

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

(21) **PI0616756-0 A2**

(22) Data de Depósito: 22/09/2006
(43) Data da Publicação: 28/06/2011
(RPI 2112)



(51) *Int.Cl.:*
C12N 15/82 2006.01
A01H 5/00 2006.01

(54) Título: **CONSTRUÇÃO GENÉTICA PARA A HIPER-EXPRESSÃO DE UMA PROTEÍNA DE DOMÍNIO-F EM PLANTAS; CÉLULA DE PLANTA TRANSGÊNICA; CÉLULA DA PLANTA ARABIDOPSIS THALIANA RECOMBINANTE; PLANTA; MÉTODO DE MODIFICAÇÃO DA PRODUÇÃO DE ÓLEO EM UMA PLANTA; E MÉTODO DE AUMENTO NA PRODUÇÃO DE ÓLEO NA ARABIDOPSIS THALIANA**

(57) Resumo: CONSTRUÇÃO GENÉTICA PARA A HIPER-EXPRESSÃO DE UMA PROTEÍNA DE DOMÍNIO-E EM PLANTAS; CÉLULA DE PLANTA TRANSGÊNICA; CÉLULA DA PLANTA ARABIDOPSIS THALIANA RECOMBINANTE; PLANTA; MÉTODO DE MODIFICAÇÃO DA PRODUÇÃO DE ÓLEO EM UMA PLANTA; E MÉTODO DE AUMENTO NA PRODUÇÃO DE ÓLEO NA ARABIDOPSIS THALIANA. Descrevem-se construções genéticas, células de plantas transgênicas e plantas transgênicas, bem como os métodos associados para aumentar a produção de óleo em uma planta utilizando-se seqüências genéticas de domínio-F.

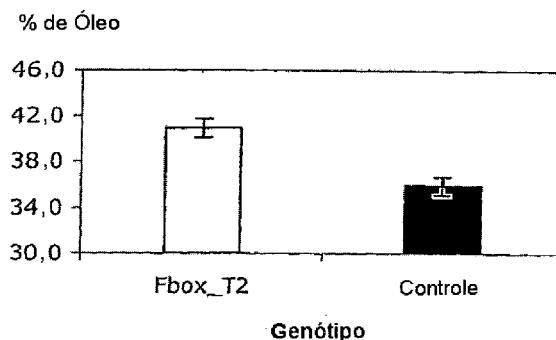
(73) Titular(es): Michigan State University

(72) Inventor(es): John B. Ohlrogge, Sari A. Ruuska, Yonghua Li

(74) Procurador(es): Trench, Rossi e Watanabe

(86) Pedido Internacional: PCT US2006037111 de 22/09/2006

(87) Publicação Internacional: WO 2007/038345 de 05/04/2007



PI.0616756-0

25/04/08

(1,2)
Nº de Cópias
(3)

"CONSTRUÇÃO GENÉTICA PARA A HIPER-EXPRESSÃO DE UMA
PROTEÍNA DE DOMÍNIO-F EM PLANTAS; CÉLULA DE PLANTA
TRANSGÊNICA; CÉLULA DA PLANTA *ARABIDOPSIS THALIANA*
RECOMBINANTE; PLANTA; MÉTODO DE MODIFICAÇÃO DA PRODUÇÃO DE
5 ÓLEO EM UMA PLANTA; E MÉTODO DE AUMENTO NA PRODUÇÃO DE ÓLEO
NA *ARABIDOPSIS THALIANA*".

REFERÊNCIA CRUZADA A PEDIDOS DE PATENTE AFINS

Este pedido de patente reivindica o benefício de
prioridade para o pedido de patente provisório nº US
10 60/720.424, depositado em 26 de setembro de 2005.

CAMPO TÉCNICO DA INVENÇÃO

Esta invenção refere-se ao campo de agricultura e
manipulação de plantas. Particularmente, esta invenção se
refere a métodos e composições para aumentar o teor de óleo
15 em plantas transgênicas, incluindo o teor de óleo de
sementes de *Arabidopsis thaliana*.

ANTECEDENTES DA INVENÇÃO

Os óleos vegetais representam uma fonte renovável
de carbono altamente reduzido. A produção mundial atual de
20 óleos vegetais é estimada em 87 milhões de toneladas
métricas com um valor aproximado de alguns 40 bilhões de
dólares norte-americanos. A maioria do óleo vegetal
atualmente vai diretamente para o consumo humano, e 25% da
ingestão calórica humana provém de ácidos graxos vegetais
25 (Broun et al., Ann. Rev. Nutr. 19:197-216 (1999)). Além da
sua importância na nutrição humana, os ácidos graxos
vegetais são também ingredientes importantes de produtos
não-alimentícios tais como sabões, detergentes,
lubrificantes, biocombustíveis cosméticos, e tintas (vide

Ohlrogge, *Plant Physiol.* 104:821-6 (1994)). Embora a demanda por óleos vegetais tenha aumentado regularmente, a capacidade de produção para atender a esta demanda é mais do que adequada e os preços dos óleos vegetais permaneceram
5 abaixo ou perto de 0,60 dólares norte-americanos por quilograma. Este baixo custo de produção estimulou interesse no uso de óleos vegetais como alternativas renováveis às matérias-primas químicas derivadas do petróleo.

10 Os ácidos graxos são a fonte mais abundante de cadeias de carbono reduzido disponíveis na natureza e têm diversos usos, desde alimentos a matérias-primas industriais. As plantas representam uma fonte renovável significativa de ácidos graxos porque muitas espécies os
15 acumulam na forma de triacil-glicerina como os principais componentes de armazenamento em sementes. Com o surgimento da tecnologia de transformação de plantas, a engenharia metabólica de ácidos graxos de sementes oleaginosas se tornou possível e os óleos de plantas transgênicas
20 representam alguns dos primeiros sucessos no desenho de produtos vegetais modificados. Por exemplo, a transferência do gene de tioesterase de uma planta da baía da Califórnia para sementes de plantas que não acumulam laurato (12:0), *Arabidopsis* and *Brassica napus* (colza) resultou na alteração
25 do processo de alongamento de cadeias acila de ácidos graxos, para produzir laurato até 24% e 58% de ácidos graxos totais das sementes, respectivamente (vide Voelker et al., *Sci.* 257:72-4 (1992); e Voelker et al., *Plant J.* 9:229-41 (1996), respectivamente). Assim sendo, a transferência de

um único gene para uma planta pode alterar dramaticamente o tipo de ácidos graxos produzidos.

Entretanto, esse sucesso com um único gene é a exceção à regra (para obter uma revisão, vide Thelen e Ohlrogge, *Metabol. Engineer.* 4:12-21 (2002)). Além disso, para ser economicamente útil para consumo humano e usos industriais, um aumento real no teor de ácidos graxos, ao invés de apenas uma mudança no tipo de ácido graxo produzido, seria altamente desejável. Embora a produção de malonil-CoA por acetil-CoA carboxilase seja uma etapa reguladora fundamental na síntese *de novo* de ácidos graxos, as tentativas para aumentar a taxa desta etapa aparentemente limitadora da taxa por engenharia genética atingiram no máximo um sucesso moderado. Além disso, a superexpressão de várias enzimas sintase de ácidos graxos individuais não resultou em um maior fluxo da biossíntese de ácidos graxos (revisado em Thelen e Ohlrogge, *Metabol. Engineer.* 4:12-21 (2002)).

Conseqüentemente, permanece existindo uma necessidade de se obter estratégias de engenharia genética, que aumentem a quantidade de ácido graxo produzida por plantas.

SUMÁRIO DA INVENÇÃO

A invenção baseia-se, em parte, na descoberta inesperada de que a superexpressão de genes de F-box em plantas resulta na maior produção de óleo vegetal. Particularmente, a superexpressão da proteína F-box de *Arabidopsis thaliana* (N^{os} de Acesso do GenBank NM_111499 (DNAC) e NP_566277 (proteína) usando o promotor faseolina

específico de semente, sementes produzidas com um fenótipo de teor mais alto de óleo (vide Figuras 1-5).

Conseqüentemente, em um aspecto, a invenção fornece uma construção genética para a expressão de uma
5 proteína F-box em plantas. Genericamente, tais construções genéticas da invenção incluem uma seqüência promotora de gene; e uma seqüência que codifica uma proteína F-box da planta, que está ligada funcionalmente à seqüência promotora do gene. Em algumas modalidades, a proteína F-box é uma
10 proteína F-box vegetal ou codificada por um gene de F-box da planta. Em outras modalidades, a proteína F-box é uma proteína F-box não-vegetal ou codificada por um gene de F-box não-vegetal, tal como uma proteína ou gene de F-box animal (por exemplo, um mamífero). Em algumas modalidades,
15 a seqüência promotora do gene é um promotor específico de semente. Em modalidades particularmente úteis, a seqüência promotora do gene é faseolina. Em certas modalidades particularmente úteis, a construção genética compreende um vetor de expressão pBBV-PHAS.

20 Em outra modalidade, as construções genéticas da invenção compreendem uma seqüência que codifica a proteína F-box da planta tem uma seqüência de polipeptídeo que PE pelo menos 75% idêntica à seqüência de polipeptídeo de SEQ ID NO: 1. Em outras modalidades, a proteína F-box da planta
25 tem uma seqüência de polipeptídeo que é pelo menos 90% idêntica à seqüência de polipeptídeo de SEQ ID NO: 1. Em ainda outras modalidades úteis, as construções genéticas da invenção compreendem uma proteína F-box vegetal que tem a seqüência de polipeptídeo de SEQ ID NO: 1. Em modalidades

específicas, a proteína F-box da planta é codificada pela sequência de ácidos nucleicos de SEQ ID NO: 2. Em outras modalidades, a proteína F-box é codificada por um ácido nucleico que hibridiza para a sequência de ácidos nucleicos de SEQ ID NO: 2. Em modalidades particularmente úteis, a proteína F-box é codificada por um ácido nucleico que hibridiza sob condições severas para a sequência de ácidos nucleicos de SEQ ID NO: 2.

Em outro aspecto, a invenção fornece células vegetais transgênicas que compreendem qualquer uma das construções genéticas da invenção descritas acima. Em certas modalidades úteis, a célula vegetal transgênica é uma *Arabidopsis thaliana*. Em outras modalidades, a célula vegetal transgênica é uma planta de cultivo agrícola, tal como milho ou trigo. Em outras modalidades, a célula vegetal transgênica é uma cultura agrícola produtora de óleo, tal como soja, palma, colza ou girassol.

Em outro aspecto, a invenção fornece uma célula vegetal de *Arabiposis thaliana* recombinante, que tem uma construção genética heteróloga que inclui uma sequência promotora de gene ligada funcionalmente a uma sequência codificadora de F-box. Em certas modalidades particularmente úteis, a sequência de F-box codificada corresponde ao polipeptídeo de F-box vegetal de SEQ ID NO: 1 e o promotor é faseolina.

Em ainda outro aspecto, a invenção fornece uma planta que tem uma célula vegetal transgênica da invenção, como descrito acima, ou uma sua parte, propágulo ou

progênie, compreendendo uma das construções genéticas da invenção.

Em ainda outro aspecto, a invenção fornece um método para modificar a produção de óleo em uma planta, primeiramente, incorporando de forma estável no genoma da planta uma construção genética da invenção, como descrito acima, pra produzir uma planta transformada, e depois, regenerar a planta transformada, de tal modo que a expressão da construção genética incorporada modifique a produção de óleo na planta. Em modalidades específicas, a planta modificada é *Arabidopsis thaliana*. Em outras modalidades, a planta é uma planta de cultura agrícola, tal como milho ou trigo. Em outras modalidades, a planta é uma cultura agrícola produtora de óleo, tal como soja, palma, colza ou girassol.

BREVE DESCRIÇÃO DOS DESENHOS

A Figura 1 é uma representação gráfica de experimentos que demonstram o teor de óleo aumentado de linhagens transgênicas de F-box-faseolina, em comparação com linhagens de controle com vetores. Os dados ilustrados são a média de 18 linhagens (T2) e as barras de erros são SE.

A Figura 2 é uma representação gráfica de experimentos que demonstram FAMES totais de sementes de linhagens transgênicas de F-box, em comparação com o tipo selvagem (WT), e outras linhagens de controle com vetores. As barras de erros são baseadas em SE em 18 linhagens: 3 duplicatas para cada linhagem.

A Figura 3 é uma representação gráfica de experimentos que demonstram FAMES totais de sementes de

linhagens transgênicas versus WT e linhagens de controle com vetores. Os pontos grandes são as duas linhagens com teor de óleo mais alto e foram denominadas Fbox2 e Fbox3 escolhidas posteriormente para a análise da próxima geração.

5 A Figura 4 é uma representação gráfica de experimentos que ilustram o espectro de ácidos graxos e a composição das razões molares de ácidos graxos de linhagens transgênicas de F-box versus plantas de controle.

10 A Figura 5 é uma representação gráfica da determinação do peso das sementes. Cerca de 400 sementes foram contadas quanto a sementes do tipo selvagem e transformantes com genes de F-box. As sementes contadas foram guardadas em um dessecador de um dia para o outro, e depois determinadas usando uma balança analítica. As barras
15 de erros baseiam-se em SE em 6 réplicas.

 A Figura 6 é uma representação gráfica que demonstra FAMES totais por semente nas sementes com F-box T3 versus WT e linhagens de controle com vetores. As barras de erros são baseadas em SE em dados obtidos de 12 plantas para
20 linhagens com F-box2 e F-box3; 6 plantas para WT e linhagens de controle com vetores. As duas linhagens neste caso denominadas F-box2 e F-box3 são os dois pontos de dados indicados por pontos grandes na Figura 3.

 A Figura 7 é uma representação gráfica de
25 experimentos que demonstram FAMES totais de sementes de linhagens transgênicas (F-box2 e F-box3) versus WT e linhagens de controle com vetores.

 A Figura 8 (A) é uma representação gráfica de experimentos que demonstram uma análise da razão C:N para

duas linhagens independentes F-box2 e F-box3. As barras de erros são baseadas em SE em seis réplicas. A Figura 8 (B) é uma representação gráfica que demonstra a correlação positiva entre teor de óleo e razão C:N. As colunas vazadas
5 são superexpressores F-box2 e F-box3; as colunas cinza são sementes de *Arabidopsis* do tipo selvagem.

A Figura 9 (A) é uma representação gráfica de experimentos que demonstram uma análise de FAMES das linhagens de F-box-KO.

10 A Figura 9 (B) é uma representação gráfica de experimentos que demonstram o espectro de ácidos graxos de linhagens KO e WT. As barras de erros são baseadas em SE em nove réplicas independentes.

A Figura 10A é uma representação esquemática da
15 sequência de polipeptídeo de uma proteína F-box de *Arabidopsis thaliana* correspondente ao N° de Acesso do GenBank NP_566277 (SEQ ID NO: 1).

A Figura 10B é uma representação esquemática da sequência de nucleotídeos de uma proteína F-box de
20 *Arabidopsis thaliana* correspondente ao N° de Acesso do GenBank NM_111499 (SEQ ID NO: 2). Os códons de iniciação e terminação previstos da estrutura de leitura aberta da proteína F-box estão sublinhados.

A Figura 11 é uma representação esquemática da
25 sequência de ácidos nucléicos (SEQ ID NO: 6) de pGATE-PHASE: F-box, com as seqüências codificadoras de F-box inseridas ilustradas em letras maiúsculas.

A Figura 12 é uma representação diagramática do vetor específico de semente pGATE-PHASE: F-box, ilustrando o

sítio de clonagem atrás do promotor específico de semente, faseolina.

DESCRIÇÃO DETALHADA DA INVENÇÃO

A patente e a literatura aqui referidas
5 estabelecem o conhecimento disponível para os versados nessas técnicas. As patentes norte-americanas, pedidos de patente concedidos, os pedidos de patente estrangeiros publicados, e as referências, incluindo as seqüências da base de dados do GenBank, aqui citadas são aqui incorporadas
10 como referência até o mesmo grau em que se cada uma tivesse sido especificamente e individualmente indicada para ser incorporada como referência.

4.1. Informações Gerais

A invenção baseia-se, em parte, na descoberta
15 inesperada de que a superexpressão de proteínas F-box vegetais aumenta o teor de óleo de sementes vegetais. Particularmente, a superexpressão da proteína F-box de *Arabidopsis thaliana* (N^{os} de Acesso do GenBank NM_111499 (DNAC) e NP_566277 (proteína), usando o promotor faseolina
20 específico de semente, produziu um fenótipo de teor mais alto de ácidos graxos (vide Figuras 1-5). Conseqüentemente, genericamente, a invenção fornece plantas transgênicas que incorporam genes heterólogos de F-box, que aumentam o teor de óleo de um ou mais estágios do desenvolvimento de uma
25 planta vascular.

Os ácidos graxos armazenados em sementes vegetais são usualmente compostos não-ramificados com um número par de carbonos na faixa entre 12 e 22 e com 0 a 3 ligações duplas *cis*. Entretanto, existem inúmeras variações neste

tema na natureza, particularmente com relação a grupos funcionais adicionais tais como hidroxila, epóxi, ciclopropeno ou acetileno. As plantas representam um grande reservatório de diversidade de ácidos graxos, sintetizando pelo menos 200 tipos diferentes de ácidos graxos (vide van de Loo et al., (1993) "Unusual fatty acids" in "Plant Lipids" (Editor: T. Moore), páginas 91-126, CRC, Boca Raton, FL, EUA). O uso humano, entretanto, ficou predominantemente restrito a uns poucos seletos ácidos graxos que se acumulam em plantas domesticadas. As quatro culturas de sementes oleaginosas mais importantes são em ordem descendente, soja, palma oleaginosa, colza, e girassol, as quais, em conjunto, representam 65% da produção mundial atual de óleos vegetais. Os ácidos graxos abundantes produzidos nestas principais sementes oleaginosas compreendem apenas quatro das mais de 200 possibilidades, a saber, linoleato, palmitato, laurato, e oleato.

Com relação à nomenclatura de lipídeos, uma anotação simples abreviada baseada no comprimento da molécula e no número e posições de ligações duplas foi desenvolvida para designar os ácidos graxos. Por exemplo, o ácido graxo, ácido oléico, monoinsaturado (ácido octadecenóico) é designado 18:1. O primeiro valor, 18, representa o número de carbonos. O segundo valor, 1, indica o número de ligações duplas. Além disso, a posição das ligações duplas, contando a partir do grupo carboxila, é designada por delta (Δ) e o ácido oléico pode ser designado mais inteiramente como 18:1 Δ 9. As ligações duplas em ácidos graxos de ocorrência natural são quase que exclusivamente

isômeros *cis*, e usualmente nenhuma designação para o tipo de ligação dupla é utilizada, a menos que ele seja um isômero *trans*, como em 16:1 Δ^3t . Alguns autores designam também as posições das ligações duplas em relação ao carbono do radical metila terminal. Assim sendo, um ácido graxo ômega 3 contém uma ligação dupla 3 carbonos a partir da extremidade metila do ácido graxo (por exemplo, 18:3 $\Delta^9, 12, 15$ é um ácido graxo ômega 3). A posição na qual um ácido graxo é esterificado para o esqueleto glicerol de glicerolipídeos é designada *sn*-3 (o terminal hidroxila que é fosforilado no glicerol-3-fosfato), *sn*-2 (a hidroxila central), e *sn*-1 (a hidroxila terminal que não é fosforilada).

No caso de usos comestíveis e industriais, um aumento no teor de óleo das sementes é desejável e tem sido uma meta importante da engenharia de sementes oleaginosas. Entretanto, para ser economicamente útil, essa mudança não deve ser à custa do volume de produção global das sementes ou à perda de outros componentes de alto valor. Por exemplo, a soja é a maior fonte de óleo vegetal, compreendendo 30% do mercado mundial, e constitui agora mais de 80% de todos óleos vegetais dietéticos nos Estados Unidos. Embora denominada uma semente oleaginosa, a soja contém apenas 18-22% de óleo, baseado no peso seco das sementes, e é cultivada principalmente como uma farinha com alto teor protéico para rações animais. Assim sendo, aumentar o óleo na soja na maioria dos casos não será útil se este aumento advir à custa da proteína de soja de alto valor que impulsiona a economia da cultura.

Comparativamente, outras culturas de sementes oleaginosas (exceto o algodão) são cultivadas principalmente devido ao seu óleo e produzem sementes com 40-60% de óleo. A ampla faixa de porcentagem de óleo das sementes, observada na natureza, sugere que este caminho poderia ser suscetível à engenharia metabólica, particularmente em sementes oleaginosas de "baixo teor de óleo", desde que os mecanismos fundamentais que controlam o teor de óleo sejam identificados.

10 4.2 Definições

Por conveniência, o significado de certos termos e expressões empregados no relatório descritivo, nos exemplos, e nas reivindicações apensadas, é fornecido abaixo.

O termo "agonista", como aqui utilizado, significa um agente que mimetiza ou regula de forma ascendente (por exemplo, potencializa ou suplementa) uma bioatividade. Por exemplo, um agonista de F-box pode ser uma proteína F-box do tipo selvagem ou derivado dela que tem pelo menos uma bioatividade da proteína F-box do tipo selvagem. Um agonista pode ser também um composto que aumenta a interação de um polipeptídeo bioativo com outra molécula, por exemplo, um receptor. Os agonistas podem ser qualquer classe de molécula, tal como uma molécula pequena, incluindo um ácido nucléico, proteína, lipídeo ou combinação delas.

25 O termo "alelo", que é aqui utilizado de forma intercambiável com "variante alélica", refere-se a formas alternativas de um gene ou partes dele. Os alelos ocupam o mesmo locus ou posição em cromossomas homólogos. Quando um indivíduo tem dois alelos idênticos de um gene, o indivíduo

é dito como sendo homozigótico para o gene ou alelo. Quando um indivíduo tem dois alelos diferentes de um gene, o indivíduo é dito como sendo heterozigótico para o gene ou alelo. Os alelos de um gene específico podem diferir entre si em um único nucleotídeo, ou vários nucleotídeos, e podem incluir substituições, deleções, e inserções de nucleotídeos. As variações de seqüências que ocorrem freqüentemente incluem mutações de transição (isto é, purina para substituições de purina e pirimidina para substituições de pirimidina, por exemplo, A para G ou C para T), mutações transversas (isto é, purina para pirimidina e pirimidina para substituições de purina, por exemplo, A para T ou C para G), e alteração nas respectivas seqüências de DNA repetitivas (por exemplo, expansões e contrações de repetições de trinucleotídeos e seqüências de repetições tandem). Um alelo de um gene pode estar também na forma de um gene que contém uma mutação. O termo "variante alélica de uma região polimórfica de um gene" refere-se a uma região de um gene do locus que tem uma ou várias diferenças de seqüências de nucleotídeos encontradas nessa região do gene em outros indivíduos.

Como aqui utilizado, o termo "F-box" ou "proteína F-box" refere-se às seqüências de aminoácidos de proteínas envolvidas em proteólise, incluindo, porém sem limitações, proteínas envolvidas no complexo ubiquitina-ligase obtido de qualquer espécie, incluindo espécies vegetais, de qualquer origem, seja natural, sintética, semi-sintética ou recombinante. F-box é uma seqüência de 35-45 aminoácidos e permite que as proteínas F-box entrem dentro de complexos

com outros componentes da degradação de proteínas dependentes de ubiquitina (por exemplo, Skp1). Assim sendo, as proteínas F-box podem se ligar a Skp1, e podem conter um modelo que apresente uma similaridade de sequência com Grr1 ou Cdc4 ou com a proteína F-box de *Arabidopsis thaliana* ilustrada na Figura 10A. Este modelo estrutural conservado está incluído nos alinhamentos de seqüências ilustrados na Figura 7 da patente nº US 6.573.094, cujo teor é aqui incorporado como referência em sua totalidade (isto é, os resíduos de aminoácidos que são compartilhados pelas proteínas F-box ilustradas). Entretanto, não se pretende que o termo seja limitado às seqüências exatas aqui enunciadas. Em algumas modalidades, as proteínas F-box podem compreender ainda modelos adicionais, particularmente modelos envolvidos na interação proteína-proteína. Estes modelos adicionais estão incluídos, porém não-limitados a repetições ricas em leucina, e WD-40. In certas modalidades, a proteína F-box é de uma planta vascular, enquanto que em modalidades particularmente úteis, a proteína F-box é de *Arabidopsis thaliana*.

O termo "antagonista", como aqui utilizado, refere-se a um agente que regula de forma descendente (por exemplo, suprime ou inibe) pelo menos a bioatividade. Um antagonista pode ser um composto que inibe ou diminui a interação entre uma proteína e outra molécula, por exemplo, um substrato. Conseqüentemente, um antagonista útil é um composto que inibe ou diminui a ligação a um substrato, e desta forma, bloqueia a função da enzima. Um antagonista pode ser também um composto que regula de forma descendente

a expressão de um gene ou genes, ou que reduz a quantidade de um produto gênico transladado. O antagonista da bioatividade-alvo pode ser uma forma negativa dominante de um polipeptídeo que possui essa bioatividade, por exemplo, os antagonistas de F-box incluiriam uma forma de um polipeptídeo de F-box que é capaz de interagir com outros componentes da via dependente de ubiquitina, mas que interfere com a função do complexo resultante (isto é, uma forma negativa dominante da bioatividade-alvo). Um antagonista pode ser também um ácido nucléico *antisense*, ou uma ribozima capaz de interagir especificamente um RNAm que codifica a bioatividade-alvo. Ainda outros antagonistas são moléculas que se ligam a uma bioatividade-alvo e inibem sua ação. Tais moléculas incluem peptídeos tais como aqueles que se ligam a um sítio ativo de uma enzima e impedem que ele interaja com o substrato. Ainda outros antagonistas da bioatividade-alvo incluem anticorpos que interagem especificamente com um epítipo do polipeptídeo-alvo, de tal modo que a ligação interfira com a função biológica do polipeptídeo. Em ainda outra modalidade útil, o antagonista é uma molécula pequena, tal como uma molécula capaz de inibir a interação entre uma enzima-alvo e seu substrato.

"Propagação assexuada" refere-se à produção de progênie regenerando uma planta inteira a partir de retalhos de folhas, retalhos de pedúnculos, retalhos de raízes, células vegetais individuais (protoplastos) e calos.

O termo "sítio catalítico" refere-se à parte de uma molécula que é capaz de se ligar a um reagente e melhorar a velocidade de uma reação. Os sítios catalíticos

podem estar presentes em polipeptídeos ou proteínas, enzimas, compostos orgânicos, compostos organometálicos, metais, e similares. Um sítio catalítico pode ser constituído de partes separadas presentes em uma ou mais

5 cadeias do polipeptídeo ou compostos. Estas partes catalíticas separadas se associam para formar uma parte maior de um sítio catalítico. Um sítio catalítico pode ser formado por um polipeptídeo ou proteína que está ligada a um metal.

10 Os termos "células", "células hospedeiras" ou "células hospedeiras recombinantes" são aqui utilizados de forma intercambiável. Deve-se entender que tais termos referem-se não apenas à célula específica do indivíduo, mas também à progênie ou progênie potencial de tal célula.

15 Devido ao fato de que certas modificações podem ocorrer em gerações sucessivas, em virtude de mutação ou influências ambientais, tal progênie, de fato, não ser idêntica à célula parental, mas ainda assim estão incluídas dentro do âmbito do termo como aqui utilizado.

20 Um "polipeptídeo quimérico" ou "polipeptídeo de fusão" é uma fusão de uma primeira sequência de aminoácidos que codifica um dos polipeptídeos em questão com uma segunda sequência de aminoácidos que define um domínio (por exemplo, parte de polipeptídeo) estranho e não substancialmente

25 homólogo a qualquer domínio do polipeptídeo em questão. Um polipeptídeo quimérico pode apresentar um domínio estranho que é encontrado (embora em um polipeptídeo diferente) em um organismo que também expressa o primeiro polipeptídeo, ou ele pode ser uma fusão "entre espécies", "intergênica",

etc., de estruturas polipeptídicas expressadas por diferentes tipos de organismos. Genericamente, um polipeptídeo de fusão pode ser representado pela fórmula genérica X-polipeptídeo-Y, onde polipeptídeo representa uma
5 primeira ou referida proteína ou polipeptídeo, e X e Y estão independentemente ausentes ou representam seqüências de aminoácidos que não estão relacionadas à primeira seqüência em um organismo, incluindo mutantes de ocorrência natural.

Como aqui utilizado, o termo "variações
10 conservativamente modificadas" de uma seqüência de ácidos nucléicos específica refere-se aos ácidos nucléicos que codificam seqüências de aminoácidos idênticas ou essencialmente idênticas, ou onde o ácido nucléico não codifica uma seqüência de aminoácidos, para seqüências
15 essencialmente idênticas. Por causa da degeneração do código genético, um grande número de ácidos nucléicos funcionalmente idênticos codifica qualquer dado polipeptídeo. Por exemplo, os códons CGU, CGC, CGA, COG, AGA, e AGG codificam o aminoácido arginina. Assim sendo, em
20 cada posição na qual uma arginina é especificada por um códon, o códon pode ser alterado para qualquer um dos códons correspondentes descritos, sem alterar o polipeptídeo codificado. Tais variações de ácidos nucléicos são "variações silenciosas," que são uma espécie de "variações
25 conservativamente modificadas". Cada seqüência de ácidos nucléicos neste caso, que codifica um polipeptídeo, descreve também cada variação silenciosa possível. Os versados nessas técnicas devem reconhecer que cada códon em um ácido nucléico (exceto AUG, que é normalmente o único códon para

metionina) pode ser modificado para produzir uma molécula funcionalmente idêntica por técnicas usuais. Conseqüentemente, cada "variação silenciosa" de um ácido nucléico que codifica um polipeptídeo está implícita em cada
5 seqüência descrita. Além disso, os versados nessas técnicas devem reconhecer que substituições, deleções ou adições individuais que alteram, adicionam ou deletam um único aminoácido ou uma pequena porcentagem de aminoácidos (tipicamente menos do que 5%, mais tipicamente menos do que
10 1%) em uma seqüência codificada, são "variações conservativamente modificadas", onde as alterações resultam na substituição de um aminoácido por um aminoácido quimicamente similar. Tabelas de substituições conservativas que fornecem aminoácidos funcionalmente
15 similares são bem conhecidas nessas técnicas. Os seguintes seis grupos contêm aminoácidos que são substituições conservativas entre si: 1) Alanina (A), Serina (S), Treonina (T); 2) Ácido Aspártico (D), Ácido Glutâmico (E); 3) Asparagina (N), Glutamina (Q); 4) Arginina (R), Lisina (K);
20 5) Isoleucina (I), Leucina (L), Metionina (M), Valina (V); e 6) Fenilalanina (F), Tirosina (Y), Triptofano (W).

Como aqui descrito, as seqüências podem ser otimizadas para a expressão em uma célula hospedeira específica usada para produzir a proteína (por exemplo, uma
25 célula vegetal, tal como um tomate, ou um sistema de clonagem e expressão, tal como uma célula de levedura). Similarmente, "substituições conservativas de aminoácidos", em um ou poucos aminoácidos em uma seqüência de aminoácidos são substituídos por aminoácidos diferentes com propriedades

altamente similares (vide a Seção de definições acima), são também facilmente identificadas como sendo altamente similares a uma seqüência de aminoácidos específica ou a uma seqüência de ácidos nucléicos específica que codifica um aminoácido. Tais variações conservativamente substituídas de qualquer seqüência específica são uma característica da presente invenção.

Um "complexo de distribuição" deve significar um meio de assesto (por exemplo, uma molécula que resulta em ligações de afinidade mais alta de uma proteína, polipeptídeo ou peptídeo, para a superfície da célula-alvo e/ou maior captação celular ou nuclear por uma célula-alvo). Os exemplos de meios de assesto incluem: esteróis (por exemplo, colesterol), lipídeos (por exemplo, lipídeo catiônico, virossoma ou lipossoma), vírus (por exemplo, vírus do mosaico do tabaco) ou agentes de ligação específica à célula-alvo (por exemplo, ligantes reconhecidos por receptores específicos da célula-alvo). Os complexos úteis são suficientemente estáveis *in vivo* para impedir o desacoplamento significativo antes da internalização pela célula-alvo. Entretanto, o complexo é clivável sob condições apropriadas dentro da célula, de tal modo que o gene, proteína, polipeptídeo ou peptídeo seja liberado de uma forma funcional.

Como é bem sabido, os genes podem existir em uma única cópia ou em múltiplas cópias dentro do genoma de um indivíduo. Tais genes em duplicata podem ser idênticos ou podem ter certas modificações, incluindo substituições, adições ou deleções de nucleotídeos, todas elas ainda

codificando polipeptídeos que têm substancialmente a mesma atividade. O termo "seqüência de DNA que codifica um polipeptídeo-alvo" pode, assim, se referir a um ou mais genes em um indivíduo específico. Além disso, podem existir
5 certas diferenças em seqüências de nucleotídeos entre organismos individuais, que são denominadas alelos. Tais diferenças alélicas podem ou não resultar em diferenças na seqüência de aminoácidos do polipeptídeo codificado, e ainda assim codificar um polipeptídeo com a mesma atividade
10 biológica.

As expressões "rompimento do gene" e "rompimento assestado" ou qualquer expressão similar refere-se à interrupção específica do sítio de uma seqüência de DNA nativa de modo a impedir a expressão desse gene na célula,
15 em comparação com a cópia do tipo selvagem do gene. A interrupção pode ser causada por deleções, inserções, ou modificações do gene, ou qualquer combinação.

O termo "sítio enzimático" refere-se à parte de uma molécula de proteína que contém um sítio catalítico. A
20 maioria dos sítios enzimáticos apresenta uma especificidade seletiva de substrato muito alta. Um sítio enzimático pode compreender duas ou mais partes do sítio enzimático em segmentos diferentes da mesma cadeia do polipeptídeo. Estas partes do sítio enzimático são associadas entre si para
25 formar uma parte maior de um sítio enzimático. Uma parte de um sítio enzimático pode ser também um metal.

O termo "enzima" refere-se a uma proteína, polipeptídeo, molécula de RNA peptídica, ou proteína multimérica capaz de acelerar ou produzir por ação

catalítica alguma mudança em um substrato pelo qual ela é freqüentemente específica.

O termo "epítopo" refere-se a uma parte de uma molécula que é reconhecida especificamente por um produto de
5 imunoglobulina. Ela é referida também como o determinante ou determinante imunogênico.

Como aqui utilizado, o termo "imunoglobulina" é uma proteína multimérica que contém partes imunologicamente ativas de uma cadeia pesada de imunoglobulina e de uma
10 cadeia leve de imunoglobulina acopladas de forma covalente entre si e capazes de se combinarem especificamente com o antígeno.

Como aqui utilizado, um fragmento Fab é uma proteína multimérica que consiste na parte de uma molécula
15 de imunoglobulina que contém as partes imunologicamente ativas de uma cadeia pesada de imunoglobulina e de uma cadeia leve de imunoglobulina acopladas de forma covalente entre si e capazes de se combinarem especificamente com o antígeno. Os fragmentos Fab são preparados tipicamente por
20 digestão proteolítica de moléculas de imunoglobulina substancialmente intactas com papaína, usando métodos bem conhecidos nessas técnicas. Entretanto, um fragmento Fab pode ser preparado também expressando em uma célula hospedeira apropriada as partes desejadas da cadeia pesada
25 de imunoglobulina e da cadeia leve de de imunoglobulina, usando métodos bem conhecidos nessas técnicas.

Como aqui utilizado, um fragmentos F[v]: uma proteína multimérica que consiste nas partes imunologicamente ativas de uma região variável da cadeia

pesada de imunoglobulina e uma região da variável da cadeia leve de imunoglobulina, acopladas de forma covalente entre si e capazes de se combinarem especificamente com o antígeno. Os fragmentos F[v] são preparados tipicamente expressando em uma célula hospedeira apropriada as partes desejadas da região variável da cadeia pesada de imunoglobulina e da região variável da cadeia leve de imunoglobulina, usando métodos bem conhecidos nessas técnicas.

10 Como aqui utilizado, o termo "gene" ou "gene recombinante" refere-se a um ácido nucléico que compreende uma estrutura de leitura aberta que codifica um polipeptídeo da presente invenção, incluindo seqüências de éxons e (opcionalmente) íntrons. Um "gene recombinante" refere-se a um ácido nucléico que codifica tais polipeptídeos reguladores, que podem incluir opcionalmente seqüências de íntrons que são derivadas de um DNA cromossômico. Os genes recombinantes exemplificativos incluem aqueles que codificam uma atividade de polipeptídeo F-box.

20 Como aqui utilizado, o termo "DNA heterólogo" ou "ácido nucléico heterólogo" inclui um DNA que não ocorre naturalmente como parte do genoma no qual ele está presente ou que é encontrado em um local ou locais no genoma que diferem daqueles nos quais ele ocorre na natureza. O DNA heterólogo não é endógeno para a célula dentro da qual ele é introduzido, mas foi obtido a partir de outra célula. Genericamente, embora não necessariamente, tal DNA codifica o RNA e proteínas que não são normalmente produzidas pelas células na qual ele é expressado. O DNA heterólogo pode ser

referido também como DNA estranho. Qualquer DNA que os versados nessas técnicas reconheceria ou consideraria como heterólogo ou estranhopara a célula na qual ele é expressado está aqui englobado pelo termo DNA heterólogo. Os exemplos
5 de DNA heterólogo incluem, porém sem limitações, DNA isolado que codifica uma proteína F-box.

O termo "homologia" ou "identidade" ou "similaridade" refere-se à similaridade de seqüências entre dois peptídeos ou entre duas moléculas de ácidos nucleicos.
10 A homologia pode ser determinada comparando uma posição em cada seqüência que pode ser alinhada com propósito de comparação. Quando uma posição na seqüência comparada é ocupada pela mesma base ou aminoácido, então as moléculas são idênticas nessa posição. Um grau de homologia ou
15 similaridade ou identidade entre seqüências de ácidos nucleicos é função do número de nucleotídeos idênticos ou equivalentes em posições compartilhadas pelas seqüências de ácidos nucleicos. Um grau de identidade de seqüências de aminoácidos é função do número de aminoácidos idênticos em
20 posições compartilhadas pelas seqüências de aminoácidos. Um grau de homologia ou similaridade de seqüências de aminoácidos é função do número de aminoácidos, isto é, estruturalmente relacionados, em posições compartilhadas pelas seqüências de aminoácidos. Uma seqüência "não-
25 relacionada" ou "não-homóloga" compartilha menos do que 40% de identidade, ou menos do que 25% de identidade, com uma das seqüências da presente invenção.

O termo "inativação", com relação a genes da célula hospedeira, significa que a produção de um produto

gênico funcional é impedida ou inibida. A inativação pode ser conseguida por deleção do gene, mutação do promotor, de tal modo que a expressão ao ocorra, ou a mutação da seqüência codificadora, de tal modo que o produto gênico
5 fique inativo (de forma constitutiva ou indutiva). A inativação pode ser parcial ou total.

O termo "interagir", como aqui utilizado, significa incluir relações ou associações detectáveis (por exemplo, interações bioquímicas) entre moléculas, tal como a
10 interação entre proteína-proteína, proteína-ácido nucléico, ácido nucléico-ácido nucléico, e proteína-molécula pequena ou ácido nucléico-molécula pequena na natureza.

O termo "isolado", como aqui utilizado com relação a ácidos nucléicos, tais como DNA ou RNA, refere-se a
15 moléculas separadas de outros DNAs, ou RNAs, respectivamente, que estão presentes na origem natural da macromolécula. Por exemplo, os ácidos nucléicos que codificam os polipeptídeos em questão incluem tipicamente não mais do que 10 quilobases (kb) da seqüência de ácidos
20 nucléicos que naturalmente flanqueia imediatamente esse gene no DNA genômico, ou não mais do que 5 kb dessas seqüências flanqueadoras de ocorrência natural, e mais utilmente menos do que 1,5 kb da seqüência flanqueadora de ocorrência natural. O termo "isolado", como aqui utilizado, refere-se
25 também a um ácido nucléico ou polipeptídeo que é substancialmente isento de material celular material, material viral, ou meio de cultura, quando produzido por técnicas de DNA recombinante, ou precursores químicos ou outros produtos químicos quando sintetizados quimicamente.

Além disso, um "ácido nucléico isolado" inclui fragmentos de ácido nucléico que não são de ocorrência natural, pois os fragmentos não seriam encontrados no estado natural. O termo "isolado" é também aqui utilizado para se referir a polipeptídeos que estão isolados de outras proteínas celulares e engloba polipeptídeos purificados e recombinantes.

O termo "knock-out" refere-se à supressão parcial ou completa da expressão de um gene endógeno. Isso é realizado genericamente deletando uma parte do gene ou substituindo uma parte por uma segunda sequência, mas pode ser causado também por outras modificações do gene, tal como a introdução de códons de interrupção, a mutação de aminoácidos críticos, a remoção de uma junção de íntrons, etc.

O termo "marcador" ou "sequência marcadora" ou expressão similar significa qualquer gene que produz um genótipo selecionável ou um fenótipo selecionável. Ele inclui exemplos tais como o gene de espectinomicina (spcR), neo gene (neoR), gene da proteína fluorescente verde (GFP) gene, gene de TK, gene de β -galactosidase, etc. A sequência marcadora pode ser qualquer sequência conhecida pelos versados nessas técnicas, que serve para estes propósitos, embora tipicamente a sequência marcadora deva ser uma sequência que codifica uma proteína que confere um traço selecionável, tal como um gene de resistência a antibiótico, ou uma enzima que pode ser detectada e que não é tipicamente encontrada na célula. A sequência marcadora pode incluir também regiões reguladoras, tal como um promotor ou

intensificador que regula a expressão dessa proteína. Entretanto, é possível também transcrever o marcador usando seqüências reguladoras endógenas. Em uma modalidade da presente invenção, o marcador facilita a separação de

5 células transfectadas das não-transfectadas por seleção de células ativadas por fluorescência, por exemplo, pelo uso de um anticorpo marcado de forma fluorescente ou a expressão de uma proteína fluorescente, tal como GFP. Outras seqüências de DNA que facilitam a expressão de genes marcadores também

10 podem ser incorporadas nas construções de DNA da presente invenção. Estas seqüências incluem, porém sem limitações, sinais de iniciação e terminação da transcrição, sinais de translação, sinais de modificação pós-translacionais, junções de íntrons removidos, sítios de ligação de

15 ribossomas, e sinais de poliadenilação, para mencionar alguns. A seqüência marcadora pode ser usada também para apensar a seqüência ao gene-alvo. Por exemplo, ela pode ser usada para adicionar um códon de interrupção à translação de F-box truncada. O uso de marcadores selecionáveis é bem

20 conhecido nessas técnicas e não precisa ser aqui detalhado. O termo "modulação", como aqui utilizado, refere-se à regulação ascendente (isto é, ativação ou estimulação (por exemplo, agonizando ou potencializando)) e à regulação descendente (isto é, inibição ou supressão (e.g., por

25 exemplo, antagonizando ou inibindo)).

Um "gene mutado" ou "mutação" refere-se a uma forma alélica de um gene, que é capaz de alterar o fenótipo de um indivíduo que tem o gene mutado em relação a um indivíduo que não tem o gene mutado. Caso um indivíduo seja

homozigótico para esta mutação para ter um fenótipo alterado, a mutação é dita como sendo recessiva. Caso uma cópia do gene mutado seja suficiente para alterar o fenótipo do indivíduo, a mutação é dita como sendo dominante. Caso
5 um indivíduo tenha uma cópia do gene mutado e tenha um fenótipo que é intermediário entre aquele de um homozigótico e aquele de um indivíduo heterozigótico (para esse gene), a mutação é dita como sendo co-dominante.

Como aqui utilizado, o termo "ácido nucléico"
10 refere-se a polinucleotídeos tal como o ácido desoxirribonucléico (DNA), e, quando apropriado, ácido ribonucléico (RNA). O termo deve ser entendido como incluindo formas de filamento único ou filamento duplo de ácido nucléico, e, como equivalentes, análogos de RNA ou
15 DNA. Tais análogos de ácidos nucléicos podem ser constituídos de análogos de nucleotídeos, e, conforme aplicável à modalidade que está sendo descrita, podem ser polinucleotídeos de filamento único (tal como *sense* ou *antisense*) ou filamento duplo.

20 A frase "seqüência de nucleotídeos complementar à seqüência de nucleotídeos enunciada em SEQ ID NO: x" refere-se à seqüência de nucleotídeos do filamento complementar de um filamento de ácido nucléico que tem SEQ ID NO: x. O termo "filamento complementar" é aqui utilizado de forma
25 intercambiável com o termo "complemento". O complemento de um filamento de ácido nucléico pode ser o complemento de um filamento codificador ou o complemento de um filamento não-codificador. Quando se faz referência a ácidos nucléicos de filamento duplo, o complemento de um ácido nucléico que tem

SEQ ID NO: x refere-se ao filamento complementar do filamento que tem SEQ ID NO: x ou a qualquer ácido nucléico que tem a seqüência de nucleotídeos do filamento complementar de SEQ ID NO: x. Quando se faz referência a um
5 ácido nucléico de filamento único que tem a seqüência de nucleotídeos SEQ ID NO: x, o complemento deste ácido nucléico é um ácido nucléico que tem uma seqüência de nucleotídeos que é complementar àquela de SEQ ID NO: x. As seqüências de nucleotídeos e suas seqüências complementares
10 são sempre fornecidas na direção de 5' para 3'.

A frase "ligada operacionalmente" refere-se a uma ligação funcional entre um promotor e uma segunda seqüência, onde a seqüência promotora inicia a transcrição do RNA correspondente à segunda seqüência.

15 O termo "identidade percentual" refere-se à identidade de seqüências entre duas seqüências de aminoácidos ou entre duas seqüências de nucleotídeos. Cada identidade pode ser determinada comparando uma posição em cada seqüência que pode ser alinhada com o propósito de
20 comparação. Quando uma posição equivalente nas seqüências comparadas é ocupada pela mesma base ou aminoácido, então as moléculas são idênticas naquela posição; quando o sítio equivalente ocupado pelo mesmo resíduo de aminoácido ou por um aminoácido similar (por exemplo, similar na natureza
25 estérica e/ou eletrônica), então as moléculas podem ser referidas como homólogas (similares) naquela posição. A expressão como uma porcentagem de homologia/similaridade ou identidade refere-se a uma função do número de aminoácidos idênticos ou similares em posições compartilhadas pelas

seqüências comparadas. Vários algoritmos e/ou programas de alinhamento podem ser usados, incluindo FASTA, BLAST ou ENTREZ. FASTA e BLAST estão disponíveis como parte do pacote de análise de seqüências GCG (Universidade de Wisconsin, Madison, Wisconsin), e podem ser usados, por exemplo, com ajustes de omissões. ENTREZ está disponível no National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD. Em uma modalidade, a identidade percentual de duas seqüências pode ser determinada pelo programa GCG com peso de intervalos de 1, por exemplo, cada intervalo de aminoácido é ponderado como se ele fosse um único desemparelhamento de aminoácido ou nucleotídeo entre as duas seqüências.

O termo "planta" inclui plantas integrais, órgãos de plantas (por exemplo, folhas, pedúnculos, flores, raízes, etc.), sementes e células vegetais e a progênie das mesmas. A classe de plantas que podem ser usadas no método da invenção tão ampla quanto a classe de plantas superiores suscetíveis a técnicas de transformação, incluindo plantas monocotilédones e dicotilédones, bem como certas plantas inferiores tais como algas. Ela inclui plantas de uma série de níveis de ploidia, incluindo poliplóides, diplóides e haplóides. As plantas de cultura úteis da invenção incluem, porém sem limitações, milho (*Zea mays*), *Brassica* spp. (e.g., canola (*B. napus*), *B. rapa*, *B. juncea*), particularmente as espécies de *Brassica* úteis como fontes de óleo de sementes, alfafa (*Medicago sativa*), arroz (*Oryza sativa*), centeio (*Secale cereale*), sorgo (*Sorghum bicolor*, *Sorghum vulgare*), painço (por exemplo, milheto-pérولا (*Pennisetum glaucum*),

painço (*Panicum miliaceum*), painço-português (*Setaria italica*), capim-pé-de-galinha (*Eleusine coracana*), girassol (*Helianthus annuus*), açafroa (*Carthamus tinctorius*), trigo (*Triticum aestivum*), soja (*Glycine max*), tabaco (*Nicotiana tabacum*), batata (*Solanum tuberosum*), amendoim (*Arachis hypogaea*), algodão (*Gossypium barbadense*, *Gossypium hirsutum*), batata-doce (*Ipomoea batatas*), mandioca (*Manihot esculenta*), café (*Coffea* spp.), coco (*Cocos nucifera*), abacaxi (*Ananas comosus*), árvores cítricas (*Citrus* spp.), cacau (*Theobroma cacao*), chá-da-índia (*Camellia sinensis*), banana (*Musa* spp.), abacate (*Persea americana*), figo (*Ficus casica*), goiaba (*Psidium guajava*), manga (*Mangifera indica*), oliva (*Olea europaea*), mamão (*Carica papaya*), caju (*Anacardium occidentale*), macadâmia (*Macadamia integrifolia*), amêndoa (*Prunus amygdalus*), beterraba (*Beta vulgaris*), cana-de-açúcar (*Saccharum* spp.), aveia, cevada, vegetais, plantas ornamentais, e coníferas.

O termo "planta" refere-se ainda às seguintes classes de espécies vegetais:

20 Dicotilédones (dicot): uma planta florescente cujos embriões têm duas metades de sementes ou cotilédones. Os exemplos de dicotilédones são: tabaco; tomate; os legumes incluindo alfafa; carvalho; bordo; rosa; hortelãs; abóboras; margaridas; nozes; cactos; violetas; e botão-de-ouro.

25 Monocotilédones (monocot): uma planta florescente cuos embriões têm um cotilédone ou folha de semente. Os exemplos não-limitativos de monocotilédones são: lírios; gramíneas; milho; grãos, incluindo aveia, trigo e cevada; orquídeas; íris; cebolas e palmas.

"Planta inferior" refere-se a uma planta não-florescente, incluindo samambaias, gimnospermas, coníferas, cavalinhas, lycopódios, hepáticas, ceratofiláceas, musgos, algas vermelhas, algas marrons, gametófitos, esporófitos de
5 pteridófitos, e algas verdes.

Genericamente, as plantas de interesse incluem plantas granosas que produzem sementes de interesse, plantas com sementes oleaginosas, e plantas leguminosas. As sementes de interesse incluem sementes granosas, tais como
10 milho, trigo, cevada, arroz, sorgo, centeio, painço, etc. As plantas com sementes oleaginosas incluem algodão, soja, açafrão, girassol, brássica, milho, alfafa, palma, coco, linho, mamona, oliva, etc. As plantas leguminosas incluem favas e ervilhas. As favas incluem guar, alfarroba, feno-
15 grego, soja, feijoeiros, feijão-fradinho, feijão-da-índia, feijão-de-lima, feijão-fava, lentilhas, grão-de-bico, etc.

O termo "promotor" refere-se a uma região de seqüências de ácidos nucléicos, localizada a montante e/ou a jusante do início da transcrição, que auxiliam no
20 reconhecimento, ligação e/ou iniciação da RNA polimerase ou outras proteínas de transcrição que iniciam a transcrição de um gene associado. Um "promotor vegetal" é um promotor capaz de iniciar a transcrição em células vegetais. Um
"promotor leucina aminopeptidase vegetal" é um promotor
25 derivado de um gene de leucina aminopeptidase, por exemplo, clonando, isolando ou modificando de forma recombinante um promotor nativo a partir de um gene de leucina aminopeptidase.

Um "ácido nucléico recombinante" compreende ou é codificado por um ou mais ácidos nucléicos derivados de de um ácido nucléico construído artificialmente. Por exemplo, o ácido nucléico pode compreender ou ser codificado por um
5 ácido nucléico clonado formado unindo ácidos nucléicos heterólogos, como enunciado, por exemplo, em Berger e Kimmel, "Guide to Molecular Cloning Techniques", "Methods In Enzymology", Volume 152 Academic Press, Inc., San Diego, CA (Berger), e em Sambrook et al., "Molecular Cloning - A
10 Laboratory Manual" (2ª edição) Volumes 1-3 (1989) (Sambrook), e em "Current Protocols In Molecular Biology", editores Ausubel, F. M., et al., Greene Publishing Associates, Inc. e John Wiley & Sons, Inc., (suplemento de 1996) (Ausubel). Alternativamente, o ácido nucléico pode
15 ser sintetizado quimicamente.

Como aqui utilizado, o termo "construção de gene repórter" é um ácido nucléico que inclui um "gene repórter" ligado operacionalmente a seqüências reguladoras transcricionais. A transcrição do gene repórter gene é
20 controlada por estas seqüências. As seqüências reguladoras transcricionais incluem as regiões do promotor e outras regiões reguladoras, tais como seqüências intensificadoras, que modulam a atividade do promotor, ou seqüências reguladoras que modulam a atividade ou eficiência da RNA
25 polimerase que reconhece o promotor, ou as seqüências reguladoras são reconhecidas por moléculas efetoras.

Como aqui utilizado, o termo "ácido nucléico" refere-se a polinucleotídeos ou oligonucleotídeos, tal como o ácido desoxirribonucléico (DNA), e, quando apropriado, o

ácido ribonucléico (RNA). O termo deve ser entendido também como incluindo equivalentes, análogos de RNA ou DNA preparados a partir de análogos de nucleotídeos, e conforme aplicável à modalidade que está sendo descrita, 5 polinucleotídeos de filamento único (*sense* ou *antisense*) e de filamento duplo.

Como aqui utilizado, o termo "promotor" significa uma sequência de DNA que regula a expressão de uma sequência de DNA selecionada ligada operacionalmente ao promotor, e 10 que efetua a expressão da sequência de DNA selecionada em células. O termo engloba promotores "específicos de tecidos", isto é, promotores que efetuem a expressão da sequência de DNA selecionada apenas em células específicas (por exemplo, células de um tecido específico). O termo 15 cobre também os assim denominados promotores "leaky" que regulam a expressão de um DNA selecionado principalmente em um tecido, mas causam a expressão também em outros tecidos. O termo engloba também promotores não-específicos de tecidos e promotores que expressam de forma constitutiva ou que são 20 induzíveis (isto é, os níveis de expressão podem ser controlados).

Os termos "proteína", "polipeptídeo" e "peptídeo" são utilizados de forma intercambiável quando se referem a um produto gênico.

25 O termo "proteína recombinante" refere-se a um polipeptídeo da presente invenção que é produzido por técnicas de DNA recombinante, onde, genericamente, o DNA que codifica um polipeptídeo específico é inserido em um vetor de expressão apropriado, o qual, por sua vez, é usado para

transformar uma célula hospedeira a fim de produzir a proteína heteróloga. Além disso, a frase "derivado de", com relação a um gene-alvo recombinante, pretende incluir dentro do significado de "proteína recombinante" as proteínas que
5 têm uma sequência de aminoácidos de um polipeptídeo-alvo nativo, ou uma sequência de aminoácidos similar a ela, que é gerada por mutações, incluindo substituições e deleções (incluindo truncamento) de uma forma de ocorrência natural do polipeptídeo.

10 Como aqui utilizado, o termo "células recombinantes" inclui quaisquer células que foram modificadas pela introdução de DNA heterólogo. Células de controle incluem células que são substancialmente idênticas às células recombinantes, mas não expressam uma ou mais das
15 proteínas codificadas pelo DNA heterólogo, por exemplo, não incluem ou expressam um gene de F-box recombinante.

O termo "molécula pequena", como aqui tilizado, pretende se referir a uma composição que tem peso molecular menor do que cerca de 5 kD, e mais preferivelmente, menor do
20 que cerca de 4 kD. As moléculas pequenas podem ser ácidos nucleicos, peptídeos, polipeptídeos, peptidomiméticos, carboidratos, lipídeos ou outras moléculas orgânicas (que contêm carbono) ou inorgânicas. Muitas empresas farmacêuticas têm bibliotecas extensas de misturas químicas
25 e/ou biológicas, freqüentemente extratos fúngicos, bacterianos ou algais, que podem ser triados com qualquer um dos ensaios da invenção para identificar compostos que modulam uma bioatividade-alvo.

Como aqui utilizado, o termo "hibridiza especificamente" ou "detecta especificamente" refere-se à capacidade de uma molécula de ácido nucléico da invenção hibridizar para aproximadamente 6, 12, 20, 30, 50, 100, 150, 200, 300, 350, 400 ou 425 nucleotídeos consecutivos de um gene, mais preferivelmente um gene de F-box vegetal.

O termo "substancialmente homóloga", quando utilizado com relação a seqüências de aminoácidos, refere-se a seqüências que são substancialmente idênticas ou similares em seqüência, gerando uma homologia em conformação, e assim sendo, uma atividade biológica similar. O termo pretende inferir uma evolução comum das seqüências.

Como aqui utilizado, o termo "transfecção" significa a introdução de um ácido nucléico, por exemplo, por intermédio de um vetor de expressão, em uma célula receptora por transferência gênica mediada por ácido nucléico. Os métodos para transformação conhecidos nessas técnicas incluem qualquer meio elétrico, magnético, físico, biológico ou químico. Como aqui utilizado, o termo "transfecção" inclui técnicas específicas tais como eletroporação, magnetoporação, tratamento com Ca^{++} , injeção, bombardeio, infecção retroviral e lipofecção, dentre outros. O termo "transformação", como aqui utilizado, refere-se a um processo no qual o genótipo de uma célula é mudado como resultado da captação celular de DNA ou RNA exógeno, e, por exemplo, a célula transformada expressa uma forma recombinante de um polipeptídeo-alvo ou, no caso da expressão *anti-sense* a partir do gene transferido, a

expressão de uma forma de ocorrência natural do polipeptídeo-alvo é rompida.

Como aqui utilizado, o termo "transgene" significa uma sequência de ácidos nucléicos (que codifica, por exemplo, um dos polipeptídeos-alvo, ou um transcrito 5 antissense a ele) que foi introduzida em uma célula. Um transgene poderia ser parcial ou inteiramente heterólogo, isto é, estranho para o animal ou células transgênicas na qual ele é introduzido, ou, é homólogo a um gene endógeno do animal ou célula transgênica dentro da qual ele é 10 introduzido, mas que é desenhado para ser inserido ou é inserido no genoma do animal, de modo a alterar o genoma da célula dentro da qual ele é inserido (por exemplo, ele é inserido em um local que difere daquele do gene natural ou sua inserção resulta em *knockout*). Um transgene pode estar 15 também presente em uma célula na forma de um epissoma. Um transgene pode incluir uma ou mais sequências reguladoras transcricionais e qualquer outro ácido nucléico, tais como íntrons, que podem ser necessários para a expressão ideal de um ácido nucléico selecionado. 20

Uma "planta transgênica" refere-se a qualquer planta na qual uma ou mais das células da planta contêm um ácido nucléico introduzido por meio de intervenção humana, tal como por técnicas transgênicas bem conhecidas nesse 25 campo científico. O ácido nucléico é introduzido na célula, direta ou indiretamente, pela introdução em um precursor da célula, por meio de manipulação genética deliberada, tal como por microinjeção ou por infecção com um vírus recombinante. O termo "manipulação genética" não inclui

hibridização clássica, mas ao invés disso é direcionada para a introdução de uma molécula de DNA recombinante. Esta molécula pode ser integrada dentro de um cromossoma, ou ela pode ser DNA replicativo de forma extracromossômica. Nas

5 plantas transgênicas típicas aqui descritas, o transgene faz com que as células expressem uma forma recombinante dos polipeptídeos-alvo, por exemplo, formas agonísticas ou antagonísticas. Entretanto, as plantas transgênicas nas quais o gene-alvo recombinante é silencioso também estão

10 contempladas, como por exemplo, construções recombinantes dependentes FLP ou CRE. Além disso, o termo "planta transgênica" inclui também os animais recombinantes nos quais o rompimento do gene de um ou mais genes vegetais é causada por intervenção humana, incluindo técnicas de

15 recombinação e *antisense*. Uma "planta transgênica" é, ainda, uma que foi geneticamente modificada para conter e expressar seqüências de DNA heterólogas, seja como moléculas de RNA reguladoras ou como proteínas. Como aqui especificamente exemplificado, uma planta transgênica é

20 modificada geneticamente para conter e expressar pelo menos uma seqüência de DNA heteróloga ligada operacionalmente e sob o controle regulador de seqüências de controle da transcrição que funcionam em células ou tecidos vegetais ou em plantas integrais. Como aqui utilizado, o termo "planta

25 transgênica" refere-se também à progênie da planta transgênica inicial cuja progênie contém e são capazes de expressar a seqüência codificadora heteróloga sob o controle regulador das seqüências de controle da transcrição expressáveis em plantas, aqui descritas. As sementes que

contêm embriões transgênicos estão englobadas dentro desta definição, como o são retalhos e outros materiais da planta para propagação vegetativa de uma planta transgênica.

Quando a expressão vegetal de um gene heterólogo
5 ou sequência codificadora de interesse é desejada, esta
sequência codificadora é ligada operacionalmente na
orientação sense a um promotor apropriado e vantajosamente
sob o controle regulador de sequências de DNA que regulam
quantitativamente a transcrição de uma sequência a jusante
10 em células ou tecidos vegetais ou em plantas, na mesma
orientação que o promotor, de tal modo que seja produzido um
RNAm sense (isto é, funcional para a expressão
translacional). Um sinal da terminação da transcrição, por
exemplo, como sinal de poliadenilação, funcional em uma
15 célula vegetal fica vantajosamente colocado a jusante da
sequência codificadora de resistência a metais ou
organometálicos, e um marcador selecionável que pode ser
expressado em uma planta, pode ser ligado de forma covalente
à unidade de expressão induzível, de tal modo que depois
20 disso esta molécula de DNA seja introduzida em uma célula ou
tecido vegetal, sua presença possa ser selecionada e as
células ou tecidos vegetais não assim transformados sejam
exterminados ou impedidos de crescer. Na presente invenção,
a sequência codificadora de resistência ao mercúrio pode
25 servir como um marcador selecionável para transformação de
células ou tecidos vegetais. Quando a expressão de um gene
constitutivo é desejada, os promotores apropriados
expressáveis em plantas incluem os promotores 35S ou 19S do
vírus do mosaico da couve-flor, os promotores "nos", "ocs"

ou "mas" de plasmídeos Ti de *Agrobacterium tumefaciens*, e outros conhecidos nessas técnicas. Quando a expressão específica de tecidos da seqüência codificadora de resistência a metais expressável em plantas for desejada, os
5 versados nessas técnicas devem escolher entre inúmeras seqüências bem conhecidas para mediar essa forma de expressão gênica. Promotores regulados ambientalmente também são conhecidos nessas técnicas, e os versados nessas técnica podem escolher entre seqüências reguladoras da
10 transcrição para conseguir o resultado desejado.

O termo "seqüência reguladora transcricional" é um termo genérico utilizado neste relatório descritivo inteiro para se referir a seqüências de DNA, tais como sinais de iniciação, intensificadores e promotores que induzem ou
15 controlam a transcrição de seqüências codificadoras de proteínas com as quais elas estão ligadas operacionalmente. Em algumas modalidades, a transcrição de um gene recombinante fica sob o controle de uma seqüência promotora (ou outra seqüência reguladora transcricional) que controla
20 a expressão do gene recombinante em um tipo de célula no qual a expressão é intencionada. Deve-se entender que o gene recombinante pode ficar sob o controle de seqüências reguladoras transcricionais que são iguais ou diferentes daquelas seqüências que controlam a transcrição da forma de
25 ocorrência natural da proteína.

O termo "vetor" refere-se a uma molécula de ácido nucléico capaz de transportar outro ácido nucléico ao qual ele foi ligado. Um tipo de vetor útil é um episssoma, isto é, um ácido nucléico capaz de replicação extracromossômica.

São particularmente úteis os vetores capazes de replicação e/ou expressão autônoma de ácidos nucleicos aos quais eles estão ligados. Os vetores capazes de direcionar a expressão de genes aos quais eles estão ligados operacionalmente são aqui referidos como "vetores de expressão". Genericamente, os vetores de expressão de utilidade em técnicas de DNA recombinante estão freqüentemente na forma de "plasmídeos" que se referem genericamente a alças circulares de DNA de filamento duplo, as quais, na sua forma de vetor não estão ligadas ao cromossoma. No presente relatório descritivo, "plasmídeo" e "vetor" são utilizados de forma intercambiável, pois a forma mais comumente usada de vetor. Entretanto, a invenção pretende incluir outras formas de vetores de expressão que servem para funções equivalentes e que se tornam conhecidos nessas técnicas subsequente-mente.

O termo "alelo do tipo selvagem" refere-se a um alelo de um gene, que, quando presente em duas cópias em um indivíduo resulta em um fenótipo do tipo selvagem. Podem existir vários alelos do tipo selvagem diferentes de um gene específico, pois certas mudanças de nucleotídeos em um gene podem não afetar o fenótipo de um indivíduo que tem duas cópias do gene com as cópias de nucleotídeos.

4.3 Sequências de Ácidos Nucleicos de F-box e Polipeptídeos

As seqüências de de ácidos nucleico e proteínas F-Box da invenção incluem a proteína F-box de *Arabidopsis thaliana* (N^{os} de Acesso do GenBank NM_111499 (DNAC) e NP_566277 (proteína), vide Figuras 10B e 10A, respectivamente).

Outros ácidos nucléicos e polipeptídeos equivalentes include aqueles discerníveis a partir da sinopse que se segue desta família de proteínas F-box (At3g06240).

5 SIMILARIDADES DE PROTEÍNAS SELECIONADAS

Comparação de seqüências em UniGene com proteínas sustentadas por um genoma completo. Os alinhamentos podem sugerir a função de um gene.

A. *thaliana* referência: NP_566277.1 - proteína
10 expressada [Arabidopsis thaliana] 100 % / 416 aa (vide ProtEST)

EXPRESSÃO GÊNICA

Tecidos e estágios de desenvolvimento a partir da expressão gênica das seqüências deste gene. Ligações a
15 outros recursos de expressão de NCBI.

Fontes de DNAc: mistura de siliqua e flor ; síliquas verdes; raízes; semente; tecido vegetativo adulto; botões de flores; flores e botões; flores abertas sem inflorescência ou mis velhas.

20 Perfis de GEO: Perfis da expressão gênica na base de dados do NCBI Gene Expression Omnibus.

POSIÇÃO NO MAPEAMENTO

Localização genômica especificada por mapeamento de transcritos, mapeamento de híbridos por radiação,
25 mapeamento genético ou mapeamento citogenético.

Cromossoma de Arabidopsis: III

SEQÜÊNCIAS

As Sequências que representam este gene; RNAm's, ESTs, e previsões do gene fundamentadas por seqüências transcritas.

NM_111499.3 Proteína da família F-box de
5 Arabidopsis thaliana (At3g06240) RNAm, completo cds P

AY084423.1 clone de Arabidopsis thaliana 108003
RNAm, seqüência completa P

RNAm At3g06240 de Arabidopsis thaliana AK118303.1
para proteína desconhecida, cds completo, clone: RAFL19-58-
10 P14 P

RNAm (At3g06240) da família de proteínas F-box
putativas do clone U50970 de Arabidopsis thaliana
BT006048.1, cds completo P

Seqüência completa do DNAc de comprimento inteiro
15 de Arabidopsis thaliana BX824678.1 a partir do clone
GSLTPGH64ZA06 do calo tratado com hormônio da cepa col-0 de
Arabidopsis thaliana (agrião-amargo) P

Seqüência completa do DNAc de comprimento inteiro
de Arabidopsis thaliana BX829355.1 a partir do clone
20 GSLTSIL94ZF11 da siliqua da cepa col-0 de Arabidopsis
thaliana (agrião-amargo) P

Seqüências EST (19)

BP842389.1 Clone RAFL21-54-K05 5' lido

BP825839.1 Clone RAFL22-69-013 mistura de siliqua
25 e flor 5' lido

BP824403.1 Clone RAFL22-09-H01 mistura de siliqua
e flor 5' lido

BP817574.1 Clone RAFL22-45-F24 mistura de siliqua
e flor 5' lido

BP652696.1 Clone RAFL19-94-A01 misturas de siliqua
e flor 3' lido

BP644487.1 Clone RAFL19-65-J12 mistura de siliqua
e flor 3' lido

5 BP643288.1 Clone RAFL19-61-H05 mistura de siliqua
e flor 3' lido

CF773451.1 Clone 17B03 flores abertas sem
inflorescência ou mais velhas

10 BX838532.1 Clone GSLTFB50ZF02 flores e botões 5'
lido P

BX836858.1 Clone GSLTLS79ZF01 tecido vegetativo
adulto 5' lido

BX837508.1 Clone GSLTLS79ZE01 tecido vegetativo
adulto 5' lido P

15 AU239163.1 Clone RAFL19-58-P14 mistura de siliqua
e flor 5' lido P

AU230440.1 Clone RAFL19-58-P14 mistura de siliqua
e flor 3' lido

20 AV562416.1 Clone SQ169f12F síliquas verdes 3'
lido A

AV551881.1 Clone RZ15g09R raízes 5' lido

AV540406.1 Clone RZ15g09F raízes 3' lido P

AV534311.1 Clone FB078a02F botões de flores 3'
lido P

25 BE525819.1 Clone 600034345R1 semente 5' lido

BE523193.1 Clone M35B4 semente 5' lido P.

Proteínas F-box Homólogas

A invenção fornece ainda genes codificadores de
proteínas F-box, que são homólogas àquelas aqui descritas

especificamente (por exemplo, a proteína F-box de *Arabidopsis thaliana* ilustrada na Figura 10A, e a sequência codificadora correspondente demarcada na Figura 10B). Por exemplo, as sequências de nucleotídeos equivalentes incluem
5 as sequências que diferem em uma ou mais substituições, adições ou deleções de nucleotídeos, tais como as variações intragênero; e incluem também sequências que diferem da sequência de nucleotídeos que codifica a parte da proteína aqui representada em virtude da degeneração do código
10 genético. Os ácidos nucléicos equivalentes incluem também as sequências de nucleotídeos que hibridizam sob condições severas (isto é, equivalentes a cerca de 20-27° C abaixo da temperatura de fusão (T_m) do DNA duplex formado em sal cerca de 1 M) para uma sequência de nucleotídeos de um gene
15 heterólogo da invenção.

Os ácidos nucléicos particularmente úteis codificam polipeptídeos que compreendem uma sequência de aminoácidos que é pelo menos 70% idêntica, 80% idêntica ou 85% idêntica a uma sequência de aminoácidos da invenção. Os
20 ácidos nucléicos que codificam polipeptídeos, particularmente polipeptídeos que retêm um atividade de um dos genes heterólogos em questão que conferem uma atividade de F-box, e que compreendem uma sequência de aminoácidos que é pelo menos cerca de 90%, ou pelo menos cerca de 95%, e
25 mais preferivelmente pelo menos cerca de 98-99% idêntica a uma sequência de aminoácidos da invenção, também estão dentro do âmbito da invenção.

Outro aspecto da invenção fornece um ácido nucléico que hibridiza sob condições de alta ou baixa

severidade para um ácido nucléico que codifica um polipeptídeo idêntico ou homólogo a uma sequência de aminoácidos da invenção. As condições de severidade apropriadas, que promovem a hibridização do DNA, por exemplo, 6,0 x cloreto de sódio/citrato de sódio (SSC) a 5 cerca de 45 °C., em seguida uma lavagem com 2,0 x SSC a 50 °C, são conhecidas pelos versados nessas técnicas ou podem ser encontradas em "Current Protocols in Molecular Biology", John Wiley & Sons, NY (1989) 6.3.1-6.3.6. Por exemplo, a 10 concentração de sais na etapa de lavagem pode ser selecionada entre uma baixa severidade de cerca de 2,0 x SSC a 50 °C até uma alta severidade de cerca de 0,2 x SSC a 50 °C. Além disso, a temperatura na etapa de lavagem pode ser aumentada de condições de baixa severidade à temperatura 15 ambiente, cerca de 22 °C, até alta severidade a cerca de 65 °C.

Os ácidos nucléicos isolados que codificam uma proteína heteróloga da presente invenção, mas que diferem das sequências de nucleotídeos aqui referidas, em virtude da 20 degeneração do código genético, também estão dentro do âmbito da invenção. Tais ácidos nucléicos são entendidos como sendo capazes de codificar polipeptídeos funcionalmente equivalentes (isto é, um polipeptídeo que tem pelo menos uma parte da atividade biológica de uma proteína codificada 25 pelas sequências enumeradas). Por exemplo, inúmeros aminoácidos são designados por mais do que um triplete. Os códons que especificam o mesmo aminoácido (por exemplo, CAU e CAC são sinônimos de histidina) podem resultar em mutações "silenciosas" que não afetam a sequência de aminoácidos da

proteína. Entretanto, espera-se que polimorfismos de seqüências de DNA, que realmente levam a mudanças nas seqüências de aminoácidos da proteína devem existir, mesmo dentro da mesma espécie. Os versados nessas técnicas devem
5 avaliar que estas variações em um ou mais nucleotídeos (até cerca de 3-4% dos nucleotídeos) de um gene que codifica uma proteína podem existir entre células individuais de uma dada espécie, por exemplo, entre uma população de células de *Arabidopsis thaliana*, em virtude a uma variação alélica
10 natural. Qualquer uma e todas essas variações de nucleotídeos e os polimorfismos de aminoácidos resultantes estão dentro do âmbito desta invenção.

Os fragmentos das partes que codificam ácidos nucléicos da F-box heteróloga em questão, tais como os
15 fragmentos que retêm a capacidade de interagir com a via proteolítica de ubiquitina, também estão dentro do âmbito da invenção. Como aqui utilizados, tais fragmentos se referem a seqüências de nucleotídeos que têm menos nucleotídeos do que a seqüência codificadora do gene, mas ainda assim
20 incluem o suficiente da seqüência codificadora de modo a codificar um polipeptídeo com pelo menos uma parte da atividade da proteína de comprimento inteiro.

Em ainda outra modalidade, os genes reguladores recombinantes pode incluir ainda seqüências de nucleotídeos
25 adicionais. Por exemplo, o gene recombinante pode incluir seqüências de nucleotídeos de um fragmento da PCR gerado ampliando o gene a partir de uma biblioteca de DNA genômico, por exemplo, seqüências não-codificadoras de 5' e 3' de qualquer um dos genes em questão.

Os ácidos nucléicos dentro do âmbito da invenção podem conter também seqüências ligantes, sítios de endonucleases de restrição modificados e outras seqüências úteis para clonagem molecular, expressão ou purificação dos polipeptídeos recombinantes.

Como indicado pelos exemplos enunciados abaixo, um ácido nucléico que codifica uma das proteínas em questão pode ser obtido a partir do RNAm presente em uma amostra de células eucarióticas, tais como aquelas de uma planta vascular. Será possível também obter ácidos nucléicos que codificam s proteínas em questão a partir do DNA genômico obtido a partir de tais células. Por exemplo, um gene que codifica uma das proteínas F-box em questão pode ser clonado a partir de um DNAc ou uma biblioteca genômica de outras espécies vegetais de acordo com os protocolos aqui descritos, bem como aqueles conhecidos genericamente nessas técnicas. Por exemplo, um DNAc que codifica uma proteína heteróloga pode ser obtido isolando o RNAm total de uma planta, gerando DNAc's de filamento duplo a partir do RNAm total, clonando o DNAc em um plasmídeo ou vetor bacteriófago apropriado, e isolando os clones que expressam a proteína em questão, usando qualquer uma das inúmeras técnicas conhecidas, por exemplo, sondas de oligonucleotídeos, análise *Western blot*, ou complementação. Os genes que codificam proteínas relacionadas também podem ser clonados usando técnicas estabelecidas de reação de polimerase em cadeia, de acordo com as informações sobre as seqüências de nucleotídeos fornecidas pela invenção. O ácido nucléico da invenção pode ser DNA ou RNA.

Além disso, a sequência de nucleotídeos determinada a partir da clonagem dos genes heterólogos em questão permitirá a geração de sondas desenhadas para uso na identificação do DNA transgênico heterólogo, bem como para
5 detectar a presença do RNAm heterólogo correspondente. Por exemplo, os ácidos nucléicos em questão podem ser usados após o assesto transgênico para confirmar a presença e a integridade da sequência introduzida bem como a quantidade e a especificidade da expressão na progênie transgênica. Por
10 exemplo, a presente invenção fornece uma sonda/iniciador que compreende um oligonucleotídeo substancialmente purificado, onde o oligonucleotídeo compreende uma região da sequência de nucleotídeos que hibridiza sob condições severas para pelo menos 10, 25, 50, ou 100 nucleotídeos consecutivos da
15 sequência *sense* ou *antisense* de um dos ácidos nucléicos em questão, ou seus mutantes de ocorrência natural. Em algumas modalidades, a sonda/iniciador compreende ainda um grupo marcador anexado a à sonda/iniciador para ser detectado; por exemplo, o grupo marcador é selecionado no grupo que
20 consiste em radioisótopos, compostos fluorescentes, enzimas, e co-fatores de enzimas.

Outras Proteínas F-box

Ainda outras proteínas F-box para uso na invenção são conhecidas nessas técnicas e estão descritas
25 especificamente alhures (vide, por exemplo, patentes nºs US 6.573.094 e 6.232.081, cujos teores são aqui incorporados como referência em sua totalidade). Outras proteínas F-box vegetais conhecidas incluem a proteína F-box TIR1, um receptor de auxina (vide Dharmasiri *et al.*, *Nature*. 435:441-

5 (2005); nº de Acesso do GenBank Q570C0; GI: 68053009), bem como três proteínas F-box adicionais, denominadas AFB1, 2, e 3, que também regulam a resposta de auxina (vide Dharmasiri *et al.*, *Dev. Cell* 9:109-119 (2005)). Na realidade, o genoma de *Arabidopsis* sozinho codifica quase 700 proteínas F-box (vide Gagne *et al.*, *Proc. Natl. Acad. Sci. USA* 99:11519-24 (2002)). Ainda outras proteínas F-box vegetais são conhecidas nessas técnicas e estão disponíveis para os versados nessas técnicas usando métodos apenas rotineiros em
10 clonagem gênica.

Vetores

Esta invenção fornece também vetores de expressão que incluem uma sequência de nucleotídeos que codifica um dos polipeptídeos em questão, e ligada operacionalmente a
15 pelo menos uma sequência reguladora. Ligada operacionalmente pretende significar que a sequência de nucleotídeos está ligada a uma sequência reguladora de uma maneira que permite a expressão da sequência de nucleotídeos. As sequências reguladoras vegetais são
20 reconhecidas nessas técnicas. Conseqüentemente, o termo "sequência reguladora" inclui promotores, intensificadores e outros elementos de controle da expressão. As sequências reguladoras vegetais exemplificativas estão descritas em Yusibo *et al.*, *Curr. Top. Micro. Immun.* 240:81-94 (1999) e
25 Hood *et al.*, *Adv. Exp. Med. Biol.* 464:127-47 (1999). Por exemplo, qualquer uma entre uma série de sequências de controle da expressão que controlam a expressão de uma sequência de DNA quando ligada operacionalmente a ela pode ser usada nesses vetores para expressar sequências de DNA

que codificam as proteínas e os ácidos nucléicos da invenção, que conferem o traço genético de *Zostera*. Tais seqüências úteis de controle da expressão incluem, por exemplo, o promotor constitutivo de ubiquitina do milho (promotor ubi) (Christensen *et al.*, *Plant Mol. Biol.* 18:675-89 (1992); Cornejo *et al.*, *Plant Mol. Biol.* 23:567-81 (1993)) e a seqüência terminadora PinII da batata (An *et al.*, *Plant Cell* 1:115-22 (1989)). Outras seqüência de controle da expressão são aquelas derivadas de vírus vegetais tais como: o promotor 35S, que é derivado de seqüências do vírus do mosaico da couve-flor; e o promotor da proteína de revestimento TMV, tal como aquele contido no vetor de clonagem designado "30B" que é derivado do vírus do mosaico do tabaco. Estão incluídas também em certos aspectos da invenção as seqüências reguladoras transcricionais não-vegetais, tais como os promotores precoces e tardios de SV40, promotor precoce imediato de adenovírus ou citomegalovírus, os sistema lac, o sistema trp, o sistema TAC ou TRC, promotor T7 cuja expressão é direcionada por T7 RNA polimerase, as maiores regiões operadoras e promotoras do fago lambda, as regiões de controle para a proteína de revestimento fd, o promotor para 3-fosfoglicerato cinase ou outras enzimas glicolíticas, os promotores de fosfatase ácida, por exemplo, Pho5, os promotores dos fatores de reprodução "a" de leveduras, o promotor poliedro dos sistema de baculovírus e outras seqüências que reconhecidamente controlam a expressão de genes de células procarióticas ou eucarióticas ou seus vírus, e várias combinações deles. De-se entender que o desenho do vetor de expressão pode

dependem de fatores tais como a escolha da célula hospedeira a ser transformada e/ou do tipo de proteína que se deseja expressar. Além disso, o número de cópias do vetor, a capacidade de controlar o número e a expressão de quaisquer
5 outras proteínas codificadas pelo vetor, tais como marcadores de antibióticos, também devem ser considerados.

A construção recombinante da presente invenção pode incluir um marcador selecionável para propagação da construção. Por exemplo, uma construção a ser propagada em
10 bactérias contém, de preferência, um gene de resistência a antibiótico, tal como aquele que confere resistência à canamicina, tetraciclina, estreptomicina, ou cloranfenicol. Os vetores apropriados para propagar a construção incluem plasmídeos, cosmídeos, bacteriófagos ou vírus, para
15 mencionar apenas alguns.

Além disso, as construções recombinantes podem incluir genes de marcadores expressáveis em plantas selecionáveis ou que podem ser triados, para isolar, identificar ou rastrear células vegetais transformadas por
20 estas construções. Os marcadores selecionáveis incluem, porém sem limitações, genes que conferem resistências a antibióticos (por exemplo, resistência à canamicina ou higromicina) ou resistência a herbicidas (por exemplo, resistência a sulfoniluréia, fosfinotricina, ou glifosato).
25 Os marcadores que podem ser triados incluem, porém sem limitações, os genes que codificam beta-glicuronidase (Jefferson, *Plant Mol. Biol. Rep.* 5:387-405 (1987)), luciferase (Ow et al., *Sci.* 234:856-859 (1986)), produtos

gênicos B e C1 que regulam a produção do pigmento antocianina (Goffet *et al.*, *EMBO J.* 9:2517-2522 (1990)).

A presente invenção pode utilizar também o sistema de *Agrobacterium* para transformar plantas, sendo que as
5 construções de DNA recombinante compreendem adicionalmente pelo menos a sequência limítrofe direita do T-DNA que flanqueia as sequências de DNA a serem transformadas na célula vegetal. Em algumas modalidades, as sequências a serem transferidas flanqueadas pelas sequências limítrofes
10 direita e esquerda do T-DNA. O desenho e a construção apropriada desses vetores de transformação baseados em T-DNA são bem conhecidas pelos versados nessas técnicas.

Esta invenção refere-se também a uma célula hospedeira transfectada com uma gene recombinante para que
15 ela possa expressar uma proteína recombinante da presente invenção. A célula hospedeira pode ser qualquer célula procariótica ou eucariótica. Por exemplo, uma proteína F-box vegetal da presente invenção pode ser expressada em células bacterianas, tal como *E. coli*, células de insetos,
20 levedura, ou células de mamíferos. Outras células hospedeiras apropriadas são conhecidas pelos versados nessas técnicas.

Outro aspecto da presente invenção refere-se a formas recombinantes das proteínas vegetais em questão. O
25 termo "proteína recombinante" refere-se a uma proteína da presente invenção que é produzida por técnicas de DNA recombinante, onde genericamente o DNA que codifica a proteína é inserido em um vetor de expressão apropriado, o qual, por sua vez, é usado para transformar uma célula

hospedeira para produzir a proteína heteróloga. Além disso, o termo "derivado de", com relação a um gene recombinante que codifica as proteínas em questão, pretende incluir dentro do significado de "proteína recombinante" as

5 proteínas que têm uma sequência de aminoácidos da forma nativa (ou "autêntica") da proteína vegetal, ou uma sequência de aminoácidos similar a ela, que é gerada por mutação de modo a incluir substituições e/ou deleções em relação a uma forma de ocorrência natural da proteína. Para

10 ilustrar, as proteínas recombinantes úteis na presente invenção, além daquelas que têm uma sequência de aminoácidos das proteínas nativas, são aquelas proteínas recombinantes que têm sequências de aminoácidos que são pelo menos 70% homólogas, ou 80% homólogas, e mais preferivelmente, 90%

15 homólogas a uma sequência de aminoácidos da presente invenção. Um polipeptídeo que tem uma sequência de aminoácidos que é pelo menos cerca de 95%, ou pelo menos cerca de 98%, e mais preferivelmente, idêntica a uma das sequências de polipeptídeo da invenção, também estão dentro

20 do âmbito da invenção. Assim sendo, a presente invenção refere-se a proteínas recombinantes que são derivadas, for exemplo, de genes de F-box de *Arabidopsis thaliana* e que têm sequências de aminoácidos evolutivamente relacionadas a uma sequência codificada por um gene ortólogo de outra proteína

25 vegetal, onde "evolutivamente relacionada a" refere-se a polipeptídeos que têm sequências de aminoácidos que surgiram naturalmente (por exemplo, por variância alélica), bem como variantes mutacionais das proteínas reguladoras que são derivadas, por exemplo, por mutagênese combinatória.

4.4 Plantas Transgênicas e Células Vegetais

As técnicas para incorporar de forma estável construções genéticas no genoma de plantas-alvo são bem conhecidas nesse campo científico e incluem introdução mediada por *Agrobacterium tumefaciens*, eletroporação, fusão de protoplastos, injeção em órgãos reprodutivos, injeção em embriões imaturos, introdução de projéteis em alta velocidade, e similares. A escolha da técnica dependerá da planta-alvo a ser transformada. Por exemplo, plantas dicotilédones e certas monocotilédones e gimnospermas pode ser transformadas pela tecnologia de plasmídeo *Agrobacterium* Ti (como descrito, por exemplo, por Bevan, *Nucl. Acids Res.* 12:8711-8721 (1984)). Os alvos para a introduções de construções genéticas da presente invenção incluem tecidos, tal como o tecido de folhas, células dissociadas, protoplastos, sementes, embriões, regiões meristemáticas; cotilédones, hipocótilos, e similares.

Depois que as células são transformadas, as células que têm a construção genética da invenção incorporada no seu genoma podem ser selecionadas por meio de um marcador, tal como o marcador de resistência à canamicina discutido acima. As células transgênicas podem ser então cultivadas em um meio apropriado para regenerar plantas integrais, usando técnicas conhecidas nesse campo científico. No caso de protoplastos, a célula é deixada se reformar sob condições osmóticas apropriadas. No caso de sementes ou embriões, um meio de germinação ou iniciação de calo apropriado é empregado. No caso de explantes, usa-se um meio de regeneração apropriado. A regeneração de plantas

está bem estabelecida para muitas espécies. Para obter uma revisão da regeneração de árvores florestais, vide Dunstan *et al.* A embriogênese somática em plantas lenhosas está descrita em Thorpe, editora T. A., "In Vitro Embryogenesis of Plants", Volume 20 em "Current Plant Science and Biotechnology in Agriculture", Capítulo 12, páginas 471-540, 1995. Os protocolos específicos para a regeneração de abeto-vermelho estão descritos por Roberts *et al.*, "Somatic Embryogenesis of Spruce" em: "Synseed Applications of synthetic seed to crop improvement", editor Redenbaugh, K., CRC Press, Capítulo 23, páginas 427-449, 1993. As plantas transformadas resultantes podem ser reproduzidas de forma sexuada ou assexuada, usando métodos bem conhecidos nessas técnicas, para dar gerações sucessivas de plantas transgênicas.

Como discutido acima, a produção de RNA em células vegetais-alvo pode ser controlada pela escolha da sequência promotora. Uma planta-alvo pode ser transformada com mais do que uma construção genética da presente invenção, modulando desta forma a atividade de mais do que um polipeptídeo, afetando a atividade dos polipeptídeos em mais do que um tecido, ou afetando a atividade de polipeptídeos em mais do que um tempo de expressão. Similarmente, uma construção genética pode ser montada contendo mais do que uma estrutura de leitura aberta que codifica um polipeptídeo da invenção ou mais do que uma região não-codificadora de um gene que codifica tal polipeptídeo. Os polinucleotídeos da presente invenção podem ser empregados também em combinação

com outras seqüências conhecidas que codificam os polipeptídeos envolvidos na sinalização de células vegetais.

Os polinucleotídeos isolados da presente invenção podem ser empregados como sondas para isolar outras
5 seqüência do DNA de F-box relacionadas, incluindo aquelas de outras espécie vegetais, usando técnicas bem conhecidas pelos versados nessa área, tais como as técnicas usadas rotineiramente de hibridização de DNA e PCR.

Os polinucleotídeos, polipeptídeos e anticorpos e
10 os anticorpos para esses polipeptídeos podem ser usados para triar moléculas que interagem com tais polinucleotídeos e/ou polipeptídeos e que, desta forma, modulam a sinalização de células. As técnicas para realizar esses ensaios são bem conhecidas nessa área. Similarmente, os polinucleotídeos e
15 polipeptídeos da presente invenção podem ser empregados em estudos desenhados para elucidar o mecanismo de vias de sinalização de células.

A invenção fornece inúmeros métodos para a transformação de plantas com um gene ou genes de F-box
20 heterólogos que contribuem para a produção de óleos na planta receptora, cujos detalhes estão descritos adicionalmente abaixo. Em algumas modalidades, o gene heterólogo é introduzido por transformação, e o gene introduzido é expressado de forma estável durante a vida da
25 planta e é capaz ainda de ser transmitido para a progênie da planta. Genericamente, é desejável que o transgene seja integrado no DNA nuclear, embora o genoma do plastídeo possa ser um alvo apropriado para algumas construções.

A transformação da cultura e outras plantas pode ser efetuada por inúmeros métodos conhecidos no campo de biotecnologia de plantas. O método para a transformação variará com a espécie da planta a ser transformada e com o padrão desejado e a estabilidade da expressão do transgene. Por exemplo, métodos de bombardeio de partículas demonstraram ser eficazes na transformação de muitas espécies vegetais, incluindo aquelas anteriormente consideradas recalcitrantes para transformação. Este método é usado comumente na transformação de plantas monocotilédones, tal como o milho. Outro método de transformação de plantas disponível é a transferência gênica mediada por *Agrobacterium*, que é usado comumente para transformar culturas dicotilédones.

Ainda outros métodos disponíveis para transformação de plantas não se baseiam em cultura de tecidos para a recuperação de plantas transgênicas, permitindo desta forma a produção de transgênicos a partir de espécies vegetais para as quais não existe qualquer método confiável de cultura de tecidos. Por exemplo, o microassesto de DNA ligado a partículas para dentro do tecido meristemático de brotos produz partes fluorescentes transgênicas a partir das quais surgem sementes transgênicas (Sautter et al., *Biotech.* 9:1080-85 (1991)). Sementes transgênicas podem ser criadas também por eletroforese de DNA em tecido meristemático (Griesbach, *Plant Sci.* 102:81-89 (1994); Burchi et al., *J. Genet. Breeding* 49: 163-8 (1995)). Este método demonstrou ser bem sucedido na transformação de várias espécies vegetais, incluindo orquídeas, crisântemos,

cravos, lisianto, pimentas, e mesmo espécies de plantas lenhosas, tal como ameixa *Plumus domestica*).

Genericamente, a invenção fornece métodos e reagentes para a engenharia genética de uma planta hospedeira-alvo, tal como uma planta cultivada, com um ácido nucléico heterólogo que codifica uma função de proteína F-box, que contribui para a produção de óleos na planta recebedora. Um método para transformação faz uso da bactéria de solo comum supramencionada *Agrobacterium* (vide Birch, *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 48:297-326 (1997)). Este método envolve o vetor de transferência de DNA (T-DNA) modificado que carrega o fragmento de ácido nucléico desejado entre as regiões limítrofes do T-DNA (regiões de repetições diretas específicas com 25 pares de bases). O vetor resultante é transferido para dentro de um hospedeiro de *Agrobacterium* e a planta hospedeira-alvo é inoculada com a bactéria recombinante transformada. Os produtos gênicos virulentos de *Agrobacterium*, então, reconhecem ativamente, excisam, transportam e integram a região do T-DNA dentro do genoma da planta hospedeira.

As técnicas de transformação mediadas por *Agrobacterium tumefaciens*, incluindo desarmamento e uso de vetores binários, estão bem descritas na literatura científica (vide, por exemplo, Horsch, et al., *Science* 233:496-8 (1984), e Fraley, et al., *Proc. Nat. Acad. Sci. USA* 80:4803 (1983)). A transformação mediada por *Agrobacterium* é um método particularmente útil para transformação de dicotilédones.

A faixa de hospedeiras naturais de *Agrobacterium* é limitada e assim sendo esta abordagem para transformação não é praticável em algumas plantas hospedeiras-alvo, particularmente culturas de cereais e outras espécies monocotilédones. Para essas culturas, a invenção fornece abordagens alternativas para transformação, tal como a captação direta de DNA nu em protoplastos ou tecidos, usando eletroporação ou bombardeio de partículas com pistola. Neste método, a co-transformação de um gene marcador selecionável junto com o gene de interesse permite o crescimento preferencial das células transformadas em cultura de células. Manipulações sucessivas da composição química do meio de cultura, especialmente hormônios vegetais, permitem a regeneração de plantas completas. Este método permitiu a recuperação de plantas geneticamente manipuladas em virtualmente todas as plantas cultivadas.

Um método representativo para transformação direta da construção do transgene é por bombardeio de partículas contra tecidos de plantas-alvo com microprojéteis de alta velocidade (vide, por exemplo, Finer et al. *Curr. Top. Micro. and Immun.* 240:60-80 (1999)). Este método utiliza um acelerador de partículas ou "gene blaster" para penetrar nas camadas superficiais externas do tecido ou protoplasto da planta (Sanford, *Trends Biotech.* 16:299-302 (1998)). A "biolística", uma combinação de "biológica" e "balística", descreve uma técnica que utiliza instrumentação para acelerar microprojéteis revestidos com DNA para dentro de células, atravessando a parede celular e a membrana celular. O microprojétil é genericamente pequeno o suficiente (0,5-

5,0 mm) para entrar na célula vegetal sem dano demais, mas ainda grande o suficiente para fazer com que a massa penetre na parede celular e conduza uma quantidade apropriada de DNA sobre sua superfície para dentro do interior da célula vegetal.

Podem ser usados inúmeros desenhos diferentes de pistolas de partículas. A base de todos estes desenhos é aplicar o DNA como um revestimento sobre pequenas partículas densas e acelerar as partículas na direção de um tecido-alvo. As partículas usualmente consistem em partículas esféricas de ouro ou tungstênio, que têm um diâmetro entre 0,5 e 5,0 mm. As partículas de ouro são quimicamente inertes, geralmente com um tamanho mais uniforme do que as partículas de tungstênio, e não produzem efeitos citotóxicos. Conseqüentemente, as partículas de ouro são genericamente mais eficazes do que as partículas de tungstênio. Idealmente, as partículas usadas para o bombardeio devem ter boa afinidade inicial pelo DNA, mas ainda assim soltarem livremente o DNA depois de estarem dentro do citoplasma ou núcleo da célula-alvos.

Para preparar os microprojéteis revestidos com DNA, as partículas de ouro ou tungstênio lavadas são misturadas com DNA plasmídico. O DNA fica ligado sobre as partículas usando métodos de precipitação com etanol ou cloreto de cálcio, que são conhecidos nessas técnicas. Espermidina pode ser adicionada à mistura, possivelmente protegendo o DNA contra degradação e/ou alterando sua conformação. Depois da precipitação, as partículas podem ser lavadas, recolocadas em suspensão e secadas ou guardadas

sobre gelo como uma suspensão aquosa até quando for necessário.

A pistola de partículas pode utilizar um macrocarreador, que sustenta ou carrega as partículas e é
5 acelerado junto com as partículas na direção do alvo. O macrocarreador é usualmente retido por uma placa ou peneira de parada, antes que ele colida com o alvo, enquanto que as partículas continuam ao longo do seu curso. Na maioria dos casos, as partículas são aceleradas sob vácuo parcial em uma
10 câmara de vácuo para reduzir o arraste pelo ar. A penetração das partículas é controlada modificando a intensidade da rajada explosiva, mudando a distância que as partículas devem percorrer para atingir o tecido-alvo, ou usando partículas de diferentes tamanhos. Um dispositivo
15 portátil comercial (a Pistola de Genes Helios) está disponível em BioRad Laboratories (Hercules, CA, EUA). Um dispositivo de bombardeio modificado com hélio, que utiliza um acúmulo contínuo de contrapressão de hélio distribuída para um disco de ruptura que transmite uma onda de choque
20 para um segundo disco ou macrocarreador que retém as partículas revestidas com DNA, também está disponível na BioRad (isto é, a unidade PDS-1000/He). Uma pistola de descarga elétrica de alta voltagem, que causa uma vaporização rápida de uma gotícula de água, a qual, por sua
25 vez, transmite uma onda de choque para uma lâmina de Mylar revestida com partículas ligadas ao DNA, também foi desenvolvida (vide McCabe e Christou, *Plant Cell Tiss. Organ. Cult.* 33:227-236 (1993)). Ainda outro dispositivo para o bombardeio de partículas é um dispositivo

microapontador, que não utiliza um macrocarreador (Sautter
et al. Bio/Technology 9:1080-5 (1991)). Este dispositivo
acelera pequenas quantidades de uma mistura de
DNA/partículas em uma corrente focalizada de nitrogênio sob
5 alta pressão. O DNA não é precipitado sobre as partículas
de ouro, mas é distribuído como uma mistura.

Vários tecidos vegetais diferentes foram usados
como alvos para a transformação mediada por bombardeio de
partículas. A seleção do tecido-alvo apropriado é
10 dependente de múltiplos fatores. Para uma análise rápida da
expressão de genes, várias construções de plasmídeos podem
ser introduzidas em diferentes tecidos e a expressão
transiente pode ser analisada rapidamente para avaliar a
atividade do promotor, sem a produção de plantas
15 transformadas de forma estável (vide, por exemplo, Iida *et al.*,
Plant Cell Rep. 14:539-44 (1995)). Quase qualquer
tecido pode ser usado para estudos de expressão transiente,
desde que a parede celular seja penetrável pelas partículas
revestidas com DNA. Por exemplo, culturas de células
20 vegetais embriogênicas foram usadas exitosamente para a
produção de plantas transformadas (vide, por exemplo, Fromm
et al., *Bio/Technology* 8:833-9 (1990)). A transformação de
meristemas apicais de brotos resulta em plantas quiméricas,
onde as células transformadas geram diretamente o tecido de
25 linhagem germinal e o DNA introduzido é então passado para
plantas progênicas. O bombardeio de tecidos meristemáticos
de brotos, e em seguida, expansão da cultura dos tecidos das
células transformadas, foram usados para produzir linhagens
de plantas transgênicas geneticamente transmissíveis

(McCable *et al.*, *Bio/Technology* 6:923-6 (1988)). Além das culturas embriogênicas e pontas de brotos, outros tecidos que foram submetidos a bombardeio de partículas incluem folhas (Klein *et al.*, *Proc. Nat. Acad. Sci.USA* 85:8502-5 (1988)), seções de raízes (Seki *et al.*, *Appl. Microbiol. Biotech.* 36: 228-30 (1991)), seções de pedúnculos (Loopstra *et al.*, *Can. J. Res.* 22: 993-6 (1992)), pólen (Twell *et al.*, *Plant Physiol.* 91:1270-4 (1989)), estilos (Clark e Sims, *Plant Physiol.* 106:25-36 (1994)), células de aleuronas de cereais (Kim *et al.* (1992) *Mol. Gen. Genet.* 232:383-93 (1992)) blastema de borla (Dupeuis e Pace, *Plant Cell Rep.* 12:607-11 (1993)). Em certos casos, o tecido vegetal selecionado para transformação mediada por bombardeio de partículas pode ser relativamente novo, pois as culturas de células prolongadas podem resultar em anormalidades que podem comprometer a utilidade da planta transgênica, tal como infertilidade da progênie transgênica subsequente (vide Rhodes *et al.*, *Biotech.* 6:56-60 (1988)).

Em certos casos, a magnitude da expressão do transgene varia acentuadamente com o local de inserção e a natureza da(s) sequência(s) inserida(s). Por exemplo, enquanto que uma transferência mediada por T-DNA resulta tipicamente na inserção de um único fragmento de DNA intacto completo em um único locus, as abordagens de transferência direta de DNA resultam freqüentemente em concatâmeros longos do DNA transferido (vide, por exemplo, Czernilofsky *et al.*, *DNA* 5:473-82 (1986)). Tais múltiplas inserções tandem estão associadas aos fenômenos de "silenciamento" transcricional em certos casos. Além disso, o local de inserção dentro do

genoma da planta freqüentemente afeta a intensidade da expressão do transgene - um fenômeno conhecido como "efeito de posição". Conseqüentemente, a invenção fornece métodos para mitigar a interferência com a expressão do transgene.

5 Por exemplo, os efeitos da posição podem ser mitigados por transgenes flanqueadores com regiões específicas associadas à matriz, que isolam a regulação transcricional contra os efeitos da cromatina circundante (vide, por exemplo, Mlynarova et al., *Plant Cell* 7:599-609 (1994)). Por

10 exemplo, regiões de anexação de armações (*scaffold*) (SARs, conhecidas também como região de anexação de matri ou MARs) podem ser incluídas na construção do vetor de transgene. Utilmente, as SARs são ligadas às regiões flanqueadoras do gene de interesse. Estas seqüências são conhecidas nessas

15 técnicas (por exemplo, uma SAR de tabaco está descrita em Breyne et al., *Plant Cell* 4:463-71 (1992); e Allen et al., *Plant Cell* 8:899-913 (1996)). Além disso, o silenciamento do transgene mediado por processos dependentes de homologia pode ser evitado utilizando linhagens de plantas

20 transgênicas que evitam padrões de inserção repetidas tandem ou invertidas, e limitando a homologia do transgene inserido com qualquer gene ou genes hospedeiros endógenos correspondentes, manipulando por engenharia genética substituições de códons pela construção de transgene, quando

25 apropriado. Quando o transgene é inserido como um fragmento de DNA intacto em um único locus, sua expressão comporta-se genericamente de uma maneira altamente consistente. Tais *loci* transgênicos apresentam a ação aditiva esperada do gene dentro dos *loci* (hemizigoto versus homozigoto) e entre *loci*

(di-híbridos entre indivíduos transgênicos homozigóticos). A perda de junção de transgenes é rara em tais linhagens transgênicas (aproximadamente uma em cada dez mil), o que é consistente com o desempenho de muitos genes vegetais endógenos. As plantas transgênicas otimizadas da invenção podem ser obtidas triando as plantas-candidatas quanto à expressão persistente do transgene em múltiplas gerações de uniões ou rodadas de propagação vegetativa.

EXEMPLOS

Os exemplos que se seguem são fornecidos a título ilustrativo e não limitativo.

Materiais e Métodos

AT3G06240: proteína da família F-box, contém o domínio Pfam:PF00646 de F-box

INICIADORES para a PCR original
data 28/02/2002

Iniciador de avanço (Fbox-For)

5'-CAC CAA ACA ATG GAG GCG ACG AAG AGA G-3'

Iniciador reverso (Fbox-Rev_1)

5'-ATC TTA GCA AAA AGC AGA CTT C-3'

Estes iniciadores foram usados para efetuar exitosamente a RT-PCR da sequência que se segue.

CACC AAACA ATGGAGGCGACGAAGAGAGAAAGACGGA

GAGAAGATGACGACGGCGAAAAAGCTTCACCGGAATCACTCGTTCTTCCA

CCAGAGATCA

TTACAGAAATTCTTCTCCGATTACCAGCCAAATCGATCGGGCGATTTCAGG

TGCGTATCAA

AGCTCTTTTGCACTTTATCGTCGGATCCAGGGTTCGCGAAGATTCACCTC

GATCTGATCC

TTCGAAACGAATCCGTAAGATCGCTCCACCGTAAGCTCATTGTGTCTTCA
CATAATCTGT
ACTCGTTAGATTTCAATTCGATCGGTGACGGAATTAGGGATTAGCGGCT
GTGGAACACA
5 ATTATCCTCTTAAAGACGATCCAAGCATTTTCTCTGAGATGATTAGGAAT
TACGTGGGGG
ACCATCTGTACGATGATCGTCGCGTGATGCTTAAGCTGAATGCGAAATCG
TATCGAAGAA
ACTGGGTTGAGATCGTTGGATCTTCCAATGGTTTAGTGTGTATCTCTCCT
10 GGTGAAGGAG
CTGTTTTCTTGTATAATCCAACCTACCGGAGATTCCAAGAGATTACCTGAA
AATTTTCGTC
CCAAATCTGTAGAATACGAAAGAGATAATTTCCAACTTATGGATTTGGT
TTCGATGGTC
15 TCACTGATGATTACAAATTGGTGAAGCTTGTTGCTACCAGTGAAGATATT
CTCGATGCTA
GTGTCTATTCCTTGAAGGCTGACTCATGGAGACGGATCTGCAATTTGAAT
TATGAGCACA
ACGATGGCTCCTACACGTCCGGTGTGCATTTCAACGGTGCGATTCACTGG
20 GTGTTACACAG
AGAGTAGGCACAACCAAAGAGTGGTTGTAGCATTTGATATTCAAACCTGAG
GAGTTTCGAG
AGATGCCAGTGCCTGATGAAGCTGAAGATTGTTCCCATAGGTTTAGCAAC
TTTGTGGTTCG
25 GAAGTCTCAATGGACGTCTCTGTGTGGTCAATAGTTGCTACGATGTGCAT
GATGATATAT
GGGTGATGAGTGAGTACGGTGAAGCTAAATCCTGGAGCAGAATTCGAATC
AACTTGTTGT

ATAGGTCGATGAAACCGCTCTGTTCTGACTAAGAACGATGAAGAGGTTCTT
CTGGAGCTTG

ATGGAGACCTGGTGTGTACAACCTTTGAAACCAATGCATCGAGTAATCTA
GGAATTTGTG

5 GGGTTAAGCTCAGTGACGGGTTTCGAGGCAAATACATACGTAGAGAGCCTC
ATATCACCCA

ACTCTTATG GTATAGAGAGCTGA GGAAGTCTGCTTTTTGCTAAGAT

Um dois produtos da PCR purificado, Fbox_likel, foi então clonado no vetor GATEWAY pENTR/D-TOPO. Este foi
10 designado como "Clone de Entrada" e foi usado para subclonar a sequência desejada nos vetores de expressão GATEWAY (Invitrogen, Carlsbad, CA, EUA, número do catálogo pENTR/D-TOPO, K240020). Outras informações sobre as sequências estão disponíveis na Invitrogen.

15 Um dos plasmídeos do clone de entrada de Fbox foi escolhido para o sequenciamento de comprimento inteiro (codificado Fbox-like -LIG1). O clone foi sequenciado completamente usando iniciadores internos e demonstrou ser ser de comprimento inteiro e sem erros da PCR.

20 A subclonagem do fragmento do gene Fbox-likel gene para o vetor de expressão GATEWAY pGATE-Phas está ilustrada na Figura 12. O vetor foi construído inserindo o Cassete Gateway Reading Frame Cassette b no sítio PmeI do vetor parental pBBV-Phas.

25 A sequência completa de pGATE-Phas está ilustrada na Figura 11 (SEQ ID NO: 6). As sequências curtas sublinhadas em cada lado da inserção foram removidas durante a reação de ligação, isto é, onde a inserção vai. A subclonagem e a transformação de *Agrobacterium* da construção

Fbox foi então completada. Antes da transformação em *Agrobacterium*, as orientações das inserções foram verificadas por seqüenciamento, e as corretas foram utilizadas.

5 Depois da transformação em *Agrobacterium*, a presença da construção foi confirmada por PCR. Finalmente, uma mistura de 3 episódios de transformação em *Agrobacterium* foram usados para transformação em *Arabidopsis*.

Resultados da Superexpressão de of F-box

10 As linhagens que têm aumento positivo no teor de óleo foram levadas para a geração T3 e/ou T4 e reanalisadas.

A triagem demonstrou que a superexpressão de F-box (At3g06240), quando impulsionada sob um promotor específico de semente, faseolina, produziu sementes com fenótipo de
 15 teor mais alto de óleo e está ilustrada nas Figuras 1-5. Além disso, os dados obtidos para a próxima geração (T3) estão indicados nas Figuras 6-7, e outra análise de outros aspectos das sementes transgênicas está ilustrada nas Figuras 8-9. Particularmente, a Figura 1 ilustra o teor de
 20 óleo de linhagens de faseolina-F-box em comparação com linhagens de controle do vetor. Os dados indicados são médias de 18 linhagens (T2) e as barras de erros são SE. A Figura 2 ilustra as FAMES totais de linhagens transgênicas de F-box em comparação com o tipo selvagem (WT), e as
 25 linhagens de controle do veotr (barras de erros são SE baseado em 18 linhagens: 3 duplicatas para cada linhagem). A Figura 3 ilustra as FAMES totais de sementes de linhagens transgênicas versus WT e linhagens de controle do vetor (os pontos grandes são as duas linhagens com teor mais alto de

óleo e foram denominadas F-box2 e F-box3, escolhidas posteriormente para análise das próximas gerações). A Figura 4 ilustra o espectro de ácidos graxos e a composição da razão molar de ácidos graxos das linhagens transgênicas de F-box versus plantas de controle. A Figura 5 ilustra a determinação do peso das sementes (cerca de 400 sementes foram contadas para sementes do tipo selvagem e transformantes com genes de F-box e as sementes contadas foram guardadas em dessecadores de um dia para o outro, e depois determinadas usando uma balança analítica e as barras de erros são SE, baseado em 6 réplicas).

Os dados na seção acima ilustram que há um aumento de >10% no teor de óleo em linhagens que superexpressam F-box, comparando com plantas do tipo selvagem. Este aumento é confirmado na base de por mg e na base por semente. Para confirmar ainda mais este resultado, é essencial determinar as mudanças no teor de óleo para a próxima geração. Além disso, há um decréscimo pequeno, mas ainda significativo, em 18:1 e 18:2 em plantas que superexpressam F-box em comparação com o tipo selvagem, e há um aumento significativo em 18:3 observado. A razão para isto não está clara.

Avaliação do gene de F-box gene na geração T3: plantas T2 de linhagens de F-box2 e F-box3 (vide Figura 3) foram desenvolvidas sob condições usuais. Para eliminar o efeito potencial de BASTA sobre o metabolismo das sementes, os transformantes foram triados por PCR ao invés de spray BASTA. As linhagens F-box2 e F-box3 lines foram

desenvolvidas com linhagens WT e de controle do vetor. A análise de FAMES de sementes T3 está ilustrada na Figura 6.

A Figura 6 ilustra FAMES totais por semente nas sementes de F-box T3 versus WT e linhagens de controle do vetor (as barras de erros são SE, baseado em dados obtidos a partir de 12 plantas para as linhagens F-box2 e F-box3; 6 plantas para WT e linhagens de controle do vetor; e as duas linhagens aqui denominadas F-box2 e F-box3 são os dois pontos de dados indicados por pontos grandes na Figura 3). A Figura 7 ilustra FAMES totais de sementes de linhagens transgênicas (F-box2 e F-box3) versus WT e linhagens de controle do vetor.

Análise da razão C:N para sementes T2: As duas linhagens F-box2 e F-box3 foram escolhidas para a análise da razão C:N. Esta análise confirmará as mudanças nos compostos de estocagem de sementes (óleo versus proteína) (vide Figura 8). A Figura 8(a) ilustra a análise da razão C:N para duas linhagens independentes, F-box2 e F-box3. As barras de erros são SE, baseado em seis réplicas. A Figura 8(b) ilustra a correlação positiva entre teor de óleo e razão C:N, confirmando ainda mais o resultado (as colunas vazadas são plantas que superexpressam F-box2 e F-box3; as colunas cinza são sementes de *Arabidopsis* do tipo selvagem).

Resultados do knock-out de F-box

Análise de KO (knock-out) do gene de F-box:
Linhagens de inserção de Salk T-DNA (SALK_048969 e SALK_048962) foram ordenadas, e as plantas homozigóticas foram captadas por PCR a partir da linhagem SALK_048969, que tem uma inserção de T-DNA no pelo menos éxona do gene. As

plantas homzigóticas, junto com as plantas do tipo selvagem (WT) foram desenvolvidas até a maturidade; as sementes foram analisadas quanto a FAMES totais (vide Figura 9).

A Figura 9 (a) ilustra a análise de FAMES de
 5 linhagens KO de F-box, e a Figura 9(b) ilustra o espectro de ácidos graxos de linhagens KO e WT (as barras de erros baseiam-se em SU em oves repicas independente). Como indicado acima, não existiram diferenças observáveis no teor de óleo entre linhagens WT e KO de F-box; entretanto, houve
 10 um aumento significativo em 18:1 e 18:2 em KO de F-box, comparando com o tipo selvagem, e houve um decréscimo significativo em 18:3 observado. Isto espelha o perfil observado com as linhagens de super-expressão (vide Figura 4).

15 Portanto, através desta análise detalhada de linhagens transgênicas de F-box, demonstrou-se que a superexpressão do gene de F-box em *Arabidopsis* pode aumentar o teor total de óleo em mais do que 10%. Na realidade, este aumento foi confirmado em duas gerações e a análise foi
 20 validada com duas abordagens diferentes (FAMES totais e análise da razão de C:N).

Equivalentes

Os versados nessas técnicas devem reconhecer, ou serem capazes de determinar, usando experimentação não mais
 25 do que rotineira, inúmeros equivalentes das modalidades específicas aqui descritas especificamente. Pretende-se que tais equivalentes estejam englobados no âmbito das reivindicações que se seguem.

PJ 0616786-0
26/03/08

REIVINDICAÇÕES

1. Construção genética para a hiper-expressão de uma proteína de domínio-F em plantas, **CARACTERIZADA** pelo fato de compreender:

- 5 (a) uma seqüência de promotor genético; e
 (b) uma seqüência que codifica uma funcionalidade protéica de domínio-F ligada à seqüência de promotor genético.

10 2. Construção genética, de acordo com a reivindicação 1, **CARACTERIZADA** pelo fato de que a proteína de domínio-F é uma proteína de domínio-F de planta.

3. Construção genética, de acordo com a reivindicação 1, **CARACTERIZADA** pelo fato de que a seqüência de promotor genético é um promotor específico de semente.

15 4. Construção genética, de acordo com a reivindicação 3, **CARACTERIZADA** pelo fato de que a seqüência de promotor genético é a faseolina.

20 5. Construção genética, de acordo com a reivindicação 1, **CARACTERIZADA** pelo fato de compreender um vetor de expressão pBBV-PHAS.

25 6. Construção genética, de acordo com a reivindicação 1, **CARACTERIZADA** pelo fato de que a proteína de domínio-F compreende uma seqüência de polipeptídeo, pelo menos, aproximadamente 75% idêntica à seqüência de polipeptídeo com ID de SEQ N°: 1.

7. Construção genética, de acordo com a reivindicação 1, **CARACTERIZADA** pelo fato de que a proteína de domínio-F compreende a seqüência de polipeptídeo , pelo

menos, aproximadamente, 90% idêntica à sequência de polipeptídeo com ID de SEQ Nº: 1.

8. Construção genética, de acordo com a reivindicação 1, **CARACTERIZADA** pelo fato de que a proteína de domínio-F de planta compreende a sequência de polipeptídeo com ID de SEQ Nº: 1.

9. Construção genética, de acordo com a reivindicação 1, **CARACTERIZADA** pelo fato de que a proteína de domínio-F é codificada por um ácido nucléico que hibridiza com a sequência de ácido nucléico com ID de SEQ Nº: 2.

10. Construção genética, de acordo com a reivindicação 1, **CARACTERIZADA** pelo fato de que a proteína de domínio-F é codificada por um ácido nucléico que hibridiza mediante rígidas condições com a sequência de ácido nucléico com ID de SEQ Nº: 2.

11. Construção genética, de acordo com a reivindicação 1, **CARACTERIZADA** pelo fato de que a proteína de domínio-F de planta é codificada pela sequência de ácido nucléico com ID de SEQ Nº: 2.

12. Célula de planta transgênica, **CARACTERIZADA** pelo fato de que compreende uma construção genética de acordo com a reivindicação 1.

13. Célula de planta transgênica, de acordo com a reivindicação 12, **CARACTERIZADA** pelo fato de que a planta é a *Arabidopsis thaliana*.

14. Célula de planta transgênica, de acordo com a reivindicação 12, **CARACTERIZADA** pelo fato de que a planta é

selecionada do grupo que consiste em feijão-soja, palmeira, semente de colza e girassol.

15. Célula da planta *Arabiposis thaliana* recombinante, **CHARACTERIZADA** pelo fato de que compreender uma
5 construção genética heteróloga tendo uma funcionalidade da sequência de promotor genético ligada a uma sequência de codificação de domínio-F de planta.

16. Célula da planta *Arabiposis thaliana* recombinante, de acordo com a reivindicação 13,
10 **CHARACTERIZADA** pelo fato de que a sequência de domínio-F de planta codifica o polipeptídeo com ID de SEQ N^o: 1 e o promotor é a faseolina.

17. Planta, **CHARACTERIZADA** pelo fato de compreender uma célula de planta transgênica, de acordo com a
15 reivindicação 12, ou uma parte propágula ou progênica dela, sendo que a parte propágula ou progênica compreende a construção genética.

18. Método de modificação da produção de óleo em uma planta, **CHARACTERIZADO** pelo fato de que compreende:
20 incorporar, de modo estável, ao genoma da planta, uma construção genética que está de acordo com a reivindicação 1, com fins de fornecer uma planta transformada; e

regenerar a planta transformada,
25 sendo que a expressão da construção genética incorporada modifica a produção de óleo na planta.

19. Método, de acordo com a reivindicação 18, **CARACTERIZADA** pelo fato de que a planta é a *Arabidopsis thaliana*.

20. Método, de acordo com a reivindicação 18, **CARACTERIZADA** pelo fato de que a planta é selecionada do grupo que consiste em feijão-soja, palmeira, semente de colza e girassol

21. Método de aumento na produção de óleo na *Arabidopsis thaliana*, **CARACTERIZADO** pelo fato de
10 compreender:

incorporar, de modo estável, ao genoma da planta *Arabidopsis thaliana*, uma construção genética tendo uma funcionalidade da seqüência de promotor genético da faseolina que codifica a ID DE SEQ Nº: 1; e

15 regenerar a planta transformada,
sendo que a expressão da construção genética incorporada aumenta a produção de óleo na planta *Arabidopsis thaliana*.

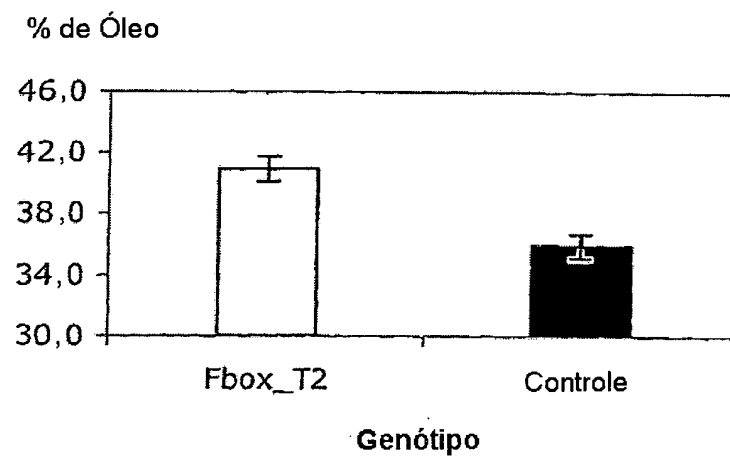


FIG. 1

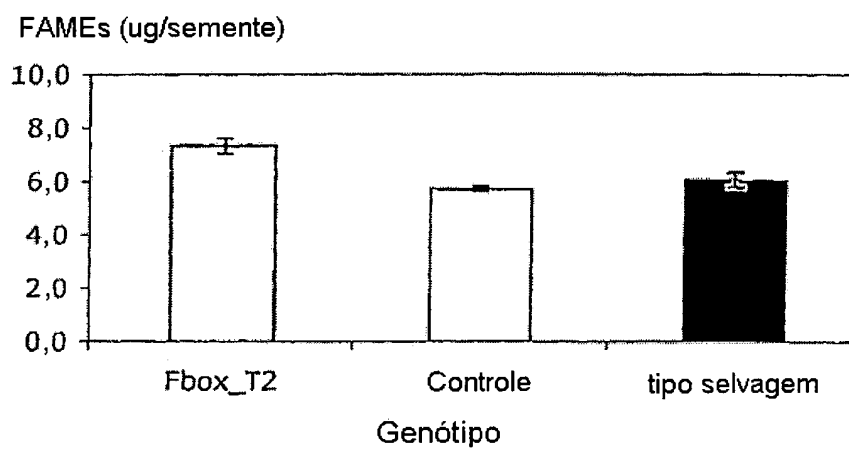


FIG. 2

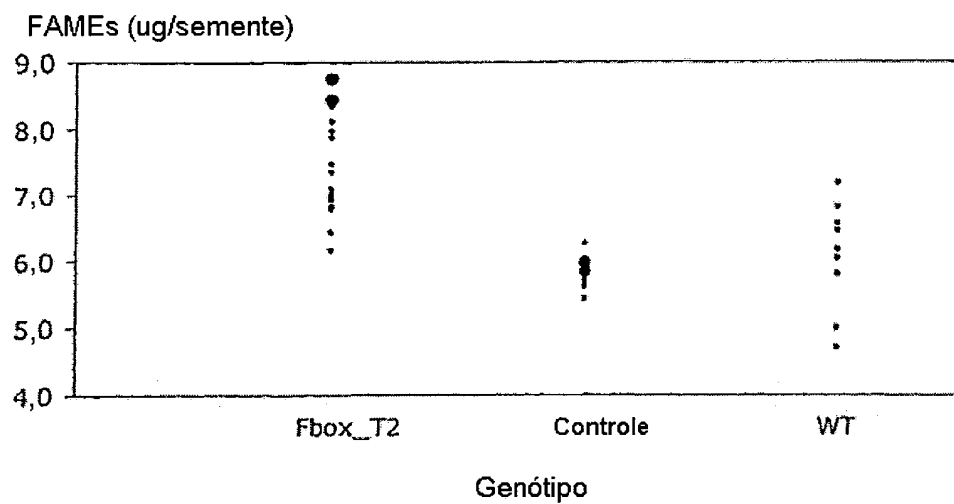


FIG. 3

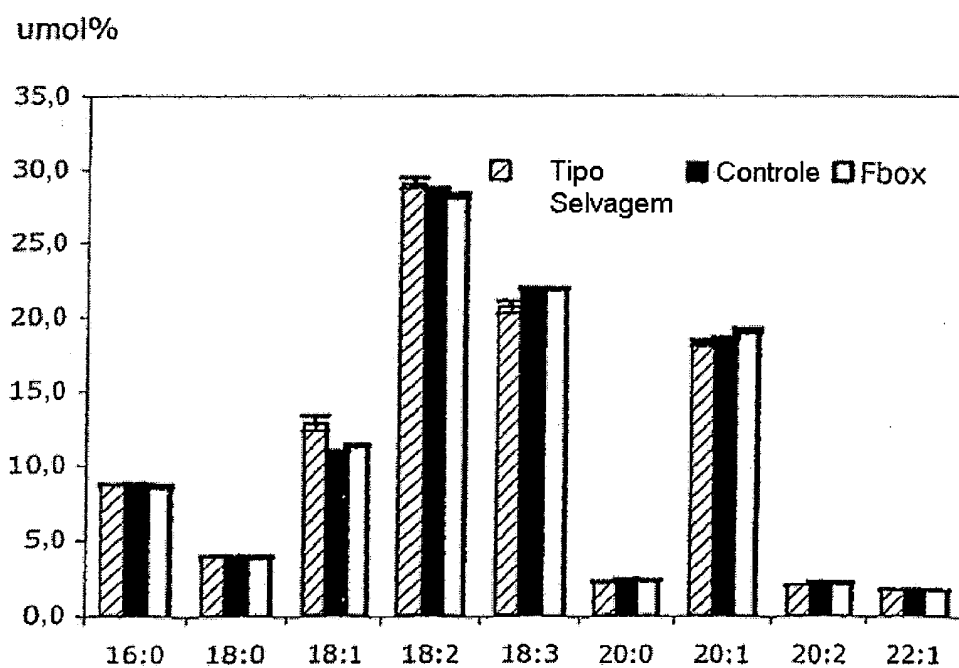


FIG. 4

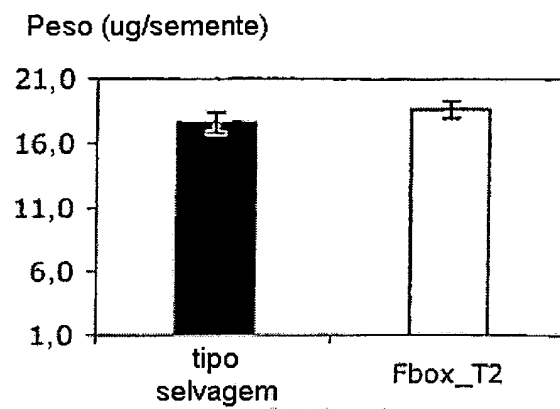


FIG. 5

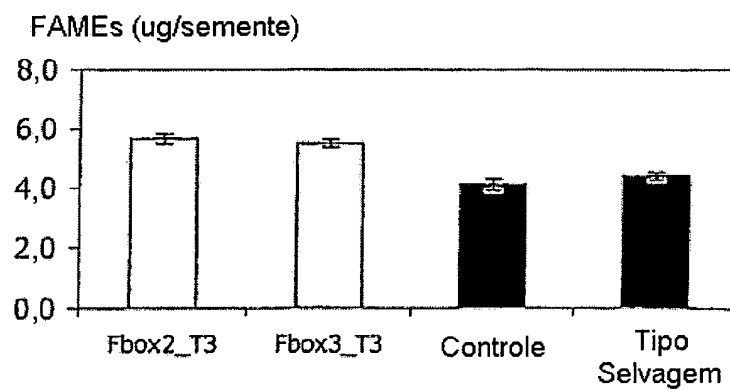


FIG. 6

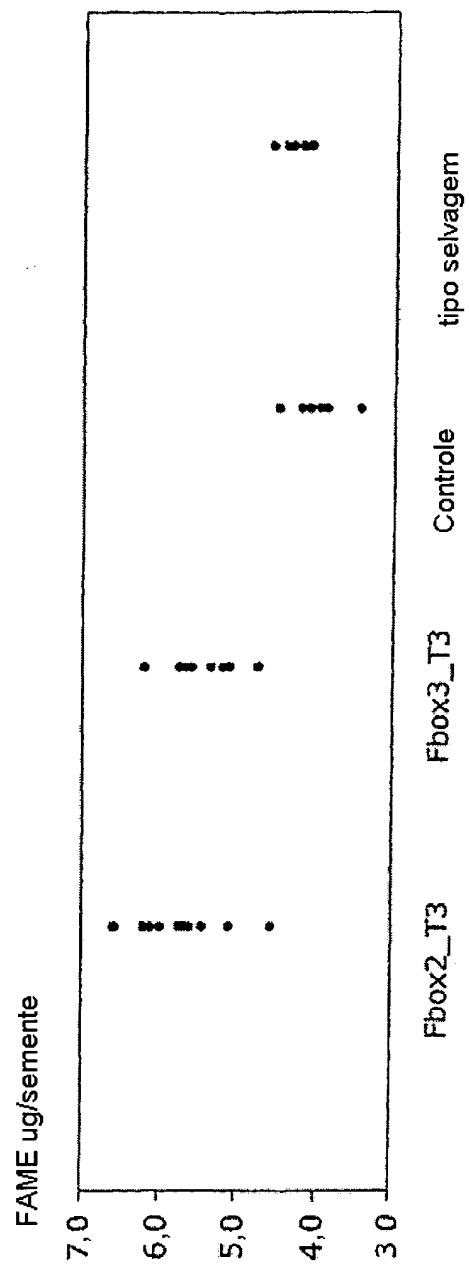


FIG. 7

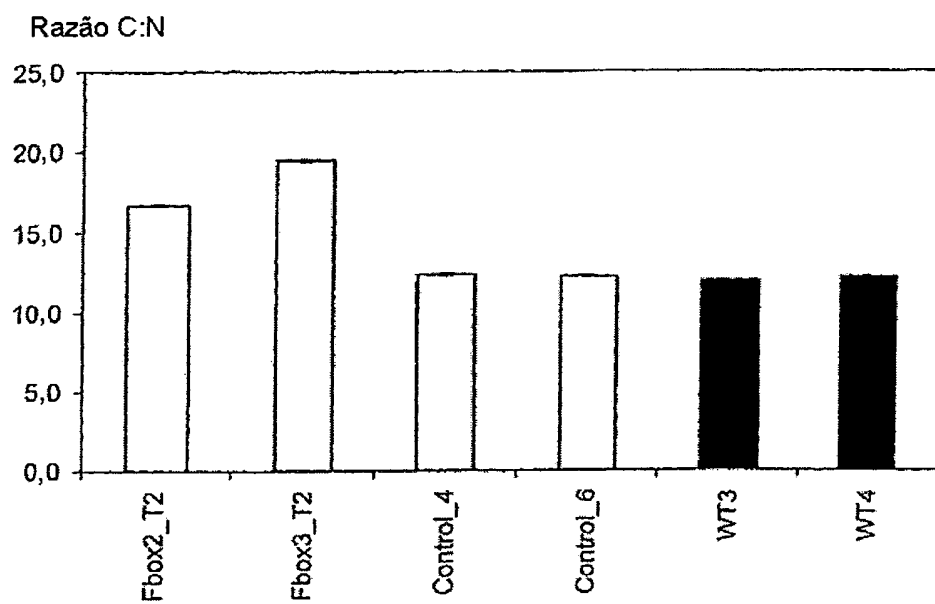


FIG. 8A

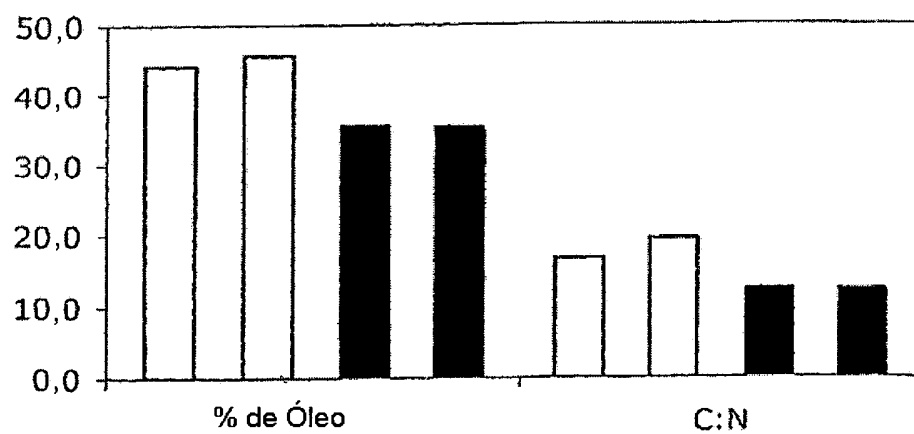


FIG. 8B

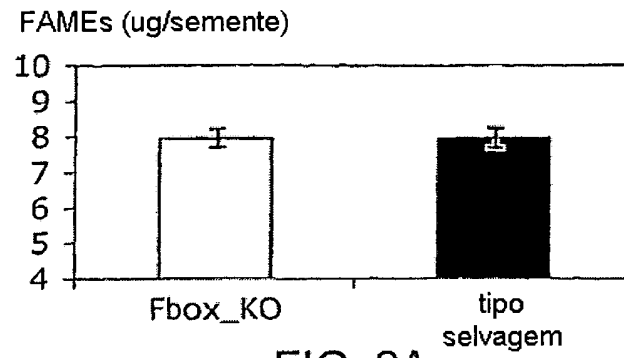


FIG. 9A

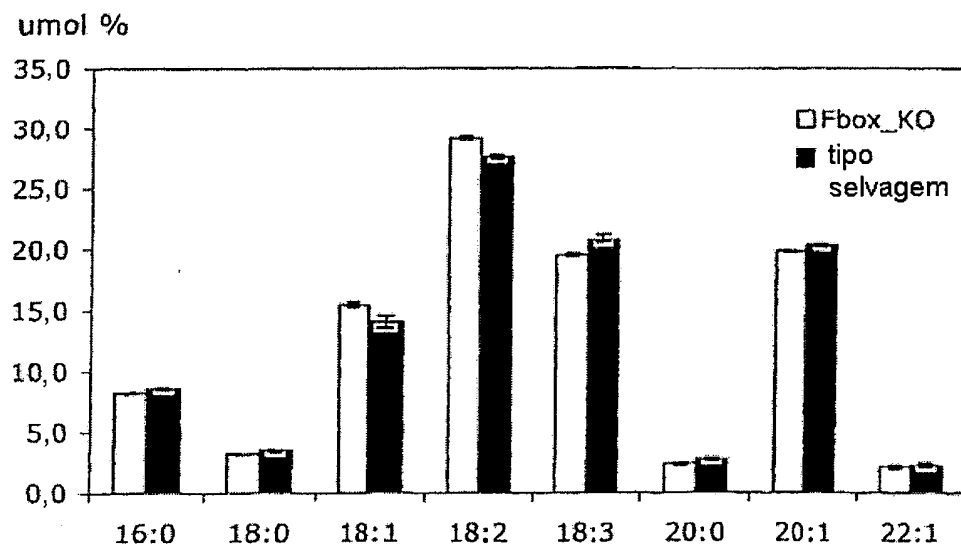


FIG. 9B

**Sequência da Proteína F-box de Arabidopsis thaliana
(No. de Acesso do Genbank NP_566277 (SEQ ID NO. 1))**

1 MKAIQILLWEA IMEATKRERR REDDDGEKAS PESLVLPPEI ITEILLRLPA KSIGRFRCVS
61 KLECTLSSDP GFAKIHLDLI LRNESVRSIH RKLIVSSHNL YSLDFNSIGD GIRDLAAVEH
121 NYPLKDDPSI FSEMIRNYVG DHLyddRRVM LKLNaksYRR NWVEIVGSSN GLVCISPGEG
181 AVFLYNPTTG DSKRLPENFR PKSVEYERDN FQTYGFGFDG LTDDYKLVKL VATSEDILDA
241 SVYSLKADSW RRICNLNVEH NDGSYTSGVH FNGAIHWVFT ESRHNQRVVV AFDIQTEEFR
301 EMPVPDEAED CSHRFSNFVV GSLNGRLCVV NSCYDVHDDI WVMSEYGEAK SWSRIRINLL
361 YRSMKPLCST KNDEEVLLLEL DGDVLVYNFE TNASSNLGIC GVKLSDGFEA NTYVESLISP
421 NSYGIES

FIG. 10A

FIG. 10B-1
FIG. 10B-2
FIG. 10B-3

Seqüência da Proteína F-box de Arabidopsis thaliana
(No. de Acesso do Genbank NP_111499 (SEQ ID NO. 2))

FIG. 10B

```

1  TTTTTCAAAT CAAATCAGAA TACATTGATT CTGTATAICT TATGAAAAA TCCATCAATT
61 TACATCAACA ATTTTATATC TAATAATTAA TTAAAGAGA AAATTTATAA AAGTTTATTA
121 GAGCAAATAA CTCAAACTCG GATTTTATAG TCGTTATGAC CCGGTTTGAC TATTGAACCG
181 TTAAACCGAG AAATTGGGAC TCAATTAGA CAACCGAAC TAGACCCGGA TCCAGTGTTA
241 GCGGGCTAGA TTAAGGTGTC GGGTCATAGC GGAGAAGCAA CCAGACGCCA ACAATGAAAG
301 CGATCCAGTT GCTGTGGGAA GCGATAATGG AGCGACGAA GAGAGAAAGA CGGAGAGAAG
361 ATGACGACGG CGAAAAAGCT TCACCGGAAT CACTCGTTCT TCCACCAGAG ATCATTACAG
421 AAATTCTTCT CCGATTACCA GCCAAATCGA TCGGGCGGAT CAGGTCCGTA TCAAAGCTCT

```

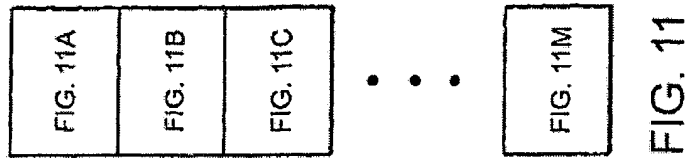
FIG. 10B-1

481 TTTGCACTTT ATCGTCAGAT CCAGGGTTCG CCAAGATTCA CCTCGATCTG ATCCTTCGAA
541 ACGAATCCGT AAGATCGCTC CACCGTAAGC TCATTGTGTC TTCACATAAT CTGTACTCGT
601 TAGATTTCAA TTCGATCCGT GACGGAATTA GGGATTAGC GGCTGTGGAA CACAATTATC
661 CTCTTAAAGA CGATCCAAGC ATTTTCTCTG AGATGATTAG GAATTACGTG GGGGACCATC
721 TGTACGATGA TCGTCGCGTG ATGCTTAAGC TGAATGCGAA ATCGTATCGA AGAAACTGGG
781 TTGAGATCGT TGGATCTTCC AATGGTTTAC TGTGTATCTC TCCTGGTGAA GGAGCTGTTT
841 TCTTGTATAA TCCAACTACC GGAGATTCCA AGAGATTACC TGAATAATTT CGTCCCAAAT
901 CTGTAGAATA CGAAAGAGAT AATTCCAAA CTTATGGATT TGGTTTCGAT GGTCTCACTG
961 AAGATTACAA ATTGGTGAAG CTTGTTGCTA CCAGTGAAGA TATCTTCGAT GCTAGTGCTT
1021 ATTCCTTGAA GGCTGACTCA TGGAGACCGA TCTGCAATTT GAATTATGAG CACAACGATG
1081 GCTCCTACAC GTCCGGGTGTG CATTTCACCG GTCCGATTCA CTGGGTGTTT ACAGAGAGTA

FIG. 10B-2

1141 GGCACAACCA AAGAGTGGTT GTAGCATTG ATATTCAAAC CGAGGAGTTT CGAGAGATGC
1201 CAGTGCCTGA TGAAGCTGAA GATTGTTCCC ATAGGTTTAG CAACTTTGTG GTCGGAAGTC
1261 TCAATGGGACG TCTCTGTGTG GTC AATAGTT GCTACGATGT GCATGATGAT ATATGGGTGA
1321 TGAGTGAGTA CGGTGAAGCT AAATCCTGGA GCAGAATTTC AATCAACTTG TTGTATAGGT
1381 CGATGAAACC GCTCTGTTCG ACTAAGAACG ATGAAGAGGT TCTTCTGGAG CTTGATGGAG
1441 ACCTGGTGT GTACAACCTT GAAACCAATG CATCGAGTAA TCTAGGAATT TGTGGGGTTA
1501 AGCTCAGTGA CGGGTTCGAG GCAAATACAT ACGTAGAGAG CCTCATATCA CCCAACTCTT
1561 ATGGTATAGA GAGCTGAGGA AGTCTGCTTT TTGCTAAGAT ATAATAAAC AACATTCGGA
1621 TTAGAAATGT TTTAGAAACA TAATCATGTA ATATGTATCA TGTAATTAAAC AACGAATGGT
1681 CAATGGGTAT TTTAAGTTTC TTCTCCT

FIG. 10B-3



(SEQ ID NO:6)

```

1  ggccgcaaca gaggtggatg gacagaccgg ttcttacacc ggactgggcg
51  cgggatagga tattcagatt gggatgggat tgagcttaaa gccggcgctg
101 agaccatgct caaggtaggc aatgtcctca gcgtcgagcc cggcatctat
151 gtcgagggca ttggtggagc gcgcttcggg gataccgtgc ttgtaactga
201 gaccggatat gaggcctca ctccgcttga tcttggcaaa gatatttgac
251 gcatttatta gtatgtgtta atttcattt gcagtgcagt attttctatt
301 cgatctttat gtaattcgtt acaattaata aatattcaaa tcagatttat
351 gactgtcatt tgtatcaaat cgtgtttaat ggataatttt attataatat
401 tgatgatata tcaatcaaaa cgtagataat aataatatatt atttaatat
451 ttgcgctgc acagtgaaaa tctatatgag attacaaaa accgacaaca
501 ttatttaaga tacatagaca ttaaccctga gactgttga cagagctcat
551 tggtaacctca gatctgggtg actggcctaa ctggccttgg aggagctggc
601 aactcaaaat ccttttgcca aaaaccaaca tcatgccatc caccatgctt
651 gtatccagct gcgcgcaatg taccocgggc tgtgtatccc aaagcctcat
701 gcaaccctaac agatggatcg tttggaaggc ctataacagc aaccacagac
751 ttaaaacctt gcgcctccat agacttaagc aaatgtgtgt acaatgtgga
801 tcctaggccc aacctttgat gcctatgtga cagtaaaaca gtactctcaa
851 ctgtccaatc gtaagcgttc ctagccttcc agggccagc gtaagccaata
901 ccagccacaa caccctcaac ctacgcaaac aaccaagggt atctatcttg

```

FIG. 11A

```

951 caacctctct agatcatcaa tccactcttg tgggtgttgt ggctctgtcc
1001 taaagttcac tgtagacgtc tcaatgtaat ggttaacgat atcacaaacc
1051 gcggccatat cagctgctgt agctggccta atctcaactg gtctcctctc
1101 cggagaagcc atggtttgga tccacaaact tacaaatttc tctgaagtgt
1151 taccctcagt acctcaaaga aaatagctta caccaatttt tcttggtttt
1201 cacaaatgcc gaacttggtt cttatatag gaaactcaa gggcaaaaat
1251 gacacggaaa aatataaaa gataagtag gggggataag attcctttgt
1301 gataagggtta cttccgccc ttacattttc caccttacct gtgtcctcta
1351 tgtctctttc acaatcaccc accttatctt cttcttttca ttgttgtcgt
1401 cagtgccttac gtcttcaaga ttcttttctt cgcctgggtc ttctttttca
1451 atttctacgt attcttcttc gtattctggc agtataggat cttgtatctg
1501 tacatttctc atttttgaac ataggttgca tatgtgccgc atattgatct
1551 gcttcttgct gagctcacat aatacttcca tagtttttcc cgtaaacatt
1601 ggattcttga tgctacatct tggataaatta cttcttgccc ggccgcgaat
1651 tcgttggttag ggtgctagga aacttgtttt tgggggtttg tataagggtt
1701 gaaacatccc tgaagtgctc cattttattt tatttattct ttgctgataa
1751 aaaaaataaaa taaaagaagc taagcacacg gtcaaccatt gctctactgc
1801 taaaagggtt atgtgtagtg ttttactgca taaattatgc agcaacaag

```

FIG. 11B

```

1851 acaactcaaa ttaaaaaatt tcctttgctt gtttttttgt tgtctctgac
1901 ttgactttct tgtggaagtt gtttgtataa ggattgggac accattgtcc
1951 ttcttaattt aattttattc ttgctgata aaaaaaaa attcatata
2001 gtgttaaata ataatttgtt aaataaccaa aaagtcaaat atgtttactc
2051 tcgtttaaat aattgagatt cgttccagca aggcbaaacg attgtataga
2101 ttatgacaa tatttacttt ttatagata aatgttatat tataataaat
2151 ttatatacat atatttatatg ttatttatta ttatttatta tttaaatcc
2201 ttcaatatat tatcaaaacca actcataatt ttttttttat ctgtaagaag
2251 caataaaatt aaatagaccc actttaagga tgatccaacc ttatacaga
2301 gtaagagagt tcaaatagta ccttttcata tacatatcaa ctaaaatatt
2351 agaaatatca tggatcaaac cttataaaga cattaaataa gtggataagt
2401 ataatatata aatgggtagt atataatata taaatggata caaactttctc
2451 tctttataat tgttatgtct ccttaacatc ctaatatata acataagtgg
2501 gtaatatata atataaaat ggagacaaac ttcttccatt ataattgtta
2551 tgtcttctta acacttatgt ctcgttcaca atgctaaagt tagaattgtt

```

FIG. 11C

```

2601 tagaaagtct tatagtacac atttggtttt gtactatttg aagcattcca
2651 taagccgtca cgattcagat gatttataat aataagagga aatttatcat
2701 agaacaataa ggtgcataga tagagtgtta atatatcata acatcctttg
2751 tttattcata gaagaagtga gatggagctc agttattata ctgttacatg
2801 gtcggataca atattccatg ctctccatga gctcttacac ctacatgcat
2851 tttagttcat actggtgacc ctcgaggcga tcgctttggc cggccattta
2901 aatggcgcg ctttgcccg gctttcctgc agggtttATC AACCACTTTG
2951 TACAAGAAAG CTGAACGAGA AACGTAAAT GATATAAATA TCAATATATT
3001 AAATTAGATT TTGCATAAAA AACAGACTAC ATAATACTGT AAAACACAAC
3051 ATATCCAGTC ACTATGGTCG ACCTGCAGAC TGGCTGTGTA TAAGGGAGCC
3101 TGACATTTAT ATTCCCCAGA ACATCAGGTT AATGGCGTTT TTGATGTCTAT
3151 TTTCGCGGTG GCTGAGATCA GCCACTTCTT CCCCAGATAAC GGAGACCGGC
3201 ACAC TGGCCA TATCGGTGGT CATCATGCCG CAGCTTTCAT CCCCAGATATG
3251 CACCACCGGG TAAAGTTTAC GGGAGACTTT ATCTGACAGC AGACGTGCAC
3301 TGGCCAGGGG GATCACCATC CGTCGCCCCG GCGTGTCAAT AATATCACTC
3351 TGTACATCCA CAAACAGAGG ATAACGGCTC TCTCTTTTAT AGGTGTAAC
3401 CTTAAACTGC ATTTCAACCAG TCCCTGTTCT CGTCAGCAAA AGAGCCGTTT
3451 ATTTCAATAA ACCGGCGGAC CTCAGCCATC CCTTCCTGAT TTTCCGCTTT
3501 CCAGCGTTTC GCACGCAGAC GACGGGCTTC ATTCTGCATG GTTGTGCTTA

```

FIG. 11D

3551 CCAGACCGGA GATATTGACA TCATATATGC CTTGAGCAAC TGATAGCTGT
 3601 CGCTGTCAAC TGTCACTGTA ATACGCTGCT TCATAGCACA CCTCTTTTG
 3651 ACATACTTCG GGTATACATA TCAGTATATA TTCTTATACC GCAAAAATCA
 3701 GCGCGCAAAAT ACGCATACTG TTATCTGGCT TTAGTAAAG CCGATCCAGA
 3751 TCTTPACGCC CCGCCCTGCC ACTCATCGCA GTACTGTTGT AATTCATTAA
 3801 GCATTCTGCC GACATGGAAG CCATCACAGA CGGCATGATG AACCTGAATC
 3851 GCCAGCGGCA TCAGCACCTT GTCGCCCTGC GTATAATATT TGCCCATGGT
 3901 GAAAACGGGG GCGAAGAAGT TGTCCATATT GCCACGTTT AAATCAAAAC
 3951 TGGTGAAACT CACCCAGGGA TTGGCTGAGA CGAAAAACAT ATTCTCAATA
 4001 AACCCTTTAG GGAATAGGC CAGGTTTCA CCGTAACACG CCACATCTTG
 4051 CGAATATATG TGTAGAAACT GCCGGAATC GTCGTGGTAT TCACTCCAGA
 4101 GCGATGAAA ACGTTTCAGTT TGCTCATGGA AACCGGTGTA ACAAGGTGA
 4151 ACACATATCC ATATCACCCAG CTCACCCGCTT TTCAATGCCA TACGGAATTC
 4201 CGGATGAGCA TTCATCAGGC GGGCAAGAAT GTGAATAAAG GCCGGATAAA
 4251 ACTTGTGCTT ATTTTCTTTT ACGGTCCTTA AAAAGGCCGT AATATCCAGC
 4301 TGAACGGTCT GGTATAGGT ACATGAGCA ACTGACTGAA ATGCCCTCAA
 4351 ATGTTCTTTA CGATGCCATT GGGATATATC AACGGTGGTA TATCCAGTGA
 4401 TTTTTCCTTC CATTTTAGCT TCCTTAGCTC CTGAAAAATCT CGCCGGATCC
 4451 TAACTCAAAA TCCACACATT ATACGAGCCG GAAGCATAAA GTGTAAAGCC
 4501 TGGGGTGCCT AATGCGGCCG CCATAGTGAC TGGATATGTT GTGTTTACAA
 4551 GTATTATGTA GTCTGTTTAT TATGCAAAAT CTAATTTAAT ATATTGATAT
 4601 TTATATCAAT TTACGTTTCT CGTTCAGCTT TTTTGFACAA ACTTGTGTAT

FIG. 11E

```

4651 aaactttaatt aaggatccta gagtagtatt gaatatgagt tgggttgggg
4701 tattatagta gtagagtagt agtactctgg atggatggat gatgaaagaa
4751 gtgagtgata ttagaggtat ttataggtat tataataagag agaaggtggt
4801 tggaaacatgc atggagatctt gggcatggga tgacacgcat atgcaggttg
4851 acgtgtgttg aagtgaagaa attgaggtgg cggaagagaa tgaatatata
4901 caggtggttg tggatgatgac gaagaaaaag gcaatgtgtt tgtgtgtggg
4951 ttgagatggg tgagccattt aaagtgcattg ttaagcacgt gttgctttgc
5001 atggcattta gacatacatg gacgcggcga tcttgatcag ccagtgacta
5051 atttgagttg gttgtgtgat tgcgttttgt ctctctgttt tgtctttttt
5101 ctttgttctt tgtcttttct ttgcgcaagc atccatgcat gaaccaaaag
5151 accacagagt gtcattggcaa cccacagtaa ttccagttac ggacttacct
5201 accaagaaaa ggtaaaagca ataagaaata tatgaaaaat agtccacct
5251 aaatcttata gtttatggtt tagggtaaca ctctaacct ctactacct
5301 acttataatt actttaact atctataaaa caatttaaac atagtagaat
5351 aataaatcaa tagtcacaaa ttcaacaatt aaacttaaaa ttaaaaaagt
5401 aatattttta ttatatctaa ttaatttttt agaagtaata ttgagtattt

```

FIG. 11F

```

5451 gatatatgaa atcttgaata tagtaactat tattaaaatt actttagaag
5501 atgtgtctcg catgtaaaag cagatcttca gttacttccg tagtgtcaaa
5551 tgggaattat agaattttgc ataacatgac ttgcttcaga aataccacaa
5601 atcattgttt ggtgaaattt tcattgtata aaaaaaatat aatgataatt
5651 ggattttttt attcaaagaa aaaaatggct agttgtgtca ctgggtgttg
5701 ctttaagatc agtcgaataa aaaaatttat gagttaaaatt ttattacttt
5751 tgaaacaact tattattatg agattttacgt ggttgaaaaa tatttgataa
5801 atatatttt aaaaataaaa atgggaaatc ctctttaagg taaagaattt
5851 gtttatattg tataattaaac atttatatga agaagaataa gaataaatca
5901 ttatgctttc taccaacgct aaaattaagt aaattatata ttcaatatg
5951 aaaatgttag actacattaa agatagacgg gacttcataa aattttatgc
6001 ggtttgaaaa tgtttaacaa ataataaatt tgtagggata tcgtgtgcgg
6051 aagcgtgata atttcaacca aagatttatga gaaattaaag taacaagtaa
6101 agtgagaatg ataccagaat ttttaggtgg aaaaacctt taaatagagg
6151 taaaaaacca ccggcgagag agccaaaact ttcactataa tgatactggg
6201 agtacaatgg cggccggggc tgcaattgat ccggtgagta atattgtacg
6251 gctaagagcg aatttggcct gtagacctca atbgcgagct ttctaatttc
6301 aaactattcg ggcctaactt ttggtgtgat gatgctgact ggcaggatat
6351 ataccgttgt aatttgagct cgtgtgaata agtcgctgtg tatgtttgtt
6401 tgattgtttc tgttgagtg cagcccat t caccggacaa gtccggctaga

```

FIG. 11G

```

6451 ttgatttagc cctgatgaac tgcgagggg aagccatctt gagcgcgaa
6501 tgggaatgga tcgaaccggg agcacaggat gacgcctaac aattcattca
6551 agccgacacc gcttcgcggc gcggcttaat tcaggagtta aacatcatga
6601 gggaagcggg gatcgccgaa gtatcgactc aactatcaga ggtagttggc
6651 gtcatcgagc gccatctcga accgacgttg ctggccgtac atttgtacgg
6701 ctccgcagtg gatggcgcc tgaagccaca cagtatatatt gatttgctgg
6751 ttacggtgac cgtaaggcct gatgaaacaa cgcggcgagc ttgatcaac
6801 gaccttttgg aaacttcggc ttcccctgga gagagcgaga ttctccgcgc
6851 tgtagaagtc accattgttg tgcacgacga catcatccg tggcgttatac
6901 cagctaagcg cgaactgcaa tttaggagaat ggcagcgcaa tgacattctt
6951 gcaggatatc tcgagccagc cacgatcgac attgatctgg ctatcttgct
7001 gacaaaagca agagaacata gcgttgcctt ggtaggtcca gcggcgagg
7051 aactctttga tccggttcct gaacaggatc tatttgaggc gctaaatgaa
7101 accttaacgc tatggaaact gccgcccgac tggctggcg atgagcgaaa
7151 tgtagtgcctt acgttgtccc gcatttggtta cagcgcagta accggcaaaa
7201 tcgcgccgaa ggatgtcgct gccgactggg caatggagcg cctgccggcc
7251 cagtatcagc ccgtcatact tgaagctagg caggcttatac ttggacaaga
7301 agatcgcttg gcctcgcgcg cagatcagtt ggaagaattt gttcactacg

```

FIG. 11H

```

7351 tgaaaggcga gatcaccaag gtagtcggca aataatgtct aacaattcgt
7401 tcaagccgac gccgcttcgc gccgaggctt aactcaagcg ttagagagct
7451 ggggaagact atgcgcgac tgttgaaggt ggttctaagc ctctactctg
7501 cgatggcatt tcgatacga ggggtacaaa ttccactaa gcgctcgggg
7551 gctgagaaag ccagtaagg aaacaactgt aggttcgagt cgcgagatcc
7601 ccggaacca aaggaagtag gttaaacccg ctccgatcag gccgagccac
7651 gccaggccga gaacatttgt tcctgtaggc atcgggattg gcggatcaaa
7701 cactaaagct actggaacga gcagaaagtc tccggccgcc agttgccagg
7751 ccgtaaaagg gagcagaggc acgggagggt gccacttgcg ggtcagcacg
7801 gttccgaacg ccatggaaac cgcgcccgcc agggccgctg cgacgccgac
7851 aggatctagc gctgcgtttg gtgtcaacac caacagcgcc acgcccgcag
7901 ttccgcaaat agccccagg accgccatca atcgtatcgg gctacctagc
7951 agagcggcag agatgaacac gaccatcagc ggctgcacag cgcctaccgt
8001 cgccgcgacc cgcccggcag gcggtagacc gaaataaaca acaagctcca
8051 gaatagcgaa atattaagtg gcccgaggat gaagatgcgc atccaccaga
8101 ttcccgttgg aatctgtcgg acgatcatca cgagcaataa acccgccggc
8151 aacgccccga gcagcatacc gccgacccct cggcctcgct gttcgggctc
8201 cacgaaaaacg ccggacagat gcgccttgtg agcgtccttg gggccgtcct
8251 cctgtttgaa gaccgacagc ccaatgatct cgcgctcgat gtaggcgccg

```

FIG. 11I

```

8301 aatgccacgg catctcgcaa ccgttcagcg aacgcctcca tgggcttttt
8351 ctctcgtgc tcgtaaacgg acccgaacat ctctggagct ttcttcaggg
8401 ccgacaatcg gatctcgcg aaatcctgca cgtcggcgcg tccaagccgt
8451 cgaatctgag ccttaatac aattgtcaat ttaatacctc tgtttatcgg
8501 cagttcgtag agcgcgccgt gcgcccgagc gatactgagc gaagcaagtg
8551 cgtcgagcag tgcccgcttg ttctgaaat gccagtaaa gctgggctgc
8601 tgaaccccca gccgggaactg accccacaag gccctagcgt ttgcaatgca
8651 ccaggtcatc attgacccag gcgtgttcca ccaggccgct gcctcgcaac
8701 tcttcgcagg cttcgccgac ctgctcgcg cacttcttca cgcgggtgga
8751 atccgatccg cacatgaggc ggaaggttcc cagcttgagc gggtaaggct
8801 cccggtgcga gctgaaatag tcgaacatcc gtccggccgt cggcgacacg
8851 ttgcgggtact tctccatat gaatttcgtg tagtggtcgc cagcaaacag
8901 cacgacgatt tcctcgtcga tcaggacctg gcaacgggac gttttcttgc
8951 cacggtccag gacgcggaag cggtcgagca gcgacaccga ttccaggtgc
9001 ccaacgcggg cggacgtgaa gccatcgcc gtccctgta ggcgcgacag
9051 gcattcctcg gccttcgtgt aataccggcc attgatcgac cagcccaggt
9101 cctggcacaag ctcgtagaac gtgaagggtga tcggctcgcc gataggggtg
9151 cgcttcgcgt actccaacac ctgctgccac accagttcgt catcgtcggc
9201 ccgcagctcg acgcccgtgt aggtgatctt cacgtccttg ttgacgtgga

```

FIG. 11J

```

9251 aaatgacctt gttttgcagc gcctcgcgcg ggattttctt gttgcgcgtg
9301 gtgaacaggg cagagcgggc cgtgctgttt ggcacgctc gcacgtgtgc
9351 cggccacggc gcaatatcga acaaggaaag ctgcatttcc ttgatctgct
9401 gcttcgtgtg tttcagcaac gcggcctgct tggcctcgct gacctgtttt
9451 gccaggtcct cgcggcggtt ttttcgcttc ttggtcgta tagttcctcg
9501 cgtgtcgatg gtcatcgact tcgccaaacc tgcgcctcc tgttcgagac
9551 gacgcgaacg ctccacggcg gccgatggcg cgggcagggc agggggagcc
9601 agttgcacgc tgtcgcgctc gatcttggcc gtagcttgc ttgacctga
9651 gccgacggac tggaaagttt cgcggggcgc acgcatgacg gtgcggccttg
9701 cgatggtttc ggcatacctc gcgggaaacc ccgcgtcgat cagttcttgc
9751 ctgtatgcct tccggtcaaa cgtccgattc attcacctc cttgcgggat
9801 tgcgccgact cagcgcgggg caatgtgccc ttattcctga ttbgaccgc
9851 ctggtgcctt ggtgtccaga taatccacct tatcggaat gaagtcggtc
9901 ccgtagaccg tctggccgtc cttctcgtae ttggtattcc gaatccttgc
9951 ctgcacgaat accagcgacc ccttgcccaa atacttgccg tgggacctcg
10001 cctgagagcc aaaacacttg atgcggaaga agtcgggtcg ctccctgcttg
10051 tcgccggcat cgttgcgcca ctcttcatta accgctatat cgaaaaattgc
10101 ttgcgggcttg ttagaattgc catgacgtac ctcggtgtca cgggtaagat

```

FIG. 11K

```

10151 taccgataaa ctggaactga ttatggcnnc tgaattcc ctcggctctg
10201 ccttgctcgt cggatgatga cttcaccagc tccgcgaagt cgctcttctt
10251 gatggagcgc atggggacgt gcttggcaat cacgcgcacc cccgggccgt
10301 tttagcggct aaaaaagtca tggctctgcc ctcgggcgga ccacgcccac
10351 catgaccttg ccaagctcgt cctgctcttc ttcgatcttc gccagcaggg
10401 cgaggatcgt ggcataccg aaccgcgcg tgcgcgggtc gtcggtgagc
10451 cagagtttca gcaggccgcc caggcgcccc aggtcgccat tgatgcgggc
10501 cagctcgcgg acgtgctcat agtccacgac gcccgtagt ttgtagccct
10551 ggcgcacggc cagcaggtag gccgacaggc tcatgccggc cgccggccgc
10601 ttttctctaa tcgctctctg ttctgtgga aggcagtaca ccttgatagg
10651 tgggctgccc ttctgtggtg gcttggtttc atcagccatc cgcttgcct
10701 catctgttac gccggcggtg gccggccagc ctcgcagagc aggatccccg
10751 ttgagcaccg ccaggtgcga ataagggaca gtgaagaagg aacacccgct
10801 cgcggttggg cctacttcac ctatcctgcc cggctgacgc cgttggatac
10851 accaaggaaa gtctacacga accctttggc aaaaatcctgt atatcgtgcg
10901 aaaaaggatg gatataccga aaaaatcgt ataatgaccc cgaagcaggg
10951 ttatgcagcg gaaaagatcc gtcgaccctt tccgacgctc accgggctgg
11001 ttgcccctgc cgctgggctg gcggccgtct atggccctgc aaacgcgcca
11051 gaaacgccgt cgaagccgtg tgcgagacac cgcgccgcc gcgcttgtgg
11101 ataccacgcg gaaaacttgg ccctcactga cagatgaggg gcggacgttg

```

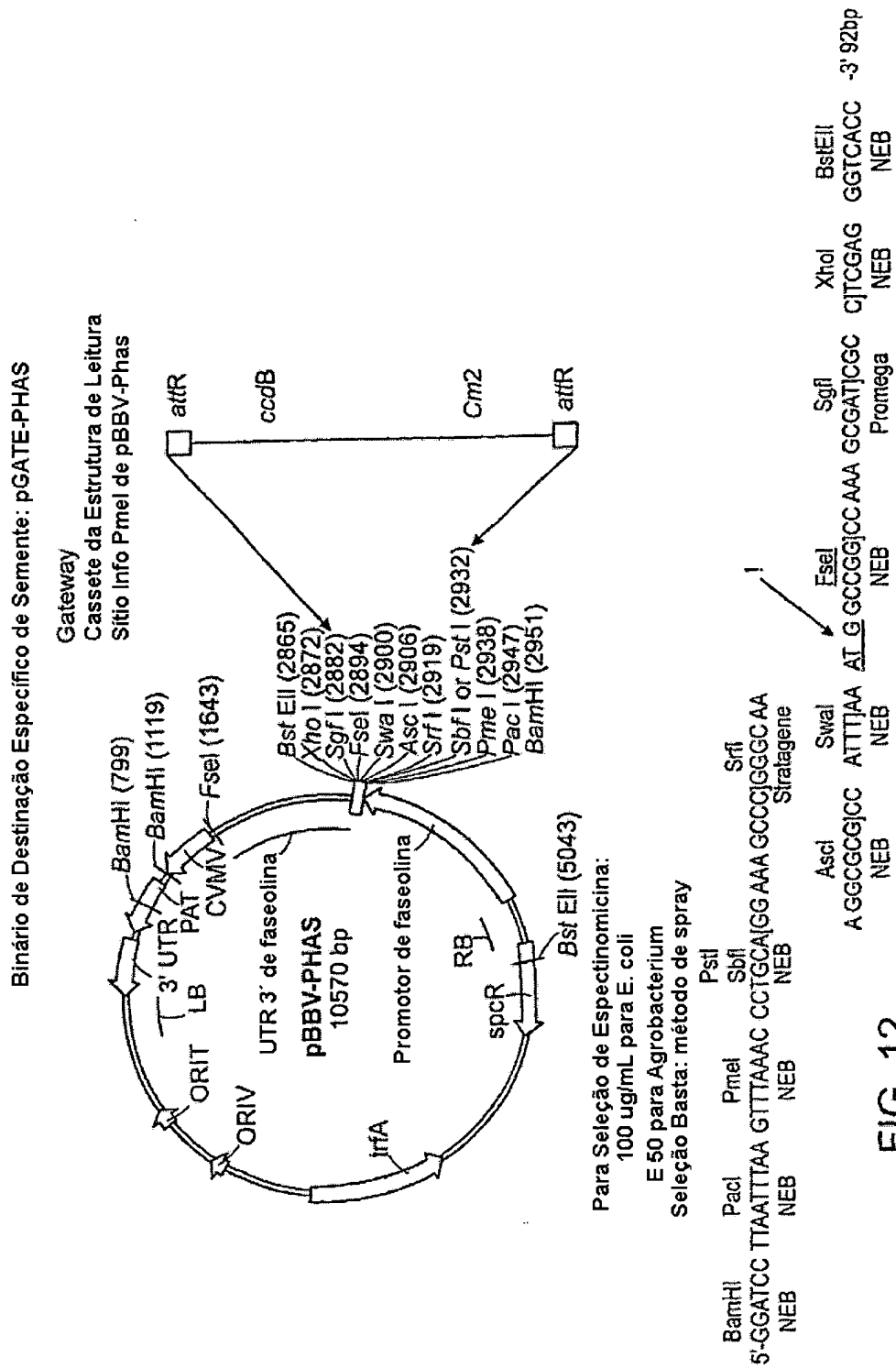
FIG. 11L

```

11151 acacttgagg ggcgactca ccggcgcg ccgttgacaga tgagggcgag
11201 gctcgatttc ggcggcgac gtggagctgg ccagctcgc aaatcgggcga
11251 aaacgcctga ttttacgca gttccaca gatgatgtgg acaagcctgg
11301 ggataagtgc cctgcggtat tgacacttga gggcgcgac tactgacaga
11351 tgagggggcg gatccttgac acttgagggg cagagtgatg acagatgagg
11401 ggcgaccta ttgacatttg aggggctgtc cacaggcaga aaatccagca
11451 tttgcaaggg ttccgcccg ttttcggcc accgctaacc tgtcttttaa
11501 cctgctttta aaccaatatt tataaacctt gtttttaacc agggctgcgc
11551 cctggcgctg gacggcgac gcgaagggg ggtgcccccc cttctcgaac
11601 cctcccggcc cgctaacgcg ggctcccat cccccaggg gctgcgcccc
11651 tcggccgcca acggcctcac ccaaaaatg gcaggccaag cttgcttggg
11701 cgttccggta cgtaccgtga acgtcggtc gattgtacct gcgttcaaat
11751 actttgcgat cgtgttgccg gcctgccccg tgcgtcggct gatctcacgg
11801 atcgactgct tctctcgcaa cgcataccga cggatgatgt ttaaaagtcc
11851 catgtggatc actccgttgc cccgtcgctc accgtgttgg ggggaagggtg
11901 cacatggctc agttctcaat ggaattatc tgcctaaccg gctcagttct
11951 gcgtagaaac caacatgcaa gctccaccgg gtgcaaaagc gcagcgcgcg
12001 caggatatat tcaattgtaa atggcttcat gtccgggaaa tctacatgga
12051 tcagcaatga gtatgatggt caatatggag aaaaagaaaag agtaattacc
12101 aattttttt caattcaaaa atgtagatgt ccgcagcgtt attataaaat
12151 gaaagtacat ttgtataaaa cgacaaatta cgatccgtcg tatttatagg
12201 cgaaagcaat aaacaaaatta ttctaattcg gaaatcttta ttctgacgtg
12251 tctacattca cgtccaaatg gggcgcgga att

```

FIG. 11M



RESUMO

"CONSTRUÇÃO GENÉTICA PARA A HIPER-EXPRESSÃO DE UMA
PROTEÍNA DE DOMÍNIO-F EM PLANTAS; CÉLULA DE PLANTA
TRANSGÊNICA; CÉLULA DA PLANTA *ARABIDOPSIS THALIANA*
5 RECOMBINANTE; PLANTA; MÉTODO DE MODIFICAÇÃO DA PRODUÇÃO DE
ÓLEO EM UMA PLANTA; E MÉTODO DE AUMENTO NA PRODUÇÃO DE ÓLEO
NA *ARABIDOPSIS THALIANA*".

Descrevem-se construções genéticas, células de
plantas transgênicas e plantas transgênicas, bem como os
10 métodos associados para aumentar a produção de óleo em uma
planta utilizando-se seqüências genéticas de domínio-F.