

(11) Número de Publicação: **PT 973920 E**

(51) Classificação Internacional:

C12N 15/82 (2007.10) **C12N 15/29** (2007.10)

C12N 15/53 (2007.10) **A01H 5/00** (2007.10)

A01H 5/10 (2007.10)

(12) **FASCÍCULO DE PATENTE DE INVENÇÃO**

(22) Data de pedido: **1998.04.09**

(30) Prioridade(s): **1997.04.09 US 831575**

(43) Data de publicação do pedido: **2000.01.26**

(45) Data e BPI da concessão: **2008.07.02**
158/2008

(73) Titular(es):

BAYER CROPSCIENCE S.A.

55 AVENUE RENÉ CASSIN 69009 LYON FR

(72) Inventor(es):

TERRY L. THOMAS

US

ZHONGSEN LI

US

(74) Mandatário:

JOSÉ EDUARDO LOPES VIEIRA DE SAMPAIO

R DO SALITRE 195 RC DTO 1250-199 LISBOA PT

(54) Epígrafe: **REGIÃO REGULADORA 5' DE OLEORESINA PARA MODIFICAR A COMPOSIÇÃO
LIPÍDICA DE SEMENTES DE PLANTAS**

(57) Resumo:

Resumo

"Região reguladora 5' de oleoresina para modificar a composição lipídica de sementes de plantas"

A presente invenção diz respeito a regiões reguladoras 5' de um gene de oleoresina de *Arabidopsis*. As regiões reguladoras 5', quando estão operacionalmente ligadas à sequência de codificação de um gene heterólogo ou a uma sequência complementar a um gene da planta nativa dirigem a expressão da sequência de codificação ou sequência complementar numa semente de planta. As regiões reguladoras são úteis em cassettes de expressão e vectores de expressão para a transformação de plantas. Também são disponibilizados processos para modular os níveis de um gene heterólogo, tal como o gene da síntese de ácidos gordos ou gene do metabolismo lipídico transformando uma planta com as ditas cassettes de expressão e vectores de expressão.

Descrição

"Região reguladora 5' de oleoresina para modificar a composição lipídica de sementes de plantas"

Antecedentes da Invenção

Tradicionalmente o teor em óleo de sementes tem sido modificado através de reprodução de plantas. A utilização de tecnologia de ADN recombinante para alterar a composição do óleo da semente pode acelerar este processo e nalguns casos alterar os óleos das sementes de um modo que não pode ser conseguido unicamente por reprodução.

A composição do óleo de *Brassica* foi significativamente alterada modificando a expressão de um número de genes do metabolismo lipídico. Tais manipulações da composição do óleo de sementes têm-se focalizado na alteração da proporção de componentes endógenos de ácidos gordos. Por exemplo, a repressão anti-sentido do gene da $\Delta 12$ -dessaturase na semente de colza transgénica resultou num aumento em ácido oleico até cerca de 83%. Topfer et al. 1995 Science 268:681-686.

Houve algumas tentativas bem sucedidas para modificar a composição do óleo de sementes em plantas transgénicas por introdução de novos genes que permitem a produção de um ácido gordo que as plantas hospedeiras não eram previamente capazes de sintetizar. Van de Loo, et al. (1995 Proc. Natl. Acad. Sci USA 92:6743-6747) foram capazes de introduzir um gene de $\Delta 12$ -hidróxilase em tabaco transgénico, resultando na introdução de um novo ácido gordo, o ácido ricinolénico no seu óleo das sementes. A acumulação reportada foi modesta em plantas que transportam construções em que a transcrição do gene da hidroxilase estava sob o controlo do promotor 35S do vírus do mosaico da couve de flor (CaMV).

De forma semelhante, as plantas do tabaco foram submetidas a engenharia genética para produzirem baixos níveis de ácido petroselinico por expressão de uma acil-ACP dessaturase a partir do coentro (Cahoon et al. 1992 Proc. Natl. Acad. Sci USA 89:11184-11188).

Os ácidos gordos de cadeia longa (C18 e maiores), possuem um valor económico significativo como alimentos, tanto do ponto de vista nutricional, como alimentos importantes do ponto de vista médico e como mercadorias industriais (Ohlrogge, J. B. 1994 Plant Physiol. 104:821-826). Ácido linoleico (18:2 $\Delta^9,12$) e ácido α -linolénico (18:3 $\Delta^9,12,15$) são ácidos gordos essenciais encontrados em muitos óleos de sementes. Os níveis destes ácidos gordos foram manipulados em culturas de sementes de oleaginosas, através de reprodução e biotecnologia (Ohlrogge, et al. 1991 Biochim. Biophys. Acta 1082:1-26; Topfer et al. 1995 Science 268:681-686). Além disso, a produção de novos ácidos gordos em óleos de sementes pode ser de utilidade considerável tanto na saúde humana como nas aplicações industriais.

Pensa-se que consumos de óleos de plantas ricos em ácido γ -linolénico (GLA) (18:3 $\Delta^6,9,12$) aliviam a hipercolesterolemia e outras desordens clínicas relacionadas que correlacionam com a susceptibilidade a doença arterial coronária (Brenner, R.R 1976 Adv. Exp Med. Biol. 83:85-101). Os benefícios terapêuticos do GLA dietético pode resultar do seu papel como precursor da síntese da prostanglandina (Weete, J. D. 1980 em Lipid Biochemistry of Fungi and Other Organisms, eds. Plenum Press, New York, pp. 59-62). O ácido linolénico (18:2) (LA) é transformado em ácido gama linolénico (18:3) (GLA) pela enzima Δ^6 -dessaturase.

Poucos óleos de sementes contêm GLA apesar dos elevados teores do ácido linolénico precursor. Isto é

devido à ausência de actividade de $\Delta 6$ -dessaturase na maior parte das plantas. Por exemplo, apenas borragem (*Borago officinalis*), primavera da noite (*Oenothera biennis*) e caril (*Ribes nigrum*) produzem quantidades significativas de ácido linoleico. Destas três espécies, apenas *Oenothera* e borragem são cultivadas como fonte comercial de GLA. Seria benéfico se óleos de sementes agronómicas pudessem ser submetidos a engenharia para produzir GLA em quantidades significativas através da introdução de um gene heterólogo da $\Delta 6$ -dessaturase. Seria também benéfico se outros produtos de expressão associados com a síntese de ácidos gordos e o metabolismo lipídico poderia ser produzidos em plantas com níveis suficientemente elevados para que a produção comercial de um produto de expressão particular possa ser realizável.

Como foi revelado na WO 96/21022, foi recentemente isolado um gene cianobacteriano da $\Delta 6$ -dessaturase. Expressão deste gene cianobacteriano no tabaco transgénico resultou numa acumulação de GLA significativa, mas de baixo nível. (Reddy et al. 1996 Nature Biotech. 14:639-642). Um pedido de patente de invenção da requerente também pendente US N° 08,366,779 revela um gene da $\Delta 6$ -dessaturase isolado a partir da planta de *Borago officinalis* e a sua expressão no tabaco sob o controlo do promotor 35S da CaMV. Uma tal expressão resultou num nível significativo, mas baixo de acumulação de GLA e de ácido octadecatetranoico (ODTA ou OTA) nas sementes. Assim, existe a necessidade de um promotor que funcione em plantas e que dirija consistentemente para um nível elevado de expressão de genes do metabolismo dos lípidos em sementes de plantas transgénicas.

As oleoresinas são proteínas abundantes de sementes associadas com a monocamada fosfolipídica da membrana de corpos oleosos. O primeiro gene da oleoresina, L3 foi

clonado a partir do milho seleccionando os clones, cujos produtos traduzidos *in vitro* foram reconhecidos por um anticorpo anti-L3 (Vance et al. 1987 J. Biol. Chem. 262:11275-11279). Posteriormente foram clonadas as diferentes isoformas dos genes da oleoresina de espécies tão diferentes como *Brassica*, soja, cenoura, pinheiro, e *Arabidopsis* (Huang, A H. C., 1992, Ann. Reviews Plant Phys. e Plant Mol. Biol. 43:177-200; Kirik et al., 1996 Plant Mol. Biol. 31:413-417; Van Rooijen et al., 1992 Plant. Mol. Biol. 18:1177-1179; Zou et al., Plant Mol. Biol. 31:429-433). Sequências de proteína de oleoresina previstas a partir destes genes são altamente conservadas, especialmente a partir do domínio hidrofóbico central. Todas estas oleoresinas possuem a característica de terem três domínios distintos. Um domínio anfipático de 40-60 aminoácidos encontra-se presente no terminal-N; um domínio completamente hidrofóbico de 68-74 aminoácidos encontra-se localizado no centro; e um domínio anfipático de hélice- α de 33-40 aminoácidos encontra-se situado no terminal-C (Huang, A H.C. 1992).

Foram ainda descritas sequências reguladoras a partir de um gene de oleoresina que dirige para uma expressão de nível elevado de genes do metabolismo dos lípidos em plantas transgênicas. De acordo com a presente invenção são descritas construções quiméricas compreendendo uma região reguladora 5' da oleoresina ligada operacionalmente à sequência de codificação para um gene do metabolismo dos lípidos, tal como um gene da $\Delta 6$ -dessaturase. Plantas transgênicas compreendendo as ditas construções quiméricas produzem níveis de GLA que se aproximam do nível encontrado naquelas poucas plantas que produzem GLA naturalmente, tal como a primavera da noite (*Oenothera biennis*).

Sumário da Invenção

A presente invenção disponibiliza cassettes de expressão e vectores de expressão compreendendo a região reguladora 5' que dirige a expressão específica na semente constituída pela sequência nucleótida apresentada na SEQ ID N° 12 ligada operacionalmente a pelo menos um ácido nucleico que codifica para um gene heterólogo da $\Delta 6$ -dessaturase.

Vectores de transformação de plantas compreendendo as cassettes de expressão e vectores de expressão são também disponibilizados, assim como células de plantas transformadas por estes vectores, e plantas e a sua descendência contendo os vectores.

Num outro aspecto é descrito um processo para produzir uma planta com níveis acrescidos de um produto de uma síntese de ácido gordo ou gene do metabolismo lipídico.

Em particular, é descrito um processo para produzir uma planta com níveis acrescidos de um gene de uma síntese de ácido gordo ou do metabolismo lipídico transformando a planta com as ditas cassettes de expressão e vectores de expressão que compreende uma região reguladora 5' de oleoresina e uma sequência de codificação para um gene de uma síntese de ácido gordo, ou de metabolismo lipídico.

Num outro aspecto é descrito um processo para co-supressão de um gene da síntese de um ácido gordo nativo, ou do metabolismo lipídico transformando uma planta com as ditas cassettes de expressão e vectores de expressão que compreendem uma região 5' reguladora de oleoresina e uma sequência de codificação para uma síntese de ácido gordo ou gene do metabolismo lipídico.

Ainda outro aspecto descreve um processo para diminuir a produção de um gene nativo da planta, tal como um gene da síntese de ácido gordo ou gene do metabolismo lipídico

transformando a planta com um vector de expressão compreendendo uma região reguladora 5' da oleoresina operacionalmente ligada a uma sequência de ácido nucleico complementar a um gene nativo da planta.

Também são descritos processos de modular os níveis de um gene heterólogo, tal como um gene da síntese de ácido gordo ou do metabolismo lipídico transformando uma planta com as ditas cassettes de expressão e vectores de expressão.

Breve Descrição das Figuras

A Fig. 1 representa o nucleótido e a sequência de aminoácidos correspondente do gene da $\Delta 6$ -dessaturase da borragem (SEQ ID N° 1). O motivo de ligação ao hemo do citocromo b5 encontra-se encaixilhado e a ligação putativa ao metal, motivos ricos em histidina (HRMs) encontram-se sublinhados. Os motivos reconhecidos pelos iniciadores (análise PCR) encontram-se sublinhados com linhas a tracejado, i.e. tgg aaa tgg aac cat aa; e gag cat cat ttg ttt cc.

A Fig. 2 é um dendograma que apresenta semelhança entre a $\Delta 6$ -dessaturase da borragem com outras dessaturases ligadas a membranas. A sequência de aminoácidos da $\Delta 6$ -dessaturase da borragem foi comparada com outras dessaturases conhecidas utilizando Gene Works (IntelliGenetics). Valores numéricos correlacionam-se com distâncias filogenéticas relativas entre os subgrupos comparados.

A Fig. 3A apresenta um perfil de cromatografia gás líquido dos ésteres metílicos dos ácidos gordos (FAMES) derivados do tecido das folhas de um tabaco selvagem do tipo "Xanthi".

A Fig. 3B apresenta um perfil de cromatografia gás líquido dos (FAMES) derivados do tecido das folhas de uma planta de

tabaco transformada com o cADN da $\Delta 6$ -dessaturase da borragem sob o controle de transcrição do promotor 35S da CaMV (pAN2). São indicados os picos correspondentes ao linoleato de metilo (18:2), γ -linolenato de metilo (18:3 γ), α -linolenato de metilo (18:3 α), e octadecatetraenoate de metilo (18:4).

A Fig. 4 representa a sequência nucleotídica e a correspondente sequência de aminoácidos da oleoresina do cADN de AtS21 (SEQ ID N°3).

A Fig. 5 é uma mapa ácido-base da proteína AtS21 prevista gerada pelo DNA Strider 1,2.

A Fig. 6 é uma representação de Kyte-Doolittle da proteína AtS21 prevista gerada pelo DNA Strider 1,2.

A Fig. 7 é um alinhamento de sequência da oleoresina isolada da *Arabidopsis*. As sequências de oleoresinas publicadas ou depositadas nas bases de dados EMBL, BCM, NCBI foram alinhadas umas com as outras utilizando o GeneWorks@2.3. Os resíduos idênticos estão encaixilhados em rectângulos. As sete sequências caem em três grupos. O primeiro grupo inclui AtS21 (SEQ ID N°:5), X91918 (SEQ ID N°6) e Z29859 (SEQ ID N°: 7). O segundo grupo inclui X62352 (SEQ ID N°:8) e Ato13 (SEQ ID N°:9). O terceiro grupo inclui X91956 (SEQ ID N°10) e L40954 (SEQ ID N°:11). As diferenças nos resíduos de aminoácidos do mesmo grupo são indicadas a sombreado. Ato2/Z54164 é idêntico a AtS21. Na realidade a sequência Ato13 (N° de acesso Z541654 na base de dados EMBL) não é revelada na base de dados EMBL. O número de acesso Z54165 designa a mesma sequência que Z54164 que é Ato12.

A Fig. 8A é uma análise de Northern do gene AtS21. Um gel de transferência de ARN contendo 10 microgramas de ARN total extraído de folhas de *Arabidopsis* (F), folhas (L), raízes (L), sementes em desenvolvimento (Se) e tegumentos

em desenvolvimento de siliqua (Si) foram hibridizados com uma sonda feita de cADN de AtS21 de comprimento completo.

A Fig. 8B é uma análise de Southern do gene AtS21. Um gel de transferência de ADN contendo 10 microgramas de ADN genómico digerido com BamHI (B), EcoRI (E), HindIII (H), SacI (S), e XbaI (X) foi hibridizado com uma sonda feita a partir de cADN de AtS21 de comprimento completo.

A Fig. 9 é uma sequência nucleótida do fragmento SacI de ADN genómico AtS21 (SEQ ID N°:12). O promotor e as sequências intrão encontram-se em maiúscula. Os fragmentos correspondentes a sequências de cADN de AtS21 encontram-se em minúsculas. O primeiro codão ATG e uma caixa TATA putativa encontram-se a sombreado. A sequência complementar ao iniciador 21P para amplificação por PCR encontra-se encaixilhada. Um elemento putativo da resposta do ácido abscísico (ABRE) e duas repetições de 14 pb encontram-se sublinhadas.

A Fig. 10 é um mapa da construção promotor AtS21 /GUS (pAN5).

A Fig. 11A representa a expressão do gene AtS21/GUS na fechadura de *Arabidopsis* e nas folhas.

A Fig. 11B representa a expressão do gene AtS21/GUS na *Arabidopsis* siliqua.

A Fig. 11C representa a expressão do gene AtS21/GUS nas sementes em desenvolvimento da *Arabidopsis*.

As Fig. 11D até 11J representam a expressão do gene AtS21/Gus nos embriões em desenvolvimento da *Arabidopsis*.

A Fig. 11K representa a expressão do gene AtS21/GUS na raiz de *Arabidopsis* e capilares das raízes de uma plântula nova.

A Fig. 11L representa a expressão do gene AtS21/GUS nos cotilédones de *Arabidopsis* e no rebento apical de uma plântula de cinco dias.

A Fig. 11M e 11N representam a expressão do gene AtS21/GUS nos cotilédones de *Arabidopsis* e nos rebentos apicais de uma plântula de 5-15 dias.

A Fig. 12A representa a expressão do gene AtS21/GUS nos embriões e endosperma do tabaco.

A Fig. 12B representa a expressão do gene AtS21/GUS nas sementes de tabaco em germinação.

A Fig. 12C representa a expressão do gene AtS21/GUS numa plântula de tabaco com 5 dias.

A Fig. 12D representa a expressão do gene AtS21/GUS em plântula de tabaco com 5 a 15 dias.

A Fig. 13A é uma análise de Northern que apresenta níveis de AtS21 mRNA em plântulas em desenvolvimento de *Arabidopsis* do tipo selvagem. A zona 1 foi carregada com ARN de sementes em desenvolvimento, a zona 2 foi carregada com ARN de sementes inibidas durante 24-48 horas, a zona 3: sementes de 3 dias; a zona 4: sementes de 4 dias; a zona 5: plântulas de 5 dias; a zona 6: plântulas de 6 dias; a zona 7: plântulas de 9 dias; a zona 8: plântulas de 12 dias. A sonda foi marcada com cADN de AtS21. A exposição foi de uma hora a -80°C .

A Fig. 13B representa a mesma transferência que a Fig. 13A apenas a exposição foi durante 24 horas a -80°C .

A Fig. 13C representa a mesma transferência que as representados nas Figs. 13A e 13B após remoção e hibridização com uma sonda genética de tubulina de *Arabidopsis*. A pequena banda em cada uma das zonas 1 e 2 é o que resta da sonda AtS21 anterior. A exposição foi realizada durante 48 horas a -80°C .

A Fig. 14 é um gráfico que compara as actividades GUS expressas pelos promotores AtS21 e 35S. As actividades GUS expressas pelo promotor AtS21 em sementes de *Arabidopsis* em desenvolvimento e folhas são apresentadas lado a lado com aquelas expressas pelo promotor 35S. As actividades GUS

expressas pelo promotor AtS21 nas sementes secas do tabaco e folhas são apresentadas do lado direito da figura. A actividade GUS na folha de tabaco é tão baixa que não aparece nenhuma coluna. "G-H" designa a fase globular até à fase coração; "H-T" designa a fase coração à fase torpedo; "T-C" designa a fase torpedo até à fase de cotilédone; "Early C" é a expressão inglesa para cotilédone prematuro; "Late C" é a expressão inglesa para cotilédone tardio. Os desvios padrões encontram-se listados na Tabela 2.

A Fig. 15A é uma análise de transferência em gel efectuada em amostras de 5 µg de ARN isolado da folha de borragem, raiz e tecido de embrião 12 dpp utilizando como sonda de hibridização cADN marcado com $\Delta 6$ -dessaturase da borragem.

A Fig. 15B representa um gráfico correspondente aos resultados da análise de Northern para a experiência apresentada na Fig. 15A.

A Fig. 16A é um gráfico que apresenta a acumulação relativa de ARN de legumina em embriões em desenvolvimento da borragem baseadas em resultados da transferência de Northern.

A Fig. 16C é um gráfico que apresenta uma acumulação de ARN de $\Delta 6$ -dessaturase em embriões em desenvolvimento da borragem baseada em resultados da transferência de Northern.

A Fig. 17 é uma análise de PCR que mostra a presença do gene delta $\Delta 6$ -dessaturase da borragem em plantas transformadas da colza. As zonas 1,3 e 4 foram carregadas com reacções PCR efectuadas com ADN de plantas transformadas com o gene delta $\Delta 6$ -dessaturase da borragem ligado à região 5' reguladora da oleoresina; a zona 2: ADN de planta transformada com o gene da delta $\Delta 6$ -dessaturase da borragem ligado à região reguladora 5' da albumina; zonas 5 e 6: ADN de plantas não transformadas; zona 7:

marcador de peso molecular (cadeia de 1 kb, Gibco); zona 8: PCR sem adição de ADN molde; zona 9: controlo com ADN de *Agrobacterium tumefaciens* EHA 105 contendo o plasmídeo pAN3 (i.e. o gene delta $\Delta 6$ -dessaturase da borragem ligado à região reguladora 5').

Descrição Detalhada da Invenção

A presente invenção disponibiliza uma cassette de expressão que compreende a região reguladora da oleoresina 5' que dirige para a expressão específica na semente constituída pela sequência nucleotídica apresentada na SEQ ID N° 12 ligada operacionalmente a pelo menos um ácido nucleico que codifica para um gene heterólogo da $\Delta 6$ -dessaturase. Uma tal cassette de expressão pode ser incorporada numa variedade de vectores de replicação autónoma de modo a construir-se um vector de expressão.

Verificou-se surpreendentemente que as plantas transformadas com os vectores de expressão da presente invenção produzem níveis de GLA que se aproximam do nível encontrado naquelas poucas espécies de plantas que produzem GLA naturalmente, tal como a primavera da noite (*Oenothera biennis*).

Quando aqui utilizado, o termo "cassette" diz respeito a uma sequência nucleotídica capaz de expressar um gene particular se o dito gene for inserido de modo a estar operacionalmente ligado a uma ou mais regiões reguladoras presentes na sequência nucleotídica. O termo "expressão específica nas sementes", quando aqui utilizada refere-se à expressão em várias porções de uma semente da planta, tal como o endosperma e o embrião.

Um ácido nucleico isolado que codifica para uma região reguladora 5' de um gene de oleoresina pode ser fornecido como se segue. Clones genómicos recombinantes da oleoresina

são isolados por rastreio de uma biblioteca de ADN genómico de planta com um cADN (ou uma sua porção) representando ARNm de oleoresina. Foram isolados vários cADNs diferentes de oleoresina. Os métodos utilizados para isolar tais cADNs, assim como os nucleótidos e as sequências correspondentes de aminoácidos foram publicados em Kirik et al. 1986 Plant Mol. Biol. 31:413-417; Zou et al. Plant Mol. Biol. 31:429-433; Van Rooijen et al. 1992 Plant Mol. Biol. 18:1177-1179.

O rastreio por subtracção virtual de uma biblioteca específica de um tecido utilizando uma sonda de cADN de polimerase iniciada aleatoriamente (RP-PCR- do inglês random primed polymerase chain (RP-PCR) é um outro método para obter um cADN de oleoresina útil para rastrear uma biblioteca de ADN genómico de planta. O rastreio por subtracção virtual refere-se a um método em que uma biblioteca de cADN é construída a partir de um tecido alvo e apresentada a uma baixa densidade de modo a que clones de cADN individuais possam ser facilmente separados. Estes clones de cADN são rastreados de forma subtractiva com quantidades condutoras (i.e., concentrações de ADN para conduzir cineticamente a reacção de hibridização) de sondas de cADN feitas de tecido ou tecidos diferentes do tecido alvo (i.e. tecido condutor). As placas hibridizadas representam genes que são expressos tanto nos tecidos alvo e nos tecidos condutores; as placas não hibridizadas representam genes que podem ser específicos do tecido alvo ou genes pouco abundantes que não podem ser detectados pela sonda de cADN condutora. Os cADNs não hibridizados são seleccionados como genes putativos específicos do tecido alvo e ainda analisados por uma sequenciação de uma passagem e hibridização de Northern.

PCR iniciada aleatoriamente (RP-PCR) envolve a síntese de grandes quantidades de sondas de cADNs a partir de uma

quantidade vestigiária de cADN molde. O método combina o poder de amplificação de PCR com uma representação de iniciação aleatória para amplificar simultaneamente e marcar um cADN de cadeia dupla numa reacção de tubo único.

Métodos considerados úteis para obter ADN recombinante genómico de oleoresina são fornecidos em Sambrook et al. 1989, em *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY, por exemplo, ou qualquer um da miríade de manuais de laboratório sobre tecnologia de ADN recombinante que estão largamente disponíveis. Para determinar as sequências nucleótídicas encontram-se disponíveis uma multiplicidade de técnicas e são conhecidas pelo vulgar perito no estado da técnica. Por exemplo, fragmentos de restrição contendo uma região reguladora da oleoresina podem ser subclonadas no local do polilinker de um vector de sequenciação, tal como pBluescript (Stratagene). Estes subclones pBluescript podem então ser sequenciados pelo método do didesóxi de dupla cadeia (Chen e Seeburg, 1985, ADN 4:165).

A região reguladora da oleoresina compreende os nucleótidos 1-1267 da Fig. 9 (SEQ ID N°:12).

A região reguladora 5' da presente invenção pode ser obtida por digestão da endonuclease, ou exonuclease de restrição de um clone genómico de oleoresina. Assim, por exemplo, o nucleótido conhecido, ou sequência de aminoácidos da região codificante de um gene isolado da oleoresina (e.g. Fig. 7) é alinhada com a sequência de ácidos nucleicos, ou a sequência de aminoácidos deduzida de um clone genómico de oleoresina isolado e da sequência flanqueadora 5' (i.e. a sequência a montante do codão de iniciação da tradução da região de codificação) do clone genómico da oleoresina isolado localizado.

A região reguladora 5' da oleoresina apresentada na SEQ ID N°:12 (nucleótidos 1-1267 da Fig. 9) pode ser gerado

tanto a partir de um clone genómico possuindo um excesso da sequência flanqueadora 5' como de uma sequência de codificação, ou ambas através de supressão mediada pela exonuclease III. Isto é conseguido fazendo digerir ADN preparado adequadamente com exonuclease III (exoIII) e removendo aliquotas em intervalos de tempo crescentes durante a digestão. Os fragmentos de ADN sucessivamente mais pequenos resultantes podem ser sequenciados para determinar o ponto final exacto das supressões. Estes são vários sistemas comerciais disponíveis que utilizam a exonuclease III (exoIII) para criar uma tal série de supressões, por ex. Promega Biotech, sistema "Erase-A-Base". Alternativamente, os iniciadores de PCR podem ser definidos para permitir a amplificação directa das ditas regiões reguladoras 5'.

Utilizando as mesmas metodologias o vulgar perito no estado da técnica pode gerar um ou vários fragmentos de supressão de nucleótidos 1-1267 como apresentado na SEQ ID N°: 12.

A identificação das sequências reguladoras 5' da oleoresina que dirigem para a expressão específica na semente compreendendo os nucleótidos 1-1267 da SEQ ID N°:12 e modificações ou os seus fragmentos de supressão pode ser conseguida, através de fusões transcricionais de sequências específicas com as sequências de codificação de um gene heterólogo, transferência do gene quimérico para um hospedeiro adequado, e detecção da expressão do gene heterólogo. O ensaio utilizado para detectar a expressão depende da natureza da sequência heteróloga. Por exemplo, os genes reporter exemplificados por cloroanfenicol acetil transferase e β -glucuronidase (GUS) são utilizados habitualmente para avaliar a competência transcricional e de tradução de construções quiméricas. Encontram-se

disponíveis ensaios padrão para detectar de forma sensível a enzima reporter num organismo transgénico.

O gene da β -glucuronidase (GUS) é útil como reporter da actividade promotora em plantas transgénicas, devido à elevada estabilidade da enzimas nas células das plantas, a falta de actividade intrínseca da β -glucuronidase nas plantas superiores e disponibilidade de um ensaio fluorimétrico quantitativo e de uma técnica de localização histoquímica. Jefferson et al. (1987 EMBO J 6:3901) estabeleceram procedimentos padrão para a detecção bioquímica e histoquímica da actividade GUS em tecidos de plantas. Os ensaios bioquímicos são efectuados misturando lisados de tecidos de plantas com 4-metilumbeliferil- β -D-glucuronido, um substrato fluorimétrico para GUS, incubando durante uma hora a 37°C e depois medindo a fluorescência da 4-metil-umbeliferona resultante. A localização histoquímica da actividade GUS é determinada incubando amostras de tecido de plantas em 5-bromo-4-cloro-3-indolil-glucuronido (X-Gluc) durante cerca de 18 horas a 37°C e observando o padrão das manchas de X-Gluc. A construção de tais genes quiméricos permite a definição de sequências reguladoras específicas e demonstra que estas sequências podem dirigir a expressão directa dos genes heterólogos de uma maneira específica na semente.

Um outro aspecto da invenção é direccionado para cassettes de expressão e vectores de expressão (também aqui referidos como "genes quiméricos") compreendendo uma região reguladora 5' de um gene de oleoresina que direcciona para uma expressão específica da semente operacionalmente ligado à sequência de codificação de um gene heterólogo de tal modo que o elemento regulador é capaz de controlar a expressão de um ácido nucleico que codifica para um gene heterólogo de $\Delta 6$ -dessaturase. Se necessário, elementos reguladores adicionais, ou partes destes elementos

suficientes para provocar a expressão resultando na produção de uma quantidade eficaz do polipéptido codificado pelo gene heterólogo são incluídos nas construções quiméricas. Exemplos de genes do metabolismo lipídico incluem dessaturases lipídicas tais como as $\Delta 6$ -dessaturases, $\Delta 12$ -dessaturases, $\Delta 15$ -dessaturases e outras dessaturases relacionadas, tais como as estearoilo-ACP dessaturases, proteínas transportadoras de acilo (ACPs), tioesterases, acetil transacilases, acetil-coA carboxilases, cetoacil-sintases, malonil transacilases, e elongases. Tais genes do metabolismo dos lipídios foram isolados e caracterizados a partir de várias bactérias diferentes e espécies de plantas. As suas sequências codificadoras de nucleótidos, assim como métodos de isolar tais sequências codificadoras são revelados na literatura publicada e estão largamente disponíveis aos peritos no estado da técnica.

Em particular os genes da $\Delta 6$ -dessaturase revelados na WO 96/21022 e no pedido de patente de invenção das requerentes pendente US N° 08/366,779 depositado a 30 de Dezembro de 1994 são contemplados como genes do metabolismo lipídico particularmente úteis na prática da presente invenção.

Os genes quiméricos da presente invenção são construídos ligando uma região reguladora 5' de um ADN genómico de oleoresina como apresentado na SEQ ID N°12 com a sequência codificadora de um gene heterólogo. A justaposição destas sequências pode ser realizada de várias maneiras. Numa concretização preferida, a ordem das sequências de 5' a 3' é uma região reguladora 5' de oleoresina (incluindo um promotor), uma sequência codificadora, e uma sequência de terminação que inclui um local de poliadenilação.

Técnicas padrão para a construção de tais genes quiméricos são bem conhecidas dos peritos no estado da técnica e podem ser encontradas em referências tais como Sambrook et al. (1989). Encontram-se disponíveis uma variedade de estratégias para ligar fragmentos de ADN, cuja escolha depende da natureza dos fragmentos de ADN terminais. Um perito vulgar no estado da técnica reconhece que para os genes heterólogos serem expressos, a construção necessita de elementos promotores e sinais para uma poliadenilação eficiente do transcrito. Neste contexto, a região 5' reguladora da oleoresina que contém a sequência promotora consenso conhecida como a caixa TATA pode ser ligada directamente a uma sequência de codificação heteróloga sem promotores.

Os fragmentos de restrição ou de supressão que contêm a caixa da oleoresina TATA são ligados numa orientação directa em relação a um gene heterólogo sem promotor, tal como a sequência de codificação da β -glucuronidase (GUS). O perito no estado da técnica irá reconhecer que as ditas regiões reguladoras 5' da oleoresina podem ser fornecidas por outros meios, por exemplo por síntese química, ou enzimática. O terminal 3' de uma sequência de codificação heteróloga está operacionalmente ligada a uma sequência de terminação compreendendo um local de poliadenilação, exemplificado por, mas não limitado a local de poliadenilação da nopalina sintase, ou o local de poliadenilação da octopina T-ADN do gene 7. Alternativamente, o local de poliadenilação pode ser fornecido pelo gene heterólogo.

Ainda são descritos processos para aumentar os níveis de genes heterólogos em sementes de plantas. De acordo com estes processos, as ditas cassettes de expressão e vectores de expressão são introduzidos numa planta para efectuarem a expressão de um gene heterólogo. Por exemplo, um método

para produzir uma planta com níveis acrescidos de um produto de uma síntese de ácido gordo, ou um gene de metabolismo lipídico é obtido transformando uma célula de planta com um vector de expressão compreendendo uma região reguladora 5' de oleoresina ligada operacionalmente a um gene de uma síntese de ácido gordo, ou do metabolismo lipídico e que regenera uma planta com níveis acrescidos do produto da dita síntese de ácido gordo, ou gene do metabolismo lipídico.

São também descritos métodos para reduzir os níveis de produção de um gene que é nativo em relação a uma planta que compreendem a transformação de uma célula de planta com um vector de expressão compreendendo a dita região reguladora da oleoresina ligada operacionalmente a uma sequência de ácido nucleico que é complementar ao gene da planta nativa. Desta forma, níveis de produto endógeno do gene nativo da planta são reduzidos através do mecanismo conhecido como regulação anti-sentido. Assim, por exemplo, níveis de um produto de um gene da síntese de ácido gordo, ou de gene de metabolismo lipídico são reduzidos transformando uma planta com um vector de expressão compreendendo uma dita região reguladora 5' da oleoresina ligada operacionalmente a uma sequência de ácido nucleico que é complementar a uma sequência de ácido nucleico que codifica para um gene de uma síntese de ácido gordo, ou para um gene do metabolismo lipídico.

A presente invenção também descreve um método de co-supressão de um gene que é nativo em relação a uma planta que compreende a transformação de uma célula de planta com um vector de expressão compreendendo uma dita região reguladora 5' de oleoresina operacionalmente ligada a uma sequência de ácido nucleico que codifica para o gene nativo da planta. Desta maneira, níveis de produto endógeno do gene da planta nativa são reduzidos através do mecanismo

conhecido como co-supressão. Assim, por exemplo, níveis de um produto de um gene da síntese de ácido gordo ou gene de metabolismo lipídico são reduzidos transformando uma planta com um vector de expressão compreendendo uma dita região reguladora 5' da oleoresina ligada operacionalmente a uma sequência de ácido nucleico que codifica para um gene nativo de síntese de ácido gordo, ou gene do metabolismo lipídico nativo da planta. Apesar do mecanismo exacto da co-supressão não ser completamente compreendido um perito no estado da técnica encontra-se familiarizado com obras publicadas que reportam as condições experimentais e resultados associados com a co-supressão (Napoli et al. 1990 *The Plant Cell* 2:270-289; Van der Krol 1990 *The Plant Cell* 2:291-299).

Para fornecer uma resposta regulada dos genes heterólogos, as plantas são transformadas com as construções do gene quimérico da invenção. Métodos de transferência genética são bem conhecidos no estado da técnica. Os genes quiméricos podem ser introduzidos nas plantas através do procedimento de transformação-regeneração de disco de folha como descrito por Horsch et al. 1985 *Science* 227:1229. Podem ser também utilizados outros métodos de transformação, tal como cultura de protoplasto (Horsch et al. 1984 *Science* 223:496, DeBlock et al. 1984 *EMBO J.* 2:2143, Barton et al. 1983. *Cell* 32:1033) e encontram-se no âmbito da invenção. Numa concretização preferida as plantas são transformadas com vectores derivados de *Agrobacterium*, tais como os descritos em Klett et al. (1987) *Annu. Ver. Plant Physiol.* 38:467. Outros métodos bem conhecidos estão disponíveis para inserir os genes quiméricos da presente invenção em células de plantas. Tais métodos alternativos incluem estratégias biolísticas (Klein et al. 1987 *Nature* 327:70),

electroporação, recebimento de ADN quimicamente induzido, e utilização de vírus, ou de pólen como vectores.

Quando necessário para o método de transformação, os genes quiméricos da presente invenção podem ser inseridos num vector de transformação da planta, i.e. o vector binário descrito por Bevan, M. 1984 Nucleic Acids Res. 12:8711-8721. Vectores de transformação da planta podem ser obtidos modificando o sistema natural de transferência genética de *Agrobacterium tumefaciens*. O sistema natural compreende grandes plasmídeos Ti (indutores de tumores) contendo um grande segmento, conhecido como T-ADN, que é transferido para plantas transformadas. Outro segmento do plasmídeo Ti, a região vir, é responsável pela transferência de T-ADN. A região de T-ADN tem como fronteira repetições terminais. Nos vectores binários modificados, os genes indutores de tumores foram suprimidos e as funções da região vir são utilizadas para transferir ADN estranho bordejado pelas sequências de fronteira de T-ADN. A região-T também contém um marcador seleccionável para a resistência a antibiótico, e um local de clonagem múltiplo para inserir sequências a transferir. Tais estirpes sujeitas a engenharia são conhecidas como estirpes de *A. tumefaciens* "desarmadas" e permitem a transferência eficiente das sequências bordejadas pela região-T para o genoma nuclear das plantas.

Discos de folhas de superfície esterilizada e outros tecidos susceptíveis são inoculados com *A. tumefaciens* contendo ADN estranho "desarmado" cultivados durante uma série de dias e depois transferidos para um meio contendo antibiótico. Os rebentos seleccionados foram então seleccionados após enraizamento em meio contendo o antibiótico apropriado e transferidos para o solo. As plantas transgênicas são polinizadas e as sementes destas plantas são colhidas e feitas crescer em meio antibiótico.

A expressão de um gene heterólogo ou reporter em sementes em desenvolvimento, plantas novas e plantas maduras podem ser monitorizadas por ensaios imunológicos, histoquímicos ou de actividade. Como aqui discutido, a escolha de um ensaio para a expressão de um gene quimérico depende da natureza da região de codificação heteróloga. Por exemplo, a análise de Northern pode ser utilizada para avaliar a transcrição se se encontrarem disponíveis sondas nucleótídicas adequadas. Se se encontrarem disponíveis anticorpos para os polipéptidos codificados pelo gene heterólogo, a análise de Western e localização imunohistoquímica pode ser utilizada para avaliar a produção e localização do polipéptido. Dependendo do gene heterólogo podem ser utilizados ensaios bioquímicos adequados. Por exemplo, as acetiltransferases são detectadas medindo a acetilação de um substrato padrão. A expressão de um gene de uma dessaturase lipídica pode ser ensaiada através da análise de ésteres metílicos de ácidos gordos (FAMES).

Outro aspecto da presente invenção disponibiliza plantas transgênicas, ou descendência destas plantas contendo os genes quiméricos da invenção. Estão contempladas tanto plantas monocotiledóneas como dicotiledóneas. As células de plantas são transformadas com os genes quiméricos através de quaisquer dos processos de transformação de plantas descritos acima. A célula da planta transformada, geralmente na forma de uma cultura de caule, disco de folha, explante, ou planta completa (através do método de infiltração com vácuo de Bectold et al. 1993 C.R. Acad. Sci. Paris, 316:1194-1199) é regenerada numa planta transgênica completa, através de métodos bem conhecidos do vulgar perito no estado da técnica (por ex. Horsh et al. 1985 Science 227:1129). Numa concretização preferida, a planta transgênica é o girassol, algodão,

colza, milho, tabaco, *Arabidopsis*, amendoim, ou soja. Uma vez que a descendências das plantas transformadas herdam os genes quiméricos, as sementes, ou estacas de plantas transformadas são utilizadas para manter a linha transgênica.

Os exemplos seguintes ilustram ainda a invenção.

Exemplo 1

Isolamento de ARN polissomal ligado à membrana e construção de biblioteca de cADN de borragem

Foram isolados polissomas ligados à membrana a partir de sementes de borragem 12 dias após a polinização (12DPP) utilizando o protocolo estabelecido para ervilhas por Larkins e Davies (1975 Plant Phys: 55:749-756). O ARN foi extraído a partir dos polissomas como descrito por Mechler (1987 Methods in Enzymology 152:241-248, Academic Press). ARN Poli-A⁺ foi isolado a partir do ARN polissomal ligado à membrana utilizando esferas Oligotex.dTTM (Qiagen).

O cADN correspondente foi feito utilizando o kit de síntese da Stratagene ZAP. A biblioteca de cADN foi construída no vector lambda ZAP II (Stratagene) utilizando o kit lambda ZAP II. A biblioteca primária foi empacotada com extracto de empacotamento Gigapack II Gold (Stratagene).

Exemplo 2

Isolamento de um cADN Δ -6 dessaturase de borragem

Protocolo de hibridização

A biblioteca de cADN de borragem amplificada foi colocado numa placa a baixa densidade (500 pfu em placas de

Petri de 150 mm). Foram reduzidos cADNs de proteína de armazenamento da semente muito comuns (subtraídos do total de cADNs) por rastreio com os cADNs correspondentes.

Sondas de hibridização para rastreio da biblioteca de cADN de borragem foram geradas utilizando a síntese de ADN de iniciação aleatória como descrito por Ausubel et al. (1994 Current Protocols in Molecular Biology, Wiley Interscience, N. Y.) e correspondia aos cADNs anteriormente identificados da proteína de armazenamento na semente abundantemente expressa. Foram removidos nucleótidos não incorporados utilizando uma coluna giratória G-50 (Boehringer Mannheim). A sonda foi desnaturada por hibridização fervendo num banho de água durante 5 minutos, e depois arrefecidas rapidamente em gelo. Filtros de nitrocelulose que transportam bacteriófago fixo recombinante foram pré-hibridizados a 60°C durante 2-4 horas numa solução de hibridização [4X SET (600 mM NaCl, 80 mM de Tris-HCl, 4 mM Na₂EDTA; pH 7.8), 5x reagente de Denhardt's (0,1% de albumina de soro de bovino, 0,1 % de Ficoll, e 0,1% de polivinilpirrolidona), 100 µg/ml de ADN de esperma de salmão desnaturado, 50 µg/ml de poliadenina e 10 µg/ml de policitidina). Isto foi substituído com uma solução de hibridização fresca à qual foi adicionadas uma sonda radioactiva desnaturada (2 ng/ml de solução de hibridização). Os filtros foram incubados a 60°C com agitação de um dia para o outro. Os filtros foram lavados sequencialmente em 4x, 2x e 1x SET (NaCl 150 mM, Tris-HCl 20 mM, Na₂EDTA 1 mM; pH 7,8) durante 15 minutos cada um a 60°C. Os filtros foram secos ao ar e depois expostos a um filme de raios-X durante 24 horas com ecrans intensificadores a -80°C.

Placas não hidrolizantes foram excisadas utilizando um protocolo de excisão da Stratagene e reagentes. As colónias bacterianas resultantes foram utilizadas para inocular

culturas líquidas e foram sequenciadas ou manualmente, ou por um sequenciador ABI automatizado.

Sequenciação aleatória de cADNs de uma semente de borragem 12 (DPP) de uma biblioteca polissomal ligada à membrana

Cada cADN correspondente a uma placa não hibridizante foi sequenciado uma vez e foi gerada uma etiqueta de sequência de 200-300 pares de bases. Toda a sequenciação foi efectuada por sequenciação em ciclos (Epicentre). Foram geradas acima de 300 etiquetas de sequências (ESTs). Cada etiqueta de sequência foi comparada com a base de dados GenBank utilizando o algoritmo BLAST (Altschul et al. 1990 J. Mol. Biol. 215:403-410). Foram identificados vários genes do metabolismo lipídico, incluindo o gene da $\Delta 6$ -dessaturase.

Pesquisas a base de dados com o clone de cADN designado mbp-65 utilizando BLASTX com a base de dados GenBank resultou uma correspondência significativa com a $\Delta 6$ -dessaturase da *Synechocystis* anteriormente isolado. Contudo, foi determinado, que mbp-65 não era um cADN de comprimento completo. Um cADN de comprimento completo foi isolado utilizando mbp-65 para rastrear a biblioteca polissomal ligada à membrana da borragem. O clone resultante foi designado pAN1 e a inserção de cADN de pAN1 foi sequenciada pelo método de sequenciação em ciclos. A sequência de aminoácidos deduzida a partir da estrutura de leitura aberta (Fig. 1 SEQ ID N°:1) foi comparada com outras dessaturases desconhecidas utilizando o programa de alinhamento de proteínas Geneworks (IntelligGenetics). Este alinhamento indicou que a inserção de cADN do pAN1 foi o gene da $\Delta 6$ -dessaturase da borragem.

O dendograma resultante (Figura 2) mostra que a Δ^{15} -dessaturase e a Δ^{12} -dessaturase compreende dois grupos. A

nova sequência de borragem isolada e a anteriormente isolada Δ^6 -dessaturase da *Synechocystis* (patente de invenção US N° 5,552,306) formado a partir de um terceiro grupo distinto. A comparação de motivos de aminoácidos comuns a dessaturases e que se pensa estarem envolvidos cataliticamente na ligação a metal ilustra a semelhança global da proteína codificada pelo gene da borragem em relação a dessaturases em geral e à Δ^6 -dessaturase da *Synechocystis* em particular (Tabela 1). Simultaneamente a comparação dos motivos na Tabela 1 indica diferenças definidas entre esta proteína e outras dessaturases de plantas. Além disso a sequência da borragem é também distinguida das dessaturases de ácidos gordos conhecidas associadas à membrana das plantas pela presença de um motivo de ligação ao hemo conservado nas proteínas do citocromo b_5 (Schmidt et al. 1994 Plant. Mol. Biol. 26:631-642) (Figure 1). Assim, enquanto que estes resultados sugerem claramente que o cADN isolado era um gene da Δ^6 -dessaturase da borragem, foi necessária confirmação adicional. Para confirmar a identidade do cADN da Δ^6 -dessaturase da borragem, a inserção do cADN a partir de pAN1 foi clonada numa cassette de expressão para uma expressão estável. O vector pBI121 (Jefferson et al. 1987 EMBO J. 6:3901-3907) foi preparado para ligação por digestão com BamHI e EcoICR I (um isoesquisómero de SacI que deixa extremidades rombas; disponível na Promega) que faz a excisão da região que codifica para GUS deixando o promotor 35S e o terminador NOS intactos. O cADN da Δ^6 -dessaturase da borragem foi excisado a partir do plasmídeo recombinante (pAN1) por digestão com BamHI e XhoI. O terminal XhoI foi tornado rombo efectuando uma reacção de enchimento catalisada pelo fragmento Klenow da ADN polimerase I. Este fragmento foi então clonado nos locais BamHI/EcoICR I do pBI121.1, resultando no plasmídeo pAN2.

Tabela 1

Comparação de motivos de aminoácidos comuns em dessaturases
ligadas à membrana

Dessaturase	Caixa lipídica	Caixa metálica 1	Caixa metálica 2
Borragem Δ^6	WIGHDAGH (SEQ. ID. NO: 15)	HNAHH (SEQ. ID. NO:21)	FQIEHH (SEQ. ID. NO:29)
Synechocystis Δ^6	NVGH DANH (SEQ. ID. NO: 16)	HNYLHH (SEQ. ID. NO:22)	HQVTHH (SEQ. ID. NO:30)
Cloroplasto Arab. Δ^{15}	VLGHDCGH (SEQ. ID. NO: 17)	HRTHH (SEQ. ID. NO:23)	HVIHH (SEQ. ID. NO:31)
Arroz Δ^{15}	VLGHDCGH (SEQ. ID. NO: 17)	HRTHH (SEQ. ID. NO:23)	HVIHH (SEQ. ID. NO:31)
Cloroplasto de Glicina Δ^{15}	VLGHDCGH (SEQ. ID. NO: 17)	HRTHH (SEQ. ID. NO:23)	HVIHH (SEQ. ID. NO:31)
Arab. fad3 (Δ^{15})	VLGHDCGH (SEQ. ID. NO: 17)	HRTHH (SEQ. ID. NO:23)	HVIHH (SEQ. ID. NO:31)
Brassica fad3 (Δ^{15})	VLGHDCGH (SEQ. ID. NO: 17)	HRTHH (SEQ. ID. NO:23)	HVIHH (SEQ. ID. NO:31)
Borragem Δ^{12} (P1-81) *	VIAHECGH (SEQ. ID. NO: 18)	HRRHH (SEQ. ID. NO:24)	HVAHH (SEQ. ID. NO:32)
Arab. fad2 (Δ^{12})	VIAHECGH (SEQ. ID. NO: 18)	HRRHH (SEQ. ID. NO:24)	HVAHH (SEQ. ID. NO:32)
cloroplasto Arab. Δ^{12}	VIGEDCAH (SEQ. ID. NO: 19)	HDRHH (SEQ. ID. NO:25)	HIPHH (SEQ. ID. NO:33)
Glicina plastid Δ^{12}	VIGHDCAH (SEQ. ID. NO: 19)	HDRHH (SEQ. ID. NO:25)	HIPHH (SEQ. ID. NO:33)
Espinafre platidial n-6	VIGHDCAE (SEQ. ID. NO: 19)	HDQHH (SEQ. ID. NO:26)	HIPHH (SEQ. ID. NO:33)
Synechocystis Δ^{12}	VVGHDCGH (SEQ. ID. NO: 20)	HDHHH (SEQ. ID. NO:27)	HIPHH (SEQ. ID. NO:33)
Anabaena Δ^{12}	VLGHDCGH (SEQ. ID. NO: 17)	HNHHH (SEQ. ID. NO:28)	HVPHH (SEQ. ID. NO:34)
*P1-81 é um cADN de comprimento completo que foi identificado por análise de EST e apresenta uma elevada semelhança com a dessaturase Δ^{12} da Arabidopsis (fad2)			

Exemplo 3

Preparação de plantas transgênicas e preparação e análises de ésteres metílicos de ácidos gordos (FAMES)

O plasmídeo de expressão pAN2 foi utilizado para transformar o tabaco (*Nicotiana tabacum*, *xanthi*) via

Agrobacterium tumefaciens de acordo com procedimentos padrão (Horsch, et al. 1985 Science 227:1229-1231; Bogue et al. 1990 Mol. Gen. Genet, 221:49-57) com excepção de que os transformantes iniciais foram seleccionados em 100 µg/ml de canamicina.

Os tecidos de plantas transgênicas foram congelados em azoto líquido e liofilizados de um dia para o outro. FAMES foram preparados de acordo com o descrito por Dahmer et al. (1989) J. Amer. Oil. Chem. Soc. 66:543-548. Nalguns casos o solvente foi evaporado novamente e os FAMES foram novamente suspensos em acetato de etilo e extraídos uma vez com água desionizada para remover quaisquer contaminantes solúveis em água. Os FAMES foram analisados utilizando um cromatógrafo gás-líquido Tracor-560 como anteriormente descrito por (Reddy et al. 1996 Nature Biotech. 14:639-642).

Como indicado na Fig. 3, folhas de tabaco transgênicas contendo cADN de borragem produzida tanto por GLA e ácido octatetraenoico (OTA) (18:4 Δ 6, 9,12, 15). Estes resultados demonstram assim que o cADN isolado codifica para uma Δ 6-dessaturase de borragem.

Exemplo 4

Expressão de Δ 6-dessaturase na borragem

A expressão nativa de Δ 6-dessaturase foi examinada por análise de Northern do ARN derivado de tecidos da borragem. O ARN foi isolado a partir de embriões de borragem em desenvolvimento seguindo o método de Chang et al. 1993 Plant Mol. Biol. Rep. 11:113-116. O ARN foi separado por electroforese em geles de formaldeído-agarose, transferido para membranas de Nylon por transferência capilar, e imobilizado cozendo a 80°C durante 30 minutos seguindo

protocolos padrão (Brown T., 1996 em Current Protocols in Molecular Biology, eds. Auselbel, et al [Greene Publishing e Wiley-Interscience, Nova Iorque] pág. 4.9. 1-4.9.14). Os filtros foram pré-incubados a 42°C numa solução contendo 50% de formamida desionizada, 5X reagente de Denhardt, 5X SSPE (NaCl 900 mM; fosfato de sódio 5,0 mM, pH 7,7; e EDTA 5 mM), SDS 0,1%, e 200 µg/ml de ADN de esperma de salmão desnaturado. Após duas horas, os filtros foram adicionados a uma solução fresca da mesma composição com a adição de sonda de hibridização radioactiva desnaturada. Neste caso, as sondas utilizadas foram cADN de legumina de borragem (Fig.16A), cADN de borragem de oleoresina (Fig. 16B), e cADN de $\Delta 6$ -dessaturase de borragem (pAN1, Exemplo 2) (Fig. 16C). A legumina de borragem e cADNs de oleoresina foram isolados por clonagem de EST e identificada por comparação com a base de dados da GenBank utilizando o algoritmo BLAST como descrito no Exemplo 2. A variação da carga foi corrigida por normalização para níveis de borragem EF1 α mARN. EF1 α mARN foi identificado correlacionando o cADN correspondente obtido por análise de EST descrita no Exemplo 2. Os filtros foram hibridizados a 42°C durante 2-20 horas, depois lavados como descrito acima (com excepção do facto da temperatura ser 65°C), secos ao ar, e expostos a um filme de raios-X.

Como representado nas Fig. 15A e 15B, a $\Delta 6$ -dessaturase é expressa basicamente na semente de borragem. As sementes de borragem atingem a maturação entre os dias 18-20 após polinização (dpp). A expressão de mARN de $\Delta 6$ -dessaturase ocorre ao longo dos intervalos de tempo recolhidos (8-20 dpp), mas parece ser máximo de 10-16 dias após polinização. Este perfil de expressão é semelhante ao observado para a oleoresina da borragem e mARNs da proteína 12S de armazenamento na semente (Figs. 16A, 16B, e 16 C).

Exemplo 5

Isolamento e caracterização de um novo cADN da oleoresina

O cADN da oleoresina (AtS21) foi isolado por rastreio de uma biblioteca de cADN por subtração virtual de uma semente em desenvolvimento de *Arabidopsis* utilizando uma sonda de cADN de reacção em cadeia da polimerase iniciada aleatoriamente (PC-PCR) do tecido da raiz.

Preparação de ARN

Foram cultivadas plantas de *Arabidopsis thaliana* Landsberg erecta sob iluminação contínua numa mistura de vermiculite/solo à temperatura ambiente (22°C). Foram dissecados silíquas 2-5 dias após florescência para recolher separadamente sementes em desenvolvimento e tegumentos de silíqua. Foram também colhidos botões de flores iniciais contendo inflorescências e flores completamente abertas, folhas e silíquas completas um ou três dias após florescência. Foram obtidas raízes de plântulas que tinham sido feitas crescer em meio líquido Gamborg B₅ (GIBCO BRL) durante duas semanas. As sementes para cultura de raiz foram esterilizadas previamente com 50% de branqueador durante cinco minutos e lavadas extensivamente com água. Todos os tecidos foram congelado em azoto líquido e armazenados a -80°C até serem usados. Os ARNs totais foram isolados seguindo um protocolo com uma extracção a quente fenol/SDS e precipitação com LiCl (Harris et al. 1978 Biochem. 17:3251-3256; Galau et al. 1981 J. Biol. Chem. 256:2551-2560). Poly A + ARN foi isolado utilizando cromatografia em coluna oligo dT de acordo com os protocolos dos fabricantes (Pharmacia ou

Stratagene) ou utilizando partículas de latex oligotex dT (Qiagen).

Construção de bibliotecas de ADN específicas de tecido

Bibliotecas de cADN de flores, de siliqua de um dia, de siliqua de três dias, de folhas, raízes e sementes em desenvolvimento, cada uma foi construída utilizando 5 μ de poly A+RN utilizando o kit de síntese ZAP cADN (Stratagene). Os cADNs foram clonados direccionalmente nos locais EcoRI e XhoI de pB1bluescript SK(-) no vector χ -ZAPII (Short et al. 1988 Nucleic Acids Res. 16:7583-7600). As placas de fagos não recombinantes foram identificadas pelo desenvolvimento da cor azul em placas NZY contendo X-gal (5-bromo-4-cloro-3-indoil- β -D-galactopiranosido) e IPTG (isopropil-1-tio- β -D-galactopiranosido). As bases não recombinantes para as bibliotecas de cADN da flor, siliqua de um dia, siliqua de três dias, folha, raiz e semente em desenvolvimento foram 2,8%, 2,3%, 3,3%, 6,5%, 2,5%, e 1,9% respectivamente.

Marcação de ADN por iniciação aleatória

As inserções de cADN de clones isolados (cADNs não hibridizados) foram excisados por dupla digestão por EcoRI/XhoI e purificadas por gel para marcação por iniciação aleatória. A mistura da reacção de Klenow continha 50 ng de moldes de ADN, Tris-HCl 10 mM, pH 7,5, MgCl₂ 5 mM, DTT 7,5 mM e 50 μ M de cada dCTP, dGTP, e dTTP, 10 μ M de iniciadores aleatórios de hexamer (Boehringer Mannheim), 50 μ Ci α -³²P-dATP, 3000 Ci/mmol, 10mCi/ml (DuPont), e 5 unidades de fragmento de Klenow de ADN polimerase (New England Biolabs). As reacções foram

efectuadas a 37°C durante uma hora. Foram utilizadas aliquotas das misturas reacionais diluídas para precipitação de TCA e análise alcalina desnaturante em gel. As sondas de hibridização foram marcadas apenas com ADN de Klenow polimerase e os dNTPs não incorporados foram removidos utilizando colunas giratórias Sephadex R G-50 (Boehringer Mannheim).

PCR iniciada aleatoriamente

Foi sintetizado cADN de cadeia dupla a partir de poly A+ ARN isolado a partir de tecido de *Arabidopsis* utilizando o sistema de síntese de cADN (Gibco BRL) com oligo dT12-18 como iniciadores. Os cADNs mais longos do que 300 pb foram enriquecidos por cromatografia Sephacryl S-400 (Stratagene). Os cADNs fraccionados foram utilizados como moldes para marcação por RP-PCR. A reacção continha Tris HCl 10 mM, pH 9,0, KCl 50 mM, Triton X-100 0,1%, 2 mM. MgCl₂, 5 unidades de Taq ADN polimerase (PROMEGA), 200 µM de dCTP, cGTP, e dTTP e diferentes concentrações de iniciadores aleatórios de hexamero α-32P dATP, 800 mCi/mole, 10 mCi/ml (DuPont), e dATP frio num volume final de 25 µl. Após 5 minutos iniciais a 95°C, foram corridas diferentes reacções com diferentes programas para otimizar as condições do cADN de RP-PCR. A não ser que indicado de outro modo, o programa seguinte foi utilizado para a maior parte das marcações de sonda cADN RP-PCR: 95°C/ 5 minutos, depois 40 ciclos de 95°C durante 30 segundos, 18°C/ 1 segundo, rampa a 30°C a uma velocidade de 0,1°C/segundo. 72°C/ 1 minuto. Produtos da RP-PCR foram extraídos com fenol/clorofórmio e precipitados com etanol, ou purificados fazendo passar através colunas giratórias Sephadex G-50 (Boehringer Mannheim).

Transferência do clone por subtração virtual

A excisão de massa χ -ZAP de bibliotecas de cADN foi efectuada por co-infecção de células hospedeiras XL1-Blue MRF com fago recombinante das bibliotecas e do fago auxiliar ExAssist (STRATAGENE). Fagemídeos excisados foram recuperados por células SOLR. Foram preparados ADNs de plasmídeo através do método de ebulição mini-prep (Holmes et al. 1981 Anal. Biochem. 114:193-197) a partir de clones isolados aleatoriamente. As inserções de cADN foram excisadas por dupla digestão por EcoRI e XhoI, e resolvidas em geles de 1% de agarose. Os ADNs foram desnaturados em NaOH 0,5 N e NaCl 1,5 M durante 45 minutos, neutralizados em Tris-HCl 0,5 M, pH 8,0, e NaCl 1,5 M durante 45 minutos e depois transferidas para membranas de nylon (Micron Separations, Inc.) em 10X SSC de um dia para o outro. Após uma hora de pré-hibridização a 65°C, uma sonda de raiz RP-cADN foi adicionada ao mesmo tampão de hibridização contendo 1% da fracção V de albumina bovina (Sigma), EDTA 1 mM, NaHPO₄ 0,5 M, pH 7,2, SDS 7%. A hibridização continuou durante 24 horas a 65°C. Os filtros foram lavados em albumina bovina a 0,5%, EDTA 1 mM, NaHPO₄ 40 mM, pH 7,2, SDS 5% durante 10 minutos à temperatura ambiente, e 3x10 minutos em EDTA 1 mM, NaHPO₄ a 40 mM, pH 7,2, SDS 1% a 65°C. Autoradiografias foram expostas a filmes de raios X (Kodak) durante 2 a cinco dias a -80°C.

A hibridização das transferências resultantes com sondas de RP-PCR de raízes, cADNs de sementes "subtraídas virtualmente" partilhadas com a população de mARN de raiz. Os cADNs de raízes restantes representando cADNs putativos específicos das sementes, incluindo aqueles que codificam para oleoresinas foram sequenciadas pelo método de sequenciação por ciclos identificando deste modo AtS21 como um clone cADN da oleoresina.

Análise da sequência de AtS21

O cADN da oleoresina tem 834 pb de comprimento incluindo uma cauda poly A de 18 pb de comprimento (Fig. 4, SEQ ID N°:2). Possui uma homologia elevada em relação aos outros genes da oleoresina da *Arabidopsis*, assim como das outras espécies. Recentemente um gene idêntico de oleoresina foi reportado (Zou, et al., 1996, Plant Mol. Biol. 31:429-433). A proteína prevista tem de comprimento 191 aminoácidos com um domínio médio altamente hidrofóbico flanqueado por um domínio hidrofílico de cada lado. A existência de dois codões de paragem a montante na estrutura e a semelhança com outros genes de oleoresina indica que este cADN é de comprimento completo. Uma vez que existem dois codões de paragem na estrutura a montante mesmo a montante do primeiro ATG, este cADN é considerado como sendo um cADN de comprimento completo (Figura 4, SEQ ID N°: 2). A proteína prevista possui três domínios distintos baseados na distribuição dos seus resíduos de aminoácidos. Tanto os domínios do terminal N e do terminal C são ricos em resíduos carregados enquanto que o domínio central é completamente hidrofóbico (Figura 5). No domínio central encontram-se localizados 20 resíduos de leucina e dispostos como repetições com uma leucina em cada 7-10 resíduos. Outros resíduos de aminoácidos não polares encontram-se também agrupados no domínio central tornando este domínio absolutamente hidrofóbico (Figura 6).

Pesquisas extensivas de diferentes bases de dados utilizando tanto cADN de AtS21 e a sua sequência prevista identificou oleoresinas da cenoura, milho, algodão, colza, *Arabidopsis*, e outras espécies de plantas. A homologia encontra-se basicamente restringida ao domínio hidrofóbico central. Foram encontradas sete sequências de oleoresinas

de *Arabidopsis*. AtS21 representa o mesmo gene que Z54164 que tem mais algumas bases na região não traduzida 5'. As sete sequências de *Arabidopsis* de oleoresina disponíveis até agora foram alinhadas umas com as outras (Figura 7). O resultado sugeriu que as sete sequências caem em três grupos. O primeiro grupo inclui AtS21 (SEQ ID N°:5), X91918 (SEQ ID N°:6), e a sequência parcial Z29859 (SEQ ID N°:7).

Uma vez que X91918 (SEQ ID N°: 6) possui apenas o último resíduo diferente de AtS21 (SEQ ID N°:5), e uma vez que Z299859 (SEQ ID N°: 7) possui apenas três resíduos de aminoácidos que são diferentes de AtS21 (SEQ ID N°:5) todas as três sequências representam provavelmente o mesmo gene. As duas sequências do segundo grupo, X62352 (SEQ ID N°:8) e Atol3 (SEQ ID N°:9), são diferentes tanto em sequência e em comprimento. Assim, não há dúvida que representam dois genes independentes. Tal como no primeiro grupo, as duas sequências do terceiro grupo, X91956 (SEQ ID N°:10) e L40954 (SEQ ID N°:11) também possuem apenas três resíduos divergentes que podem ser devidos a erros na sequência. Assim, X91956 (SEQ ID N°:10) e L40954 (SEQ ID N°: 11) representam provavelmente o mesmo gene. Ao contrário de todas as outras sequências de oleoresinas que foram previstas a partir das sequências de cADN, X62352 (SEQ ID N°:8) foi deduzida a partir de uma sequência genômica (Van Rooijen et al. 1992 Plant Mol. Biol. 18: 1177-1179). Concluindo, até agora foram identificados quatro genes diferentes de oleoresina de *Arabidopsis* e são apenas conservados no meio do centro hidrofóbico.

Análise de Northern

De modo a caracterizar o padrão de expressão do gene nativo AtS21, foi efectuada uma análise de Northern como descrito no Exemplo 4 com excepção de que a sonda foi cADN

de AtS21 (inserção pAN1) marcada com ^{32}P -dATP para uma actividade específica de 5×10^8 cpm/ μg .

Os resultados indicaram que o gene AtS21 é fortemente expresso em sementes em desenvolvimento e fracamente expresso em tegumentos de siliqua (Figura 8A). Um transcrito muito maior pode representar AtS21 não processado pré-ARNm, também foi detectado em ARN de sementes em desenvolvimento, AtS21 não foi detectado na flores, folhas raízes (Figura 8A), ou em ARNs de siliqua de um dia. Uma análise de Northern diferente revelou que AtS21 é também fortemente expresso em sementes em inibição germinativa (Figs. 13A e 13B).

Exemplo 6

Caracterização dos clones genómicos da oleoresina e isolamento do promotor da oleoresina

Os clones genómicos foram isolados por rastreio de uma biblioteca de ADN genómico de *Arabidopsis* utilizando o cADN de comprimento completo (AtS21) como sonda. Foram mapeados dois clones genómicos por digestão com enzima de restrição seguido por hibridização de Southern utilizando a metade 5' do cADN cindido por SacI como sonda. Um fragmento SacI de 2 kb foi subclonado e sequenciado (Fig. 9, SEQ ID N°:35). As duas regiões do clone genómico são idênticas à sequência de cADN. Um intrão de 395 pb separa as duas regiões.

O número de cópias do gene AtS21 no genoma de *Arabidopsis* foi determinado por hibridização de Southern de ADN genómico seguindo-se digestão com as enzimas BamHI, EcoRI, HindIII, SacI e XbaI utilizando o cADN de comprimento completo como sonda (Figura 8B). Foi detectada uma banda única em todas as zonas, excepto na digestão com SacI em que foram detectadas duas bandas. Uma vez que a

sonda cADN possui um local interno SacI, estes resultados indicam que a AtS21 é um simples gene de cópia no genoma de *Arabidopsis*. Uma vez que é conhecido que o genoma de *Arabidopsis* contém isoformas diferentes dos genes de oleoresina esta análise de Southern também demonstra que diferentes isoformas de oleoresina de *Arabidopsis* são divergentes ao nível da sequência de ADN.

As duas regiões separadas por um intrão de 395 pb do fragmento de ADN genómico são idênticos à sequência do cADN de AtS21. Pesquisas nas bases de dados utilizando a sequência promotora 5' a montante da sequência de cADN de AtS21 não identificou, qualquer sequência com homologia significativa. Além disso, a comparação da sequência promotora de AtS21 com outro promotor da oleoresina de *Arabidopsis* isolado anteriormente (Van Rooijen, et al., 1992) revelou pouca semelhança. A sequência promotora de AtS21 é rica em bases A/T, e contém 44 repetições directas que vão de 10 pb a 14 pb sendo permitido apenas um desemparelhamento

Duas repetições directas de 14 pb e um elemento de resposta putativo ABA encontram-se sublinhados na Figura 9.

Exemplo 7

Construção da cassette de expressão de Promotor AtS21/GUS e padrões de expressão em *Arabidopsis* transgénica e no tabaco

Construção da cassette de expressão de promotor AtS21/GUS

O fragmento de promotor de 1267 pb que começa na primeira G a montante do codão ATG do fragmento de ADN genómico foi amplificado utilizando PCR e foi fusionado com o gene reporter GUS para ser analisado, quanto à sua actividade. O fragmento promotor do clone genómico de AtS21 foi

amplificado por PCR utilizando o iniciador T7 GTAATACGACTCACTATAGGGC (SEQ ID N°: 13) e o iniciador 21P GGGGATCCTATACTAAACTATAGAGTAAAGG (SEQ ID N°: 14) complementar à região 5' não traduzida a montante do primeiro codão ATG (Figura 9). Um local de clonagem BamHI foi introduzido pelo iniciador 21P. O fragmento amplificado foi clonado nos locais BamHI e SacI de pBluescript KS (Stratagene). Foram sequenciados os clones individuais para avaliar possíveis mutações PCR, assim como orientação das suas inserções. O clone correcto foi digerido com BamHI e HindIII, e o fragmento do promotor excisado (1,3 kb) foi clonado nos locais correspondentes de pBI101.1 (Jefferson, R. A 1987a, Plant Mol. Biol. Rep. 5:387-405; Jefferson et al., 1987b, EMBO J. 6:3901-3907) a montante do gene GUS. O plasmídeo resultante foi designado pAN5 (Fig. 10). A construção do promotor AtS21/GUS (pAN5) foi introduzida tanto no tabaco (pelo método do disco da folha, Horsch et al., 1985; Bogue et al. 1990 Mol. Gen. Gen. 221:49-57) e no ecotipo de *Arabidopsis* Colombia através de infiltração por vácuo como descrito por Bectold, et al. (1993) C. R. Acad. Sci. Paris, 316:1194-1199. As sementes foram esterilizadas e seleccionadas em meio contendo 50 µg/ml de canamicina, 500 µg/ml de carbenicilina.

Ensaio de actividade GUS: padrões de expressão do gene reporter GUS foram revelados por coloração histoquímica (Jefferson, et al., 1987a, Plant. Mol. Biol. Rep. 5:387-405). Foram corados diferentes tecidos numa solução de substrato contendo 2 mg/ml de ácido 5-bromo-4-cloro-3-indolil- β -D-glucurónico (X-Gluc) (Research Organics, Inc.) ferrocianeto de potássio 0,5 mM, e cianeto férrico de potássio 0,5 mM em tampão fosfato de sódio 50 mM, pH 7,0 a 37°C de um dia para o outro e depois desidratado sucessivamente em 20%, 40% e 80% de etanol (Jefferson et

al., 1987). Foram tiradas fotografias utilizando um microscópio composto Axiophot (Zeiss) ou microscópio de dissecação Olympus SZH10. Os slides foram convertidos em imagens digitais utilizando um scanner de slides Spring/Scan 35LE (Polaroid) e compiladas utilizando Adobe PhotoshopTM 3.0.5 e CanvasTM 3.5.

As actividades GUS foram medidas quantitativamente medidas por fluorometria utilizando 4-MUG 2 mM (4-metilumbefiril- β -D-glucuronide) como substrato (Jefferson, et al., 1987). Sementes em desenvolvimento de *Arabidopsis* foram classificadas de acordo com as suas cores, e outros tecidos de plantas foram recolhidos e mantidos a -80°C até serem utilizados. Os tecidos de plantas foram moídos em tampão de extracção contendo fosfato de sódio 50 mM, pH 7,0, EDTA 10 mM, β -mercaptoetanol 10 mM, Triton-X-100 0,1%, e lauril sarcosina de sódio. Os detritos dos tecidos foram removidos através de uma centrifugação de 5 minutos numa microcentrífuga. Foram retiradas aliquotas do sobrenadante e misturadas com substrato e incubadas a 37°C durante 1 hora. Foram testadas três réplicas para cada amostra. As reacções foram paradas através da adição de 4 volumes de carbonato de sódio 0,2 M. A fluorescência foi lida utilizando um fluorómetro de ADN TKO-100 (Hoefer Scientific Instruments). Foram determinadas as concentrações de proteína dos extractos pelo método de Bradford (Bio Rad).

Padrões de expressão de promotor AtS21/GUS em *Arabidopsis* e tabaco transgénicos

Na *Arabidopsis*, foi detectada actividade GUS em sementes verdes e em regiões nodulares em que siliquas, folhas caulinas e ramos se juntam no caule da inflorescência (Figuras 11A e 11B). Não foi detectada nenhuma actividade GUS em qualquer folha, raiz, ou flor,

tegumento de siliqua, ou nas regiões internodulares do caule de inflorescência. Estudos detalhados da expressão GUS nas sementes em desenvolvimento revelaram que o promotor AtS21 era apenas activo em sementes verdes em que os embriões já se tinham desenvolvido para lá da fase coração (Figuras 11C e 11G). Os embriões mais novos apresentando actividade GUS que poderia ser detectada por coloração histoquímica estavam na fase precoce de torpedo. É interessante que a coloração se encontrava apenas restringida à parte inferior do embrião incluindo hipocotilédone e radical embriónico. Não foi detectada coloração nos cotilédones novos (Figuras 11D e 11E). Os cotilédones começaram a ser corados quando os embriões se encontravam na fase tardia de torpedo, ou mesmo na fase de cotilédone precoce (Figura 11F e 11H). Mais tarde, os embriões completos estavam corados e a coloração tornou-se mais intensa à medida que os embriões iam maturando (Figuras 11I e 11J). É também observado que a expressão genética GUS era restrita aos embriões. O tegumento da semente e o endosperma novo não se encontravam corados (Figura 11C).

A actividade GUS foi também detectada em plântulas em desenvolvimento. Plântulas novas de 3-5 dias estavam coradas em todo o lado. Apesar de alguns dos capilares da raiz perto do hipocotiledone estarem corados (Figura 11K) a maior parte das estruturas recentemente formadas, tais como capilares de raízes, primordia laterais da raiz, raízes apicais não foram coradas (Figuras 11L e 11N). Mais tarde a coloração era restrita aos cotilédones e hipocotiledones quando as raízes laterais cresceram a partir da raiz embriónica em enlongação. A coloração em raízes embriónicas desapareceu. Não foi observada coloração em raízes laterais recentemente formadas, nem em folhas verdadeiras nem em tricomas em folhas verdadeiras (Figuras 11M e 11N).

Padrões de expressão de promotor AtS21/GUS no tabaco são basicamente as mesmas que em *Arabidopsis*. A actividade GUS foi apenas detectada em sementes de fase tardia em diferentes regiões nodulares de plantas na maturidade. Nas sementes em germinação uma forte coloração foi detectada através de todo o embrião logo após uma hora após terem sido dissecados de sementes embebidas. O endosperma maduro que as sementes de *Arabidopsis* não possuem, mas não o tegumento da semente encontravam-se também corados (Figura 12A). As pontas da raiz de algumas plântulas novas de uma linha transgénica não estavam coradas (Figura 12B). Por outro lado, os padrões de expressão GUS em plântulas em desenvolvimento do tabaco eram as mesmas que em sementes de *Arabidopsis* (Figuras 12B, 12C, e 12D). Estruturas recentemente formadas, tais como raízes laterais e folhas verdadeiras não estavam coradas.

Níveis de mARN de AtS21 em plântulas em desenvolvimento

Uma vez que as fortes actividades observadas de promotor AtS21 /GUS tanto nas plântulas de *Arabidopsis* e do tabaco não são consistentes com a expressão específica da semente dos genes da oleoresina, a análise de Northern foi efectuada para determinar se mARN de AtS21 estava presente nas sementes em desenvolvimento em que a actividade GUS era muito forte. Os ARNs preparados a partir de plântulas em fases diferentes desde 24 horas a 12 dias foram analisadas por hibridização de Northern utilizando cADN de AtS21 como sonda. Surpreendentemente, mARN de AtS21 foram detectadas a um nível elevado comparado ao das sementes em desenvolvimento 24-48 horas em sementes embebidas. O nível de mARN caiu dramaticamente quando as plântulas novas emergiram pela primeira vez às 74 horas (Figuras 13A e 13B). Em plântulas com 96 horas e com mais tempo não foi

detectado nenhum sinal, mesmo com uma exposição mais longa (Figura 13 B). As cargas de amostras de ARN foram verificadas hibridizando a mesma transferência com uma sonda de gene de tubulina (Figura 13C) que foi isolada e identificada por análise de EST como descrito no Exemplo 2. Uma vez que mRNA de AtS21 era tão abundantes em sementes, as sondas de AtS21 residual permaneceram na transferência mesmo após uma remoção extensiva. Estes resultados indicaram que o mRNA de AtS21 detectado em sementes embebidas e em plântulas muito novas resultam do arrastamento de mRNA de AtS21 de sementes secas. Foi reportado recentemente que um mRNA de Atol2 de oleoresina (idêntico a AtS21) é mais abundante em sementes secas (Kirik et al., 1996 Plant. Mol. Biol. 31(2): 413-417). De forma semelhante, actividades GUS em plântulas eram muito provavelmente devidas ao transporte tanto da proteína da β -glucuronidase e da síntese de novo da β -glucuronidase a partir do seu mRNA transportado a partir da fase de semente seca.

Exemplo 8

Comparação de actividade entre o promotor de AtS21 e o promotor 35S

As actividades GUS nas sementes em desenvolvimento da *Arabidopsis* transgénica expressas pelo promotor AtS21 foram comparadas com aquelas expressas pelo promotor 35S na construção pBI221 (Jefferson et al. EMBO J. 6:3901-3907). As sementes foram classificadas de acordo com as suas cores (Tabela 2). A fase mais precoce foi da fase globular à fase de coração tardia, quando as sementes ainda eram brancas, mas suficientemente largas para serem dissecadas das siliquas. Actividade do promotor AtS21 foi detectada a um

nível cerca de três vezes mais baixo do que a do promotor 35S nesta fase. A actividade do promotor 35S permaneceu no mesmo baixo nível ao longo de todo o desenvolvimento do embrião. Em contraste, a actividade promotora AtS21 aumentou rapidamente à medida que o embrião passou da fase de torpedo e atingiu o nível mais elevado de 25,25 pmole 4-UM/min.µg de proteína numa fase madura (Figura 5-8). A actividade do pico do promotor AtS21 é 210 vezes mais elevada do que a sua actividade mais baixa na fase globular até à fase coração, e está perto de 100 vezes mais elevada do que a actividade do promotor 35S na mesma fase (Tabela 2). Os níveis de actividade do promotor AtS21 são semelhantes aquelas de um outro promotor da oleoresina da *Arabidopsis* expresso em *Brassica napus* (Plant et al. 1994, Plant Mol. Biol. 25:193-205. Actividade do promotor AtS21 foi também detectada a nível de base nas folhas. O elevado desvio padrão, mais elevado do que a própria média indicou que a actividade GUS foi apenas detectada nas folhas de algumas linhas (Tabela 2). Por outro lado, a actividade do promotor 35S na folha era mais do que 20 vezes mais elevada do que na semente. As comparações lado a lado de actividades entre o promotor AtS21 e o promotor 35S são indicadas na Figura 14.

Apesar da actividade promotora de AtS21 ser cerca de 3 vezes inferior na semente seca de tabaco do que na semente seca de *Arabidopsis*, a actividade absoluta GUS era ainda mais elevada do que a expressa pelo promotor 35S na folha de *Arabidopsis* (Tabela 2). Não foi observada actividade promotora de AtS21 detectável na folha de tabaco (Figura 14).

A comparação do promotor AtS21 em relação ao promotor 35S revelou que o último não é um bom promotor para expressar genes em níveis elevados em sementes em desenvolvimento. Devido às suas baixas actividades

consistentes ao longo de todo o período de desenvolvimento do embrião, o promotor 35S é útil para uma expressão consistente de baixo nível dos genes alvo. Por outro lado, o promotor AtS21 é um promotor muito forte que pode ser utilizado para expressar genes começando em embriões na fase de coração e acumulando até à fase de semente seca. O promotor 35S apesar de não ser eficiente é melhor do que o promotor AtS21 em expressar genes em embriões antes da fase coração.

Tabela 2

**Actividades GUS de construções promotores AtS21 e 35S
/GUS**

Fase da cor	Branco G-H	Branco/Amarelo H-T	Amarelo T-C	Verde claro precoce C	Verde escuro tardio C	Verde acastanhado
AtS21	0,12±0,17	1,35±1,57	6,77±1,25	18,99±3,75	21,85±4,45	25,25±4
35S	0,30±0,06	0,25±0,08	0,29±0,04	0,28±0,03	0,33±0,06	0,26±0,0
AtS21 (em tabaco)						

Abreviaturas: G fase globular; H fase coração; T, fase de torpedão; C, fase de cotilédono. As actividades GUS encontram-se em pmole os números são a média de cinco linhas independentes com desvios padrão. Foram ensaiadas três repetições para cada linha. Para três repetições da mesma linha com desvios padrão.

Exemplo 9

Expressão do gene da Δ^6 -dessaturase da borragem sob o controlo do promotor AtS21 e comparação com a expressão sob o controlo do promotor CaMV 35S

Para criar uma construção de expressão com o promotor de AtS21 conduzindo a expressão do gene da Δ^6 -dessaturase da borragem, o fragmento que codifica para GUS a partir de

pAN5 foi removido por digestão com SmaI e EcoICR I. A inserção de cADN de pAN1 (Exemplo 2) foi então excisada por digestão com XhoI (é preenchendo a saliência residual como em cima), e depois digerindo com SmaI. O fragmento resultante foi utilizado para substituir a porção excisada de pAN5 para dar pAN3.

Após a transformação do tabaco e da *Arabidopsis* seguindo os métodos do Exemplo 7, os níveis de actividade da Δ^6 -dessaturase foram monitorizados testando os ésteres metílicos dos ácidos gordos correspondentes dos seus produtos de reacção, ácido γ -linolénico (GLA) e ácido octadecatetranoico (OTA) utilizando os métodos referidos no Exemplo 3. Os níveis de GLA e OTA (Tabela 3) das sementes transgénicas foram de 6,7% de ácidos gordos C18 (média=3,1%) e 2,8% (média=1,1%), respectivamente. Não foi detectado GLA ou OTA nas folhas destas plantas. Em comparação, as plantas transgénicas promotor CaMV 35S/ Δ^6 -dessaturase produziram níveis de GLA nas sementes que vão até 3,1% de ácidos gordos C18 (média=1,3%) e nenhum OTA quantificável em sementes.

Tabela 3

Expressão Δ^6-dessaturase da borragem em plantas transgénicas				
Promotor	Planta	Semente		Folha
		GLA* Gama	OTA* Gama	GLA* Gama OTA* Gama
Vírus do mosaico da couve flor 35S	Tabaco	1,3	0,7-3,1 n.d	20 19-22 9,7 8-11
Oleoresina da <i>Arabidopsis</i>	<i>Arabidopsis</i>	3,1	0-6,7 1,1 0-2,8	n.d. n.d.
* valor médio expresso como percentagem dos ácidos gordos C18 n.d. não detectado				

Exemplo 10

Transformação da colza com uma cassette de expressão que compreende a região reguladora 5' da oleoresina ligada ao gene da delta 6-dessaturase da borragem

A colza, Cv. Westar foi transformada com a estirpe de *Agrobacterium tumefaciens* EHA105 contendo o plasmídeo pAN3 (i.e. o gene da $\Delta 6$ -dessaturase da borragem sob o controlo do promotor da oleoresina da *Arabidopsis* -Exemplo 9).

Os internodos terminais de Westar foram co-cultivados durante 2-3 dias com a estirpe induzida de *Agrobacterium tumefaciens* EHA105 (Alt-Moerbe et al. 1988 Mol. Gen. Genet. 213:1-8; James et al. 1993 Plant Cell Reports 12:559-563), depois foram transferidas para meio de regeneração (Boulter et al. 1990 Plant Science 70:91-99; Fry et al. 1987 Plant Cell Reports 6: 321-325). Os rebentos regenerados foram transferidos para meio de crescimento (Pelletier et al. 1983 Mol. Gen. Menet. 191; 244-250), e um teste da reacção em cadeia da polimerase (PCR) foi efectuado em fragmentos de folhas para avaliar a presença do gene.

O ADN foi isolado a partir de folhas de acordo com o protocolo de KM Haymes et al. (1996) Plant Molecular Biology Reporter 14(3):280-284 e novamente suspenso em 100 μ l de água, sem tratamento com RNase. Foram utilizados 50 μ l de extracto para a reacção PCR num volume final de 50 μ l. A reacção foi efectuada num termociclador Perkin-Elmer 9600 com os ciclos seguintes:

1 ciclo: 95°C, 5 minutos

30 ciclos: 95°C, 45 seg.; 52°C, 45 seg. 72°C, 1 minuto

1 ciclo: 72°C, 5 minutos

e os seguintes iniciadores (derivados das regiões próximas da caixa metálica, como indicado na Fig. 1, SEQ N°:1):

5'TGG AAA TGG AAC CAT AA 3'

5' GGA AAC AAA TGA TGC TC 3'

A amplificação do ADN revelou o fragmento PCR de par de bases esperado 549 (Figura 17).

Os rebentos positivos foram transferidos para o meio de elongação, depois para o meio de enraizamento (DeBlock et al. 1989 Plant Physiol. 91:694-701). Rebentos com um sistema radicalar bem desenvolvido foram transferidos para a estufa. Quando as plantas se encontravam bem desenvolvidas as folhas foram recolhidas para análise de Southern para avaliar o número de cópias do gene.

Foi extraído o ADN genómico de acordo com o procedimento de Bouchez et al. (1996) Plant Molecular Biology Reporter 14:115-123, digerido com as enzimas de restrição *Bgl* I e/ ou *Cla* I, separado electroforeticamente em gel de agarose (Maniatis et al. 1982, em Molecular Cloning; a Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor/NY), e preparado para transferência para membranas de nylon (Nytran membrane, Schleicher & Schuell) de acordo com as instruções do fabricante. O ADN foi então transferido para membranas de um dia para o outro por acção capilar utilizando 20XSSC (Maniatis et al. 1982). A seguir à transferência as membranas foram reticuladas por UV (Stratagene) durante 30 segundos e pré-hibridizadas durante 1 hora a 65°C em 15 ml de uma solução contendo 6XSSC, 0,5% de SDS e 2,25% p/p de leite desnatado desidratado em frascos de vidro num forno de hibridização (Appligene). As membranas foram hibridizadas de um dia para o outro na mesma solução contendo uma sonda de hibridização desnaturada marcada radioactivamente com ³²P até a uma actividade específica de

108 cpm/ μ g através do método de iniciação aleatória (com o kit Ready-ToGo obtido na Pharmacia). A sonda representa um fragmento de PCR do gene da delta 6-dessaturase da borragem (obtido nas condições e com os iniciadores detalhados acima). Após hibridização os filtros foram lavados a 65°C em 2XSSC, 0,1% de SDS durante 15 minutos e 0,2XSSC, 0,1% SDS durante 15 minutos. As membranas foram então envolvidas em Saran-Wrap e expostas a um filme Kodak XAR utilizando um écran intensificador a -70°C numa cassette à prova de luz. O tempo de exposição foi geralmente de 3 dias.

Os resultados obtidos confirmam a presença do gene. De acordo com a construção do gene, o número de bandas em cada zona de ADN digerido por *Bgl* 1 ou *Cla* 1 representa o número de genes de delta 6-dessaturase presentes no ADN genómico da planta. A digestão com *Bgl* 1 e *Cla* 1 em conjunto geram um fragmento de 3435 pb.

O termo "compreende" ou "compreendendo" é definido como especificando a presença das características declaradas, números inteiros, passos, ou componentes referidos nas reivindicações, mas não exclui a presença ou adição de uma ou mais características diferentes, números inteiros, passos, componentes, ou dos seus grupos.

Listagem das Sequências

(1) INFORMAÇÃO GERAL:

(i) REQUERENTE: Rhone Poulenc Agro Thomas, Terry L. Li, Zhongsen

(ii) TÍTULO DA INVENÇÃO: REGIÃO REGULADORA 5' DA OLEORESINA PARA A MODIFICAÇÃO DA COMPOSIÇÃO LIPÍDICA DAS SEMENTES DA PLANTA

(iii) NÚMERO DE SEQUÊNCIAS: 35

(iv) MORADA PARA CORRESPONDÊNCIA:

- (A) DESTINATÁRIO: Scully, Scott, Murphy & Presser
- (B) RUA: 400 Garden City Plaza
- (C) CIDADE: Garden City
- (D) ESTADO: NOVA IORQUE
- (E) PAÍS: USA
- (F) ZIP: 11530

(v) FORMA A SER LIDA PELO COMPUTADOR

- (A) TIPO DE MEIO: Floppy disk
- (B) COMPUTADOR: IBM PC COMPATÍVEL
- (C) SISTEMA OPERATIVO: PC-DOS/MS-DOS
- (D) SOFTWARE: Patentin Release # 1.0, Versão #1.30

(vi) DADOS DO PEDIDO:

- (A) NÚMERO DO PEDIDO: 08/831,575
- (B) DATA DO REGISTO: 9 de Abril de 1997
- (C) CLASSIFICAÇÃO:

(viii) INFORMAÇÃO SOBRE O AGENTE DE PROPRIEDADE INDUSTRIAL

- (A) NOME: DiGiglio, Frank S.
- (B) NÚMERO DO REGISTO: 31,346
- (C) REFERÊNCIA/NÚMERO DE ENDOSSE: 10203

(ix) INFORMAÇÃO SOBRE TELECOMUNICAÇÕES:

- (A) TELEFONE: (516) 742-4343
- (B) TELEFAX: (516) 742-4366

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 1:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 1684 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(ix) CARACTERÍSTICA:

(A) NOME/CHAVE: CDS

(B) LOCALIZAÇÃO: 43..1387

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:1:

```

ATATCTGCCT ACCCTCCCAA AGAGAGTAGT CATTTTTCAT CA ATG GCT GCT CAA
                                     Met Ala Ala Gln
                                     1

ATC AAG AAA TAC APT ACC TCA GAT GAA CTC AAG AAC CAC GAT AAA CCC
Ile Lys Lys Tyr Ile Thr Ser Asp Glu Leu Lys Asn His Asp Lys Pro
  5                      10                      15                      20

GGA GAT CTA TGG ATC TCG ATT CAA GGG AAA GCC TAT GAT GTT TCG GAT
Gly Asp Leu Trp Ile Ser Ile Gln Gly Lys Ala Tyr Asp Val Ser Asp
                25                      30                      35

TGG GTG AAA GAC CAT CCA GGT GGC AGC TTT CCC TTG AAG AGT CTT GCT
Trp Val Lys Asp His Pro Gly Gly Ser Phe Pro Leu Lys Ser Leu Ala
                40                      45                      50

GGT CAA GAG GTA ACT GAT GCA TTT GTT GCA TTC CAT CCT GCC TCT ACA
Gly Gln Glu Val Thr Asp Ala Phe Val Ala Phe His Pro Ala Ser Thr
                55                      60                      65

TGG AAG AAT CTT GAT AAG TTT TTC ACT GGG TAT TAT CTT AAA GAT TAC
Trp Lys Asn Leu Asp Lys Phe Phe Thr Gly Tyr Tyr Leu Lys Asp Tyr
  70                      75                      80

TCT GTT TCT GAG GTT TCT AAA GAT TAT AGG AAG CTT GTG TTT GAG TTT
Ser Val Ser Glu Val Ser Lys Asp Tyr Arg Lys Leu Val Phe Glu Phe
  85                      90                      95                      100

TCT AAA ATG GGT TTG TAT GAC AAA AAA GGT CAT ATT ATG TTT GCA ACT
Ser Lys Met Gly Leu Tyr Asp Lys Lys Gly His Ile Met Phe Ala Thr
                105                      110                      115

TTG TGC TTT ATA GCA ATG CTG TTT GCT ATG AGT GTT TAT GGG GTT TTG
Leu Cys Phe Ile Ala Met Leu Phe Ala Met Ser Val Tyr Gly Val Leu
                120                      125                      130

TTT TGT GAG GGT GTT TTG GTA CAT TTG TTT TCT GGG TGT TTC ATG GCG
Phe Cys Glu Gly Val Leu Val His Leu Phe Ser Gly Cys Leu Met Gly
  135                      140                      145

TTT CTT TGG ATT CAG AGT GGT TGG ATT GGA CAT GAT GCT GGG CAT TAT
Phe Leu Trp Ile Gln Ser Gly Trp Ile Gly His Asp Ala Gly His Tyr
  150                      155                      160

ATG GTA GTG TCT GAT TCA AGG CTT AAT AAG TTT ATG GGT ATT TTT GCT
Met Val Val Ser Asp Ser Arg Leu Asn Lys Phe Met Gly Ile Phe Ala
  165                      170                      175                      180

GCA AAT TGT CTT TCA GGA ATA AGT ATT GGT TGG TGG AAA TGG AAC CAT
Ala Asn Cys Leu Ser Gly Ile Ser Ile Gly Trp Trp Lys Trp Asn His
                185                      190                      195

```

AAT GCA CAT CAC ATT GCC TGT AAT AGC CTT GAA TAT GAC CCT GAT TTA
Asn Ala His His Ile Ala Cys Asn Ser Leu Glu Tyr Asp Pro Asp Leu
200 205 210

CAA TAT ATA CCA TTC CTT GTT GTG TCT TCC AAG TTT TTT GGT TCA CTC
Gln Tyr Ile Pro Phe Leu Val Val Ser Ser Lys Phe Phe Gly Ser Leu
215 220 225

ACC TCT CAT TTC TAT GAG AAA AGG TTG ACT TTT GAC TCT TTA TCA AGA
Thr Ser His Phe Tyr Glu Lys Arg Leu Thr Phe Asp Ser Leu Ser Arg
230 235 240

TTC TTT GTA AGT TAT CAA CAT TGG ACA TTT TAC CCT ATT ATG TGT GCT
Phe Phe Val Ser Tyr Gln His Trp Thr Phe Tyr Pro Ile Met Cys Ala
245 250 255 260

GCT AGG CTC AAT ATG TAT GTA CAA TCT CTC ATA ATG TTG TTG ACC AAG
Ala Arg Leu Asn Met Tyr Val Gln Ser Leu Ile Met Leu Leu Thr Lys
265 270 275

AGA AAT GTG TCC TAT CGA GGT CAG GAA CTC TTG GGA TGC CTA GTG TTC
Arg Asn Val Ser Tyr Arg Ala Gln Glu Leu Leu Gly Cys Leu Val Phe
280 285 290

TCG ATT TGG TAC CCG TTG CTT GTT TCT TGT TTG CCT AAT TGG GGT GAA
Ser Ile Trp Tyr Pro Leu Leu Val Ser Cys Leu Pro Asn Trp Gly Glu
295 300 305

AGA ATT ATG TTT GTT ATT GCA AGT TTA TCA GTG ACT GGA ATG CAA CAA
Arg Ile Met Phe Val Ile Ala Ser Leu Ser Val Thr Gly Met Gln Gln
310 315 320

GTT CAG TTC TCC TTG AAC CAC TTC TCT TCA AGT GTT TAT GTT GGA AAG
Val Gln Phe Ser Leu Asn His Phe Ser Ser Ser Val Tyr Val Gly Lys
325 330 335 340

CCT AAA GGG AAT AAT TGG TTT GAG AAA CAA ACG GAT GGG ACA CTT GAC
Pro Lys Gly Asn Asn Trp Phe Glu Lys Gln Thr Asp Gly Thr Leu Asp
345 350 355

ATT TCT TGT CCT CCG TGG ATG GAT TGG TTT CAT GGT GGA TTG CAA TTC
Ile Ser Cys Pro Pro Trp Met Asp Trp Phe His Gly Gly Leu Gln Phe
360 365 370

CAA ATT GAG CAT CAT TTG TTT CCC AAG ATG CCT AGA TGC AAC CTT AGG
Gln Ile Glu His His Leu Phe Pro Lys Met Pro Arg Cys Asn Leu Arg
375 380 385

AAA ATC TCG CCC TAC GTG ATC GAG TTA TGC AAG AAA CAT AAT TTG CCT
Lys Ile Ser Pro Tyr Val Ile Glu Leu Cys Lys Lys His Asn Leu Pro
390 395 400

TAC AAT TAT GCA TCT TTC TCC AAG GCC AAT GAA ATG ACA CTC AGA ACA
Tyr Asn Tyr Ala Ser Phe Ser Lys Ala Asn Glu Met Thr Leu Arg Thr
405 410 415 420

TTG AGG AAC ACA GCA TTG CAG GCT AGG GAT ATA ACC AAG CCG CTC CCG
 Leu Arg Asn Thr Ala Leu Gln Ala Arg Asp Ile Thr Lys Pro Leu Pro
 425 430 435

AAG AAT TTG GTA TGG GAA GCT CTT CAC ACT CAT GGT T AAAATTACCC
 Lys Asn Leu Val Trp Glu Ala Leu His Thr His Gly
 440 445

TTAGTTCATG TAATAATTTG AGATTATGTA TCTCCTATGT TTGTGTCCTG TCTTGGTTCT
 ACTGTGTCGA GTCATTGCAA CTGTGCTTTT ATGGTTTATT AGATGTTTTT TAATATATTT
 TAGAGTTTTT GCTTTCATCT CCATTATTTGA TGAATAAGGA GTTGCATATT GTCAATTGTT
 GTGCTCAATA TCTGATATTT TGGAAATGTAC TTTGTACCAC GTGGTTTTCA GTTGAAGCTC
 ATGTGTACTT CTATAGACTT TGTTTAAATG GTTATGTCAT GTTATTT

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 2:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 135 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:2:

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

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 3:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 834 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: cADN

(ix) CARACTERÍSTICA:

(A) NOME/CHAVE: CDS

(B) LOCALIZAÇÃO: 31..603

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:3:

```

TTAGCCTTTA CTCTATAGTT TTAGATAGAC ATG GCG AAT GTG GAT CGT GAT CGG
Met Ala Asn Val Asp Arg Asp Arg
140

CGT GTG CAT GTA GAC CGT ACT GAC AAA CGT GTT CAT CAG CCA AAC TAC
Arg Val His Val Asp Arg Thr Asp Lys Arg Val His Gln Pro Asn Tyr
145 150 155

GAA GAT GAT GTC GGT TTT GGT GGC TAT GGC GGT TAT GGT GCT GGT TCT
Glu Asp Asp Val Gly Phe Gly Gly Tyr Gly Gly Tyr Gly Ala Gly Ser
160 165 170 175

GAT TAT AAG AGT CGC GGC CCC TCC ACT AAC CAA ATC TTG GCA CTT AEA
Asp Tyr Lys Ser Arg Gly Pro Ser Thr Asn Gln Ile Leu Ala Leu Ile
180 185 190

GCA GGA GTT CCC ATT GGT GGC ACA CTG CTA ACC CTA GCT GGA CTC ACT
Ala Gly Val Pro Ile Gly Gly Thr Leu Leu Thr Leu Ala Gly Leu Thr
195 200 205

CTA GCC GGT TCG GTG ATC GGC TTG CTA GTC TCC ATA CCC CTC TTC CTC
Leu Ala Gly Ser Val Ile Gly Leu Leu Val Ser Ile Pro Leu Phe Leu
210 215 220

CTC TTC AGT CCG GTG ATA GTC CCG GCG GCT CTC ACT ATT GGG CTT GCT
Leu Phe Ser Pro Val Ile Val Pro Ala Ala Leu Thr Ile Gly Leu Ala
225 230 235

GTG ACG GGA ATC TTG GCT TCT GGT TTG TTT GGG TTG ACG GGT CTC AGC
Val Thr Gly Ile Leu Ala Ser Gly Leu Phe Gly Leu Thr Gly Leu Ser
240 245 250 255

TCG CTC TCG TGG GTC CTC AAC TAC CTC CGT GGG ACG AGT GAT ACA GTG
Ser Val Ser Trp Val Leu Asn Tyr Leu Arg Gly Thr Ser Asp Thr Val
260 265 270

```

CCA GAG CAA TTG GAC TAC GCT AAA CGG CGT ATG GCT GAT GCG GTA GGC
 Pro Glu Gln Leu Asp Tyr Ala Lys Arg Arg Met Ala Asp Ala Val Gly
 275 280 285

TAT GCT GGT ATG AAG CGA AAA GAG ATG GGT CAG TAT GTG CAA GAT AAG
 Tyr Ala Gly Met Lys Gly Lys Glu Met Gly Gln Tyr Val Gln Asp Lys
 290 295 300

GCT CAT GAG GCT CGT GAG ACT GAG TTC ATG ACT GAG ACC CAT GAG CCG
 Ala His Glu Ala Arg Glu Thr Glu Phe Met Thr Glu Thr His Glu Pro
 305 310 315

GGT AAG GCC AGG AGA GGC TCA TAAGCTAATA TAAATTGCGG GACTCAATTG
 Gly Lys Ala Arg Arg Gly Ser
 320 325

GAACGCGAT AAATGTAGTT TTACTTTTAT GTCCAGTTT CTTTCCTCTT TTAAGAATAT
 CTTTGTCTAT ATATGTGTTT GTTCGTTTTG TCTTGTCCAA ATAAAAATCC TTGTTACTGA
 AATAAGAAAT GAAATAAATA TGTTCCTTTT TTGAGATAA CCAGAAATCT CATACTATTT
 TCTAAAAAAA AAAAAAAAAA A

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 4:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 191 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:4:

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

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

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 5:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 191 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:5:

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

Leu Arg Gly Thr Ser Asp Thr Val Pro Glu Gln Leu Asp Tyr Ala Lys
 130 135 140
 Arg Arg Met Ala Asp Ala Val Gly Tyr Ala Gly Met Lys Gly Lys Glu
 145 150 155 160
 Met Gly Gln Tyr Val Gln Asp Lys Ala His Glu Ala Arg Glu Thr Glu
 165 170 175
 Phe Met Thr Glu Thr His Glu Pro Gly Lys Ala Arg Arg Gly Ser
 180 185 190

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 6:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 191 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:6:

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

Met Gly Gln Tyr Val Gln Asp Lys Ala His Glu Ala Arg Glu Thr Glu
 165 170 175
 Phe Met Thr Glu Thr His Glu Pro Gly Lys Ala Arg Arg Gly Pro
 180 185 190

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 7:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 78 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:7:

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

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 8:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 173 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N°:8:

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

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 9:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 141 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:9:

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

Phe	Ser	Trp	Leu	Tyr	Arg	His	Trp	Thr	Gly	Ser	Gly	Ser	Asp	Lys	Ile
			100					105					110		
Glu	Trp	Ala	Arg	Met	Lys	Val	Gly	Ser	Arg	Val	Gln	Asp	Thr	Lys	Tyr
		115					120					125			
Gly	Gln	His	Trp	Ile	Gly	Val	Gln	His	Gln	Gln	Val	Ser			
	130					135					140				

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 10:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 199 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:10:

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

 Tyr Asp Ile Ser Lys Pro His Asp Thr Thr Thr Lys Gly His Glu Thr
 180 185 190
 Gln Gly Gly Thr Thr Ala Ala
 195

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 11:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 199 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:11:

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

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 12:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 1267 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genômico)

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:12:

GAGCTCGATC ACACAAAGAA AACGTCAAAT GGATCATACT GGGCCCATT TGCAGACCAA
 GAGAAAGTGA GAGAGAGTTG TCCTCTCGTT ATCAAGTAAC AGTAGACCAC CACTAAACCG
 CCAANTAGCTT ATAATCAAAA TAGAAAGGTC TAATAACAGA AACAAATGAA AAGCCCTGT
 TCCATGGACT GCCTACCCGA ATTGATTGAT TCGACTAGTT TTTCTTCTTC TTTGATTAA
 ACCTCCGTAA GAAAAATGGT ACTACTAAAG CCACTCGCTA CCAAACTAA ACCATTCCAG
 ACTGTAAC TGACCAATATT TCTAAACTGT AACCAGATCT CAAACATATA AACTAATTAA
 GAACTATAAC CATTAACCGT AAAAATAAAT TTACTACAGT AAAAAATTAT ACTAATTTCA
 GCTATGATGG AATTTCAGCT CTTAAGAGTT GTGGAAATCA AGTAAACCTA AAATCCTAAT
 AATATTCTTC ATCCTTATTT TTGTTTCACA TGCATGCTGT CCAATCTGTT ATTAGCATTT
 GAAAGCCTAA AATTCATAT ACAGTACAAT AAATCTAATT AATTTTCATT ACTAATAAAA
 TGCTTCATAT ATACTCTTGT ATTTATAAAT CATCCGTAT CTTACTATA CTTTATACA
 TCATCCTACA TTCATACCTA AGCTAGCAAA GCAAACTACT AAAAGGGTCG TCAACGCAAG
 TTATTTGCTA GTTGGTGCAT ACTACACACG GCTACGGCAA CATTAAGTAA CACATTAAGA
 GGTGTTTTCT TAATGTAGTA TGTAATTAT ATTTATTTCA AACTTTGGAT TAGATATAAA
 GGTACAGGTA GATGAAAAAT ATTTGCTTAG CGGCTTGAGA TTAAGCGGAT ATAGGAGGCA
 TATATACAGC TGTGAGAAGA AGAGGGATAA ATACAAAAAG GGAAGGATGT TTTTGGCGAC
 AGAGAAAGGT AGATTAAAGTA GGCATCGAGA GGAGAGCAAT TGTAAAATGG ATGATTTGTT
 TGCTTTTGTA CGGTGGAGAG AAGAACGAAA AGATGATCAG GTAAAAATG AACTTTGGAA
 ATCATGCAAA GCCACACCTC TCCCTTCAAC ACAGTCTTAC GTGTCGTCTT CTCTTCACTC
 CATATCTCCT TTTTATTACC AAGAAATATA TGTCAATCCC ATTTATATGT ACGTTCTCTT

 AGACTTATCT CTATATACCC CCTTTAATT TGTGTGCTCT TAGCCTTAC TCTATAGTTT
 TAGATAG

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N°: 13:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:13:

GTAATACGAC TCACTATAGG GC

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N°: 14:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 32 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:14:

GGGGATCCTA TACTAAAACT ATAGAGTAAA GG

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 15:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 8 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:15:

Trp Ile Gly His Asp Ala Gly His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 16:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 8 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:16:

Asn Val Gly His Asp Ala Asn His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N°: 17:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 8 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:17:

Val Leu Gly His Asp Cys Gly His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 18:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 8 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:18:

Val Ile Ala His Glu Cys Gly His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 19:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 8 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:19:

Val Ile Gly His Asp Cys Ala His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 20:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 8 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:20:

Val Val Gly His Asp Cys Gly His

1

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 21:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 5 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:21:

His Asn Ala His His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N°: 22:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 6 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:22:

His Asn Tyr Leu His His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 23:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 5 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:23:

His Arg Thr His His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 24:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 5 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:24:

His Arg Arg His His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 25:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 5 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:25:

His Asp Arg His His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 26:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 5 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:26:

His Asp Gln His His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 27:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 5 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:27:

His Asp His His His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 28:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 5 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:28:

His Asn His His His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 29:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 5 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:29:

Phe Gln Ile Glu His His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 30:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 6 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:30:

His Gln Val Thr His His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 31:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 5 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:31:

His Val Ile His His

1

5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 32:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 5 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:32:

His Val Ala His His

1 5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 33:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 5 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:33:

His Ile Pro His His

1 5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 34:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 5 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADEIA: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:34:

His Val Pro His His

1 5

(2) INFORMAÇÃO PARA A SEQUÊNCIA ID N° 35:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 1941 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADEIA: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIÇÃO DA SEQUÊNCIA ID N°:35:

GAGCTCGATC ACACAAAGAA AACGTCAAAT GGATCATACT GGGCCCNFTT TGCAGACCAA
GAGAAAGTGA GAGAGAGTTG TCCTCTCGTT ATCAAGTAAC AGTAGACCAC CACTAAACCG
CCAANTAGCTT ATAATCAAAA TAGAANGGTC TAATAACAGA AACAAATGAA AAAGCCTTGT
TCCATGGACT GCCTACCCGA ATTGATTGAT TCGACTAGTT TTTCTTCTTC TTTGATTAAAG
ACCTCCGTAA GAAAAATGGT ACTACTAAGG CCACTCGCTA CCAAAACATA ACCATTCCAG
ACTGTAACTG GACCAATATT TCTAAACTGT AACCAGATCT CAAACATATA AACTAATTAA
GAACTATAAC CATTAAACCGT AAAAATAAAT TTAGTACAGT AAAAATTAT ACTAATTTCA
GCTATGATGG AATTTACAGT CTTAAGAGTT GTGGAAATCA AGTAAACCTA AAATCCTAAT
AATATTCTTC ATCCTTATTT TTGTTTCACA TGCATGCTGT CCAATCTGTT ATTAGCATTT
GAAAGCCTAA AATTCTATAT ACAGTACAAT AAATCTAATT AATTTTCATT ACTAATAAAA

TGCTTCATAT ATACTCTTGT ATTTATAAAT CATCCGTTAT CGTTACTATA CCTTTATACA
 TCATCCTACA TTCATACCTA AGCTAGCAAA GCAAACACT AAAAGGGTCG TCAACGCAAG
 TTATTTGCTA GTTGGTGCAT ACTACACACG GCTACGGCAA CATTAAGTAA CACATTAAGA
 GGTGTTTTCT TAATGTAGTA TGGTAATTAT ATTTATTTCA AAACCTGGAT TAGATATAAA
 GGTACAGGTA GATGAAAAAT ATTTGGTTAG CGGCTTGAGA TTAAGCGGAT ATAGGAGGCA
 TATATACAGC TGTGAGAAGA AGAGGGATAA ATACAAAAAG GGAAGGATGT TTTTGCCGAC
 AGAGAAAGGT AGATTAAGTA GGCATCGAGA GGAGAGCAAT TGTAAGATGG ATGATTTGTT
 TGGTTTTGTA CGGTGGAGAG AAGAACGAAA AGATGATCAG GTAAAAAATG AAACCTGGAA
 ATCATGCAAA GCCACACCTC TCCCTTCAAC ACAGTCTTAC GTGTCTCTTT CTCTTCACTC
 CATATCTCCT TTTTATTACC AAGAAATATA TGTCAATCCC ATTTATATGT ACGTTCTCTT
 AGACTTATCT CTATATACCC CCTTTTAATT TGTGTCTCTT TAGCCTTTAC TCTATAGTTT
 TAGATAGACA TGGCGAATGT GGATCGTGAT CGGCGTGTGC ATGTAGACCG TACTGACAAA
 CGTGTTCATC AGCCAAACTA CGAAGATGAT GTGGGTTTTG GTGGCTATGG CGGTTATGGT
 GCTGGTTCTG ATTATAAGAG TCGCGGCCCC TCCACTAACC AAGTATTTTT GTGGTCTCTT
 TAGTTTTTCT TGTGTTTTCC TATGATCAGC CTCTCCAAAC TATTTGAAGA TTTTCTGTAA
 ATTCATTTTA AACAGAAAGA TAAATAAAAT AGTGAAGAAC CATAGGAATC GTACGTTACG
 TTAATTATTT CCTTTAGTT CTTAAGTCCT AATTAGGATT CCTTTAAAAG TTGCAACAAT
 CTAATTGTTT ACAAATGAG TAAAGTTTGA AACAGATTTT TATACACCAC TTGCATATGT
 TTATCATGGT GATGCATGCT TGTAGATAA ACTCGATATA ATCAATACAT GCAGATCTTG
 GCACTTATAG CAGGAGTCCA TTGGTGGCAC ACTGCTAACC CTAGCTGGAC TCACTCTAGC
 CGGTTCCGTG ATCGGCTTGC TATCTCCAT ACCCTCTTC CTCTCTTCA GTCCCGTGAT
 AGTCCCGGCG GCTCTCACTA TTGGGCTTGC TGTGACGGGA ATCTTGGCTT CTGGTTTTGT
 TGGGTTGACG GGTCTGAGCT C

Lisboa, 4 de Agosto de 2008

Reivindicações

1. Cassete de expressão que compreende a região reguladora 5' da oleoresina que dirige a expressão específica na semente constituída pela sequência de nucleótidos apresentada na SEQ ID N° 12 ligada operacionalmente a pelo menos um ácido nucleico que codifica para um gene hererólogo da $\Delta 6$ -dessaturase.
2. Vector de expressão que compreende a cassette de expressão de acordo com a reivindicação 1.
3. Célula compreendendo a cassette de expressão de acordo com a reivindicação 1.
4. Célula compreendendo o vector de expressão de acordo com a reivindicação 2.
5. Célula de acordo com a reivindicação 3, em que a dita célula é uma célula bacteriana, ou uma célula de uma planta.
6. Célula de acordo com a reivindicação 4, em que a dita célula é uma célula bacteriana, ou uma célula de planta.
7. Planta transgênica compreendendo a cassette de expressão de acordo com a reivindicação 1.
8. Planta transgênica compreendendo o vector de expressão de acordo com a reivindicação 2.
9. Planta de acordo com a reivindicação 7 ou 8 em que a dita planta é pelo menos girassol, soja, milho, algodão, tabaco, amendoim, colza, ou *Arabidopsis*.

10. Descendência da planta de acordo com a reivindicação 7, ou 8.

11. Semente da planta de acordo com a reivindicação 7 ou 8.

12. Processo para produzir uma planta com níveis acrescidos em teor de ácido gama linolénico (GLA) que compreende:

- (a) a transformação de uma célula de planta com o vector de expressão de acordo com a reivindicação 2; e
- (b) regeneração de uma planta com níveis acrescidos de GLA a partir da dita célula da planta

13. Processo de acordo com a reivindicação 12, em que o dito gene da $\Delta 6$ -dessaturase é pelo menos um de um gene da $\Delta 6$ -dessaturase cianobacteriano, ou um gene de $\Delta 6$ -dessaturase da borragem.

14. Processo de acordo com qualquer uma das reivindicações 12 ou 13 em que a dita planta é girassol, soja, milho, tabaco, algodão, amendoim, semente oleaginosa de colza ou *Arabidopsis*.

15. Processo de induzir a produção de pelo menos um dos ácidos gama linolénico (GLA), ou ácido octadecatetranoico (OTA) numa planta deficiente ou desprovida de GLA compreendendo a transformação da dita planta com um vector de expressão de acordo com a reivindicação 2 e a regeneração de uma planta com níveis acrescidos de pelo menos um dos GLA ou OTA.

Lisboa, 4 de Agosto de 2008

FIG.1A

2 ata tct gcc tac cct ccc aaa gag agt agt cat ttt tca tca atg gct gct caa atc aag
 M A A Q I K
 62 aaa tac att acc tca gat gaa ctg aag aac cac gat aaa ccc gga gat cta tgg atc tcg
 K Y I T S D E L K N H D K P G D L W I S
 122 att caa ggg aaa gcc tat gat gtt tcg gat tgg gtg aaa gac cat cca ggt ggc agc ttt
 I Q G K A Y D V S D W V K D H P G G S F
 182 ccc ttg aag agt ctt gct ggt caa gag gta act gat gca ttt gtt gca ttc cat cct gcc
 P K S L A G Q E V T D A F V A F H P A
 242 tct aca tgg aag aat ctt gat aag ttt ttc act ggg tat tat ctt aaa gat tac tct gtt
 S T W K N L D K F F T G Y Y L K D Y S V
 302 tct gag gtt tct aaa gat tat agg aag ctt gtg ttt gag ttt tct aaa atg ggt ttg tat
 S E V S K D Y R K L V F E F S K M G L Y
 362 gac aaa aaa ggt cat att atg ttt gca act ttg tgc ttt ata gca atg ctg ttt gct atg
 D K K G H I M F A T L C F I A M L F A M

FIG.1B

422 agt gtt tat ggg gtt ttg ttt tgt gag ggt gtt ttg gta cat ttg ttt tct ggg tgt ttg
 s v y g v l f c e g v l v h l f s g c l

 482 atg ggg ttt ctt tgg att cag agt ggt tgg att gga cat gat gct ggg cat tat atg gta
 m g f l w i q s g w i g h d a g h y m v

 542 gtg tot gat tca agg ctt aat aag ttt atg ggt att ttt gct gca aat tgt ctt tca gga
 v s d s r l n k f m g i f a a n c l s g

 602 ata agt att ggt tgg tgg aaa tgg aac cat aat gca cat cac att gcc tgt aat agc ctt
 i s i g w w k w n h n a h h i a c n s l

 662 gaa tat gac cct gat tta caa tat ata cca ttc ctt gtt gtg tct tcc aag ttt ttt ggt
 e y d p d l q y i p f l v v s s k f f g

 722 tca ctc acc tct cat ttc tat gag aaa agg ttg act ttt gac tct tta tca aga ttc ttt
 s l t s h f y e k r l t f d s l s r f f

FIG.1C

782

gta agt tat caa cat tgg aca ttt tac cct att atg tgt gct gct agg ctc aat atg tat
V S Y Q H W T F Y P I M C A A R L N M Y

842

gta caa tct ctc ata atg ttg ttg acc aag aga aat gtg tcc tat cga gct cag gaa ctc
V Q S L I M L L T K R N V S Y R A Q E L

902

ttg gga tgc cta gtg ttc tcg att tgg tac ccg ttg ctt gtt tct tgt ttg cct aat tgg
L G C L V F S I W Y P L L V S C L P N W

962

ggt gaa aga att atg ttt gtt att gca agt tta tca gtg act gga atg caa caa gtt cag
G E R I M F V I A S L S V T G M Q Q V Q

1022

ttc tcc ttg aac cac ttc tct tca agt gtt tat gtt gga aag cct aaa ggg aat aat tgg
F S L N H F S S S V Y V G K P K G N N W

1082

ttt gag aaa caa acg gat ggg aca ctt gac att tct tgt cct cct tgg atg gat tgg ttt
F E K Q T D G T L D I S C P P W M D W F

FIG.1D

1142 cat ggt gga ttg caa ttc caa att gag cat cat ttg ttt ccc aag atg cct aga tgc aac
 H G G L Q F Q I E H H L F P K M P R C N
 1202 ctt agg aaa atc tcg ccc tac gtg atc gag tta tgc aag aaa cat aat ttg cct tac aat
 L R K I S P Y V I E L C K K H N L P Y N
 1262 tat gca tct ttc tcc aag gcc aat gaa atg aca ctc aga aca ttg agg aac aca gca ttg
 Y A S F S K A N E M T L R T L R N T A L
 1322 cag gct agg gat ata acc aag ccg ctc ccg aag aat ttg gta tgg gaa gct ctt cac act
 Q A R D I T K P L P K N L V W E A L H T
 1382 cat ggt taa aat tac cct tag ttc atg taa taa ttt gag att atg tat ctc cta tgt ttg
 H G *
 1442 tgt ctt gtc ttg gtt cta ctt gtt gga gtc att gca act tgt ctt tta tgg ttt att aga
 1502 tgt ttt tta ata tat ttt aga ggt ttt gct ttc atc tcc att att gat gaa taa gga gtt

FIG.1E

1562

gca tat tgt caa ttg ttg tgc tca ata tct gat att ttg gaa tgt act ttg tac cac tgt

1622

gtt ttc agt tga agc tca tgt gta ctt cta tag act ttg ttt aaa tgg tta tgt cat gtt

1682

att t

FIG. 2

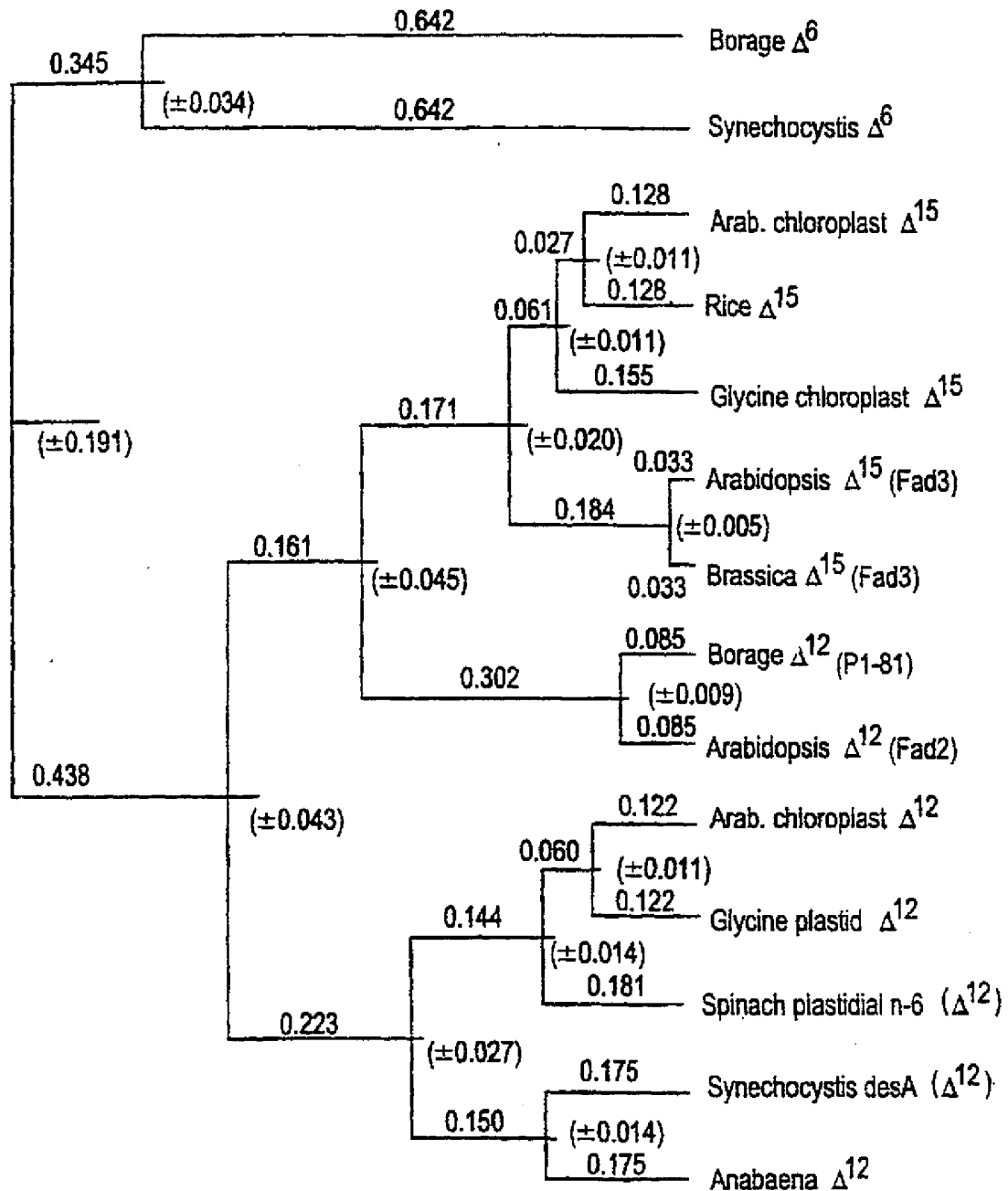


FIG. 3A

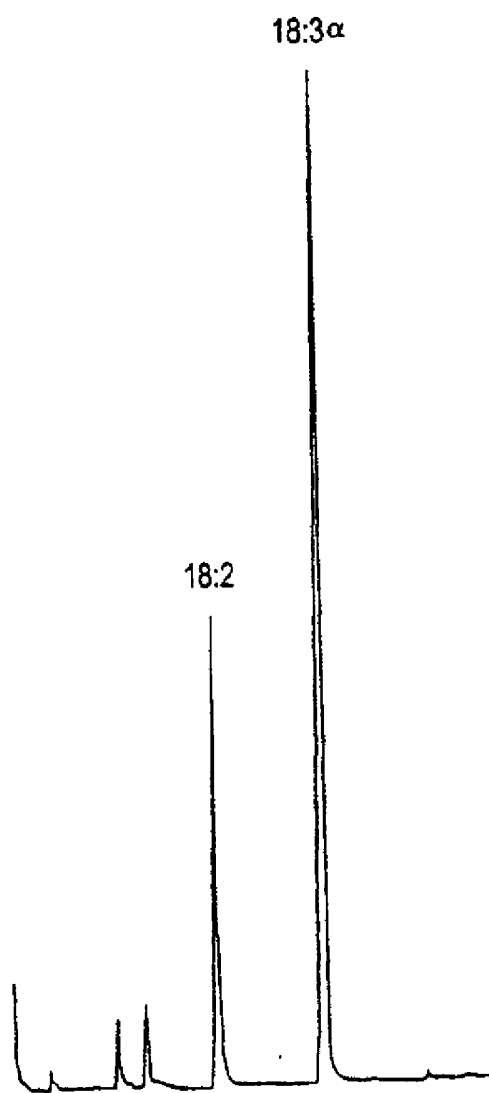


FIG. 3B

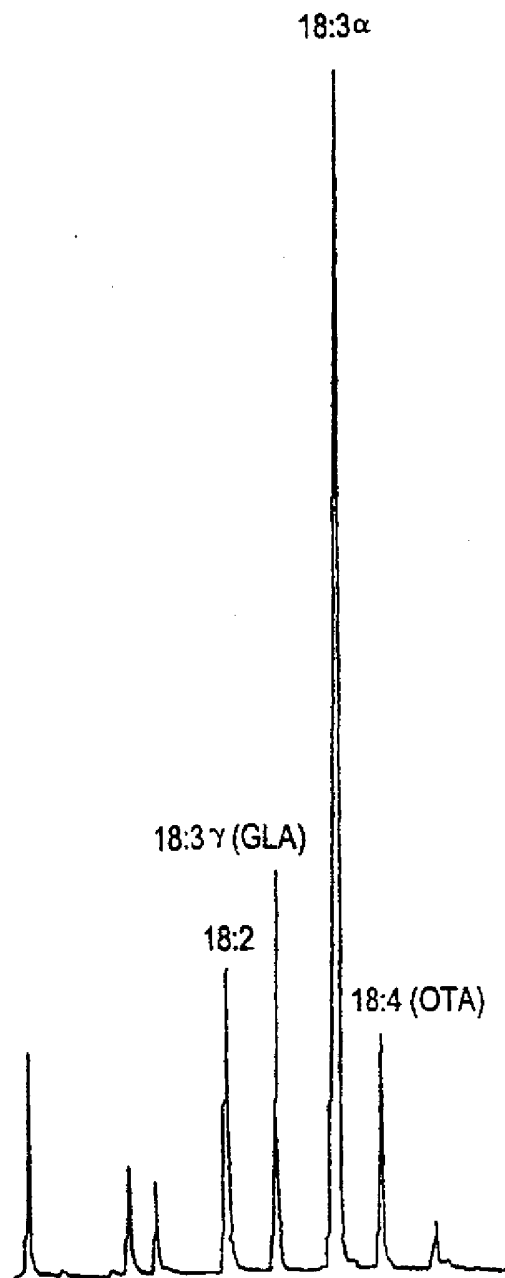


FIG.4

1 tta gcc ttt act cta tag ttt tag ata gac atg gcg aat gtg gat cgt gat cgg cgt gtg
 1 M A N V D R D R R V
 61 cat gta gac cgt act gac aaa cgt gtt cat cag cca aac tac gaa gat gat gtc ggt ttt
 11 H V D R T D K R V H Q P N Y E D D V G F
 121ggg ggc tat ggc ggt tat ggt gct ggt tct gat tat aag agt cgc ggc ccc tcc act aac
 31 G G Y G G Y G A G S D Y K S R G P S T N
 181caa atc ttg gca ctt ata gca gga gtt ccc att ggt ggc aca ctg cta acc cta gct gga
 51 Q I L A L I A G V P I G G T L L T L A G
 241ctc act cta gcc ggt tcg gtg atc ggc ttg cta gtc tcc ata ccc ctc ttc ctc ctc ttc
 71 L T L A G S V I G L L V S I P L F L L F
 301agt ccg gtg ata gtc ccg gcg gct ctc act att ggg ctt gct gtg acg gga atc ttg gct
 91 S P V I V P A A L T I G L A V T G I L A
 361tct ggt ttg ttt ggg ttg acg ggt ctg agc tcg gtc tcg tgg gtc ctc aac tac ctc cgt
 111S G L F G L T G L S S V S W V L N Y L R
 421ggg acg agt gat aca gtg cca gag caa ttg gac tac gct aaa cgg cgt atg gct gat gcg
 131G T S D T V P E Q L D Y A K R R M A D A
 481gta ggc tat gct ggt atg aag gga aaa gag atg ggt cag tat gtg caa gat aag gct cat
 151V G Y A G M K G K E M G Q Y V Q D K A H
 541gag gct cgt gag act gag ttc atg act gag acc cat gag ccg ggt aag gcc agg aga ggc
 171E A R E T E F M T E T H E P G K A R R G
 601tca taa gct aat ata aat tgc ggg agt cag ttg gaa acg cga taa atg tag ttt tac ttt
 191S *
 661tat gtc cca gtt tct ttc ctc ttt taa gaa tat ctt tgt cta tat atg tgt tcg ttc gtt
 721ttg tct tgt cca aat aaa aat cct tgt tag tga aat aag aaa tga aat aaa tat gtt ttc
 781ttt ttt gag ata acc aga aat ctc ata cta ttt tct aaa aaa aaa aaa aaa

FIG.5

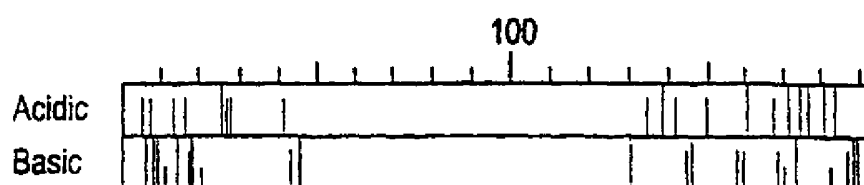


FIG.6

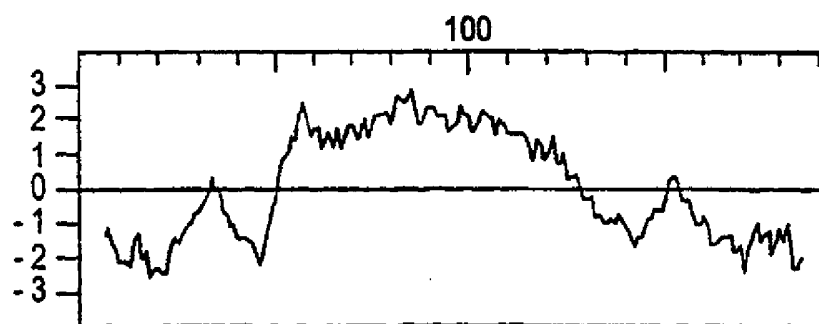


FIG.7A

AtS21	MANVDRDRRV	HVDRTDKRVH	-QPNYEDDVG	---F-G--GY	GGYGAGSDYK	43
X91918	MANVDRDRRV	HVDRTDKRVH	-QPNYEDDVG	---F-G--GY	GGYGAGSDYK	43
Z29859	-----	-----	-----	-----	-----	-----
X62352	MADTARG---	-THHDIIGRD	QYPMMGRDRD	QYQKSCRG--	-----SDY-	37
Atol3	MADQTR----	-THHE-----	---MISRST	Q-----	-----EAH-	21
X91956	MADTHR----	-VDRTDRHFQ	FQSPYEGGRG	QGQYEGDRGY	GGGGYKSMMP	45
L40954	MADTHR----	-VDRTDRHFQ	FQSPYEGGRG	QGQYEGDRGY	GGGGYKSMMP	45
AtS21	SRGPSTNQIL	ALIAGVPIGG	TLTTLAAGLTL	AGSVIGLIMS	IPFLIFSPV	93
X91918	SRGPSTNQIL	ALIAGVPIGG	TLTTLAAGLTL	AGSVIGLIMS	IPFLIFSPV	93
Z29859	-----	-----	-----	-----	-----	-----
X62352	---SKSRQIA	KAATAVTAGG	SLIVLSLTL	VGVIALTVA	TPLLVIFSPI	84
Atol3	---PKARQMV	KAATAVTAGG	SLIVLSQLTL	AGTVIALTVA	TPLLVIFSPV	68
X91956	ESGPSTQVL	SLIGVPVVG	SLIALAGLL	AGSVIGIMVA	LPLFLIFSPV	95
L40954	ESGPSTQVL	SLIGVPVVG	SLIALAGLL	AGSVIGIMVA	LPLFLIFSPV	95
AtS21	IVPAALTIGL	AVTGILASGL	FGLTGLSSVS	WVLNLRGTS	DTVPEQLDYA	143
X91918	IVPAALTIGL	AVTGILASGL	FGLTGLSSVS	WVLNLRGTS	DTVPEQLDYA	143
Z29859	-----	-----	FGLTGLSSVS	WVLQPPWAS	DTVPEQVDYA	30
X62352	IVPALITVAL	LITGELSSGG	FGIAAITVES	WLYKYATGEH	PQGSCKLDSA	134
Atol3	IVPAVITVAL	LITGELASCG	FGIAAITAFS	WLYRHHTGS-	--GSDKIENA	115
X91956	IVPAGITIGL	AMTGELASGM	FGLTGLSSIS	WVMNLRGTK	RTVPEQLEYA	145
L40954	IVPAAITIGL	AMTGELASGM	FGLTGLSSIS	WVMNLRGTR	RTVPEQLEYA	145

FIG.7B

AtS21	KRRMADAVGY AGMKGKEMQ YVQDKAHEAR ETEF-----	MTETHEPGKA 187
X91918	KRRMADAVGY AGMKGKEMQ YVQDKAHEAR ETEF-----	MTETHEPGKA 187
Z29859	KRRMADAVGY AGMKGKEMQ YVQDKAHEAR ETEF-----	MTETHEPGKA 74
X62352	RMKLGSKA-- QDLKDRA-QY YGQHTGGEH DRDRTRGGQH TT-----	173
Atol3	RMKVGSRV-- QDTK----- YGQHNIGVQH QQ-----	VS----- 141
X91956	KRRMADAVGY AGQKGKEMQ HVQNKAQDVK QYDISKPHDT TTKGHETQGG	195
L40954	KRRMADAVGY AGQKGKEMQ HVQNKAQDVK QYDISKPHDT TTKGHETQGR	195

AtS21	RRGS = Z54164/Atol2	191
X91918	RRGP	191
Z29859	RRGS	78
X62352	----	173
Atol3	----	141
X91956	TTAA	199
L40954	TTAA	199

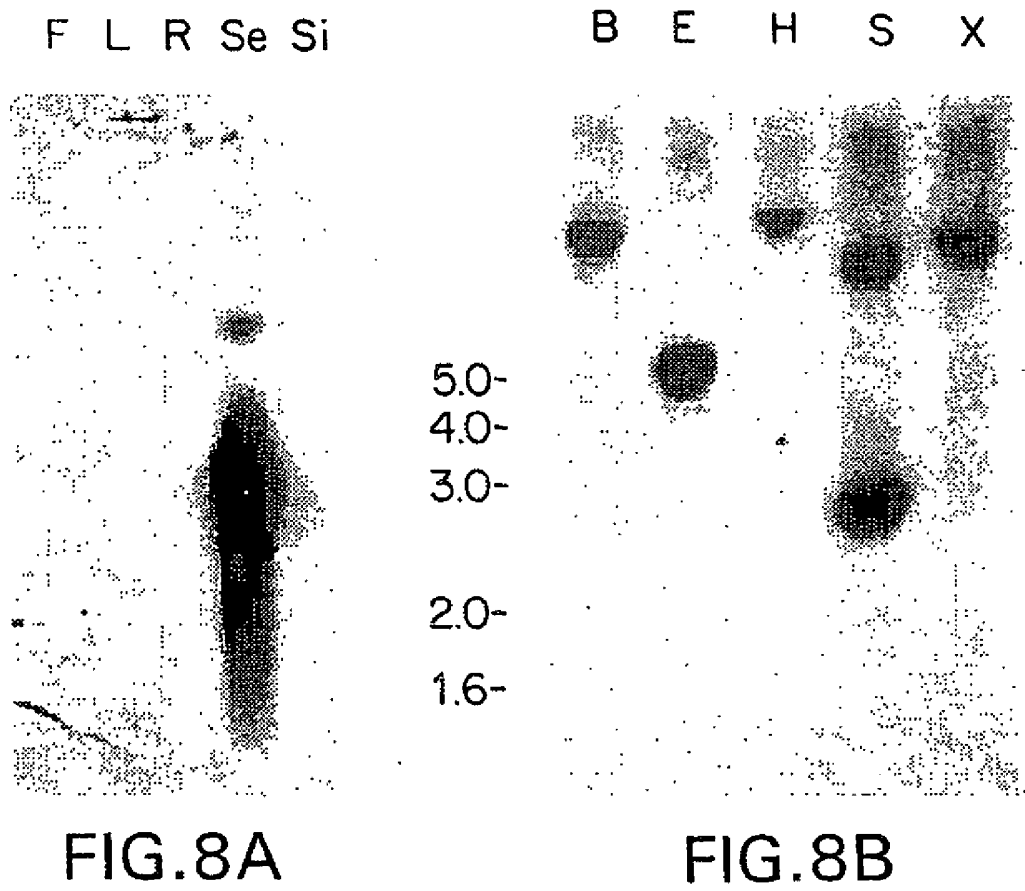


FIG.9

1 GAGCTCGATCACACAAAGAAAACGTCAATGGATCATACTGGGCCCATTTTGCAGACCAAGAGAAAGTGAGAGAGAGTTG
 81 TCCTCTCGTTATCAAGTAACAGTAGACCACCCTAAACCGCCAATAGCTTATAATCAAAATAGAAAGGTCTAATAACAGA
 161 AACAAATGAAAAGCCTTGTTCCATGGACTGCCTACCCGAATTGATTGATTGACTAGTTTTTCTTCTTCTTGATTAAAG
 241 ACCTCCGTAAGAAAAATGGTACTACTAAAGCCACTCGCTACCAAACTAAACCATTCCAGACTGTAAGTGGACCAATATT
 321 TCTAAACTGTAACCAGATCTCAAACATATAAACTAATTAAGAACTATAACCATTAAACCGTAAAAATAAATTTACTACAGT
 401 AAAAAATTATACTAATTTTCAGCTATGATGGAATTTTCAGCTCTTAAGAGTTGTGGAATCAAGTAAACCTAAAAATCCTAAT
 481 AATATTCCTCATCCTTATTTTGTTCACATGCATGCTGTCCAATCTGTTATTAGCATTTGAAAGCCTAAAAATCTATAT
 561 ACAGTACAATAAATCTAATTAATTTTCATTACTAATAAAATGCTTCATATATACTCTTGATTATATAAATCATCCGTTAT
 641 CGTTACTATACCTTTATACATCATCCTACATTCATACCTAAGCTAGCAAAGCAAACCTACTAAAAGGGTCGTCAACGCAAG
 721 TTATTTGCTAGTTGGTGCATACTACACCGCTACGGCAACATTAAAGTAACACATTAAAGAGGTGTTTTCTTAATGTAGTA
 801 TGGTAATTATATTTATTTCAAACCTTGGATTAGATATAAAGGTACAGGTAGATGAAAAATATTTGGTTAGCGGGTTGAGA
 881 TTAAGCGGATATAGGAGGCATATATACAGCTGTGAGAAGAAGAGGGATAAATACAAAAGGGAAGGATGTTTTTGCCGAC
 961 AGAGAAAGGTAGATTAAGTAGGCATCGAGAGGAGAGCAATTGTAATAATGGATGATTGTTTGGTTTGTACGGTGGAGAG
 1041 AAGAACGAAAAGATGATCAGGTAAAAAATGAAACTTGGAATCATGCAAAGCCACACCTCTCCCTTCAACACAGTCTTAC
 1121 GTGTCGTCTTCTTCTCACTCCATATCTCCTTTTTATTACCAAGAAATATATGTCAATCCCATTATATGTACGTTCTCTT
 1201 AGACTTATCTCTATATACCCCTTTTAAATTTGTGTGCTCttagcctttactctatagtttttagatagacatggcgaatgt
 1281 ggatcgatcgcggtgtgcatgtagaccgtactgacaaacgtgttcacagccaaactacgaagatgatgtcggttttg
 1361 gtggctatggcgggtatggtgctggttctgattataagagtcgcgggccctccactaaccaagTATTTTTGTGGTCTCTT
 1441 TAGTTTTTCTTGTGTTTTCTATGATCACGCTCTCCAACTATTTGAAGATTTCTGTAAATTCATTTTAAACAGAAAGA
 1521 TAAATAAAATAGTGAAGAACCATAGGAATCGTACGTTACGTTAATTATTTCCTTTTAGTTCCTTAAGTCCTAATTAGGATT
 1601 CCTTTAAAGTTGCAACAATCTAATTGTTCAAAAATGAGTAAAGTTTGAAACAGATTTTATACACCCTTGCATATGT
 1681 TTATCATGGTGATGCATGCTTGTAGATAAACTCGATATAATCAATACATGCAGatcttggcacttatagcaggagttcc
 1761 cattggtggcacactgctaaccctagctggactcactctagccggttcggtgatcggttgctagctctccatacccctct
 1821 tcctcctcttcagtcggtgatagtcgccggcggtctcactattgggcttgctgtgacgggaatcttggcttctggtttg
 1901 tttgggttgacgggtctgagctc

FIG.10

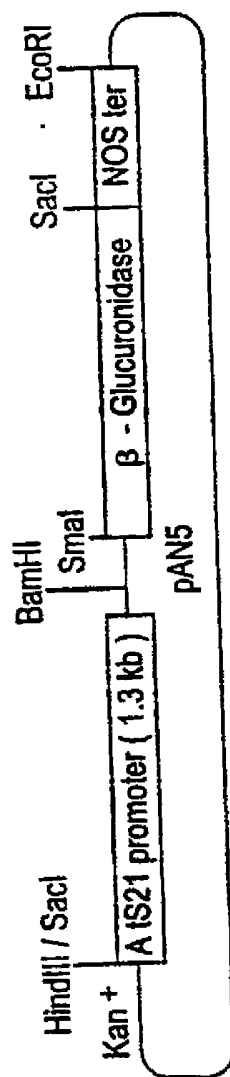


FIG.11A FIG.11B

FIG.11C

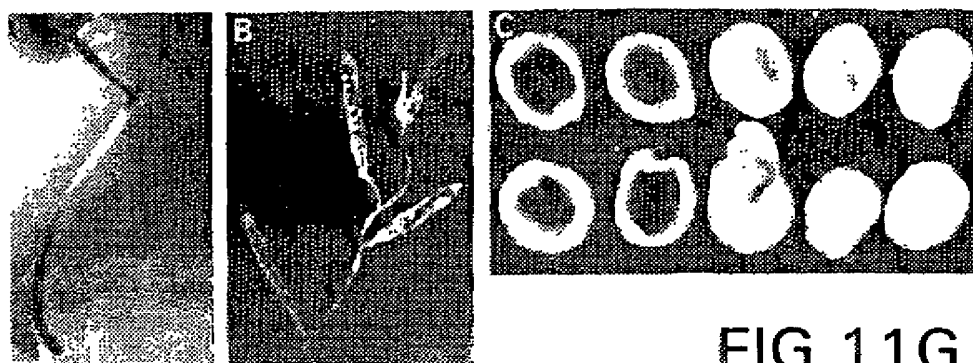


FIG.11G

FIG.11D FIG.11E FIG.11F

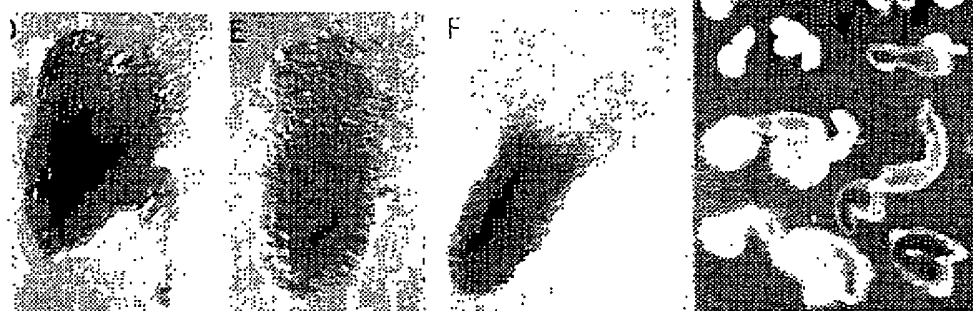


FIG.11H

FIG.11I

FIG.11J

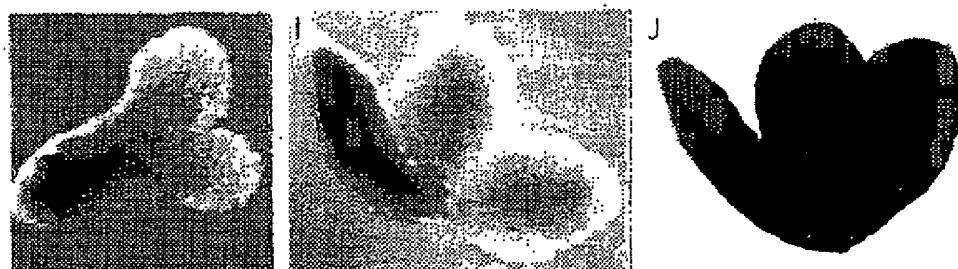


FIG.11K

FIG.11L

FIG.11M

FIG.11N



FIG.12A

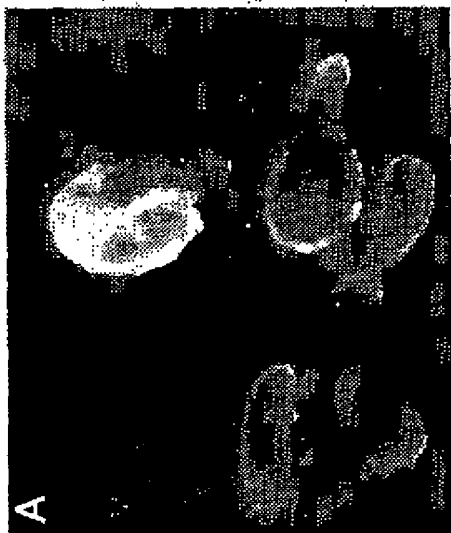


FIG.12B



FIG.12C

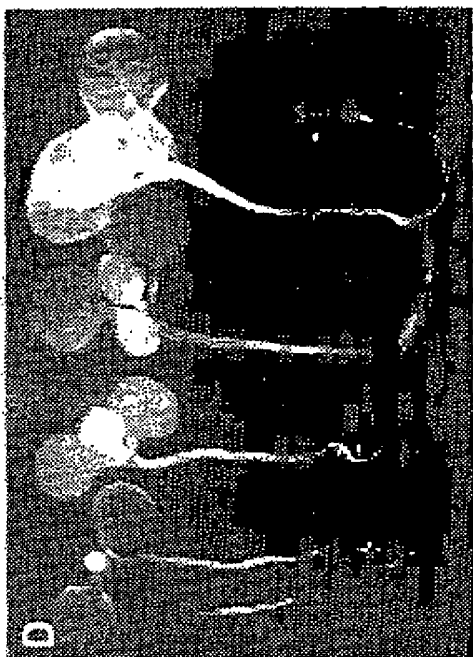


FIG.12D

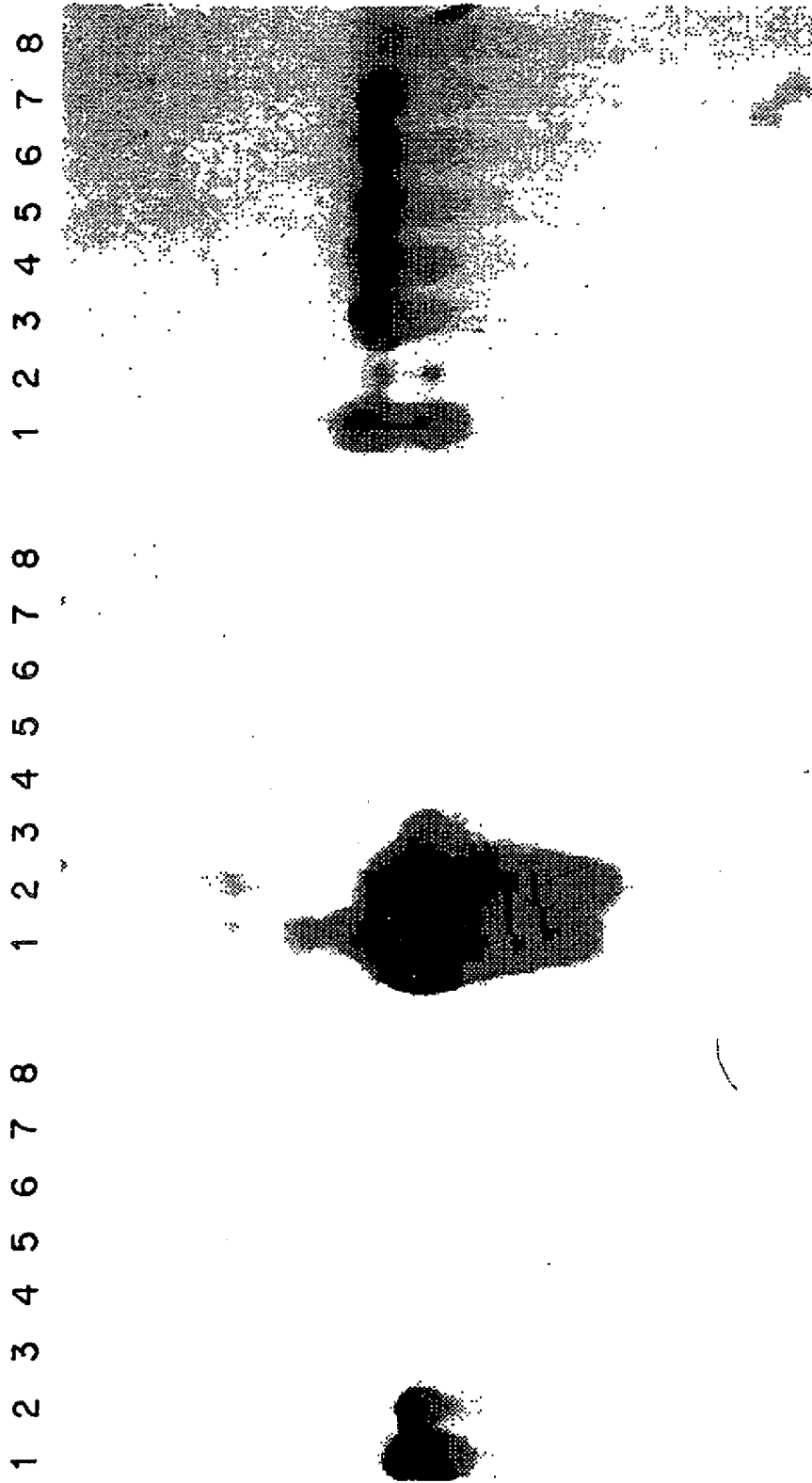


FIG.13A

FIG.13B

FIG.13C

FIG.14

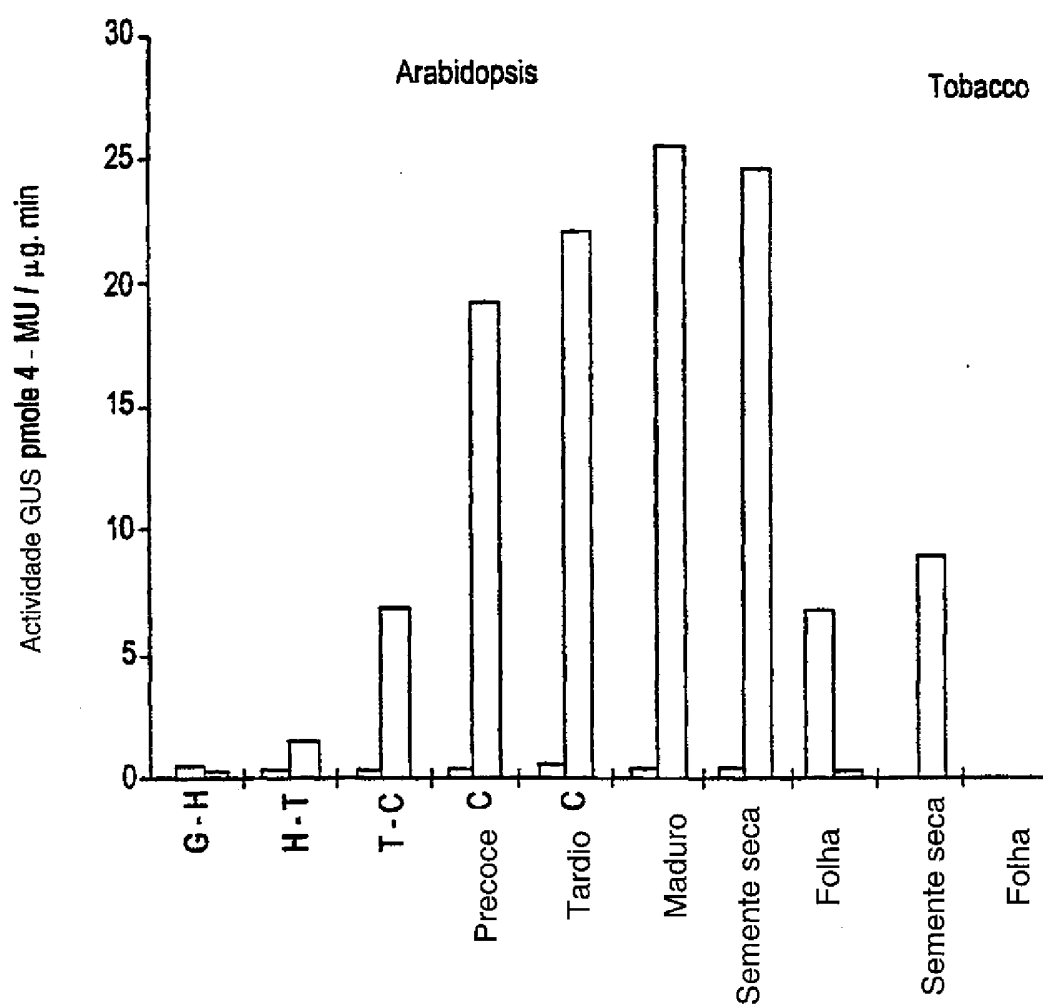
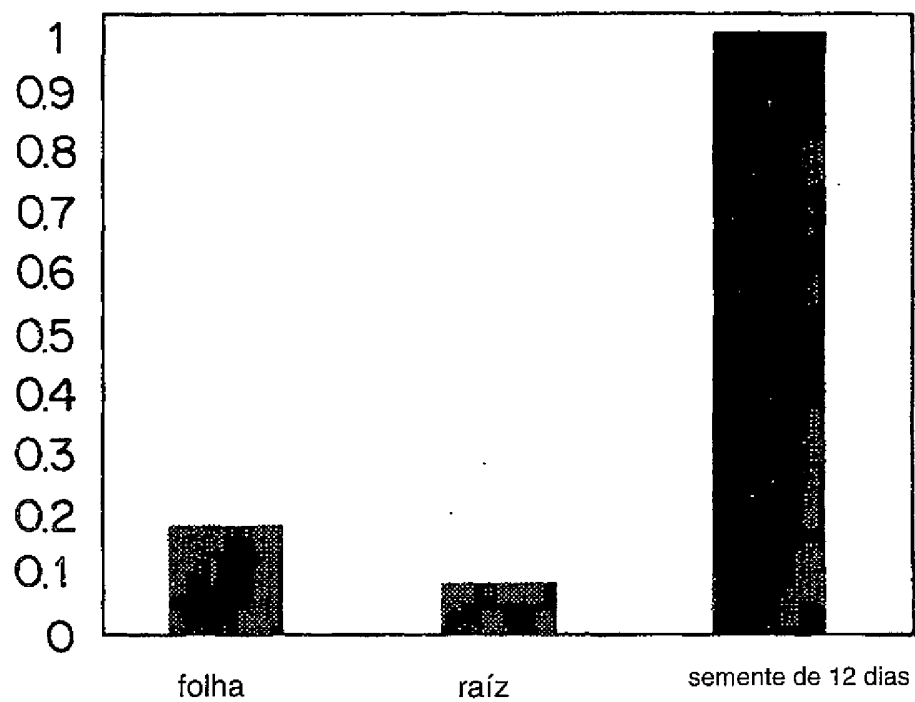


FIG.15B



Tecido da borragem

L R S



FIG.15A

FIG.16A

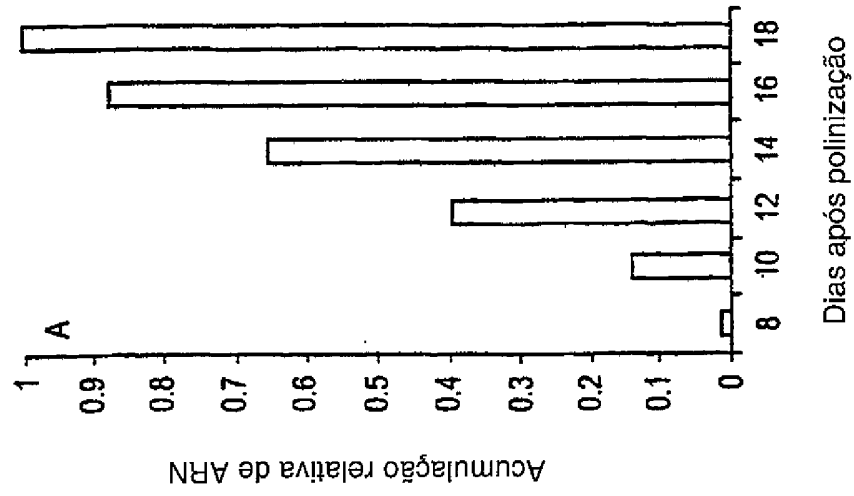


FIG.16B

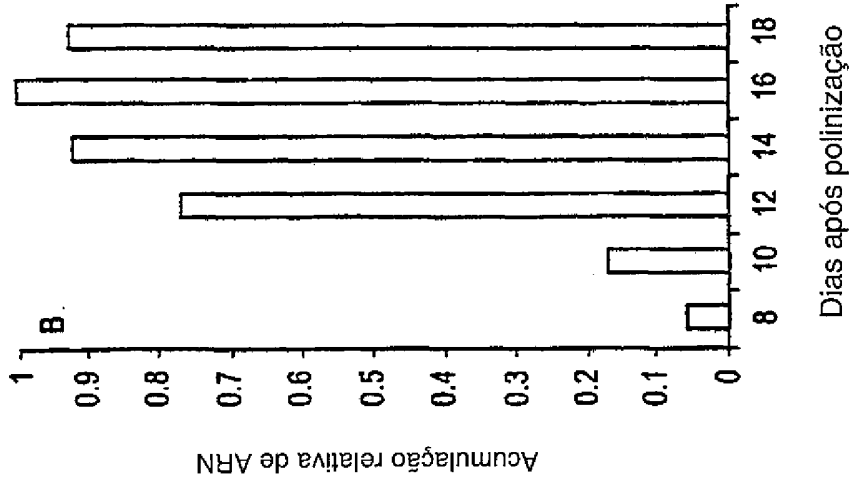


FIG.16C

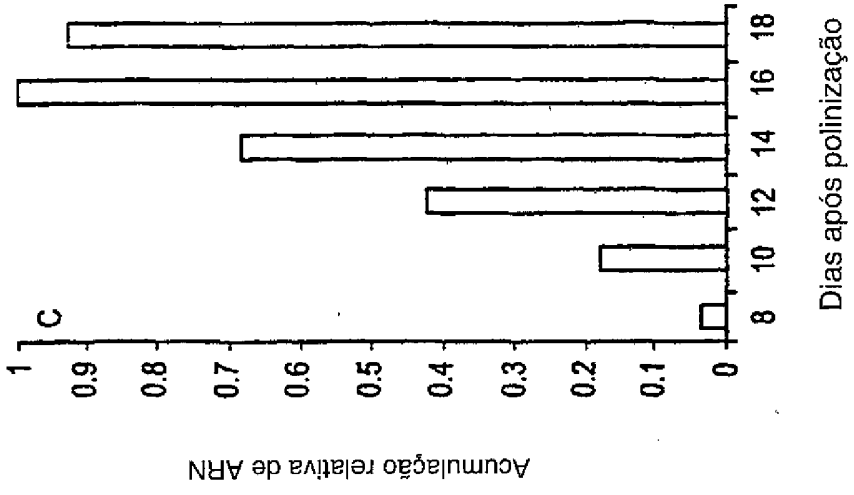


FIGURA 17

