídeo GI#	5. Ortólogos		
				Ácido nucleico GI#	Polipeptídeo GI#	Espécie
HIO2022 A	At1g21160	gil 42562216	gil 15218159	gil 42563274	gil 42563275	<i>Arabidopsis thaliana</i>
				gil 30699194	gil 30699195	<i>Arabidopsis thaliana</i>
				gil 23452070	gil 23452071	<i>Pisum sativum</i>
HIO2024 A	At5g43820	gil 18422375	gil 15239991	gil 30685153	gil 15242190	<i>Arabidopsis thaliana</i>
				gil 50918926	gil 50918927	<i>Oryza sativa</i> (grupo do cultivar japonica)
				gil 34896697	gil 34896698	<i>Oryza sativa</i> (grupo do cultivar japonica)
HIO2029 A	At4g21880	gil 30685605	gil 30685606	gil 30679876	gil 30679877	<i>Arabidopsis thaliana</i>
				gil 50905060	gil 50905061	<i>Oryza sativa</i> (grupo do cultivar japonica)
				gil 18415436	gil 15235288	<i>Arabidopsis thaliana</i>
HIO2036 A	At1g35610	gil 18399770	gil 15219660	gil 18400150	gil 15231357	<i>Arabidopsis thaliana</i>
				gil 42572370	gil 42572371	<i>Arabidopsis thaliana</i>
				gil 18399212	gil 15229741	<i>Arabidopsis thaliana</i>

TABELA 3 - continuação

1. Locus	2. Tair ID	3. Seq. de ácido nucleico GI#	4. Seq. de Polipeptídeo GI#	5. Ortólogos		
				Ácido nucleico GI#	Polipeptídeo GI#	Espécie
IO2044 A	At1g07200	gi 30680018	gi 22329388	gi 30680018	gi 22329388	<i>Arabidopsis thaliana</i>
				gi 42569462	gi 18402202	<i>Arabidopsis thaliana</i>
				gi 42568597	gi 15242850	<i>Arabidopsis thaliana</i>
HIO2062 B	At3g21530	GI:42565085	gi 42565086	gi 42570014	gi 18407510	<i>Arabidopsis thaliana</i>
				gi 34898867	gi 34898868	<i>Oryza sativa</i> (grupo do cultivar japônica)
				gi 21229478	gi 21231641	<i>Xanthomonas campestris</i> pv. <i>campestris</i> cepa ATCC 33913
HIO2070 B	At1g03230	gi 42561644	gi 18379072	gi 30678499	gi 15218740	<i>Arabidopsis thaliana</i>
				gi 285740	gi 285741	<i>Daucus carota</i>
				gi 32482805	gi 32482806	<i>Solanum tuberosum</i>
HIO2079 B	At1g71420	gi 18409850	gi 15217470	gi 18416131	gi 15242550	<i>Arabidopsis thaliana</i>
				gi 30693149	gi 30693150	<i>Arabidopsis thaliana</i>
				gi 50899005	gi 50899006	<i>Oryza sativa</i> (grupo do cultivar japônica)

TABELA 3 - continuação

1. Locus	2. Tair ID	3. Seq. de ácido nucleico GI#	4. Seq. de Polipeptídeo GI#	5. Ortólogos		
				Ácido nucleico GI#	Polipeptídeo GI#	Espécie
HIO2084 A	At3g12020	gil22331005	gil22331006	gil22326628	gil22326629	<i>Arabidopsis thaliana</i>
				gil30687505	gil30687506	<i>Arabidopsis thaliana</i>
				gil50924423	gil50924424	<i>Oryza sativa</i> (grupo do cultivar japonica)
HIO2091 B	At1g09850	gil30681401	gil18391078	gil50932272	gil50932273	<i>Oryza sativa</i> (grupo do cultivar japonica)
				gil30694200	gil18422289	<i>Arabidopsis thaliana</i>
				gil30141018	gil30141019	<i>Helianthus annuus</i>
HIO2093 A	At5g61370	gil30697519	gil15240232	gil37534457	gil37534458	<i>Oryza sativa</i> (grupo do cultivar japonica)
				gil30697448	gil15239488	<i>Arabidopsis thaliana</i>
				gil30693231	gil15229604	<i>Arabidopsis thaliana</i>
HIO2093 C	At5g61390	gil42568688	gil42568689	gil42567730	gil15240840	<i>Arabidopsis thaliana</i>
				gil22330631	gil15221225	<i>Arabidopsis thaliana</i>
				gil37534221	gil37534222	<i>Oryza sativa</i> (grupo do cultivar japonica)

TABELA 3 - continuação

1. Locus	2. Tair ID	3. Seq. de ácido nucleico GI#	4. Seq. de Polipeptídeo GI#	5. Ortólogos		
				Ácido nucleico GI#	Polipeptídeo GI#	Espécie
HIO2094 C	At3g13224	GI:30682549	GI:30682550	gi 30682552	gi 30682553	<i>Arabidopsis thaliana</i>
				gi 50944098	gi 50944099	<i>Oryza sativa</i> (grupo do cultivar japonica)
				gi 50947154	gi 50947155	<i>Oryza sativa</i> (grupo do cultivar japonica)
				GI:51964985	GI:51964986	<i>Oryza sativa</i> (grupo do cultivar japonica)
HIO2094 C	At3g13224	GI:30682552	GI:30682553	gi 30682549	gi 30682550	<i>Arabidopsis thaliana</i>
				gi 50944098	gi 50944099	<i>Oryza sativa</i> (grupo do cultivar japonica)
				gi 50947154	gi 50947155	<i>Oryza sativa</i> (grupo do cultivar japonica)
				GI:51964985	GI:51964986	<i>Oryza sativa</i> (grupo do cultivar japonica)
HIO2094 D	At3g13227	GI:30682556	gi 22331053	GI:4105789	GI:4105790	<i>Petunia</i> x híbrida

TABELA 3 - continuação

1. Locus	2. Tair ID	3. Seq. de ácido nucleico GI#	4. Seq. de Polipeptídeo GI#	5. Ortólogos		
				Ácido nucleico GI#	Polipeptídeo GI#	Espécie
				GI:50905728	GI:50905729	<i>Oryza sativa</i> (grupo do cultivar japonica)
				GI:30697586	GI:15220581	<i>Arabidopsis thaliana</i>
HIO2011 B	At4g10540	gi 18413350	gi 18413351	gi 30681486	gi 18413353	<i>Arabidopsis thaliana</i>
				gi 18413344	gi 18413345	<i>Arabidopsis thaliana</i>
				gi 30692780	gi 18398655	<i>Arabidopsis thaliana</i>
HIO2104 B	At4g33900	gi 18418316	gi 18418317	GI:18416176	gi 15236530	<i>Arabidopsis thaliana</i>
				GI:18418391	gi 15235260	<i>Arabidopsis thaliana</i>
				GI:18407622	gi 15229754	<i>Arabidopsis thaliana</i>

EXEMPLO 2

Análise da Sequência de HIO de *Arabidopsis*

As análises de sequência foram realizadas com BLAST (Altschul e outros, 1990, *J. Mol. Biol.* 215:403-410), PFAM (Bateman e outros, 1999, *Nucleic Acids Res.* 27:260-262), INTERPRO (Mulder e outros 2003 *Nucleic Acids Res.* 31, 315-318), PSORT (Nakai K e Horton P, 1999, *Trends Biochem. Sci.* 24:34-6) e/ou CLUSTAL (Thompson JD e outros, 1994, *Nucleic Acids Res.* 22:4673-4680).

EXEMPLO 3

Experimentos de Resumo

Para testar se a superexpressão dos genes nas Tabelas 1 e 2 alteram o fenótipo de composição das sementes, o conteúdo de proteínas, de proteínas que podem ser digeridas, de óleos e de fibras nas sementes das plantas transgênicas que expressam estes genes foi comparado com o conteúdo de proteínas, de proteínas que podem ser digeridas, de óleos e de fibras nas sementes das plantas de controle não transgênicas. Para fazer isso, os genes foram clonados nos vetores de transformação de plantas atrás do promotor de CsVMV constitutivo forte e do promotor PRU específico de semente. Estes constructos foram transformados em plantas de *Arabidopsis* utilizando o método de imersão floral. O vetor de transformação de plantas contém um gene, que fornece resistência a um composto tóxico e serve como um marcador que pode ser selecionado. A semente das plantas transformadas foram plantadas em meio com ágar contando o composto tóxico. Após 7 dias, as plantas transgênicas foram identificadas como plantas verdes saudáveis e foram transplantadas para o solo. As plantas de controle não transgênicas foram germinadas em meio com ágar, permitidas que crescessem durante 7 dias e então transplantadas para o solo. As mudas transgênicas e as plantas de controle não transgênicas foram transplantadas em vasos de 5,08 cm (duas polegadas) que foram colocados em posições aleatórias em uma bandeja de 25,4 cm (10 polegadas) por 50,8 cm (20 polegadas). As plantas foram crescidas até a maturidade, permitidas que sofressem autofertilização e produzissem semente. A semente foi colhida de cada

planta e seu conteúdo de óleos foi estimado através de Near Infrared (NIR) Spectroscopy utilizando métodos descritos anteriormente. O efeito de cada construção sobre a composição da semente foi verificado em pelo menos dois experimentos.

5 A Tabela 4 lista os constructos testados por causarem um aumento significativo no conteúdo de óleos, proteínas, proteínas que podem
10 ser digeridas ou um decréscimo significativo no conteúdo de fibras foi identificado através do teste de Análise de Variância em duas vias (ANOVA) em
15 um valor de $p \leq 0,05$. Estes constructos estão listados na Tabela 4. Os valores de p de ANOVA para Proteína, Óleo, Proteína Que Pode Ser Digerida e
Fibra estão listados nas colunas 4-7, respectivamente. Aqueles com valor de p significativo estão listados em negrito. Os valores médios para Proteína, Óleo, Proteína Que Pode Ser Digerida e Fibra estão listados nas colunas 8-11, respectivamente e foram calculados através da média dos valores médios determinados para as plantas transgênicas em cada experimento.

TABELA 4

1. Pseudônimo	2. Tair	3. Construção	ANOVA				Média			
			4. Proteína	5. Óleo	6. Proteína que pode ser digerida	7. Fibra	8. Proteína	9. Óleo	10. Proteína que pode ser digerida	11. Fibra
HIO2011 B	At4g10540	CsVMV::At4g10540	0,498	0,736	0,854	0,789	98,5%	100,7%	100,2%	100,3%
HIO2011 B	At4g10540	Pru::At4g10540	0,680	0,107	0,004	0,002	99,4%	103,3%	102,9%	95,9%
HIO2022 A	At1g21160	CsVMV::At1g21160	0,245	0,689	0,084	0,067	103,5%	98,2%	99,6%	100,7%
HIO2022 A	At1g21160	Pru::At1g21160	0,538	0,119	0,034	0,003	101,5%	102,3%	102,9%	96,0%
HIO2024 A	At5g43820	CsVMV::At5g43820	0,582	0,131	0,005	0,047	99,3%	102,6%	102,4%	96,9%
HIO2024 A	At5g43820	Pru::At5g43820	0,074	0,864	0,033	0,039	103,5%	99,8%	102,0%	97,3%
HIO2029 A	At4g21880	CsVMV::At4g21880	0,119	0,159	0,002	0,372	101,4%	98,3%	102,1%	99,2%
HIO2029 A	At4g21880	Pru::At4g21880	0,533	0,054	<,0001	<,0001	101,1%	102,8%	104,2%	94,5%
HIO2036 A	At1g35610	CsVMV::At1g35610	0,040	0,001	0,314	0,005	97,0%	106,5%	100,9%	96,3%
HIO2036 A	At1g35610	Pru::At1g35610	0,776	0,780	0,008	0,025	100,6%	100,6%	103,0%	97,2%
HIO2044 A	At1g07200	CsVMV::At1g07200	0,815	0,336	0,037	0,069	100,3%	101,8%	101,7%	97,7%
HIO2044 A	At1g07200	Pru::At1g07200	0,134	0,202	0,000	0,001	102,0%	102,3%	103,3%	95,7%
HIO2062 B	At3g21530	CsVMV::At3g21530	0,709	0,423	0,256	0,145	99,5%	100,7%	100,7%	99,1%
HIO2062 B	At3g21530	Pru::At3g21530	0,316	0,024	0,756	0,160	98,3%	102,8%	99,8%	98,8%
HIO2070 B	At1g03230	CsVMV::At1g03230	0,966	0,618	0,258	0,044	100,1%	101,1%	101,0%	98,1%
HIO2070 B	At1g03230	Pru::At1g03230	0,995	0,019	0,018	0,003	100,1%	102,8%	102,3%	96,8%

TABELA 4 - continuação

1. Pseudônimo	2. Tair	3. Construção	ANOVA				Média			
			4. Proteína	5. Óleo	6. Proteína que pode ser digerida	7. Fibra	8. Proteína	9. Óleo	10. Proteína que pode ser digerida	11. Fibra
HIO2079 B	At1g71420	CsVMV::At1g71420	0,297	0,033	0,083	0,058	99,0%	102,7%	101,2%	98,0%
HIO2079 B	At1g71420	Pru::At1g71420	0,114	0,810	0,008	0,002	102,6%	99,8%	102,3%	97,1%
HIO2084 A	At3g12020	CsVMV::At3g12020	0,076	0,728	0,063	0,200	102,4%	99,5%	101,6%	98,3%
HIO2084 A	At3g12020	Pru::At3g12020	0,487	0,019	0,022	0,001	98,6%	103,7%	102,0%	96,1%
HIO2091 B	At1g09850	CsVMV::At1g09850	0,041	0,065	0,760	0,830	97,2%	102,8%	99,8%	100,4%
HIO2091 B	At1g09850	Pru::At1g09850	0,354	0,170	0,475	0,748	101,9%	98,4%	101,1%	100,2%
HIO2093 A	At5g61370	CsVMV::At5g61370	0,337	0,098	0,224	0,074	98,8%	102,4%	101,0%	97,9%
HIO2093 A	At5g61370	Pru::At5g61370	0,535	0,139	<,0001	<,0001	100,8%	101,9%	103,4%	94,0%
HIO2093 C	At5g61390	CsVMV::At5g61390	0,058	0,065	0,064	0,051	97,1%	103,0%	101,1%	98,1%
HIO2093 C	At5g61390	Pru::At5g61390	0,949	0,068	0,005	0,000	100,1%	101,9%	102,7%	95,6%
HIO2094 C	At3g13224	CsVMV::At3g13224	0,710	0,065	0,186	0,070	99,5%	101,7%	101,2%	98,2%
HIO2094 C	At3g13224	Pru::At3g13224	0,034	0,173	0,025	0,125	104,5%	98,4%	102,6%	98,0%
HIO2094 D	At3g13227	CsVMV::At3g13227	0,096	0,030	0,153	0,042	97,3%	104,5%	101,2%	97,2%
HIO2094 D	At3g13227	Pru::At3g13227	0,042	0,004	0,430	<,0001	97,9%	97,8%	101,3%	105,0%
HIO2104 B	At4g33900	CsVMV::At4g33900	0,846	0,722	0,374	0,064	100,0%	99,6%	99,5%	102,5%
HIO2104 B	At4g33900	Pru::At4g33900	0,152	0,015	0,048	0,002	98,3%	103,5%	101,7%	95,6%

LISTAGEM DE SEQUÊNCIAS

<110> Agrinomics LLC
 Davies, John P.
 Ng, Hein Tsoeng

5 <120> PRODUÇÃO DE PLANTAS COM CONTEÚDO ALTERADO DE ÓLEOS,
 PROTEÍNAS OU FIBRAS
 <130> 6616-72632-27
 <150> 60/870,345
 <151> 2006-12-15

10 <160> 34
 <170> PatentIn version 3.3
 <210> 1
 <211> 3603
 <212> DNA

15 <213> Arabidopsis thaliana
 <400> 1

ctcaaatata tctctgcaac gaaaagaaat cacctgctta ctcttctcct gaaagacctg 60
 atcttctccg gtgaattgtt ttttgttccg tcaatcttcc gagtttgttt ttgcctaata 120
 tgaagttaaa cccattacag aaaaaggcta tattcatatt tgatttctat caatggatcc 180
 20 gaggaattcg agtgcgtggg gtgggggtga tggctctccc cttatctcat ccatgggagg 240
 tgttgaatca aagaatagat gctgtgtgac tgacgacgat ggctctattg gtacgaagct 300
 atgtgagaaa cctaaagctg aggagattga tttttctgct cggaaaaaga agttgaaatc 360
 gaagaaaggc gggctctgtta gctttgctct tctcgatgtt gaagatgaca aggaggaagc 420
 agaggatgga gacgacaaaa atccttctat taggagtgat gaggtgggtg gagaaacttc 480
 25 aatgagtaag aaaaagaaga aggataagag tggcaagcat ggggtatgtg aagaagatga 540
 tgatgcggac aagattcttg ctcaacatgg gataacaact actgtttcaa cagggcctga 600
 agcggagata gctctagctc aacctgagca agttgcatct gcagatgggg ctgttgacaa 660
 ggacggagaa gaaaagactg ttgagtcagc tacagcgaag aaaaagaaaa agaagaagga 720
 caaggacaag aaagcaagtg cttcttttagc tatttcttct gtggaggcta aggaggatag 780
 30 acaagggaag aaagatgtaa agattaaagt agcggaaaag aaagttccta aacatgttcg 840
 agagaagcaa gagactcttg cgcggtggaa agaagctgaa gatgggaaga agaaggaaga 900
 agaagagaga ttgaggaaag aggaagagga gcgccgtata gaggaagagc gagaaaggga 960

agctgaagag attaggcaga agagaaagat caggaaaatg gagaagaagc aagaaggcct 1020
 gattctaaca gcaaagcaaa agcgtgatgc tgctaagaat gaggctttca gaaaaaggg 1080
 gttgactgat gctggaagtc ttcttggtgc agataaaaac ggtgattcct caaaacgtcc 1140
 catatatggc aataaaaaca aattagcttg caaaaaagca aacgatcctg cttctgtcca 1200
 5 ggctaaaggt gatggggaaa ccaaggaaaa tcatgcagca gatgaaccat gtactttacc 1260
 tgacttggtt tcagttgaag atagaagggg ttgtattatc gaatcagtag acacagaaga 1320
 gacacatgaa tctgttgatg tatcacaaga gaatggtgat gaggaagatg tgtgggatgc 1380
 aaaaactaat ttaccatca aggggtgattc tgatgatgag gaggaaaaac ctcagcctgt 1440
 gttcaagaaa ggtttaaaag acactgcac aaaggcacat gattcagata gaccaacagt 1500
 10 gaagcctgga ggttctggca agccaaaaac agctgccaaa aaagctatgc ctaagggtgga 1560
 tgatactacc cggccgaaag atacttctaa aaaggatgaa ggcttagtac tgaatgaact 1620
 tgctaaagaa gtagaagaaa atcttcgctc tccaatttgc tgcacatggt gtcattgtga 1680
 ttctggtaaa acaaagctgt tggattgcat cagaggaaca aatgttcagg aagggaagc 1740
 tggagggtatt actcaacaga ttgggtgcaac ctttttcct gcagaaaaca tccgtgagag 1800
 15 gaccaaagag ttacaagcta atgctaaact caagggtgcca ggtatattgg ttatcgatac 1860
 acctggacac gagtcattca caaatctgcg gtcaaggggt tcaaatttgt gtgaccttgc 1920
 aattctagtg gttgatataa tgcgagggtc agagccaca accatagaat ctcttaattc 1980
 ttgagaaga aggaatgtaa agttcattat tgccttgaat aagggtgata ggctatatgg 2040
 gtgggaaaaa agcaagaatg ctctataag gaagacaatg atgcaacaaa cgggagatgt 2100
 20 ggttaaggaa tttaaatga ggctaaaccg agttcaaac cagttccaag aacaaggact 2160
 caacagtatg ctttattaca agaacagaga aatgggagaa actatcagta tcctacctgc 2220
 tagcgctatt agtggggaag ggattccaga tcttttactg ttcttggttc aatgggctca 2280
 gaaaacgatg gtggagaaac ttacatatgt tgacaaagt cagtgtaccg tccttgaggt 2340
 caaagttatt gaaggccatg gtattacagt tgatgttgtt ttgggtcaacg gtgtactccg 2400
 25 tgaaggatg caaattgttg ttgtgggtc acagggacca attgtaacaa ctattagatc 2460
 attgttgact ctttatccta tgaacgagat gcgggtgacg ggtacctata tgcctcacag 2520
 agaagtaaag gctgcgaggt gtatcaaaat tgctgcacag ggccttgaac acgccattgc 2580
 ttgtactgcc ttgcatgtga ttggacctaa tgaggacatg gaagaagcca agaaaaatgc 2640
 catggaagat attgagtcag ttatgaaccg cattgacaaa agtgggtgaag gcgtttatgt 2700
 30 acaagcatct acattagggg ccttggaagc attgctcgag ttcttgaagt cctcagatgt 2760
 aaaaatacct gtcagtggta ttggcatagg acctgtgcat aagaaggata tcatgaaggc 2820
 cggggtaatg ctagagaaga aaaaagagtt tgcaacgatt ttggccttg atgtgaaaat 2880

aagtgcagaa gctcgtgaac ttgcagacaa aatgggagta aaaatcttct gcgatgatac 2940
 catctattgt ctgttcgata agtttaagag ttacatagag ggtatcaaag aggagaaaaa 3000
 gaaagaaaca gcttgtgaag cagttttccc atgtatcctt cagattttac caaaccatat 3060
 ctacaatcag agggacccaa ttatccttgg agttaaagtc aatgatggaa tactaaaggt 3120
 5 tggaactccc atttgcattg ttaaaaggat tgagaatgta agagtgttta tggacattgg 3180
 gcgcgtctca tcaattaaga ataacaaca catgccagtt gattatgcaa ggaaaggaca 3240
 ggaggtggct attaagatca ttgcttctaa ccagaagaa cagaaaatgt ttgggaggca 3300
 ctttggtgta gacgacaggc ttatcagtca ctttctagt cgatccgtgg acgtaatcag 3360
 aactaattac tggaatgagt tatctaacga cgagaaggat cttgttttga gacttaaaag 3420
 10 gatattcaag atacaatgaa gatttttgag atttgagagt gagggactta aacttatgta 3480
 ttttttttct ttttttgaag tgctgatttt tttatttttt tttcgttttt aataatatag 3540
 acaaacaact cctttaatgc ccctaaatac tttatggaag caattgtctt ttgtttttgg 3600
 cta 3603

<210> 2
 15 <211> 1088
 <212> PRT
 <213> Arabidopsis thaliana
 <400> 2

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

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

Glu Asp Arg Arg Val Gly Ile Ile Glu Ser Val Asp Thr Glu Glu Thr
 370 375 380
 His Glu Ser Val Asp Val Ser Gln Glu Asn Gly Asp Glu Glu Asp Val
 385 390 395 400
 5 Trp Asp Ala Lys Thr Asn Phe Thr Ile Lys Gly Asp Ser Asp Asp Glu
 405 410 415
 Glu Glu Lys Pro Gln Pro Val Phe Lys Lys Gly Leu Lys Asp Thr Ala
 420 425 430
 Ser Lys Ala His Asp Ser Asp Arg Pro Thr Val Lys Pro Gly Gly Ser
 10 435 440 445
 Gly Lys Pro Lys Thr Ala Ala Lys Lys Ala Met Pro Lys Val Asp Asp
 450 455 460
 Thr Thr Arg Pro Lys Asp Thr Ser Lys Lys Asp Glu Gly Leu Val Leu
 465 470 475 480
 15 Asn Glu Leu Ala Lys Glu Val Glu Glu Asn Leu Arg Ser Pro Ile Cys
 485 490 495
 Cys Ile Met Gly His Val Asp Ser Gly Lys Thr Lys Leu Leu Asp Cys
 500 505 510
 Ile Arg Gly Thr Asn Val Gln Glu Gly Glu Ala Gly Gly Ile Thr Gln
 20 515 520 525
 Gln Ile Gly Ala Thr Phe Phe Pro Ala Glu Asn Ile Arg Glu Arg Thr
 530 535 540
 Lys Glu Leu Gln Ala Asn Ala Lys Leu Lys Val Pro Gly Ile Leu Val
 545 550 555 560
 25 Ile Asp Thr Pro Gly His Glu Ser Phe Thr Asn Leu Arg Ser Arg Gly
 565 570 575
 Ser Asn Leu Cys Asp Leu Ala Ile Leu Val Val Asp Ile Met Arg Gly
 580 585 590
 Leu Glu Pro Gln Thr Ile Glu Ser Leu Asn Leu Leu Arg Arg Arg Asn
 30 595 600 605
 Val Lys Phe Ile Ile Ala Leu Asn Lys Val Asp Arg Leu Tyr Gly Trp
 610 615 620

Glu Lys Ser Lys Asn Ala Pro Ile Arg Lys Thr Met Met Gln Gln Thr
 625 630 635 640
 Gly Asp Val Val Lys Glu Phe Lys Met Arg Leu Asn Arg Val Gln Asn
 645 650 655
 5 Gln Phe Gln Glu Gln Gly Leu Asn Ser Met Leu Tyr Tyr Lys Asn Arg
 660 665 670
 Glu Met Gly Glu Thr Ile Ser Ile Leu Pro Ala Ser Ala Ile Ser Gly
 675 680 685
 Glu Gly Ile Pro Asp Leu Leu Leu Phe Leu Val Gln Trp Ala Gln Lys
 10 690 695 700
 Thr Met Val Glu Lys Leu Thr Tyr Val Asp Lys Val Gln Cys Thr Val
 705 710 715 720
 Leu Glu Val Lys Val Ile Glu Gly His Gly Ile Thr Val Asp Val Val
 725 730 735
 15 Leu Val Asn Gly Val Leu Arg Glu Gly Asp Gln Ile Val Val Cys Gly
 740 745 750
 Ser Gln Gly Pro Ile Val Thr Thr Ile Arg Ser Leu Leu Thr Pro Tyr
 755 760 765
 Pro Met Asn Glu Met Arg Val Thr Gly Thr Tyr Met Pro His Arg Glu
 20 770 775 780
 Val Lys Ala Ala Gln Gly Ile Lys Ile Ala Ala Gln Gly Leu Glu His
 785 790 795 800
 Ala Ile Ala Gly Thr Ala Leu His Val Ile Gly Pro Asn Glu Asp Met
 805 810 815
 25 Glu Glu Ala Lys Lys Asn Ala Met Glu Asp Ile Glu Ser Val Met Asn
 820 825 830
 Arg Ile Asp Lys Ser Gly Glu Gly Val Tyr Val Gln Ala Ser Thr Leu
 835 840 845
 Gly Ser Leu Glu Ala Leu Leu Glu Phe Leu Lys Ser Ser Asp Val Lys
 30 850 855 860
 Ile Pro Val Ser Gly Ile Gly Ile Gly Pro Val His Lys Lys Asp Ile
 865 870 875 880

[illegible]

<400> 3

atgcttagga gatggaacat agttattgaa tcgctacgtc gcgttcatag cagtgattta 60
 gaaaacctaa tttcggcttc cctatacaat cgaacacttt gcacagcctc agaatcattg 120
 aatcatggtg ttgttgatga gagttatggt ctagctgaat tatcttcatt gcttccaatt 180
 5 tcatctaata aaacctcagt gtcaaaggaa gacagtagct caaagaatca agtagctatt 240
 gattcgtttt tgtcagcaga ggacaaactt agaggtgtgt ttcttcagaa attgaaaggc 300
 aagtctgcta ttcagaagag tttagagtagt cttggtattg gtttgagtat tgacattggt 360
 gccgatgtat tgaatcgagg gaatttaagc ggcgaagcga tggtaacggt ctttgattgg 420
 gcggttcgtg agcctggtgt gactaaagat gttggtagct acagtgtgat ttaagagcg 480
 10 ttaggaagga gaaagctttt ctcttttatg atggatgttc taaaggaat ggtatgtgaa 540
 ggtgttaatc ctgatctgga gtgtttaact attgctatgg atagctttgt tagggttcat 600
 tatgtgcgta gagcgataga gttgtttgag gaaagtgaag gttttggggg gaaatgtagt 660
 actgagtctt tcaatgcatt gttacggtgt ctatgtgagc gttcacatgt gagtgtgca 720
 aagtcagtat ttaatgctaa gaaaggtaat ataccttttg atagttgtag ttataacatt 780
 15 atgattagtg ggtggtcaaa gctcggtag gttgaagaga tggaaaagggt tttgaaggag 840
 atggtggaga gtggatttgg tcctgactgt ttgtcttata gccatcttat tgaagggttg 900
 ggaagaactg gtcgtatcaa tgattctggt gagatctttg ataataataa acacaagggt 960
 aatgttcctg atgcaaatgt ttacaatgct atgatctgta atttcatatc tgctcgagat 1020
 tttgatgagt ctatgaggta ctaccgaagg atgctggatg aggaatgtga acctaacctg 1080
 20 gaaacctatt ccaagcttgt ttctgggctc attaaaggcc gtaagggttc tgatgcgctt 1140
 gagatattcg aagagatgtt atcaagaggg gttcttccca ccacagggtt tgttacctct 1200
 ttcctcaagc ctctttgcag ctatggtcca ccacacgctg caatgggttat ctatcagaaa 1260
 tcaagaaaag ccgggtgtag gatatcagaa agcgcatata agctgttgct taagcggcta 1320
 tcaagatttg gaaaatgcgg gatgctgttg aacgtatggg atgaaatgca agagtctgga 1380
 25 tatccttctg acgtggaagt ctacgagtat attgtggatg ggctctgtat cattgggcat 1440
 ctagagaatg ctgtgcttgt gatggaagaa gctatgcgta aagggttttg cccaaaccgc 1500
 tttgtttata gtaggcttag cataattagg aaatggcagc agaagttcag aaaagggaag 1560
 gaggagatgg aaaagtggga ggcacttcaa gttcgatggg tgtccctctt caggaacgct 1620
 tcttctatca tccagagatt acaggaaatg caaaatcatg ggagctatgg agctttgaga 1680
 30 tgcctgaaag ggattgaaga cgcagttgtg cagcaacaga tgggtcaact ggagtctcta 1740
 ttgcgttcca tgaaaaacgt tatggaggaa ttctggggct gtgttttgac atttgagaag 1800
 ctacatcggg acgggtttaca gttacgtgaa attgaatcaa gcaagaggcg agtagaggaa 1860

cgaattggag taaaaccgtg cattactgat tgcttggagg gactttctat tctttatgac 1920
 atgcatcaat cagatcctgg tgatctaaac gcgctacaat acctcgtggt tgatcagccg 1980
 aatatcccga aagatgaagt acaacacatt ttcgatgtca tattcgcgga agagatcaaa 2040
 tga 2043

5 <210> 4
 <211> 680
 <212> PRT
 <213> Arabidopsis thaliana
 <400> 4

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

		180		185		190	
	Met	Asp	Ser	Phe	Val	Arg	Val
				His	Tyr	Val	Arg
				Arg	Arg	Ala	Ile
				Glu	Leu		
		195		200		205	
	Phe	Glu	Glu	Ser	Glu	Ser	Phe
				Gly	Val	Lys	Cys
				Ser	Thr	Glu	Ser
5		210		215		220	
	Asn	Ala	Leu	Leu	Arg	Cys	Leu
				Cys	Glu	Arg	Ser
				His	Val	Ser	Ala
				Ala			
	225			230		235	
				240			
	Lys	Ser	Val	Phe	Asn	Ala	Lys
				Lys	Gly	Asn	Ile
				Pro	Phe	Asp	Ser
				Cys			
				245		250	
				255			
10	Ser	Tyr	Asn	Ile	Met	Ile	Ser
				Gly	Trp	Ser	Lys
				Leu	Gly	Glu	Val
				Glu			
				260		265	
				270			
	Glu	Met	Glu	Lys	Val	Leu	Lys
				Glu	Met	Val	Glu
				Ser	Gly	Phe	Gly
				Pro			
				275		280	
				285			
	Asp	Cys	Leu	Ser	Tyr	Ser	His
				Leu	Ile	Glu	Gly
				Leu	Gly	Arg	Thr
				Gly			
15		290		295		300	
	Arg	Ile	Asn	Asp	Ser	Val	Glu
				Ile	Phe	Asp	Asn
				Ile	Lys	His	Lys
				Gly			
	305			310		315	
				320			
	Asn	Val	Pro	Asp	Ala	Asn	Val
				Tyr	Asn	Ala	Met
				Ile	Cys	Asn	Phe
				Ile			
				325		330	
				335			
20	Ser	Ala	Arg	Asp	Phe	Asp	Glu
				Ser	Met	Arg	Tyr
				Tyr	Arg	Arg	Met
				Leu			
				340		345	
				350			
	Asp	Glu	Glu	Cys	Glu	Pro	Asn
				Leu	Glu	Thr	Tyr
				Ser	Lys	Leu	Val
				Ser			
				355		360	
				365			
	Gly	Leu	Ile	Lys	Gly	Arg	Lys
				Val	Ser	Asp	Ala
				Leu	Glu	Ile	Phe
				Glu			
25		370		375		380	
	Glu	Met	Leu	Ser	Arg	Gly	Val
				Leu	Pro	Thr	Thr
				Gly	Leu	Val	Thr
				Ser			
				385		390	
				395			
				400			
	Phe	Leu	Lys	Pro	Leu	Cys	Ser
				Tyr	Gly	Pro	Pro
				His	Ala	Ala	Met
				Val			
				405		410	
				415			
30	Ile	Tyr	Gln	Lys	Ser	Arg	Lys
				Ala	Gly	Cys	Arg
				Ile	Ser	Glu	Ser
				Ala			
				420		425	
				430			
	Tyr	Lys	Leu	Leu	Leu	Lys	Arg
				Leu	Ser	Arg	Phe
				Gly	Lys	Cys	Gly
				Met			

	435	440	445
	Leu Leu Asn Val Trp Asp Glu Met Gln Glu Ser Gly Tyr Pro Ser Asp		
	450	455	460
	Val Glu Val Tyr Glu Tyr Ile Val Asp Gly Leu Cys Ile Ile Gly His		
5	465	470	475
	Leu Glu Asn Ala Val Leu Val Met Glu Glu Ala Met Arg Lys Gly Phe		480
		485	490
	Cys Pro Asn Arg Phe Val Tyr Ser Arg Leu Ser Ile Ile Arg Lys Trp		
	500	505	510
10	Gln Gln Lys Phe Arg Lys Gly Lys Glu Glu Met Glu Lys Trp Glu Ala		
	515	520	525
	Leu Gln Val Arg Trp Val Ser Leu Phe Arg Asn Ala Ser Ser Ile Ile		
	530	535	540
	Gln Arg Leu Gln Glu Met Gln Asn His Gly Ser Tyr Gly Ala Leu Arg		
15	545	550	555
	Cys Leu Lys Gly Ile Glu Asp Ala Val Val Gln Gln Gln Met Gly Gln		560
		565	570
	Leu Glu Ser Leu Leu Arg Ser Met Lys Asn Val Met Glu Glu Phe Trp		
	580	585	590
20	Gly Cys Val Leu Thr Phe Glu Lys Leu His Arg Asp Gly Leu Gln Leu		
	595	600	605
	Arg Glu Ile Glu Ser Ser Lys Arg Arg Val Glu Glu Arg Ile Gly Val		
	610	615	620
	Lys Pro Cys Ile Thr Asp Cys Leu Glu Gly Leu Ser Ile Leu Tyr Asp		
25	625	630	635
	Met His Gln Ser Asp Pro Gly Asp Leu Asn Ala Leu Gln Tyr Leu Val		640
		645	650
	Val Asp Gln Pro Asn Ile Pro Lys Asp Glu Val Gln His Ile Phe Asp		
	660	665	670
30	Val Ile Phe Ala Glu Glu Ile Lys		
	675	680	
<210>	5		

<211> 2589
 <212> DNA
 <213> Arabidopsis thaliana
 <400> 5

```

5  cgttctaggc cttgattcga ttttgtaa acgttgtaaat gtgaaaaaac aatggcttta 60
   aaggcgaaga aaagcttaag cgcgatattt cgctttgcga tgagaaacgc tgccaaaact 120
   gccgaagttg aagccaaagc cgtagccgga gattctctga tttcagacac tggaccagtg 180
   tcaatgtcct taataccgga ttttctctca tcgcttaagt ctccggaaga ttcttccaac 240
   attgggtcaag attttggtga accggcactt gcgggtcagg tctcgtctgc gttagatggt 300
10 aaatctattg gatataatct tcttcgagaa agaaatggcg agaaagtttt tccgaagaac 360
   gtattggaaa ccttagattt actgtctagg aaaggttggt cgtcgcaaaa gccaaagcag 420
   gcttcacgtg gtaggatgct gacagaaaat taccagaaca aacagtctga gattatggag 480
   aagcttgcca agggatgtgt tcggaagttg ggaactgaga ctatgtttga ggtgttgact 540
   aaaatgggga aagaagcggg tgagaaagaa tacaatgcaa tgactaaact ttgcatacaa 600
15 cgtgcgagga gaagcaatga tgcagaatat gctcttgatc aaattgggaa ggctattgaa 660
   catctcaaag aaatgagaca gcttgattc tctattggag aaggagctta tggccctttt 720
   ttttaagtact tagttgatat ggaaatggta gcggagtcc aaattctcaa ggattttatt 780
   aaggaggcat gtcctgagtc ttgtgggaga cttgtctact atgaaatgct tctttggatc 840
   caagtcaacg acgaagagaa gatccacaaa ctttgcaata aagttgatga tagtgggcta 900
20 agtttatcaa tcctacaaga tggagtggag acagtttcaa atctcatatt cagctatgca 960
   acctgcatcc caaattcaac ggtagaggat gccattttca agtttaacaa gttgcatgaa 1020
   gaactagata tcgtgccttc atctacatct tatgagaatc ttgtcagtta tctttgtggc 1080
   tcaaatgagg tggtcactgc tcttgatata gttgaaaata tgtgtgaagc aggtctggta 1140
   atatcagcaa acatcctgca ctcatgtta caggccattg agcagattct tgagtttaat 1200
25 ttggtgcaaa gaatctattc aataatgagc aacaagagtg tgaagccaaa tagtgagact 1260
   ttcaggaaat ctataaattt gtgcataaga atcaaagata ctctagcaat gaagtatgaa 1320
   tatgctgttt ctacagAAC tttttgttat ccaaccataa caatttgcta tcaacctatt 1380
   ggttccgtgc agtttgaaagg tgcataaac atgcttggtg atttaaagaa tttcaatttg 1440
   gcaccgaact ctagcatgta caattctata atggcaggat atttccggga gaaaaaagt 1500
30 aacagtgcct taaaggtcct caaggaaatg aaagaagctg atgttaaacc tgattcggtg 1560
   acttttagct atctaataaa ctattgtggc gaggaggcga ctattgCGAA gtattataag 1620
   gagatgaagc aagcaggagt tgaagttaat aagcacgttt acatgtccct tgttaaagca 1680

```

tatgcatcat gcggacagtt tgaaaaggca aaacaggtgc tcatggacct ggaagttccg 1740
 gctaaggatc acaatgagct gaaaagtgtta cttatatctg ctctggcatc aaatggaaat 1800
 ataacagagg ctttaagtat ttatgaagaa atgaagaaac tccgctgccc cgtagagcca 1860
 aaagctatct tatctcttat agagaactct gattcaaacg cagagctggg aacgttggtt 1920
 5 gaactcactc atgagctgcg tgattctaaa ttttgatag atggtttctt caaaattatc 1980
 gtgtttgctg ttcgcaacaa tagatcaagc tctattctgg atttgttaga gcagaccaa 2040
 aaccacttgt ccaaggatga cgtcggtgtg gaatattggt ttgaagaggt gtttaagtca 2100
 atagcggaaa cagagtctag tgatgtgaag gtagggctgg acttggtgag ttttatgaaa 2160
 gaagagcttg agttatgtcc ttcaagaaaa tgtctggatt ttcttctaca tgcttggtgc 2220
 10 aatgccaaag ataagcagag cgctctactg gtatgggaag agtaccaatg tgcagaactt 2280
 ccttacaatg tcatcaacta tttaaggcaa gagctctttt atctcttaca tttaaaaggg 2340
 atgcccgtat tggtcaccgc ctcaaccgct cctgcccagc ttgagtaccc atccaggtca 2400
 cagttcaagg ctacaatgat atccggtgaa gtcaggtatg ttacctcgtc ttggtgcaaa 2460
 ccttacacaa agctcacctt aactctactg tttgttggtg aggtggttgc tattgtacca 2520
 15 gataagtcag atggaacttc aggtccgac attatgatct gcacggttcc taaaccatta 2580
 agcaagtga 2589
 <210> 6
 <211> 845
 <212> PRT
 20 <213> Arabidopsis thaliana
 <400> 6
 Met Ala Leu Lys Ala Lys Lys Ser Leu Ser Ala Ile Phe Arg Phe Ala
 1 5 10 15
 Met Arg Asn Ala Ala Lys Thr Ala Glu Val Glu Ala Lys Ala Val Ala
 25 20 25 30
 Gly Asp Ser Leu Ile Ser Asp Thr Gly Pro Val Ser Met Ser Leu Ile
 35 40 45
 Pro Asp Phe Leu Ser Ser Leu Lys Ser Pro Glu Asp Ser Ser Asn Ile
 50 55 60
 30 Gly Gln Asp Phe Gly Glu Pro Ala Leu Ala Gly Gln Val Ser Ser Ala
 65 70 75 80
 Leu Asp Val Lys Ser Ile Gly Tyr Asn Leu Leu Arg Glu Arg Asn Gly

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

	340	345	350
	Ile Val Glu Asn Met Cys Glu Ala Gly Leu Val Ile Ser Ala Asn Ile		
	355	360	365
	Leu His Ser Leu Leu Gln Ala Ile Glu Gln Ile Leu Glu Phe Asn Leu		
5	370	375	380
	Val Gln Arg Ile Tyr Ser Ile Met Ser Asn Lys Ser Val Lys Pro Asn		
	385	390	395
	Ser Glu Thr Phe Arg Lys Ser Ile Asn Leu Cys Ile Arg Ile Lys Asp		
	405	410	415
10	Thr Leu Ala Met Lys Tyr Glu Tyr Ala Val Ser His Arg Thr Phe Cys		
	420	425	430
	Tyr Pro Thr Ile Thr Ile Cys Tyr Gln Pro Ile Gly Ser Val Gln Phe		
	435	440	445
	Glu Gly Ala Tyr Asn Met Leu Gly Asn Leu Lys Asn Phe Asn Leu Ala		
15	450	455	460
	Pro Asn Ser Ser Met Tyr Asn Ser Ile Met Ala Gly Tyr Phe Arg Glu		
	465	470	475
	Lys Lys Val Asn Ser Ala Leu Lys Val Leu Lys Glu Met Lys Glu Ala		
	485	490	495
20	Asp Val Lys Pro Asp Ser Val Thr Phe Ser Tyr Leu Ile Asn Tyr Cys		
	500	505	510
	Gly Glu Glu Ala Thr Ile Ala Lys Tyr Tyr Lys Glu Met Lys Gln Ala		
	515	520	525
	Gly Val Glu Val Asn Lys His Val Tyr Met Ser Leu Val Lys Ala Tyr		
25	530	535	540
	Ala Ser Cys Gly Gln Phe Glu Lys Ala Lys Gln Val Leu Met Asp Leu		
	545	550	555
	Glu Val Pro Ala Lys Asp His Asn Glu Leu Lys Ser Val Leu Ile Ser		
	565	570	575
30	Ala Leu Ala Ser Asn Gly Asn Ile Thr Glu Ala Leu Ser Ile Tyr Glu		
	580	585	590
	Glu Met Lys Lys Leu Arg Cys Pro Val Glu Pro Lys Ala Ile Leu Ser		

	595	600	605
	Leu Ile Glu Asn Ser Asp Ser Asn Ala Glu Leu Gly Thr Leu Val Glu		
	610	615	620
	Leu Thr His Glu Leu Arg Asp Ser Lys Phe Trp Ile Asp Gly Phe Phe		
5	625	630	635 640
	Lys Ile Ile Val Phe Ala Val Arg Asn Asn Arg Ser Ser Ser Ile Leu		
	645	650	655
	Asp Leu Leu Glu Gln Thr Lys Asn His Leu Ser Lys Asp Asp Val Gly		
	660	665	670
10	Val Glu Tyr Trp Phe Glu Glu Val Phe Lys Ser Ile Ala Glu Thr Glu		
	675	680	685
	Ser Ser Asp Val Lys Val Gly Leu Asp Leu Val Ser Phe Met Lys Glu		
	690	695	700
	Glu Leu Glu Leu Cys Pro Ser Arg Lys Cys Leu Asp Phe Leu Leu His		
15	705	710	715 720
	Ala Cys Val Asn Ala Lys Asp Lys Gln Ser Ala Leu Leu Val Trp Glu		
	725	730	735
	Glu Tyr Gln Cys Ala Glu Leu Pro Tyr Asn Val Ile Asn Tyr Leu Arg		
	740	745	750
20	Gln Glu Leu Phe Tyr Leu Leu His Leu Lys Gly Met Pro Val Leu Phe		
	755	760	765
	Thr Ala Ser Thr Ala Pro Ala Gln Leu Glu Tyr Pro Ser Arg Ser Gln		
	770	775	780
	Phe Lys Ala Thr Met Ile Ser Gly Glu Val Arg Tyr Val Thr Ser Ser		
25	785	790	795 800
	Trp Cys Lys Pro Tyr Thr Lys Leu Thr Leu Thr Leu Leu Phe Val Val		
	805	810	815
	Glu Val Val Ala Ile Val Pro Asp Lys Ser Asp Gly Thr Ser Gly Pro		
	820	825	830
30	Ile Ile Met Ile Cys Thr Val Pro Lys Pro Leu Ser Lys		
	835	840	845
<210>	7		

<211> 1839
 <212> DNA
 <213> Arabidopsis thaliana
 <400> 7

```

5  atggggttttg aagatggacc aagatgtctt atgtctcatc cagctcaccc ttctcataac 60
   ttgtccgttc gatacagatt caactttcgt atcggttgct ttacatgtgg cgggaaagca 120
   gctgctacga cgccggattg tctctactac tactgcacca cctgcgaggt gacgttccac 180
   aatgggttgcc atcaacgtcc aagaaggata acacatcctt atcacctcca acatcctctt 240
   actctcttct atcgaaaccc cgaaaccaga gttatatcca acatcatccc tgatgcttgt 300
10 cctggtaaaa ttgaggatac atcaggtcca gagaagtatg aattcgtgga tatcgtcccg 360
   tataagtctg atatcatatt caacaaatgc acttggtgtg caaaagattt caaaggatgat 420
   tgggttttatc gttgtttgat ctgtagtttt tgcttggtatc tctcttggtgc ggcaactctt 480
   ccacttctta ctattacaaa cccgaaaagc catcaccact ctctcgtctt cttaccccgg 540
   ccattattag ttccatgtga tgcttggtggg ttggttgacg gggttgaacc aagttatgct 600
15 tgtttccaat gtaattacat ggttcatcag aattgcatag acttaccctg agtcataaaa 660
   atcacgcgtc acccgacccg gctctctcat actccttatt gttcatcttt gacttcgtca 720
   tgccaaatct gctataaaga gggtgatatc aagtatgggc aatattcttg ccatctccaa 780
   gattgttttt atgtagtcca ttctaaatgt gcaacacatg aaaatgtttg ggatgggaag 840
   gaactcgagt gggaaatcga gtctgatgaa actgaagata ttcaccttt taggaacct 900
20 ggtgatgggt tcataaagca tttttgtcac aaacaccgtt tgaagctcaa gaatcatgat 960
   ggtgctcgag acacagaaaa gcaatgtcga gcgtgcatat atccaattgt ttctaccaa 1020
   ttttaccatt gcaagaaatg caattactct ctccacgagg tatgtgctgg cctttctcga 1080
   aaactggatc atgcattgca caatcacact cttatcctat ctccatctcc cggaaagttt 1140
   tgttgttcag cttgttcccg agaatccact gggttcagtt acatatgctc taataaagg 1200
25 tgccaagatt ttgttctaga tgttcgatgc atctcagtgc tagagtattt catccataga 1260
   agtcacgaac atcccatttt tatctccaca tcctacaaca gtaaagacga gatcctctgc 1320
   aaagtttgca agaaaagatg tttgggggct catctacagt gtacgttatg cgagtttact 1380
   atgtgttact catgtgctat cattccagat gaaatacact acaaatttga caagcatcct 1440
   ctcacccttt cttgtgggga aagcgcggt aacacgtatt ggtgcgaggt gtgtgaaaaa 1500
30 caattggatc caaaagaatg gttctacaca tgtaacaaat gttgcatcac catccacctt 1560
   cattgcatat ttggatcttc tgtctttatg aaacctgggt ccatatttcg tgattattat 1620
   ggaaagggtc aagtttttcg aaacaatagt aacactcgac aactctgtta tatgtgtcac 1680

```

```

aaccgttgta ccggtttgat cttctacgaa ggctacagaa ggaatgctac atattactac 1740
aatcatagca atcgatcaac tcatagaatg attttttgct ctttggaatg tgagatatta 1800
gggatgagaa aacgtcgtct catccgtctt ccgaggtaa 1839
<210>      8
5  <211>    612
   <212>    PRT
   <213>    Arabidopsis thaliana
   <400>    8
Met Gly Phe Glu Asp Gly Pro Arg Cys Leu Met Ser His Pro Ala His
10  1              5              10              15
Pro Ser His Asn Leu Ser Val Arg Tyr Arg Phe Asn Phe Arg Ile Gly
      20              25              30
Cys Phe Thr Cys Gly Gly Lys Ala Ala Ala Thr Thr Pro Asp Cys Leu
      35              40              45
15  Tyr Tyr Tyr Cys Thr Thr Cys Glu Val Thr Phe His Asn Gly Cys His
      50              55              60
Gln Arg Pro Arg Arg Ile Thr His Pro Tyr His Leu Gln His Pro Leu
65              70              75              80
Thr Leu Phe Tyr Arg Asn Pro Glu Thr Arg Val Ile Ser Asn Ile Ile
20              85              90              95
Pro Asp Ala Cys Pro Gly Lys Ile Glu Asp Thr Ser Gly Pro Glu Lys
      100              105              110
Tyr Glu Phe Val Asp Ile Val Pro Tyr Lys Ser Asp Ile Ile Phe Asn
      115              120              125
25  Lys Cys Thr Trp Cys Ala Lys Asp Phe Lys Gly Asp Trp Phe Tyr Arg
      130              135              140
Cys Leu Ile Cys Ser Phe Cys Leu Asp Leu Ser Cys Ala Ala Thr Leu
145              150              155              160
Pro Leu Leu Thr Ile Thr Asn Pro Lys Ser His His His Ser Leu Val
30              165              170              175
Phe Leu Pro Arg Pro Leu Leu Val Pro Cys Asp Ala Cys Gly Leu Val
      180              185              190

```

	Asp	Gly	Leu	Glu	Pro	Ser	Tyr	Ala	Cys	Phe	Gln	Cys	Asn	Tyr	Met	Val
	195							200					205			
	His	Gln	Asn	Cys	Ile	Asp	Leu	Pro	Arg	Val	Ile	Lys	Ile	Thr	Arg	His
	210						215					220				
5	Pro	His	Arg	Leu	Ser	His	Thr	Pro	Tyr	Cys	Ser	Ser	Leu	Thr	Ser	Ser
	225						230				235					240
	Cys	Gln	Ile	Cys	Tyr	Lys	Glu	Val	Asp	Ile	Lys	Tyr	Gly	Gln	Tyr	Ser
						245				250					255	
	Cys	His	Leu	Gln	Asp	Cys	Phe	Tyr	Val	Val	His	Ser	Lys	Cys	Ala	Thr
10				260					265					270		
	His	Glu	Asn	Val	Trp	Asp	Gly	Lys	Glu	Leu	Glu	Trp	Glu	Ile	Glu	Ser
			275					280					285			
	Asp	Glu	Thr	Glu	Asp	Ile	Ser	Pro	Phe	Arg	Asn	Leu	Gly	Asp	Gly	Phe
		290					295					300				
15	Ile	Lys	His	Phe	Cys	His	Lys	His	Arg	Leu	Lys	Leu	Lys	Asn	His	Asp
	305						310				315					320
	Gly	Ala	Arg	Asp	Thr	Glu	Lys	Gln	Cys	Arg	Ala	Cys	Ile	Tyr	Pro	Ile
						325				330					335	
	Val	Ser	His	Gln	Phe	Tyr	His	Cys	Lys	Lys	Cys	Asn	Tyr	Ser	Leu	His
20				340					345					350		
	Glu	Val	Cys	Ala	Gly	Leu	Ser	Arg	Lys	Leu	Asp	His	Ala	Leu	His	Asn
			355					360					365			
	His	Thr	Leu	Ile	Leu	Ser	Pro	Ser	Pro	Gly	Lys	Phe	Cys	Cys	Ser	Ala
		370					375					380				
25	Cys	Ser	Arg	Glu	Ser	Thr	Gly	Phe	Ser	Tyr	Ile	Cys	Ser	Asn	Lys	Gly
	385						390				395					400
	Cys	Gln	Asp	Phe	Val	Leu	Asp	Val	Arg	Cys	Ile	Ser	Val	Leu	Glu	Tyr
						405				410					415	
	Phe	Ile	His	Arg	Ser	His	Glu	His	Pro	Ile	Phe	Ile	Ser	Thr	Ser	Tyr
30				420					425					430		
	Asn	Ser	Lys	Asp	Glu	Ile	Leu	Cys	Lys	Val	Cys	Lys	Lys	Arg	Cys	Leu
				435				440					445			

Gly Ala His Leu Gln Cys Thr Leu Cys Glu Phe Thr Met Cys Tyr Ser
 450 455 460
 Cys Ala Ile Ile Pro Asp Glu Ile His Tyr Lys Phe Asp Lys His Pro
 465 470 475 480
 5 Leu Thr Leu Ser Cys Gly Glu Ser Ala Asp Asn Thr Tyr Trp Cys Glu
 485 490 495
 Val Cys Glu Lys Gln Leu Asp Pro Lys Glu Trp Phe Tyr Thr Cys Asn
 500 505 510
 10 Lys Cys Cys Ile Thr Ile His Leu His Cys Ile Phe Gly Ser Ser Val
 515 520 525
 Phe Met Lys Pro Gly Ser Ile Phe Arg Asp Tyr Tyr Gly Lys Val Gln
 530 535 540
 Val Phe Arg Asn Asn Ser Asn Thr Arg Gln Leu Cys Tyr Met Cys His
 545 550 555 560
 15 Asn Arg Cys Thr Gly Leu Ile Phe Tyr Glu Gly Tyr Arg Arg Asn Ala
 565 570 575
 Thr Tyr Tyr Tyr Asn His Ser Asn Arg Ser Thr His Arg Met Ile Phe
 580 585 590
 20 Cys Ser Leu Glu Cys Glu Ile Leu Gly Met Arg Lys Arg Arg Leu Ile
 595 600 605
 Arg Leu Pro Arg
 610
 <210> 9
 <211> 2037
 25 <212> DNA
 <213> Arabidopsis thaliana
 <400> 9
 tggtcccttt ggaggggttct tctcatcaac atcaaatttc agagttccat tgagtagcac 60
 agtgaatcag acgctctcta gatgccacct ctgcaacgag aagtatttgc aagaagtagc 120
 30 cgctgttctc aaggccggtt cgagtctttc tctggctgac aaatgttccg agaagtttagc 180
 cccatggcta cgggctattg agaccaaaga agacaaggga ataacaggca gcagtaagtg 240
 caggcttttag atgacgctaa tacatcagcc tcgcaaaccg ctgctctaca gaagaaatgg 300

gacaacatat gccaaagtat ccatcacact ccggcgtttc ctaaacttgg ttttcagtcg 360
 gtgagtcgc agttcccagt tcagactgag aagagtgtga gaactcctac aagctatttg 420
 gagacgccta aactgctgaa tccgccaatc tcaaagccaa aacctatgga ggatcttacg 480
 gcatcgggtga ctaaccgcac agtgagtttg cttttgagct gtgttactac agattttggg 540
 5 tttgggagtaa tctatgcac caaaaaccag gaatcaaaaa caacgaggga gaaaccgatg 600
 ctggtgactc taaactcttc tttagaacat acatatcaga aagatttcaa gtctctcaga 660
 gaaatactct ctcgtaaagt tgcctggcag accgaagctg taaatgccat aagccaaatt 720
 atctgcggat gcaaaaccga ctccacgcga agaaaccaag caagcggaat ttggctggct 780
 cttcttggac ccgataaagt ggggaagaag aaagtggcga tgactcttct tgaagtcttc 840
 10 tttggtggtga aagtcaatta catatgtgta gattttgggg cagagcattg ttcccttgat 900
 gacaaattca gaggcaaac agtggtggat tacgtaaccg gtgagttatc taggaaacca 960
 cactctgttg ttttactcga aaacgtggaa aaagctgagt tcccggatca gatgagattg 1020
 tctgaagctg tgagtacggg gaaaatccgt gatttgcag gaagagtgat tagtatgaaa 1080
 aatgtgattg ttgttgtagac gtctgggatt gccaaggata atgccactga ccatgttatt 1140
 15 aaacctgtga agtttcctga ggagcaagtt ctacgcgcga gaagctggaa actgcagata 1200
 aagctaggag atgctactaa atttggggta aataagagaa aatatgagct agaaacagcg 1260
 caacgtgcag tgaagggtgca acgttcatat ctggatctga atcttccagt gaatgaaaca 1320
 gaatttagcc ctgatcatga ggcagaggac agggacgctt ggttcgatga attcattgaa 1380
 aaagtagatg gaaaagtgac gttcaaaccg gttgatttcg atgagttagc caagaacatt 1440
 20 caagagaaga ttggttcaca ttttgagcgg tgctttggat ccgaaacaca tctagaactt 1500
 gataaagaag tgatccttca gattctggcg gcttcattgt catcattatc atcgggcgaa 1560
 gaagaaggga gaacaatagt tgatcagtg atgcaaacag ttcttgctcg aagctttgct 1620
 gaagcaaac agaagtacgg ttcgaatccc atgttggcg tgaagctggt tgcttcttct 1680
 agcggcttag cttccggagt agaattgccg gcgaagggtg atgtgatatg gtgacatgca 1740
 25 tatatataaa tgaggtaata ataaagagtt ataactctca gattgttttt ctcatggaaa 1800
 gttttttgct tgtgtaaatt acaatgtaaa atggctaagc tatcacttgg tagatagggt 1860
 tgtttaggag taccaatttg ccacatcttt tttggcattt tgagtttttag gtatttctcc 1920
 taaaattttc cccaacttcc aaatttagat gaattaaacc atcttctctg taaccgagtt 1980
 tgttgtatat tatgttacgt ttgtcgaaat atagaaacga tgctttccac gccgatc 2037
 30 <210> 10
 <211> 422
 <212> PRT

<213> Arabidopsis thaliana

<400> 10

```

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

```

Trp Lys Leu Gln Ile Lys Leu Gly Asp Ala Thr Lys Phe Gly Val Asn
 245 250 255
 Lys Arg Lys Tyr Glu Leu Glu Thr Ala Gln Arg Ala Val Lys Val Gln
 260 265 270
 5 Arg Ser Tyr Leu Asp Leu Asn Leu Pro Val Asn Glu Thr Glu Phe Ser
 275 280 285
 Pro Asp His Glu Ala Glu Asp Arg Asp Ala Trp Phe Asp Glu Phe Ile
 290 295 300
 Glu Lys Val Asp Gly Lys Val Thr Phe Lys Pro Val Asp Phe Asp Glu
 10 305 310 315 320
 Leu Ala Lys Asn Ile Gln Glu Lys Ile Gly Ser His Phe Glu Arg Cys
 325 330 335
 Phe Gly Ser Glu Thr His Leu Glu Leu Asp Lys Glu Val Ile Leu Gln
 340 345 350
 15 Ile Leu Ala Ala Ser Trp Ser Ser Leu Ser Ser Gly Glu Glu Glu Gly
 355 360 365
 Arg Thr Ile Val Asp Gln Trp Met Gln Thr Val Leu Ala Arg Ser Phe
 370 375 380
 Ala Glu Ala Lys Gln Lys Tyr Gly Ser Asn Pro Met Leu Gly Val Lys
 20 385 390 395 400
 Leu Val Ala Ser Ser Ser Gly Leu Ala Ser Gly Val Glu Leu Pro Ala
 405 410 415
 Lys Val Asp Val Ile Trp
 420
 25 <210> 11
 <211> 1496
 <212> DNA
 <213> Arabidopsis thaliana
 <400> 11
 30 tgaaaagttc aaatctttgt ttgaaaact ctcaaagtgt taatcttttg tagctgtttc 60
 catttttttga agcagaataa cagagagatt acatgtttgtg tgtgttttaga aggaagcttg 120
 gttgcctttt ttcaagactt agatgggtta tcaagaaacg ggtgagggct agagttatag 180

tcaggagggtt caggaaagca aggtggagag ctcgaagaaa agagagtccg gaatctgaag 240
 tttcttctat ccatctgagc tcaaactctg gaagacatat tcgggttgca acgtttaacg 300
 tggctatggt ctcacttgcg cctgttgtagc aaaccatgga ggaaacagct tttcttggtc 360
 atctagatag cagcaacatc acatgccctt ctccaaaagg gattctgaag cagtctccat 420
 5 tacattcttc tgcagtcagg aaaccaaagg tttgcattaa tctcccagac aatgagattt 480
 ccttagcaca gagctacagc tttctgagta tggtagagaa tgataacgac ggggaaggaaa 540
 acagaggaag tctctcgatg agatctcccg tgtgtcttcc ctcttggttg tgggaccaag 600
 aaagctttaa cggatacagc agcagaagaa gcatagctga attgttaaga gagttggatg 660
 cagacatctt ggctttacaa gacgtgaaag cagaggaaga gacacttatg aaaccattat 720
 10 cagatttggc ttccgcctta ggtatgaaat acgtgtttgc agagagctgg ggcctgagt 780
 atggaaacgc catactatca aaatggccca taaagaagtg gagagttcag agaattgctg 840
 atgtagatga ttttaggaat gtgttgagg tgactgtaga gatcccatgg gctggagatg 900
 taaatgtgta ttgtactcaa cttgaccatc tagatgagaa ctggcgtatg aagcagatcg 960
 atgcaattac tcgaggagac gaatctcccc acatcctcct tggaggcttg aattctcttg 1020
 15 atggctctga ttattccata gcaagatgga accacatcgt aaagtattac gaggattctg 1080
 gaaagccgac gcctagagtt gaagtgatga ggtttctcaa gggaaaagga tacttggatt 1140
 caaaagaatt tgcaggagaa tgtgagcctg tagtcattat tgccaaagga caaaatgtac 1200
 aaggaacatg caagtacggg acacgggtcg attacattct ggcattctca gagtcacctt 1260
 acgaatttgt tcctggctcg tactcagttg tttcttccaa aggtacttct gatcatcata 1320
 20 tagtcaaagt cgatttagtt ataaccaaag agcgatctcg aggttaacttc aaacattcac 1380
 gcaagaaagc aaagcagaag atctttcaga taaaagctaa tttgatgtcg aaagatacat 1440
 ggaaactagg aaatctgatg tcttcataaa gtttggacac tgattaagaa acaaca 1496
 <210> 12
 <211> 458
 25 <212> PRT
 <213> Arabidopsis thaliana
 <400> 12
 Met Leu Cys Val Phe Arg Arg Lys Leu Gly Cys Leu Phe Ser Arg Leu
 1 5 10 15
 30 Arg Trp Val Ile Lys Lys Arg Val Arg Ala Arg Val Ile Val Arg Arg
 20 25 30
 Phe Arg Lys Ala Arg Trp Arg Ala Arg Arg Lys Glu Ser Pro Glu Ser

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

		290		295		300			
	Gly Leu Asn Ser Leu Asp Gly Ser Asp Tyr Ser Ile Ala Arg Trp Asn 305 310 315 320								
5	His Ile Val Lys Tyr Tyr Glu Asp Ser Gly Lys Pro Thr Pro Arg Val 325 330 335								
	Glu Val Met Arg Phe Leu Lys Gly Lys Gly Tyr Leu Asp Ser Lys Glu 340 345 350								
	Phe Ala Gly Glu Cys Glu Pro Val Val Ile Ile Ala Lys Gly Gln Asn 355 360 365								
10	Val Gln Gly Thr Cys Lys Tyr Gly Thr Arg Val Asp Tyr Ile Leu Ala 370 375 380								
	Ser Pro Glu Ser Pro Tyr Glu Phe Val Pro Gly Ser Tyr Ser Val Val 385 390 395 400								
	Ser Ser Lys Gly Thr Ser Asp His His Ile Val Lys Val Asp Leu Val 405 410 415								
15	Ile Thr Lys Glu Arg Ser Arg Gly Asn Phe Lys His Ser Arg Lys Lys 420 425 430								
	Ala Lys Gln Lys Ile Phe Gln Ile Lys Ala Asn Leu Met Ser Lys Asp 435 440 445								
20	Thr Trp Lys Leu Gly Asn Leu Met Ser Ser 450 455								
	<210>	13							
	<211>	1502							
	<212>	DNA							
25	<213>	Arabidopsis thaliana							
	<400>	13							
	tcctcattca tctctaacaa tggcgctctt tcgaatcatc atctttctccg ttctttcttct	60							
	ctccatcttc tcaacttcat catctgccca accatctttc cgccccaaag ctctccttct	120							
	cccagtaaca aaagaccat ctaccctcca gtacacaacc gtcatacaacc aacgcacacc	180							
30	tctcgtcccc gcttcggtcg tattcgacct cggtggtcga gaattctggg tcgactgtga	240							
	ccaaggctac gtctccacca cttaccgctc tcctcgctgc aactccgctg tctgctcacg	300							
	cgcgggctcc atcgcttgcg gcacatqctt ctctcctccg agacctggct qtagtaacaa	360							

```

cacttgcggc gcttttccag ataactccat caccggatgg gcaacctcgg gcgaatttgc 420
ttagacggtt gtgtctatcc agtcactaa cggatccaat ccgggtcggg tcgttaaaat 480
ccccaatcta atattcagtt gcggatcaac gtctcttctt aaaggactcg ctaaaggagc 540
cgttggtatg gctggaatgg gacgtcacia catcggctta ccgttgcagt tcgccgccgc 600
5 gtttagcttc aaccgcaaat tcgccgtgtg tcttacttcc ggtagaggcg tcgccttctt 660
cggcaacgga ccttacgttt tcctccccgg gatccagatc tcaaggcttc aaaagacacc 720
gcttctcatc aatccaggga ccactgtttt tgagttttca aaaggcgaga agtcgccgga 780
gtattttatc ggcgtgacgg cgattaagat cgtcgagaaa acacttcaa tcgatccaac 840
gcttttgaag ataaacgcaa gtactggaat aggaggaacc aaaatcagct ccgtcaatcc 900
10 ttacacggtg ttggagtcac cgatctacaa agcttttacg tcggagttaa ttagacaagc 960
agcggcgagg agcatcaaga gagtagcgtc ggtgaaaccg ttcggcgcggt gtttcagcac 1020
gaagaacggt ggcgtcacgc gcctaggata cgccgtaccg gagattcagc tcgtgcttca 1080
tagcaaagac gtcgtttgga gaatctttgg agctaactcg atggtgagtg tcagtgatga 1140
cgatcatctgt ttgggtttcg ttgacggagg agtcaacca ggagcctctg tggatgatcg 1200
15 agggttccag ttggaggata atttgatcga atttgatttg gcgagtaaca aatttggggt 1260
tagttccacg ttattgggcc gtcaaaactaa ctgcgccaac tttaatttca cttccactgc 1320
ttgattaaat ttagttgatg cactttctcc aaaataaatg tgtttcagta aactaaata 1380
aatgtattgt attatgttta atttgatcag gagccaatgg aatcatttct aagcattttt 1440
atcattgat ctgaataatt atgtaatgag tgagtatttg attttaaaaa tttaaattta 1500
20 gc 1502
<210> 14
<211> 434
<212> PRT
<213> Arabidopsis thaliana
25 <400> 14
Met Ala Ser Ser Arg Ile Ile Ile Phe Ser Val Leu Leu Leu Ser Ile
1 5 10 15
Phe Ser Leu Ser Ser Ser Ala Gln Pro Ser Phe Arg Pro Lys Ala Leu
20 25 30
30 Leu Leu Pro Val Thr Lys Asp Pro Ser Thr Leu Gln Tyr Thr Thr Val
35 40 45
Ile Asn Gln Arg Thr Pro Leu Val Pro Ala Ser Val Val Phe Asp Leu

```

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

	305		310		315		320	
	Arg Val Ala Ser Val Lys Pro Phe Gly Ala Cys Phe Ser Thr Lys Asn							
		325		330		335		
5	Val Gly Val Thr Arg Leu Gly Tyr Ala Val Pro Glu Ile Gln Leu Val							
		340		345		350		
	Leu His Ser Lys Asp Val Val Trp Arg Ile Phe Gly Ala Asn Ser Met							
		355		360		365		
	Val Ser Val Ser Asp Asp Val Ile Cys Leu Gly Phe Val Asp Gly Gly							
		370		375		380		
10	Val Asn Pro Gly Ala Ser Val Val Ile Gly Gly Phe Gln Leu Glu Asp							
		385		390		395		400
	Asn Leu Ile Glu Phe Asp Leu Ala Ser Asn Lys Phe Gly Phe Ser Ser							
		405		410		415		
	Thr Leu Leu Gly Arg Gln Thr Asn Cys Ala Asn Phe Asn Phe Thr Ser							
15		420		425		430		
	Thr Ala							
	<210>	15						
	<211>	2238						
	<212>	DNA						
20	<213>	Arabidopsis thaliana						
	<400>	15						
	atgattacaa gcttatctca aatctctttt gggacattga ggcgcttttg ctcttctgtt						60	
	ttaccttctg cactgaaacg cgaatttggt gaaggattgc gtacgttagt aaggtctggt						120	
	gatatccgtc gagctgtttc cttatttctat tctgcgccgg tcgagcttca gtctcaacag						180	
25	gcttacgcag ctctgttcca agcatgtgca gagcaacgca atctcctcga tggaatcaac						240	
	ttgcaccacc acatgctctc acatccctat tggtattcac agaacgttat acttgctaata						300	
	ttcctcatta atatgtatgc caaatgtggg aatatcttat acgcacgaca ggtgtttgat						360	
	acaatgcctg agaggaatgt tgtttcattg actgcattga tcacaggcta tgtacaggct						420	
	gggaatgagc aagagggatt ctgtcttttc tcgtctatgt tatcccattg tttccctaac						480	
30	gagtttacgc tgtcaagcgt gttaacttca tgtcgatatg agcctggaaa gcaggtacat						540	
	ggacttgctc tgaaacttgg ttacattgt tcgatctatg tggcgaatgc agttatctct						600	
	atgtacggta ggtgtcatga tqgtqccgca qcatatqaag cttgqactgt gtttgaagcc						660	

ataaagttta agaatttggt tacttggaac tcaatgattg ctgcttttca gtgctgcaat 720
 cttgggaaaa aggccattgg cgtttttatg cggatgcata gtgatggagt gggatttgat 780
 cgtgccacgt tgcttaatat atgttcttcc ttgtacaaaa gctctgactt agttcctaata 840
 gaggtttcaa agtggtgtct tcaactgcat tctctgactg taaaatccgg gttggtgact 900
 5 cagactgaag tggcta~~ctgc~~ attgattaaa gtttattcag aaatgttaga agactataact 960
 gactgctaca agctttttat ggaaatgagt cactgcaggg acattgttgc ttggaatggg 1020
 attataacag cgtttgcagt gtatgatccg gagagagcta ttcattctctt tggtcagtta 1080
 cgtcaggaga agctaagccc tgactggtat actttctcaa gtgtactaaa agcttgtgca 1140
 ggacttgtca ctgctcgcca tgctttgtct atacacgctc aagtgatcaa ggggtggtttt 1200
 10 ttggctgaca cgggtgctgaa caactcatta atccatgcgt atgccaagtg tggctctctt 1260
 gacttgtgta tgcgtgtatt tgacgatatg gattcgcggg atgttgtttc ctggaactca 1320
 atgctcaagg cttactcatt acatggccaa gtagactcaa ttctaccagt ctttcagaaa 1380
 atggacatca accccgattc agcgactttc atagcccttc tctcggcttg cagccacgca 1440
 ggacgagtag aagaaggatt gagaatattc agatccatgt ttgagaaacc tgaaactctt 1500
 15 cctcagctaa atcactacgc ttgtgtgatk gatatgctaa gccgagctga gcggttttgct 1560
 gaggcagaag aggttatcaa gcagatgcca atggaccag atgctgtagt ttggattgca 1620
 ttattgggat catgtagaaa gcacggtaac acccggttg gtaagttagc ggcagataaa 1680
 cttaaagagc tagtcgaacc cacaaactca atgagttaca tccagatgtc aaacatatac 1740
 aacgcagaag gtagctttaa cgaagccaat cttagcataa aggagatgga aacatggaga 1800
 20 gtacggaaag aaccagactt aagctggaca gagataggaa acaaggttca cgaattcgca 1860
 tcgggaggaa gacaccgtcc ggacaaggaa gccgtatata gagaattgaa gagactgatt 1920
 agttggttaa aggagatggg ttacgttcca gaaatgaggt cagcttcaca agatatagaa 1980
 gacgaagaac aggaggaaga taatttatta catcacagcg agaaacttgc gttagcgttt 2040
 gcggttatgg agggaagaaa atcgagtgat tgtggtgtga atttgatata gattatgaag 2100
 25 aacacaagga tctgtattga ttgccataac ttcataaat tagcttctaa gcttttggga 2160
 aaagagattc ttatgagaga ctcgaatcgg tttcaccatt tcaaagattc ttcttgctcc 2220
 tgtaatgact actggtaa 2238
 <210> 16
 <211> 745
 30 <212> PRT
 <213> Arabidopsis thaliana
 <400> 16

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

Val Gly Phe Asp Arg Ala Thr Leu Leu Asn Ile Cys Ser Ser Leu Tyr
 260 265 270
 Lys Ser Ser Asp Leu Val Pro Asn Glu Val Ser Lys Cys Cys Leu Gln
 275 280 285
 5 Leu His Ser Leu Thr Val Lys Ser Gly Leu Val Thr Gln Thr Glu Val
 290 295 300
 Ala Thr Ala Leu Ile Lys Val Tyr Ser Glu Met Leu Glu Asp Tyr Thr
 305 310 315 320
 Asp Cys Tyr Lys Leu Phe Met Glu Met Ser His Cys Arg Asp Ile Val
 10 325 330 335
 Ala Trp Asn Gly Ile Ile Thr Ala Phe Ala Val Tyr Asp Pro Glu Arg
 340 345 350
 Ala Ile His Leu Phe Gly Gln Leu Arg Gln Glu Lys Leu Ser Pro Asp
 355 360 365
 15 Trp Tyr Thr Phe Ser Ser Val Leu Lys Ala Cys Ala Gly Leu Val Thr
 370 375 380
 Ala Arg His Ala Leu Ser Ile His Ala Gln Val Ile Lys Gly Gly Phe
 385 390 395 400
 Leu Ala Asp Thr Val Leu Asn Asn Ser Leu Ile His Ala Tyr Ala Lys
 20 405 410 415
 Cys Gly Ser Leu Asp Leu Cys Met Arg Val Phe Asp Asp Met Asp Ser
 420 425 430
 Arg Asp Val Val Ser Trp Asn Ser Met Leu Lys Ala Tyr Ser Leu His
 435 440 445
 25 Gly Gln Val Asp Ser Ile Leu Pro Val Phe Gln Lys Met Asp Ile Asn
 450 455 460
 Pro Asp Ser Ala Thr Phe Ile Ala Leu Leu Ser Ala Cys Ser His Ala
 465 470 475 480
 Gly Arg Val Glu Glu Gly Leu Arg Ile Phe Arg Ser Met Phe Glu Lys
 30 485 490 495
 Pro Glu Thr Leu Pro Gln Leu Asn His Tyr Ala Cys Val Ile Asp Met
 500 505 510

Leu Ser Arg Ala Glu Arg Phe Ala Glu Ala Glu Glu Val Ile Lys Gln
 515 520 525
 Met Pro Met Asp Pro Asp Ala Val Val Trp Ile Ala Leu Leu Gly Ser
 530 535 540
 5 Cys Arg Lys His Gly Asn Thr Arg Leu Gly Lys Leu Ala Ala Asp Lys
 545 550 555 560
 Leu Lys Glu Leu Val Glu Pro Thr Asn Ser Met Ser Tyr Ile Gln Met
 565 570 575
 Ser Asn Ile Tyr Asn Ala Glu Gly Ser Phe Asn Glu Ala Asn Leu Ser
 10 580 585 590
 Ile Lys Glu Met Glu Thr Trp Arg Val Arg Lys Glu Pro Asp Leu Ser
 595 600 605
 Trp Thr Glu Ile Gly Asn Lys Val His Glu Phe Ala Ser Gly Gly Arg
 610 615 620
 15 His Arg Pro Asp Lys Glu Ala Val Tyr Arg Glu Leu Lys Arg Leu Ile
 625 630 635 640
 Ser Trp Leu Lys Glu Met Gly Tyr Val Pro Glu Met Arg Ser Ala Ser
 645 650 655
 Gln Asp Ile Glu Asp Glu Glu Gln Glu Glu Asp Asn Leu Leu His His
 20 660 665 670
 Ser Glu Lys Leu Ala Leu Ala Phe Ala Val Met Glu Gly Arg Lys Ser
 675 680 685
 Ser Asp Cys Gly Val Asn Leu Ile Gln Ile Met Lys Asn Thr Arg Ile
 690 695 700
 25 Cys Ile Asp Cys His Asn Phe Met Lys Leu Ala Ser Lys Leu Leu Gly
 705 710 715 720
 Lys Glu Ile Leu Met Arg Asp Ser Asn Arg Phe His His Phe Lys Asp
 725 730 735
 Ser Ser Cys Ser Cys Asn Asp Tyr Trp
 30 740 745
 <210> 17
 <211> 3093

<212> DNA

<213> *Arabidopsis thaliana*

<400> 17

atggcatcta gacaaggatc aaaatcgaga aaagcagggt taaaggggtgc tgattctact 60
 5 gcatcttcga caacctcgtc ttcgaagctt tatcaagaga cttctataga tggccatagc 120
 tcccctgctt cttcatctgc tcaaagcaag cagcagttct tctcaccaga tcccttacca 180
 caaactgctc agcgggtctaa agaaaacgtc acagtgacag ttcgctttcg cccactcagt 240
 ccaagggaaa ttcgccaagg ggaggaggtt gcatggtatg cagatgggga aactatcgtg 300
 cgaaatgagc ataatccgac aatagcttat gcctatgacc gtgtcttttg acctacaacc 360
 10 acaacacgca atgtctatga tattgctgca caccatgttg ttaatggggc tatggagggg 420
 ataaatggga ccattttcgc atatggagtg acaagcagtg gaaagacaca cacaatgcat 480
 ggtgaccaga gatctcctgg tattataccg ttagcagtga aagatgcttt cagcatcatc 540
 caagagactc ctaatcgaga atttcttctg cgtatttcct acatggaaat atataatgag 600
 gttgtcaatg atttgttgaa tccagcagga cacaatttga ggatcagaga agacaagcag 660
 15 ggaacctttg tcgaagggat taaagaagaa gttgttttat cccctgctca tgcgctttct 720
 cttattgcag ctggagaaga gcaacgacat gttggatcca caaattttaaa tttgctcagc 780
 agccggagcc atacaatatt tacgttgacg atagaaagta gtcccttagg tgacaagagt 840
 aaaggtgaag ctgtacacct ctcaactg aacctcggtg atctggcagg ttccgagagt 900
 tcaaagggtt aaaccagtgg tgtaagacgc aaagaaggat catatataaa taaaagtttg 960
 20 ctgacgttag ggactgtgat atcaaagctt acggatgtga gggcttcaca tgtaccatac 1020
 agagactcta agttaaccag gatccttcag tcctcattga gtggatcatga ccgagtatct 1080
 ctcatattgta ccgtcactcc tgcataagc agttcggaag aaacacacaa tacactgaaa 1140
 tttgctcatc gtgcaaagca tattgagatt caagctgaac aaaacaagat aattgatgag 1200
 aaatcattaa tcaagaagta ccaacgcgag attcgtcaac tgaaggaaga gttggaacag 1260
 25 cttaaacagg aaattgtacc agttcctcag ctgaaggata ttggtgcaga tgatattgtt 1320
 cttctgaagc agaagctaga agatgggtcaa gtcaagctgc aatctagact cgaagaagag 1380
 gaagaagcta aagcagctct cctgagtcga atccaaagat tgacgaaatt aattttggtg 1440
 tcgactaaaa atccacaagc atctcgatta cctcatcggt ttaatccgcg gaggagacat 1500
 tcattttgggg aagaagagct tgcttaccta ccatacaaga ggcgggacat gatggacgat 1560
 30 gagcaacttg acctgtatgt ttctgtggag ggaaatcatg agattagaga taatgcgtat 1620
 agagaagaaa aaaagaccg caagcacgga ttgttaaatt ggttaaagcc taagaaaaa 1680
 gatcacagtt caagtgccag cgaccagtca agtgtggtca aatctaacag cacaccatca 1740

actcctcaag ggggaggcag tcattctgcac acagaatcaa gactttcaga aggctcgcct 1800
 ttgatggaac aactctcaga gcctagggaa gacagagaag ctttagaaga cagttcccat 1860
 gaaatggaga ttccggagac tagcaataaa atgagcgatg agttggatct tttgagggaa 1920
 cagaaaaaga tattatctga agaggcggca ctgcaattaa gttcttttaa acggatgtca 1980
 5 gatgaagctg caaagtcccc tcaaatgag gagattaatg aggaaattaa agtactcaat 2040
 gacgacatca aagcaaagaa tgaccagatt gctacgttgg agagacagat aatggatttt 2100
 gttatgacat ctcattgaggc attggataaa tctgatatca tgcaggcagt tgctgagctg 2160
 agggatcaac ttaatgagaa atcttttgaa ctcgaggtga tcgaattggt tctgtttatg 2220
 caatttttct tcattctatgc agaagttaga gatggttttg ctattgcctg gataagatta 2280
 10 ttcatttttag ctcaatatta tctctctca tacctgaata aagtttcttc tgtgtatatt 2340
 ataaaatacc agtattcaaa acgccagact atattaacaa tgaagctaag ggttaaagct 2400
 gcagataatc gcatcattca gcaaacactc aatgaaaaga catgtgaatg tgaagtgttg 2460
 caagaagaag ttgcaaatct aaagcagcag ctctctgagg ccctggaact agcacaggga 2520
 actaagatta aagagctgaa acaagacgcc aaggaactaa gtgaatcaaa ggagcagtta 2580
 15 gaactcagaa acaggaaact tgcagaagag agttcatatg caaaaggctt tgcattcagca 2640
 gctgcagttg agctcaaggc attatccgaa gaagttgcaa agcttatgaa tcagaacgag 2700
 agactagcag ctgagctggc aacacagaag agcccaatcg cacaacgaaa taagacagga 2760
 acaacaacaa atgtaaggaa caatggaaga agagagagtc ttgcaaagag gcaagaacat 2820
 gacagcccat caatggagct gaagagagaa ctgaggatga gtaaggaaac tgagctatcg 2880
 20 tatgaagctg cacttggtga aaaagaacaa agagaagctg agcttgaaag gatattagaa 2940
 gaaacaaaac agagagaagc atatctggag aatgagcttg ctaatatgtg gggtcttggt 3000
 tctaagctga gaagatctca aggagcagat tcagaaatct ctgattctat atcagagacg 3060
 cgacaaacgg aacaaacgga aggttcattt tga 3093
 <210> 18
 25 <211> 1030
 <212> PRT
 <213> Arabidopsis thaliana
 <400> 18
 Met Ala Ser Arg Gln Gly Ser Lys Ser Arg Lys Ala Gly Leu Lys Gly
 30 1 5 10 15
 Ala Asp Ser Thr Ala Ser Ser Thr Thr Ser Ser Ser Lys Leu Tyr Gln
 20 25 30

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

Gln Leu Asn Leu Val Asp Leu Ala Gly Ser Glu Ser Ser Lys Val Glu
 290 295 300
 Thr Ser Gly Val Arg Arg Lys Glu Gly Ser Tyr Ile Asn Lys Ser Leu
 305 310 315 320
 5 Leu Thr Leu Gly Thr Val Ile Ser Lys Leu Thr Asp Val Arg Ala Ser
 325 330 335
 His Val Pro Tyr Arg Asp Ser Lys Leu Thr Arg Ile Leu Gln Ser Ser
 340 345 350
 Leu Ser Gly His Asp Arg Val Ser Leu Ile Cys Thr Val Thr Pro Ala
 10 355 360 365
 Ser Ser Ser Ser Glu Glu Thr His Asn Thr Leu Lys Phe Ala His Arg
 370 375 380
 Ala Lys His Ile Glu Ile Gln Ala Glu Gln Asn Lys Ile Ile Asp Glu
 385 390 395 400
 15 Lys Ser Leu Ile Lys Lys Tyr Gln Arg Glu Ile Arg Gln Leu Lys Glu
 405 410 415
 Glu Leu Glu Gln Leu Lys Gln Glu Ile Val Pro Val Pro Gln Leu Lys
 420 425 430
 Asp Ile Gly Ala Asp Asp Ile Val Leu Leu Lys Gln Lys Leu Glu Asp
 20 435 440 445
 Gly Gln Val Lys Leu Gln Ser Arg Leu Glu Glu Glu Glu Glu Ala Lys
 450 455 460
 Ala Ala Leu Leu Ser Arg Ile Gln Arg Leu Thr Lys Leu Ile Leu Val
 465 470 475 480
 25 Ser Thr Lys Asn Pro Gln Ala Ser Arg Leu Pro His Arg Phe Asn Pro
 485 490 495
 Arg Arg Arg His Ser Phe Gly Glu Glu Glu Leu Ala Tyr Leu Pro Tyr
 500 505 510
 Lys Arg Arg Asp Met Met Asp Asp Glu Gln Leu Asp Leu Tyr Val Ser
 30 515 520 525
 Val Glu Gly Asn His Glu Ile Arg Asp Asn Ala Tyr Arg Glu Glu Lys
 530 535 540

Lys Thr Arg Lys His Gly Leu Leu Asn Trp Leu Lys Pro Lys Lys Arg
 545 550 555 560
 Asp His Ser Ser Ser Ala Ser Asp Gln Ser Ser Val Val Lys Ser Asn
 565 570 575
 5 Ser Thr Pro Ser Thr Pro Gln Gly Gly Gly Ser His Leu His Thr Glu
 580 585 590
 Ser Arg Leu Ser Glu Gly Ser Pro Leu Met Glu Gln Leu Ser Glu Pro
 595 600 605
 Arg Glu Asp Arg Glu Ala Leu Glu Asp Ser Ser His Glu Met Glu Ile
 10 610 615 620
 Pro Glu Thr Ser Asn Lys Met Ser Asp Glu Leu Asp Leu Leu Arg Glu
 625 630 635 640
 Gln Lys Lys Ile Leu Ser Glu Glu Ala Ala Leu Gln Leu Ser Ser Leu
 645 650 655
 15 Lys Arg Met Ser Asp Glu Ala Ala Lys Ser Pro Gln Asn Glu Glu Ile
 660 665 670
 Asn Glu Glu Ile Lys Val Leu Asn Asp Asp Ile Lys Ala Lys Asn Asp
 675 680 685
 Gln Ile Ala Thr Leu Glu Arg Gln Ile Met Asp Phe Val Met Thr Ser
 20 690 695 700
 His Glu Ala Leu Asp Lys Ser Asp Ile Met Gln Ala Val Ala Glu Leu
 705 710 715 720
 Arg Asp Gln Leu Asn Glu Lys Ser Phe Glu Leu Glu Val Ile Glu Leu
 725 730 735
 25 Phe Leu Phe Met Gln Phe Phe Phe Ile Tyr Ala Glu Val Arg Asp Gly
 740 745 750
 Phe Ala Ile Ala Trp Ile Arg Leu Phe Ile Leu Ala Gln Tyr Tyr Leu
 755 760 765
 Ser Ser Tyr Leu Asn Lys Val Ser Ser Val Tyr Ile Ile Lys Tyr Gln
 30 770 775 780
 Tyr Ser Lys Arg Gln Thr Ile Leu Thr Met Lys Leu Met Val Lys Ala
 785 790 795 800

	Ala Asp Asn Arg Ile Ile Gln Gln Thr Leu Asn Glu Lys Thr Cys Glu		
	805	810	815
	Cys Glu Val Leu Gln Glu Glu Val Ala Asn Leu Lys Gln Gln Leu Ser		
	820	825	830
5	Glu Ala Leu Glu Leu Ala Gln Gly Thr Lys Ile Lys Glu Leu Lys Gln		
	835	840	845
	Asp Ala Lys Glu Leu Ser Glu Ser Lys Glu Gln Leu Glu Leu Arg Asn		
	850	855	860
	Arg Lys Leu Ala Glu Glu Ser Ser Tyr Ala Lys Gly Leu Ala Ser Ala		
10	865	870	875
	Ala Ala Val Glu Leu Lys Ala Leu Ser Glu Glu Val Ala Lys Leu Met		
	885	890	895
	Asn Gln Asn Glu Arg Leu Ala Ala Glu Leu Ala Thr Gln Lys Ser Pro		
	900	905	910
15	Ile Ala Gln Arg Asn Lys Thr Gly Thr Thr Thr Asn Val Arg Asn Asn		
	915	920	925
	Gly Arg Arg Glu Ser Leu Ala Lys Arg Gln Glu His Asp Ser Pro Ser		
	930	935	940
	Met Glu Leu Lys Arg Glu Leu Arg Met Ser Lys Glu Arg Glu Leu Ser		
20	945	950	955
	Tyr Glu Ala Ala Leu Gly Glu Lys Glu Gln Arg Glu Ala Glu Leu Glu		
	965	970	975
	Arg Ile Leu Glu Glu Thr Lys Gln Arg Glu Ala Tyr Leu Glu Asn Glu		
	980	985	990
25	Leu Ala Asn Met Trp Val Leu Val Ser Lys Leu Arg Arg Ser Gln		
	Gly		
	995	1000	1005
	Ala Asp Ser Glu Ile Ser Asp Ser Ile Ser Glu Thr Arg Gln Thr		
	1010	1015	1020
30	Glu Gln Thr Glu Gly Ser Phe		
	1025	1030	
<210>	19		

<211> 1638
 <212> DNA
 <213> Arabidopsis thaliana
 <400> 19

5 ttctaggcca aaactatctc cacgattcta acaaacaaaa attacagaaa tgtcaatgtc 60
 atcgtcctcc tttatctctc tcactttctt cttccttctt ctcgtttctt catcgtcatc 120
 ttccgatgac atctctgagc tattcgacga ctggtgccag aaacatggca aaacgtacgg 180
 ttcagaagaa gagaggcaac agaggattca aatctttaaa gataatcacg acttcgttac 240
 gcaacacaat ctcatcacta acgctactta ctctctctct ctcaatgctt tcgcggatct 300
 10 gactcaccac gagttcaagg cttctcgtct tggctctctt gtttccgctc cgtcgggtgat 360
 aatggctagt aaggacaga gtctcgggtg tagcgtaaag gttccagatt ctggtgattg 420
 gagaaagaag ggagctgtta ctaatgtcaa agatcaagga agctgcgggg catgttggtc 480
 gttctcggcg actggagcta tggaggaat taaccagata gtaacaggag atctcatcag 540
 cctctctgag caggaactga ttgattgtga caagtcctac aatgctggct gtaatggtgg 600
 15 tctcatggac tatgcttttg aatttgtcat taaaaccat ggaatagaca cagagaaaga 660
 ctatccttat caagaacgtg atggcacctg taagaaggat aagttgaaac aaaaggttgt 720
 gacaattgat agctatgctg gtgtaaaatc aaatgatgag aaagcgtaa tggaggctgt 780
 agcggctcag ccagttagt tggcatctg tggagcgag agagcgtttc agctatattc 840
 tagtggaata ttctctggcc cgtgttcgac atcattggat catgctgtgc tcatagtagg 900
 20 atatggttca cagaatggtg ttgattattg gattgtgaaa aactcatggg gaaaaagttg 960
 gggaatggat gggtttatgc acatgcagcg taacactgaa aattcagacg gcgtttgcgg 1020
 aatcaacatg cttgcttcat atcccatcaa gacacatcca aaccgcctc caccgtcccc 1080
 tcctggccct acaaaatgca accttttcac gtattgttca tctggagaga cttgttgctg 1140
 tgcaagagaa ttgtttggtt tatgcttctc atggaaatgc tgcgagatag aatccgctgt 1200
 25 gtgttgcaag gatggtcgtc attgttgtcc acatgattac cctgtttgtg atacaaccag 1260
 aagtctctgc ctaagaaaa ctggttaatt cactgcgatc aagcccttct ggaagaagaa 1320
 ttcttcaaaa caacttgcc gatttgagga atgggttatg tgagagaagt ttacgcaca 1380
 tgttacagaa aattcaaact catccataca aaacctcgt agattcgatt atctgcgagg 1440
 ggattatttg tagctgttat tgagatagat atatacgata tcatcacgga tgtattttta 1500
 30 gtctctcatt tggatgtata caacttctga gtcaataaag gacacttgca ggacaagcac 1560
 agtatttgta tctgcaagac acataaagta ataagagtat catctatgta tattatagag 1620
 agactaagct tcactggt 1638

<210> 20
 <211> 437
 <212> PRT
 <213> Arabidopsis thaliana
 5 <400> 20
 Met Ser Met Ser Ser Ser Ser Phe Ile Ser Leu Thr Phe Phe Phe Leu
 1 5 10 15
 Leu Leu Val Ser Ser Ser Ser Ser Ser Asp Asp Ile Ser Glu Leu Phe
 20 25 30
 10 Asp Asp Trp Cys Gln Lys His Gly Lys Thr Tyr Gly Ser Glu Glu Glu
 35 40 45
 Arg Gln Gln Arg Ile Gln Ile Phe Lys Asp Asn His Asp Phe Val Thr
 50 55 60
 Gln His Asn Leu Ile Thr Asn Ala Thr Tyr Ser Leu Ser Leu Asn Ala
 15 65 70 75 80
 Phe Ala Asp Leu Thr His His Glu Phe Lys Ala Ser Arg Leu Gly Leu
 85 90 95
 Ser Val Ser Ala Pro Ser Val Ile Met Ala Ser Lys Gly Gln Ser Leu
 100 105 110
 20 Gly Gly Ser Val Lys Val Pro Asp Ser Val Asp Trp Arg Lys Lys Gly
 115 120 125
 Ala Val Thr Asn Val Lys Asp Gln Gly Ser Cys Gly Ala Cys Trp Ser
 130 135 140
 Phe Ser Ala Thr Gly Ala Met Glu Gly Ile Asn Gln Ile Val Thr Gly
 25 145 150 155 160
 Asp Leu Ile Ser Leu Ser Glu Gln Glu Leu Ile Asp Cys Asp Lys Ser
 165 170 175
 Tyr Asn Ala Gly Cys Asn Gly Gly Leu Met Asp Tyr Ala Phe Glu Phe
 180 185 190
 30 Val Ile Lys Asn His Gly Ile Asp Thr Glu Lys Asp Tyr Pro Tyr Gln
 195 200 205
 Glu Arg Asp Gly Thr Cys Lys Lys Asp Lys Leu Lys Gln Lys Val Val

210 215 220
 Thr Ile Asp Ser Tyr Ala Gly Val Lys Ser Asn Asp Glu Lys Ala Leu
 225 230 235 240
 Met Glu Ala Val Ala Ala Gln Pro Val Ser Val Gly Ile Cys Gly Ser
 5 245 250 255
 Glu Arg Ala Phe Gln Leu Tyr Ser Ser Gly Ile Phe Ser Gly Pro Cys
 260 265 270
 Ser Thr Ser Leu Asp His Ala Val Leu Ile Val Gly Tyr Gly Ser Gln
 275 280 285
 10 Asn Gly Val Asp Tyr Trp Ile Val Lys Asn Ser Trp Gly Lys Ser Trp
 290 295 300
 Gly Met Asp Gly Phe Met His Met Gln Arg Asn Thr Glu Asn Ser Asp
 305 310 315 320
 Gly Val Cys Gly Ile Asn Met Leu Ala Ser Tyr Pro Ile Lys Thr His
 15 325 330 335
 Pro Asn Pro Pro Pro Ser Pro Pro Gly Pro Thr Lys Cys Asn Leu
 340 345 350
 Phe Thr Tyr Cys Ser Ser Gly Glu Thr Cys Cys Cys Ala Arg Glu Leu
 355 360 365
 20 Phe Gly Leu Cys Phe Ser Trp Lys Cys Cys Glu Ile Glu Ser Ala Val
 370 375 380
 Cys Cys Lys Asp Gly Arg His Cys Cys Pro His Asp Tyr Pro Val Cys
 385 390 395 400
 Asp Thr Thr Arg Ser Leu Cys Leu Lys Lys Thr Gly Asn Phe Thr Ala
 25 405 410 415
 Ile Lys Pro Phe Trp Lys Lys Asn Ser Ser Lys Gln Leu Gly Arg Phe
 420 425 430
 Glu Glu Trp Val Met
 435
 30 <210> 21
 <211> 1862
 <212> DNA

<213> Arabidopsis thaliana

<400> 21

```

atgaaatctg caaatntag tctttaata gaaatctgaa tttcatgttg ctgaaagaat 60
ttacaacgaa gttgtaaata ctctgagaaa catgtaagtc aggagagagt cccttctatg 120
5 gtttatatatt ggcggctttt aagaagcata cgcatatcaa gtgtttgatg aaattcctca 180
gagaagatga tgtccaccac agtcagatta aacagggttta cgtatctaac gagtactgca 240
aaactgacac gttacttttg ctctcaccat ttagttgacc gatcggaac ggcgttgcat 300
gaagttattc ggatagtttc ttctccagtt ggtggtctag atgacctga agaaaacttg 360
aaccagggtt ctgttctcc atcatcaaac ctgtttaccc aagtgattga gtcttgcaag 420
10 aacgagactt caccgagaag gctgttgagg ttttctcct ggtcgtgtaa aagcttgga 480
tctagtttac acgacaagga attcaactat gtcctaagag ttttggtga gaagaaggat 540
cacacagcta tgcagatact cttatcagat ctccaggaagg agaatcgggc gatggacaag 600
caaacgttta gcattgtcgc tgaaactctg gtaagggtg gtaaggaaga agatgcaata 660
ggtatcttca agatcttgga caaatcttct tgcccacaag atgggttcac tgtgacggct 720
15 attataagcg ctctttgttc cagaggacat gtgaagagag cattgggagt tatgcatcat 780
cacaaggatg taatatccgg gaatgaattg tctgtgtata gaagtctttt attcgggttg 840
tcggtacaaa gaaacgtgaa ggaagcaaga agagttatac aagatatgaa atcagcaggg 900
attacaccgg atttgttctg tttcaactcg ttgctgacct gcctttgcga aagaaacgtg 960
aatcgaaacc cgtctgggct tgttcctgaa gctttgaata ttatgttgga gatgaggctt 1020
20 taaaaatcc aacctacgtc tatgagttac aacattttgc tctcttgttt gggaaggact 1080
aggagggatg gagagtcctg ccagatttta gaacagatga aaagatcggg atgtgatccg 1140
gatacagggg gctattactt cgttgtgagg gtcctgtatt tgacaggaag gtttggtaaa 1200
ggcaacaaaa ttgtggatga gatgatcgag agaggattta gaccgaacg caaatcttat 1260
tatgatctga tcggggttct ttgtggggtc gagcgggtta actttgctct tcaattgttt 1320
25 gagaagatga agagaagctc agtgggtggt tatggtcagg tttatgatct gctcatacca 1380
aaactatgta aaggaggaaa ctttgagaaa ggaagagagc tttgggaaga ggcattgtca 1440
attgatgtta cacttagttg ctccattagt ttgcttgatc catctgtcac agagggtttt 1500
aaacctatga agatgaaaga agaagctgca atgggtggatc gtcgtgctct taacttgaag 1560
attcacgcta gaatgaacaa aacaaagcca aaactgaaac taaaaccgaa acgcaggagc 1620
30 aagacaaaga agaagaatct gcaacattga ttgatttctt ctaccattg ttaagaacat 1680
ctgatatagc aggcaatgta tcacaagaaa ctcaagtaag atgttctgct ttttatgtat 1740
tagtagtaat gcagaagcaa aggacatgag attcttttgt tgttctcttt gtccattga 1800

```

```

ttcttattac gaaaaatatg gggtctgaag aatacaaaac aagagacaag aatattgtga 1860
tc                                                                 1862

<210>      22
<211>      487
5  <212>      PRT
    <213>      Arabidopsis thaliana
    <400>      22

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

```

	195	200	205
	Ser Val Tyr Arg Ser Leu Leu Phe Gly Trp Ser Val Gln Arg Asn Val		
	210	215	220
	Lys Glu Ala Arg Arg Val Ile Gln Asp Met Lys Ser Ala Gly Ile Thr		
5	225	230	235 240
	Pro Asp Leu Phe Cys Phe Asn Ser Leu Leu Thr Cys Leu Cys Glu Arg		
	245	250	255
	Asn Val Asn Arg Asn Pro Ser Gly Leu Val Pro Glu Ala Leu Asn Ile		
	260	265	270
10	Met Leu Glu Met Arg Ser Tyr Lys Ile Gln Pro Thr Ser Met Ser Tyr		
	275	280	285
	Asn Ile Leu Leu Ser Cys Leu Gly Arg Thr Arg Arg Val Arg Glu Ser		
	290	295	300
	Cys Gln Ile Leu Glu Gln Met Lys Arg Ser Gly Cys Asp Pro Asp Thr		
15	305	310	315 320
	Gly Ser Tyr Tyr Phe Val Val Arg Val Leu Tyr Leu Thr Gly Arg Phe		
	325	330	335
	Gly Lys Gly Asn Gln Ile Val Asp Glu Met Ile Glu Arg Gly Phe Arg		
	340	345	350
20	Pro Glu Arg Lys Phe Tyr Tyr Asp Leu Ile Gly Val Leu Cys Gly Val		
	355	360	365
	Glu Arg Val Asn Phe Ala Leu Gln Leu Phe Glu Lys Met Lys Arg Ser		
	370	375	380
	Ser Val Gly Gly Tyr Gly Gln Val Tyr Asp Leu Leu Ile Pro Lys Leu		
25	385	390	395 400
	Cys Lys Gly Gly Asn Phe Glu Lys Gly Arg Glu Leu Trp Glu Glu Ala		
	405	410	415
	Leu Ser Ile Asp Val Thr Leu Ser Cys Ser Ile Ser Leu Leu Asp Pro		
	420	425	430
30	Ser Val Thr Glu Val Phe Lys Pro Met Lys Met Lys Glu Glu Ala Ala		
	435	440	445
	Met Val Asp Arg Arg Ala Leu Asn Leu Lys Ile His Ala Arg Met Asn		

	450		455		460	
	Lys Thr Lys Pro Lys Leu Lys Leu Lys Pro Lys Arg Arg Ser Lys Thr					
	465		470		475	480
	Lys Lys Lys Asn Leu Gln His					
5		485				
	<210>	23				
	<211>	1691				
	<212>	DNA				
	<213>	Arabidopsis thaliana				
10	<400>	23				
	agttatcttc atacttcata taaatctctg ctcatttttcg ataaaaagtt gaatcctttt	60				
	cacattttcc gctcaattcg atcttacatg ttaacaatgg ttggtctaac tccagctgag	120				
	gatcgttcag agattgcctt tttcgacgtc gagacgacaa ttccgttccg agttgggtcag	180				
	gggtacgcga ttctggagtt tggatcgatt ctggtttgcc ccaagaagct tgtggagctg	240				
15	aaaaactact ctgttctggt tcgtcctgct aatctcaacc tcatcacgcc tcggtccgtg	300				
	aagtgaacg gcatcaaacg cgaagacgtc gaatcagctc ttacctttgc agatatacgc	360				
	gacaccgtct acgatattct ccatggaagg atatgggcag gtcataacat attgaagttc	420				
	gattgtccaa ggataagaga agcatttgct gagattggac gagatccacc agagccaaag	480				
	ggcaccattg actccttggc tttgcttact caaaggtttg ggagaagagc tggatgatag	540				
20	aagatggcga ctttggctac atactttgga cttggaaacc agacgcatag gagtctggac	600				
	gatgtaagga tgaacttcga ggttctcaag tattgtgccca ctgttttggt cttggagtca	660				
	agtcttccag atgaacttat agagaactcg gttactacta ctactcctga aactagttca	720				
	agaaggcgta ggactataaa aaaatctcct ctgcaatctc ccactgatca gcagacagga	780				
	gagaacatga caactatacc tattttatca tttgtttcat ccgcagaagc tcaaacggat	840				
25	ccatttgata tgagcacttt gaggaacgaa atagcccctg aagttcttca atcagatggt	900				
	cccatggaag aagaacaaaa tcagcagtct gagactgttg cctctgaagg taccgggtgac	960				
	caggaaggat tcatggagct agataaaata tctgtttcaa gtatcagagc aactcatgtc	1020				
	ccgctatatg acgggagcca aacaatgaag ctgcaactct tccttgggga cagacccctt	1080				
	cagctccatt gtcctcgtct gaaagtacgg tttggaatca atgggaaatt tatggataaa	1140				
30	gctggaagac ggaggttgaa ctttgtaatt gatctgtatc caagtctctg caacgtactt	1200				
	caagaatgcg acagtgtctg acagacaata tctgttgatt caggcagtgg ttctgactgg	1260				
	aaccctctag tcattccaat gaaaggattt ctgaactgtc ccactgcaag gatacatata	1320				

```

cccacagagt taaatggaga tatagatcgg tacgcagctg agatacacca gaaagagttc 1380
tctggagcta ctgctactca gaaactcatt tctagcaacc ccaaggctga agagattgag 1440
tctttgttga atccgagaac tggtcttgac gctttcttgt ctctggagcc atatgattat 1500
cagcagagag ccggaatccg tctagttgct agaaagcttg ttatacatta gtggacctta 1560
5 ctcaatagca ccacagagaa ttaaagaaat cagagtttag tggactgtag agatttattt 1620
atgtagtcat gaaagaaaca cagaactggt gttcctgaca gtagatatac agctgattat 1680
gttcttagtt t 1691
<210>      24
<211>      487
10 <212>      PRT
    <213>      Arabidopsis thaliana
    <400>      24
Met Leu Thr Met Val Gly Leu Thr Pro Ala Glu Asp Arg Ser Glu Ile
1              5              10              15
15 Ala Phe Phe Asp Val Glu Thr Thr Ile Pro Phe Arg Val Gly Gln Gly
    20              25              30
Tyr Ala Ile Leu Glu Phe Gly Ser Ile Leu Val Cys Pro Lys Lys Leu
    35              40              45
Val Glu Leu Lys Asn Tyr Ser Val Leu Val Arg Pro Ala Asn Leu Asn
20    50              55              60
Leu Ile Thr Pro Arg Ser Val Lys Cys Asn Gly Ile Lys Arg Glu Asp
65              70              75              80
Val Glu Ser Ala Leu Thr Phe Ala Asp Ile Ala Asp Thr Val Tyr Asp
    85              90              95
25 Ile Leu His Gly Arg Ile Trp Ala Gly His Asn Ile Leu Lys Phe Asp
    100             105             110
Cys Pro Arg Ile Arg Glu Ala Phe Ala Glu Ile Gly Arg Asp Pro Pro
    115             120             125
Glu Pro Lys Gly Thr Ile Asp Ser Leu Ala Leu Leu Thr Gln Arg Phe
30    130             135             140
Gly Arg Arg Ala Gly Asp Met Lys Met Ala Thr Leu Ala Thr Tyr Phe
145             150             155             160

```

Gly Leu Gly Asn Gln Thr His Arg Ser Leu Asp Asp Val Arg Met Asn
 165 170 175
 Phe Glu Val Leu Lys Tyr Cys Ala Thr Val Leu Phe Leu Glu Ser Ser
 180 185 190
 5 Leu Pro Asp Glu Leu Ile Glu Asn Ser Val Thr Thr Thr Thr Pro Glu
 195 200 205
 Thr Ser Ser Arg Arg Arg Arg Thr Ile Lys Lys Ser Pro Leu Gln Ser
 210 215 220
 Pro Thr Asp Gln Gln Thr Gly Glu Asn Met Thr Thr Ile Pro Ile Leu
 10 225 230 235 240
 Ser Phe Val Ser Ser Ala Glu Ala Gln Thr Asp Pro Phe Asp Met Ser
 245 250 255
 Thr Leu Arg Asn Glu Ile Ala Pro Glu Val Leu Gln Ser Asp Val Pro
 260 265 270
 15 Met Glu Glu Glu Gln Asn Gln Gln Ser Glu Thr Val Ala Ser Glu Gly
 275 280 285
 Thr Gly Asp Gln Glu Gly Phe Met Glu Leu Asp Lys Ile Ser Val Ser
 290 295 300
 Ser Ile Arg Ala Thr His Val Pro Leu Tyr Asp Gly Ser Gln Thr Met
 20 305 310 315 320
 Lys Leu Gln Leu Phe Leu Gly Asp Arg Pro Leu Gln Leu His Cys Pro
 325 330 335
 Arg Leu Lys Val Arg Phe Gly Ile Asn Gly Lys Phe Met Asp Lys Ala
 340 345 350
 25 Gly Arg Arg Arg Leu Asn Phe Val Ile Asp Leu Tyr Pro Ser Leu Cys
 355 360 365
 Asn Val Leu Gln Glu Cys Asp Ser Ala Ala Gln Thr Ile Ser Val Asp
 370 375 380
 Ser Gly Ser Gly Ser Asp Trp Asn Pro Leu Val Ile Pro Met Lys Gly
 30 385 390 395 400
 Phe Leu Asn Cys Pro Thr Ala Arg Ile His Ile Pro Thr Glu Leu Asn
 405 410 415

Gly Asp Ile Asp Arg Tyr Ala Ala Glu Ile His Gln Lys Glu Phe Ser
 420 425 430
 Gly Ala Thr Ala Thr Gln Lys Leu Ile Ser Ser Asn Pro Lys Ala Glu
 435 440 445
 5 Glu Ile Glu Ser Leu Leu Asn Pro Arg Thr Val Leu Asp Ala Phe Leu
 450 455 460
 Ser Leu Glu Pro Tyr Asp Tyr Gln Gln Arg Ala Gly Ile Arg Leu Val
 465 470 475 480
 Ala Arg Lys Leu Val Ile His
 10 485
 <210> 25
 <211> 2179
 <212> DNA
 <213> Arabidopsis thaliana
 15 <400> 25
 gaagagacat cagatcttat caaaacccta gctttcgggtg tctctctcga tccttttagct 60
 caatcgaaaa ctgaatttag gagctgcgat ggggtccaga tctcgtaatg ataatttcca 120
 gtccggcgat ggcgctagtc ctgggaaaat attcattgga ggattgcaca aagacacgac 180
 taataactgtg ttcaataagc actttggcaa gtacggggag attacagatt ctgtcattat 240
 20 gcgagatcga catactggac aacctagggg tttcggtttc atcaccttcg ctgacccttc 300
 tgttgttgac aaagttatcg aagatactca tgtcatcaat ggcaaacagg ttgagatcaa 360
 aaggactatt ccgaaagggtg ctggtggcaa tcagtcaaag gacatcaaaa caaagaagat 420
 tttcgttggt ggtataccct caacggttac agaagatgaa cttaaggact tctttgccaa 480
 gtatgggaac gtggtggagc accagggttat tcgagaccat gaaacaaaca ggtccagagg 540
 25 ttttggtttt gtaatttttg atagcgaaga agtcgtagat gaattattgt ccaaaggaaa 600
 catgatcgat atggcagaca ctgaggaag cttgtacaaa tggggattag aattttggtt 660
 catgatgaca tatttatctt ggtttcaatt tgatattttc tccttggtcc tttcattata 720
 cacgtttaac ttagttagaa ttctaccac tagcttacca aatgtaagga gcttgtcaca 780
 gtaagatagg aaatccctgc tatagtttgg gtccttagca tagctagggt aattaagttg 840
 30 ctttggttga ggtggagatc aagaaagctg aaccaaagaa atctctaaac cgttcacctc 900
 cttcttacgg tagtcaccct agaggtcgtt cttctaacga tagttatgca agttatgggtg 960
 gaccttacgg aggttttgat ggagggtatg gtcacctcc aggtccatt aggtcacatg 1020

gaggtcctgc tagtagatat gctgggggat atgggtatgg tcgtggcagt gtaggccctg 1080
 aatttgagg aggttataat aattatgggg gtggtagttt aggaggttac cggaatgaac 1140
 cacccttgg ctattctagt cgttttggac cctatgggag tgggtttggt ggtgagggat 1200
 atggccgtgg aggagaagga gcctatttgg gatatcctcg tggaggaggg gaaggctatg 1260
 5 ggggatatgg tgggcctggt tacggcggtg catatgaatc tggtgccca ggtggcagtt 1320
 atgaaggagc aggtggtcca tatgggaggg gttatagtag cagtagtcga taccatccgt 1380
 atgcaaggta gggtcgaggt aatagcaatc agttggagct atagaagctt ctctcacagg 1440
 ttcctctaga ttgttccaat tgctctaggt gcattctaga gttatcatgt agcgccaaca 1500
 tcaggaaggg ttgctaagga caggaagagc atcaaaaagc ttttattagt catcttttag 1560
 10 ttgtgtgttt acctaattta agtggtgtct tatctctata acttagattg tgttatgtag 1620
 acttttagcat atttgggctt ggagtttggt tctcactggc acaagaatga tttatcattt 1680
 ggtagtcttg ttagatgatac aattactcgc ttagcttagt tttcacaatc atacgaatta 1740
 gcttaagaag ctctgcaaaa tgtgcactca tctattcttt agagctcaag ttgaatgctt 1800
 tatgtggtct cttgtatctc tgtacatttt gttgtgctga acaaagggtt gacggcagtt 1860
 15 agtgcattgt taagtaggca ttttttgtc agctgcatac gtaggttcta gctacaaatt 1920
 ggaattgtat gtgtttgaaa acgtaaaatc aattataaat gtttttgaaa atcatctata 1980
 caagcatggt tatttggttc ttgtttcctt attttttcgg gtaagggcca cggataattg 2040
 gttagacatt ggttctaata atgtgggttt tgttatctgt aatttctcgt caaatgatg 2100
 tactgtatca ttaaatgaat atcacttatt gtacatgctt ttataagtta attgaaatcc 2160
 20 aaactctttt acgttttaa 2179
 <210> 26
 <211> 231
 <212> PRT
 <213> Arabidopsis thaliana
 25 <400> 26
 Met Gly Ser Arg Ser Arg Asn Asp Asn Phe Gln Ser Gly Asp Gly Ala
 1 5 10 15
 Ser Pro Gly Lys Ile Phe Ile Gly Gly Leu His Lys Asp Thr Thr Asn
 20 25 30
 30 Thr Val Phe Asn Lys His Phe Gly Lys Tyr Gly Glu Ile Thr Asp Ser
 35 40 45
 Val Ile Met Arg Asp Arg His Thr Gly Gln Pro Arg Gly Phe Gly Phe

	50		55		60	
	Ile Thr Phe Ala Asp Pro Ser Val Val Asp Lys Val Ile Glu Asp Thr					
	65		70		75	80
	His Val Ile Asn Gly Lys Gln Val Glu Ile Lys Arg Thr Ile Pro Lys					
5		85		90		95
	Gly Ala Gly Gly Asn Gln Ser Lys Asp Ile Lys Thr Lys Lys Ile Phe					
		100		105		110
	Val Gly Gly Ile Pro Ser Thr Val Thr Glu Asp Glu Leu Lys Asp Phe					
		115		120		125
10	Phe Ala Lys Tyr Gly Asn Val Val Glu His Gln Val Ile Arg Asp His					
		130		135		140
	Glu Thr Asn Arg Ser Arg Gly Phe Gly Phe Val Ile Phe Asp Ser Glu					
		145		150		155
	Glu Val Val Asp Glu Leu Leu Ser Lys Gly Asn Met Ile Asp Met Ala					
15		165		170		175
	Asp Thr Gln Val Ser Leu Tyr Lys Trp Gly Leu Glu Phe Trp Phe Met					
		180		185		190
	Met Thr Tyr Leu Ser Trp Phe Gln Phe Asp Ile Phe Ser Leu Phe Leu					
		195		200		205
20	Ser Leu Tyr Thr Phe Asn Leu Val Arg Ile Leu Pro Thr Ser Leu Pro					
		210		215		220
	Asn Val Arg Ser Leu Ser Gln					
		225		230		
	<210>	27				
25	<211>	1959				
	<212>	DNA				
	<213>	Arabidopsis thaliana				
	<400>	27				
	aattagatga agagacatca gatcttatca aaaccctagc tttcgggtgtc tctctcgatc				60	
30	ctttagctca atcgaaaact gaatttagga gctgcgatgg gttccagatc tcgtaatgat				120	
	aatttccagt ccggcgatgg cgctagtcct gggaaaatat tcattggagg attgcacaaa				180	
	gacacgacta atactgtgtt caataagcac tttggcaagt acggggagat tacagattct				240	

	gtcattatgc gagatcgaca tactggacaa cctagggggtt tcggtttcat caccttcgct	300
	gacccttctg ttgttgacaa agttatcgaa gatactcatg tcatcaatgg caaacagggtt	360
	gagatcaaaa ggactattcc gaaagggtgct ggtggcaatc agtcaaagga catcaaaaaca	420
	aagaagatfff tcgttggtgg tataccctca acggttacag aagatgaact taaggacttc	480
5	tttgccaagt atgggaacgt ggtggagcac cagggttattc gagaccatga aacaaacagg	540
	tccagagggtt ttggttttgt aatttttgat agcgaagaag tcgtagatga attattgtcc	600
	aaaggaaaca tgatcgatat ggcagacact cagggtggaga tcaagaaagc tgaaccaaag	660
	aaatctctaa accgttcacc tccttcttac ggtagtcacc ctagagggtcg ttcttctaac	720
	gatagttagt caagttatgg tggaccttac ggaggttttg atggagggtta tggatcatcct	780
10	ccagggtccca ttaggtcaca tggaggtcct gctagtagat atgctggggg atatgggtat	840
	ggtcgtggca gtgtaggccc tgaatttggg ggaggttata ataattatgg ggggtgtagt	900
	ttaggagggtt accggaatga accacccctt ggctattcta gtcgttttgg accctatggg	960
	agtgggtttg gtggtgaggg atatggccgt ggaggagaag gagcctatff gggatatcct	1020
	cgtggaggag ggggaaggcta tgggggatat ggtgggcctg gttacggcgg tgcatatgaa	1080
15	tctggtggcc cagggtggcag ttatgaagga gcagggtggtc catatgggag ggggttatagt	1140
	agcagtagtc gataccatcc gtatgcaagg tagggtcgag gtaatagcaa tcagttggag	1200
	ctatagaagc ttctctcaca gggtcctcta gattgttcca attgctctag gtgcattcta	1260
	gagttatcat gtagcgccaa catcaggaag gggtgctaag gacaggaaga gcatcaaaaa	1320
	gcttttatta gtcatttttt agttgtgtgt ttacctaat taagtgttgt cttatctcta	1380
20	taacttagat tgtgttatgt agactttagc atatttgggc ttggagtttg tttctcactg	1440
	gcacaagaat gatttatcat ttggtagtct tgtagatga tcaattactc gcttagctta	1500
	gttttcacaa tcatacgaat tagcttaaga agctctgcaa aatgtgcact catctattct	1560
	ttagagctca agttgaatgc tttatgtggt ctcttgatc tctgtacatt ttgttgtgct	1620
	gaacaaagggt ttgacggcag ttagtgcattg tgtaagtagg catatftttg tcagctgcat	1680
25	acgtagggtc tagctacaaa ttggaattgt atgtgtttga aaacgtaaaa tcaattataa	1740
	atgtttttga aaatcatcta tacaagcatg gttatttgggt tcttgtttcc ttattttttc	1800
	gggtaagggc cacggataat tggtagaca ttggttctaa taatgtgggt tttgttatct	1860
	gtaattttctc gtcaaatga tgtactgtat cattaaatga atatcactta ttgtacatgc	1920
	ttttataagt taattgaaat ccaaactctt ttacgtttt	1959
30	<210>	28
	<211>	358
	<212>	PRT

<213> *Arabidopsis thaliana*

<400> 28

```

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

```

Arg Tyr Ala Gly Gly Tyr Gly Tyr Gly Arg Gly Ser Val Gly Pro Glu
 245 250 255
 Phe Gly Gly Gly Tyr Asn Asn Tyr Gly Gly Gly Ser Leu Gly Gly Tyr
 260 265 270
 5 Arg Asn Glu Pro Pro Leu Gly Tyr Ser Ser Arg Phe Gly Pro Tyr Gly
 275 280 285
 Ser Gly Phe Gly Gly Glu Gly Tyr Gly Arg Gly Gly Glu Gly Ala Tyr
 290 295 300
 Leu Gly Tyr Pro Arg Gly Gly Gly Glu Gly Tyr Gly Gly Tyr Gly Gly
 10 305 310 315 320
 Pro Gly Tyr Gly Gly Ala Tyr Glu Ser Gly Gly Pro Gly Gly Ser Tyr
 325 330 335
 Glu Gly Ala Gly Gly Pro Tyr Gly Arg Gly Tyr Ser Ser Ser Ser Arg
 340 345 350
 15 Tyr His Pro Tyr Ala Arg
 355
 <210> 29
 <211> 500
 <212> DNA
 20 <213> Arabidopsis thaliana
 <400> 29
 gaaggaaaaa gaaaaaacac aaatggaagg gaaagaggaa gaagcaaaga ggaggatgat 60
 gatgaagatc aaacttaagg agacagaagc agacgtaagc caagacgaca ttaaaagaca 120
 agaagcttac aacatgtctc cacgcgtacg tcgtggtggt ggtggtggcg gtggctcggg 180
 25 aggtatgagc aagtcatcga gcgtgaggca aaactgttta tgcgcaccaa ctacacaccc 240
 tggctccttc aggtgccgtt accaccgtcg taacgctgga ctcgggatgt ctcgtggcac 300
 atccgttccc tctaaccttt ccatgttggc tgggtggagac tctaattctc ccaagtgatg 360
 attaatatat attatgtgtt tcttctcatc aatttaaaat tctcgtatgt gtctgtgtat 420
 atacaaaaat gcataagtcg agaaatctct ggatattcac tttgctgttt gcacaaatga 480
 30 aataaattca aatgatcgtg 500
 <210> 30
 <211> 111

```

<212>      PRT
<213>      Arabidopsis thaliana
<400>      30
Met Glu Gly Lys Glu Glu Glu Ala Lys Arg Arg Met Met Met Lys Ile
5  1              5              10              15
Lys Leu Lys Glu Thr Glu Ala Asp Val Ser Gln Asp Asp Ile Lys Arg
              20              25              30
Gln Glu Ala Tyr Asn Met Ser Pro Arg Val Arg Arg Gly Gly Gly Gly
              35              40              45
10 Gly Gly Gly Ser Val Gly Met Ser Lys Ser Ser Ser Val Arg Gln Asn
              50              55              60
Cys Leu Cys Ala Pro Thr Thr His Pro Gly Ser Phe Arg Cys Arg Tyr
65              70              75              80
His Arg Arg Asn Ala Gly Leu Gly Met Ser Arg Gly Thr Ser Val Pro
15              85              90              95
Ser Asn Leu Ser Met Leu Ala Gly Gly Asp Ser Asn Ser Pro Lys
              100              105              110

<210>      31
<211>      2328
20 <212>      DNA
<213>      Arabidopsis thaliana
<400>      31
atgaagagtt gcagaacctt aatttttgtg gcaataattc tcaacgggct gagtactttc 60
gttgacatg ccggtgctga gagcaagggt cacatagttt atttgggtga gaagcagcac 120
25 gatgatcccg agtttgtcac ggaatctcat catcggatgt tgtggtcact tcttggaagt 180
aaagaggatg cccatagttc aatgggtccat agttaccgac atggcttctc gggttttgca 240
gccaaactta ccaagtccca agccaagaag cttgcagatt tacctgaagt tggtcacgtc 300
acgccggata gcttctacca actagataca actcggactt gggactactt aggccctttcc 360
gtcgccaatc ctaaaaatct tctaaatgac actaatatgg gtgaagaagt tatcattggc 420
30 attgtagact caggagtttg gccagaatct gaagtcttta atgacaatgg gattggacca 480
gtgccaagcc actggaaagg aggctgtgta tcaggagaga atttcacttc ctctcagtgc 540
aacaaaaagc ttataggagc caaatacttc atcaatgggt ttcttgcgac gcacgaaagc 600

```

ttcaactcca cagaatcact tgactttatt tcccctagag accgtagcgg acatggcacg 660
 catgtagcca ccatagcggg tggttcctac gtgcctagca taagctacaa ggggtagct 720
 ggagggactg tgagaggtgg ggcaccacgt gctcgtatag cgatgtacaa agcttggtgg 780
 tatctagatc gttttgacat aaatacatgt tcactctgtg atatcttgaa agctatggat 840
 5 gaagctatgc atgatggtgt tgatgttttg tccttgctta tcggctatcg ttttccttac 900
 ttcccagaaa ctgatgttcg cgccgtgata gccacgggag cattccatgc tgttttaaaag 960
 ggtatcacag ttgtttgttc gggaggtaac tctggctctg cggctcagac tgtgggaaac 1020
 acggctccat ggattttgac agtggctgca actacttttag accgatcttt tcccacacct 1080
 atcacacttg ggaacaataa actgatattg ggtcaagcaa tgtacaccgg tccagaactt 1140
 10 ggcttcacta gcttggttta cccagagaat ccaggaaca gcaacgagag tttttctggt 1200
 gactgtgagc ttctatTTTT caactccaat catacaatgg cggggaaagt cgtgttatgc 1260
 ttcacaacat caacacgtta catcactgta tcaagtgtg tgtcttatgt gaaggaagct 1320
 ggtggtcttg gcgtaattgt agcaagaaac cccggcgaca atctttcacc atgtgaagat 1380
 gattttcctt gtgttgctgt tgactacgag cttgggacag atatactttt atacatacgt 1440
 15 tccacaggat tacctgttgt gaagatacaa ctttctaaaa cacttggttg acaaccagtg 1500
 ggcacaaagg ttgcagattt ctcatcaaga ggacctaat caattgagcc tgccatctc 1560
 aaaccggata tagcagcacc aggagtgagc atattagcgg ctacaactac caataaaact 1620
 ttcaacgacc gaggattcat tttcttgta gggacatcaa tggcagctcc taccatttca 1680
 ggagttgttg ctcttctcaa agctctgcac cgtgattggt ctctgctgc cattagatca 1740
 20 gccattgtca ccacagcttg gagaacagat ccatttgag agcagatatt tgcagaaggg 1800
 tcacctcgga agctagctga tccgtttgac tatggtggag gccttggtgaa tccagagaaa 1860
 gctgcaaaac ctggtcttgt atatgacttg ggacttgaag actatgttct ctacatgtgc 1920
 tctgttggtt acaacgagac atcaatctct caactgtcg gaaaaggaa agtttggtcg 1980
 aatcccaaac catctgttct tgattttaac ttgccttcca tcacaatccc aaacctcaaa 2040
 25 gatgaagtca ctctcaccag aacctcact aacgttgag aacttgagtc ggtctataaa 2100
 gttgtgatcg agccaccgat aggcattcaa gtgactgtga caccagagac gcttctgttt 2160
 aactctacaa ccaaaagggt ctcttttaaa gtcaaagttt cgaccacaca caaatcaac 2220
 acaggttact tctttggaag cttgacttgg agtgactctc tacataatgt gaccattcct 2280
 ctctctgtga gaactcaaat ttgcagaac tactacgatg aaaactga 2328
 30 <210> 32
 <211> 775
 <212> PRT

<213> Arabidopsis thaliana

<400> 32

```

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

```

	Gly	Gly	Thr	Val	Arg	Gly	Gly	Ala	Pro	Arg	Ala	Arg	Ile	Ala	Met	Tyr
					245					250					255	
	Lys	Ala	Cys	Trp	Tyr	Leu	Asp	Arg	Phe	Asp	Ile	Asn	Thr	Cys	Ser	Ser
					260				265					270		
5	Ala	Asp	Ile	Leu	Lys	Ala	Met	Asp	Glu	Ala	Met	His	Asp	Gly	Val	Asp
					275				280					285		
	Val	Leu	Ser	Leu	Ser	Ile	Gly	Tyr	Arg	Phe	Pro	Tyr	Phe	Pro	Glu	Thr
					290				295				300			
10	Asp	Val	Arg	Ala	Val	Ile	Ala	Thr	Gly	Ala	Phe	His	Ala	Val	Leu	Lys
	305					310					315					320
	Gly	Ile	Thr	Val	Val	Cys	Ser	Gly	Gly	Asn	Ser	Gly	Pro	Ala	Ala	Gln
					325					330					335	
	Thr	Val	Gly	Asn	Thr	Ala	Pro	Trp	Ile	Leu	Thr	Val	Ala	Ala	Thr	Thr
					340				345					350		
15	Leu	Asp	Arg	Ser	Phe	Pro	Thr	Pro	Ile	Thr	Leu	Gly	Asn	Asn	Lys	Leu
					355				360					365		
	Ile	Leu	Gly	Gln	Ala	Met	Tyr	Thr	Gly	Pro	Glu	Leu	Gly	Phe	Thr	Ser
					370				375				380			
20	Leu	Val	Tyr	Pro	Glu	Asn	Pro	Gly	Asn	Ser	Asn	Glu	Ser	Phe	Ser	Gly
	385					390					395					400
	Asp	Cys	Glu	Leu	Leu	Phe	Phe	Asn	Ser	Asn	His	Thr	Met	Ala	Gly	Lys
					405					410					415	
	Val	Val	Leu	Cys	Phe	Thr	Thr	Ser	Thr	Arg	Tyr	Ile	Thr	Val	Ser	Ser
					420					425				430		
25	Ala	Val	Ser	Tyr	Val	Lys	Glu	Ala	Gly	Gly	Leu	Gly	Val	Ile	Val	Ala
					435				440					445		
	Arg	Asn	Pro	Gly	Asp	Asn	Leu	Ser	Pro	Cys	Glu	Asp	Asp	Phe	Pro	Cys
					450				455				460			
30	Val	Ala	Val	Asp	Tyr	Glu	Leu	Gly	Thr	Asp	Ile	Leu	Leu	Tyr	Ile	Arg
	465					470					475					480
	Ser	Thr	Gly	Leu	Pro	Val	Val	Lys	Ile	Gln	Pro	Ser	Lys	Thr	Leu	Val
					485					490					495	

Gly Gln Pro Val Gly Thr Lys Val Ala Asp Phe Ser Ser Arg Gly Pro
 500 505 510
 Asn Ser Ile Glu Pro Ala Ile Leu Lys Pro Asp Ile Ala Ala Pro Gly
 515 520 525
 5 Val Ser Ile Leu Ala Ala Thr Thr Thr Asn Lys Thr Phe Asn Asp Arg
 530 535 540
 Gly Phe Ile Phe Leu Ser Gly Thr Ser Met Ala Ala Pro Thr Ile Ser
 545 550 555 560
 Gly Val Val Ala Leu Leu Lys Ala Leu His Arg Asp Trp Ser Pro Ala
 10 565 570 575
 Ala Ile Arg Ser Ala Ile Val Thr Thr Ala Trp Arg Thr Asp Pro Phe
 580 585 590
 Gly Glu Gln Ile Phe Ala Glu Gly Ser Pro Arg Lys Leu Ala Asp Pro
 595 600 605
 15 Phe Asp Tyr Gly Gly Gly Leu Val Asn Pro Glu Lys Ala Ala Lys Pro
 610 615 620
 Gly Leu Val Tyr Asp Leu Gly Leu Glu Asp Tyr Val Leu Tyr Met Cys
 625 630 635 640
 Ser Val Gly Tyr Asn Glu Thr Ser Ile Ser Gln Leu Val Gly Lys Gly
 20 645 650 655
 Thr Val Cys Ser Asn Pro Lys Pro Ser Val Leu Asp Phe Asn Leu Pro
 660 665 670
 Ser Ile Thr Ile Pro Asn Leu Lys Asp Glu Val Thr Leu Thr Arg Thr
 675 680 685
 25 Leu Thr Asn Val Gly Gln Leu Glu Ser Val Tyr Lys Val Val Ile Glu
 690 695 700
 Pro Pro Ile Gly Ile Gln Val Thr Val Thr Pro Glu Thr Leu Leu Phe
 705 710 715 720
 Asn Ser Thr Thr Lys Arg Val Ser Phe Lys Val Lys Val Ser Thr Thr
 30 725 730 735
 His Lys Ile Asn Thr Gly Tyr Phe Phe Gly Ser Leu Thr Trp Ser Asp
 740 745 750

Ser Leu His Asn Val Thr Ile Pro Leu Ser Val Arg Thr Gln Ile Leu
 755 760 765

Gln Asn Tyr Tyr Asp Glu Asn
 770 775

5 <210> 33
 <211> 1140
 <212> DNA
 <213> Arabidopsis thaliana
 <400> 33

10 atggaatatg gagaagaacc ctcgattaa agatttctaa tgcttcctga tgatttggtg 60
 tttaactgct tagctcgcgt ctcaagatta cattatccaa ctctctcatt agtctccaaa 120
 aaatttcgct ttcttcttgc ttcaaaggag ctttaccaaa cccgaatcct cttaggcgga 180
 acggagagtt gtctttacgt gtgcgtacga ctccataccg attctgaaca actacattgg 240
 ttcattatct accagggacc aaatagctcc aaaaaagttt tgggtcccgat ttcattctccc 300

15 aattttactt ccgcagccct gccgggtttt gtagtggtcg gtcattgagat atatgccatt 360
 ggcggtggat cagagaataa aaacgcttcg ataaatgcaa ctggatcaaa gacttataac 420
 gctttgtcta gcgttatggt tatggactct cgttctcaca catggcgtga agctccaagc 480
 atgcgggtgg cccgtgtggt tccatctgct tgtaccctcg atgggagaat atatgttacg 540
 ggaggctgcg aaaatctcaa ttcaatgaat tggatggaga tttttgatac gaagactcaa 600

20 acttgggagt ttttacagat ccctagcgag gaggtatgca aaggctctga gtacttaagc 660
 ataagttatc aaagaaccgt ctatgtaggg tctagggaaa aagatgtcac ttacaagatg 720
 cataaaggta aatggagagg ggcagacata tgcctgaatc atggatggag tctggatcca 780
 tcatcttggt gtgtgataga gaacgtgttc taccgttgca gtcttgggga tgtacgttgg 840
 tatgacctca agaaaagaga atgggcagct ttgaagggtt tggaaggact gcctactttt 900

25 actaattatt atagaaattt taaatcggcg gatcattgtg gaaaattggc gatttcttgg 960
 gaggagtatg tgttagtga tgacgaaaca aagatttggt gtgcggaaat tgcgattcaa 1020
 aaacgccaaa atggagaaat ttggggaacg cttgagtggg ttgacaatgt gtttatatcc 1080
 agtggaccaaa accgtcatgt ggatttatta gtgaatgctc ttactgcaac agtttggtga 1140

<210> 34

30 <211> 379
 <212> PRT
 <213> Arabidopsis thaliana

<400> 34
 Met Glu Tyr Gly Glu Glu Pro Ser Ile Lys Arg Phe Leu Met Leu Pro
 1 5 10 15
 Asp Asp Leu Val Phe Asn Cys Leu Ala Arg Val Ser Arg Leu His Tyr
 5 20 25 30
 Pro Thr Leu Ser Leu Val Ser Lys Lys Phe Arg Phe Leu Leu Ala Ser
 35 40 45
 Lys Glu Leu Tyr Gln Thr Arg Ile Leu Leu Gly Gly Thr Glu Ser Cys
 50 55 60
 10 Leu Tyr Val Cys Val Arg Leu His Thr Asp Ser Glu Gln Leu His Trp
 65 70 75 80
 Phe Ile Ile Tyr Gln Gly Pro Asn Ser Ser Lys Lys Val Leu Val Pro
 85 90 95
 Ile Ser Ser Pro Asn Phe Thr Ser Ala Ala Leu Pro Gly Phe Val Val
 15 100 105 110
 Val Gly His Glu Ile Tyr Ala Ile Gly Gly Gly Ser Glu Asn Lys Asn
 115 120 125
 Ala Ser Ile Asn Ala Thr Gly Ser Lys Thr Tyr Asn Ala Leu Ser Ser
 130 135 140
 20 Val Met Val Met Asp Ser Arg Ser His Thr Trp Arg Glu Ala Pro Ser
 145 150 155 160
 Met Arg Val Ala Arg Val Phe Pro Ser Ala Cys Thr Leu Asp Gly Arg
 165 170 175
 Ile Tyr Val Thr Gly Gly Cys Glu Asn Leu Asn Ser Met Asn Trp Met
 25 180 185 190
 Glu Ile Phe Asp Thr Lys Thr Gln Thr Trp Glu Phe Leu Gln Ile Pro
 195 200 205
 Ser Glu Glu Val Cys Lys Gly Ser Glu Tyr Leu Ser Ile Ser Tyr Gln
 210 215 220
 30 Arg Thr Val Tyr Val Gly Ser Arg Glu Lys Asp Val Thr Tyr Lys Met
 225 230 235 240
 His Lys Gly Lys Trp Arg Gly Ala Asp Ile Cys Leu Asn His Gly Trp

REIVINDICAÇÕES

1. Planta transgênica, que compreende um vetor de transformação de plantas que compreende uma sequência de nucleotídeos que codifica ou é complementar a uma sequência que codifica um polipeptídeo HIO
5 que compreende a sequência de aminoácidos que é apresentada na SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32 ou SEQ ID NO: 34 ou um ortólogo ou um parálo-
10 go da mesma, através do que a planta transgênica possui um fenótipo de maior conteúdo de óleo ou um fenótipo de maior qualidade de farinha, em relação às plantas de controle.

2. Planta transgênica de acordo com a reivindicação 1, que é selecionada do grupo que consiste em plantas de espécies de *Brassica*, in-
15 cluindo canola e semente de colza, soja, milho, girassol, algodão, cacau, açafrão, dendezeiro, coqueiro, linho, mamona, amendoim, trigo, aveia e arroz.

3. Parte da planta obtida da planta como definida na reivindicação 1.

20 4. Parte da planta de acordo com a reivindicação 3, que é uma semente.

5. Farinha, produto alimentício ou alimento produzido partindo da semente como definida na reivindicação 4.

25 6. Processo de produção de óleo que compreende o cultivo da planta transgênica como definida na reivindicação 1, e a recuperação do óleo da dita planta.

7. Processo de acordo com a reivindicação 6, em que o óleo é recuperado de uma semente da planta.

30 8. Processo de produção de maior conteúdo de óleo em uma planta, o dito processo compreendendo:

a) a introdução nas células progenitoras da planta de um vetor de transformação de plantas que compreende uma sequência de nucleotídeos

- b) que codifica ou é complementar a uma sequência que codifica um polipeptídeo HIO que compreende a sequência de aminoácidos que é apresentada na SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32 ou SEQ ID NO: 34 ou um ortólogo ou um parálogo da mesma e
- c) o cultivo das células progenitoras transformadas para produzir uma planta transgênica, em que a dita sequência de polinucleotídeo é expressa e a dita planta transgênica exibe um fenótipo de maior conteúdo de óleo em relação às plantas de controle.

9. Planta obtida através de um processo como definido na reivindicação 8.

10. Planta de acordo com a reivindicação 9, que é selecionada do grupo que consiste em plantas das espécies de *Brassica*, incluindo canola e semente de colza, soja, milho, girassol, algodão, cacau, açafrão, dendêzeiro, coqueiro, linho, mamona, amendoim, trigo, aveia e arroz.

11. Planta de acordo com a reivindicação 9, em que a planta é selecionada do grupo que consiste em uma planta crescida partindo das ditas células progenitoras, uma planta que é a progênie direta de uma planta crescida partindo das ditas células progenitoras e uma planta que é a progênie indireta de uma planta crescida partindo das ditas células progenitoras.

12. Processo de produção de uma planta que possui um maior conteúdo de óleo, que compreende a identificação de uma planta que possui um alelo em seu gene HIO que resulta no maior conteúdo de óleo comparada a plantas que não possuem o alelo e a produção de progênie da dita planta identificada, em que a progênie gerada herda o alelo e possui o fenótipo de alto conteúdo de óleo.

13. Processo de acordo com a reivindicação 12 que emprega metodologia de gene/QTL candidato.

14. Processo de acordo com a reivindicação 12 que emprega a metodologia de TILLING.

15. Produto alimentício, farinha, grão, alimento ou semente que compreende um polipeptídeo codificado pela sequência de ácido nucleico que é apresentada na SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31 ou SEQ ID NO: 33.

16. Produto alimentício, farinha, grão, alimento ou semente que compreende um polipeptídeo que compreende uma sequência de aminoácidos que é apresentada na SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32 ou SEQ ID NO: 34 ou um ortólogo ou um parálogo da mesma.

17. Planta transgênica de acordo com a reivindicação 1, em que um fenótipo de maior qualidade de farinha compreende um aumento na energia que pode ser metabolizada disponível na farinha produzida partindo das sementes da planta transgênica, em relação às plantas de controle.

18. Planta transgênica de acordo com a reivindicação 17, em que um aumento na energia que pode ser metabolizada disponível compreende um conteúdo alterado de proteínas e/ou de fibras nas sementes da planta transgênica.

19. Planta transgênica de acordo com a reivindicação 18, em que o conteúdo de proteínas é aumentado e/ou o conteúdo de fibras é diminuído.

20. Planta transgênica de acordo com a reivindicação 17, em que um aumento na energia que pode ser metabolizada disponível compreende um menor conteúdo de fibras nas sementes da planta transgênica.

21. Processo de produção de farinha, que compreende o cultivo da planta transgênica como definida na reivindicação 1, e a recuperação da farinha da planta,

produzindo assim a farinha.

22. Processo de acordo com a reivindicação 21, em que a fari-

nha é produzida partindo de sementes da planta.

23. Processo de produção de um fenótipo de maior qualidade de farinha em uma planta, o dito processo compreendendo:

5 a) a introdução nas células progenitoras da planta de um vetor de transformação de plantas que compreende uma sequência de nucleotídeos que codifica ou é complementar a uma sequência que codifica um polipeptídeo HIO que compreende a sequência de aminoácidos apresentada na SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ
10 ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32 ou SEQ ID NO: 34 ou um ortólogo ou um parálogo da mesma e

b) o cultivo das células progenitoras transformadas para produzir uma planta transgênica, em que a sequência de nucleotídeos é expressa e a
15 planta transgênica exibe um fenótipo de maior qualidade de farinha em relação às plantas de controle,

produzindo assim o fenótipo de maior qualidade de farinha na planta.

20 24. Planta obtida através de um processo como definido na reivindicação 23.

25 25. Planta de acordo com a reivindicação 24, que é selecionada do grupo que consiste em plantas das espécies de *Brassica*, incluindo canola e semente de colza, soja, milho, girassol, algodão, cacau, açafrão, dendezeiro, coqueiro, linho, mamona, amendoim, trigo, aveia e arroz.

26 26. Planta de acordo com a reivindicação 24, em que a planta é selecionada do grupo que consiste em uma planta crescida partindo das ditas células progenitoras, uma planta que é a progênie direta de uma planta crescida partindo das ditas células progenitoras e uma planta que é a progênie indireta de uma planta crescida partindo das ditas células progenitoras.

30 27. Processo de produção de uma planta que possui um fenótipo de maior qualidade de farinha que compreende a identificação de uma planta que possui um alelo em seu gene HIO que resulta na melhor qualidade de

farinha, comparada a plantas que não possuem o alelo e a produção da progênie da dita planta identificada, em que a progênie gerada herda o alelo e possui a qualidade de farinha melhorada,

- gerando assim uma planta que possui um fenótipo de maior qualidade de farinha.

28. Processo de acordo com a reivindicação 27 que emprega a metodologia de gene/QTL candidato.

29. Processo de acordo com a reivindicação 27 que emprega a metodologia de TILLING.

RESUMO

Patente de Invenção: **"PRODUÇÃO DE PLANTAS COM CONTEÚDO ALTERADO DE ÓLEOS, PROTEÍNAS OU FIBRAS"**.

- 5 A presente invenção refere-se a plantas que exibem um fenótipo de maior quantidade de óleo ou um fenótipo de maior qualidade de farinha devido à expressão alterada de um ácido nucleico de HIO. A invenção se direciona adicionalmente a processos de produção de plantas com um fenótipo de maior quantidade de óleo ou com um fenótipo de maior qualidade de farinha.