



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

(11) PI 0717355-5 B1

(22) Data do Depósito: 10/10/2007

(45) Data de Concessão: 16/01/2018



(54) Título: MÉTODO PARA OBTENÇÃO DE UMA PLANTA TRANSGÊNICA DE NICOTINA, MÉTODO PARA OBTENÇÃO DE UMA SEMENTE; CONSTRUTO DE ÁCIDO NUCLÉICO RECOMBINANTE; MÉTODO DE REDUÇÃO DA CONVERSÃO DE NICOTINA EM NORNICOTINA EM UMA PLANTA DE NICOTINA

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

(30) Prioridade Unionista: 13/10/2006 US 11/580,765

(73) Titular(es): NORTH CAROLINA STATE UNIVERSITY. UNIVERSITY OF KENTUCKY RESEARCH FOUNDATION

(72) Inventor(es): RALPH E. DEWEY; BALAZS SIMINSZKY; STEVEN W. BOWEN; LILY GAVILANO

**MÉTODO PARA OBTENÇÃO DE UMA PLANTA TRANSGÊNICA DE
NICOTIANA, MÉTODO PARA OBTENÇÃO DE UMA SEMENTE;
CONSTRUTO DE ÁCIDO NUCLEÍCO RECOMBINANTE; MÉTODO DE
REDUÇÃO DA CONVERSÃO DE NICOTINA EM NORNICOTINA EM UMA
5 PLANTA DE NICOTIANA**

CAMPO DA INVENÇÃO

A invenção se refere a composições e métodos para reduzir o nível de nornicotina e seus metabólitos, *N'*-nitrosonornicotina, em uma planta que é membro do gênero *Nicotiana*, especialmente composições e métodos para inibir a expressão ou
10 função de um polipeptídeo do citocromo P450, envolvido na conversão metabólica da nicotina em nornicotina.

ANTECEDENTES DA INVENÇÃO

O alcalóide predominante encontrado nas variedades comerciais de tabaco é a nicotina, que corresponde normalmente a 90 – 95% do total do pool de alcalóide. A
15 fração de alcalóide remanescente é formada principalmente por três alcalóides de piridina: nornicotina, anabasina, e anatabina. A nornicotina é gerada diretamente da nicotina através da atividade da enzima nicotina *N*-demetilase (Figura 1). A nornicotina representa geralmente menos de 5% do total do reservatório de alcalóide de piridina, mas através de um processo chamado de “conversão”, as plantas de
20 tabaco que produzem inicialmente quantidades muito baixas de nornicotina dão origem a uma progênie que “converte” metabolicamente uma grande percentagem de nicotina da folha em nornicotina. Em plantas de tabaco que convertem geneticamente (chamadas de “conversoras”), a grande maioria da produção de nornicotina ocorre durante a senescência e cura da folha madura (Wernsman and Matzinger (1968) *Tob.*
25 *Sci.* 12:226-228). Os tabacos do tipo Burley são particularmente propensos à conversão genética, com altas taxas que chegam a 20% por geração em alguns cultivares.

Durante a cura e processamento da folha de tabaco, uma parte da nornicotina é metabolizada ao composto *N'*-nitrosonornicotina (NNN; Figure 1), uma nitrosamina
30 específica do tabaco (TSNA, em inglês, tobacco-specific nitrosamine) a qual tem se mostrado carcinogênica em animais de laboratório (Hecht and Hoffmann (1990) *Cancer Surveys* 8:273-294; Hoffmann *et al.* (1994) *J. Toxicol. Environ. Health* 41:1-

52; Hecht (1998) *Chem. Res. Toxicol.* 11:559-603). Em tabacos curados em fumeiros, constatou-se que as TSNAs são formadas predominantemente através da reação de alcalóides com quantidades insignificantes de óxidos de nitrogênio presentes em gases de combustão formados pelos sistemas de aquecimento de queima direta, encontrados em celeiros de cura tradicionais (Peele and Gentry (1999) "Formation of Tobacco-specific Nitrosamines in Flue-cured Tobacco," CORESTA Meeting, Agro-Phyto Groups, Suzhou, China). O aperfeiçoamento desses celeiros de cura com trocadores de calor elimina virtualmente a mistura de gases de combustão com o ar da secagem, e reduz acentuadamente a formação de TSNAs em tabacos curados dessa maneira (Boyette and Hamm (2001) *Rec. Adv. Tob. Sci.* 27:17-22.). Em contraste, nos tabacos do tipo Burley curados ao ar, a formação de TSNA ocorre principalmente através da reação dos alcalóides do tabaco com nitrito, um processo catalisado por micróbios provenientes da folha (Bush *et al.* (2001) *Rec. Adv. Tob. Sci.* 27:23-46). Até agora, as tentativas de reduzir as TNSAs através da modificação das condições de cura mantendo ao mesmo tempo padrões de qualidade aceitáveis não têm sido bem sucedidas para tabacos curados ao ar.

Em tabacos do tipo Burley, tem sido encontrada uma correlação positiva entre o conteúdo de nornicotina da folha e a quantidade de NNN acumulada no produto curado (Bush *et al.* (2001) *Rec. Adv. Tob. Sci.* 27:23-46; Shi *et al.* (2000) *Tob. Chem. Res. Conf.* 54:Abstract 27). No entanto, manter níveis mínimos de nornicotina tem sido difícil por causa do fenômeno da conversão, que resulta na introdução contínua de plantas com alta produção de nornicotina dentro de populações Barley crescidas para fins comerciais. A minimização do número de plantas tipo Burley que acumulam altos níveis de nornicotina tradicionalmente tem sido responsabilidade dos cultivadores de planta e produtores de sementes. Apesar da possibilidade da percentagem de plantas conversoras crescidas basicamente em campos de fazendeiros serem reduzida através da invasão das plantas conversoras por outras espécies, durante a propagação de estoques de sementes, esse processo é custoso, demanda tempo, e é imperfeito.

Estudos prévios têm mostrado que uma vez que a planta é convertida, o traço de alta nornicotina é herdado como um gene dominante simples (Griffith *et al.* (1955) *Science* 121:343-344; Burk and Jeffrey (1958) *Tob. Sci.* 2:139-141; Mann *et al.* (1964) *Crop Sci.* 4:349-353). No entanto, a natureza desse gene é atualmente

desconhecida. Nos cenários mais simples, o locus da conversão pode representar um gene não funcional de nicotina *N*-demetilase que recupera sua função em plantas conversoras, possivelmente através da mobilização de um elemento transponível indutor de mutação. Alternativamente, o locus conversor pode decodificar uma

5 proteína que inicia uma cascata de eventos, que ao final capacita a planta a metabolizar a nicotina em nornicotina, o que pode significar que genes múltiplos estão envolvidos.

Apesar de poderem existir um ou muitos genes associados ao processo de conversão, está claro que o(s) gene(s) codificadores de polipeptídeos que têm

10 atividade de nicotina demetilase desempenha(m) um papel essencial nesse processo. Embora a incapacidade de purificar a nicotina *N*-demetilase ativa a partir de extratos brutos tenha impedido o isolamento e identificação dessa enzima, há evidência de que um membro da superfamília de monooxigenases citocromo P450 possa estar envolvido (Hao and Yeoman (1996) *Phytochem.* 41:477-482; Hao and Yeoman

15 (1996) *Phytochem.* 42:325-329; Chelvarajan *et al.* (1993) *J. Agric. Food Chem.* 41:858-862; Hao and Yeoman (1998) *J. Plant Physiol.* 152:420-426). Contudo, esses estudos não são conclusivos, uma vez que os inibidores clássicos de P450, monóxido de carbono e tetcilase, têm falhado em diminuir a atividade da enzima a taxas comparáveis a outras reações mediadas por P450 (Chelvarajan *et al.* (1993) *J. Agric.*

20 *Food Chem.* 41:858-862).

Além disso, os citocromos P450 são proteínas transmembrana onipresentes que participam do metabolismo de uma ampla série de compostos (revisados por Schuler (1996) *Crit. Rev. Plant Sci.* 15:235-284; Schuler and Werck-Reichhart (2003) *Annu. Rev. Plant Biol.* 54:629-667). Os exemplos de reações bioquímicas mediadas

25 por citocromos P450 incluem hidroxilações, demetilações, e epoxidações. Em plantas, as famílias de genes do citocromo P450 são muito grandes. Por exemplo, o exame da sequência do genoma total revelou 272 genes previstos de citocromo P450 em *Arabidopsis* e no mínimo 455 genes únicos de citocromo P450 em arroz (ver, por exemplo, Nelson *et al.* (2004) *Plant Physiol.* 135(2):756-772). Embora o citocromo

30 P450 esteja implicado em um papel na conversão metabólica da nicotina em nornicotina, a identificação de membros chave que participam dessa família de proteínas ainda é um desafio.

Salvo o fato de servir como precursora para NNN, estudos recentes sugerem que a nornicotina encontrada em produtos do tabaco pode ter consequências indesejáveis para a saúde. Dickerson e Janda demonstraram que a nornicotina causa a glicação de proteínas aberrantes dentro da célula (Dickerson and Janda (2002) *Proc. Natl. Acad. Sci. USA* 99:15084-15088). Constatou-se que concentrações de proteínas modificadas por nornicotina são muito mais altas no plasma de fumantes quando comparadas aos não-fumantes. Além disso, esse mesmo estudo mostrou que a nornicotina pode modificar covalentemente fármacos esteróides comumente prescritos como a prednisona. Tais modificações têm o potencial de alterar tanto a

5
10

eficácia como a toxicidade desses fármacos.

Em vista das dificuldades associadas à conversão e efeitos indesejáveis à saúde relacionados à acumulação de nornicotina, é desejável a criação de métodos aperfeiçoados para redução do conteúdo da nornicotina em variedades de tabaco, especialmente o tabaco Burley. Tais métodos não atenuariam apenas potenciais

15

consequências negativas à saúde causadas pela nornicotina em si, como descrito acima, mas poderiam também reduzir concomitantemente os níveis de NNN.

RESUMO DA INVENÇÃO

Apresentam-se composições e métodos para redução do conteúdo de

20

nornicotina em plantas que são membros do gênero *Nicotiana*. As composições incluem polinucleotídeos isolados de citocromo P450 e polipeptídeos que estão envolvidos na conversão metabólica da nicotina em nornicotina em plantas, particularmente nas espécies de *Nicotiana*. Os polinucleotídeos isolados incluem uma sequência de nucleotídeos apresentados na ID SEQ N°:1, 3, 5, 7, 9, ou 11, uma

25

sequência de nucleotídeos que codifica um polipeptídeo conforme mostrado na ID SEQ N°:2, 4, 6, 8, 10, ou 12, e fragmentos e variantes dessas. Polipeptídeos isolados da invenção abrangem uma sequência de aminoácidos mostrada na ID SEQ N°:2, 4, 6, 8, 10, ou 12, uma sequência de aminoácidos codificada pela sequência de nucleotídeos mostrada em ID SEQ N°:1, 3, 5, 7, 9, ou 11, e fragmentos e variantes

30

dessas.

Os polinucleotídeos da invenção têm uso na supressão da expressão de um citocromo P450 que está envolvido na conversão metabólica da nicotina em nornicotina em uma planta, incluindo os citocromos P450 da presente invenção.

Desse modo, as composições incluem ainda cassetes de expressão abrangendo uma sequência inibidora que é capaz de inibir a expressão ou função de um polipeptídeo de citocromo P450 da invenção, no qual a sequência inibidora está ligada operacionalmente a um promotor que é funcional em uma célula de planta. Em
5 algumas concretizações, a sequência inibidora inclui a sequência apresentada nas IDs SEQ N°:1, 3, 5, 7, 9, 11, 13, 14, 15, ou 16, ou um complemento ou fragmento dessas. As composições também incluem plantas transformadas e partes de plantas que incluem um cassete de expressão da presente invenção, opcionalmente incorporado de modo estável dentro do genoma da planta. Adicionalmente são apresentados produtos
10 de tabaco, incluindo gomas de mascar de tabaco, tabaco em pó, cigarros, tabaco de cachimbo, e charutos, que possuem um nível reduzido de nornicotina, e sua nitrosamina relacionada, a *N'*-nitrosonornicotina.

Os métodos da invenção abrangem a inibição da expressão ou função de um polipeptídeo do citocromo P450 da presente invenção. Em algumas concretizações,
15 um cassete de expressão abrangendo uma sequência inibidora, que tem como alvo a expressão ou função de um polipeptídeo do citocromo P450 da presente invenção, é introduzida dentro da planta ou parte da planta de interesse, na qual a expressão da sequência inibidora produz um polinucleotídeo ou polipeptídeo que inibe a expressão ou função de um polipeptídeo do citocromo P450 da invenção. Em uma
20 concretização, a sequência inibidora inclui a sequência apresentada nas ID SEQ N°:1, 3, 5, 7, 9, 11, 13, 14, 15, ou 16, ou um complemento ou fragmento dessas.

Os métodos da invenção têm uso na produção da plantas de *Nicotiana* que possuem níveis reduzidos de nornicotina e seu metabólito, a nitrosamina *N'*-nitrosonornicotina, dentro dos tecidos da folha e caule. Quando colhidos, os tecidos
25 da folha e do caule dessas plantas podem ser utilizados para produzir tabaco com níveis reduzidos de nornicotina e sua nitrosamina específica do tabaco, e portanto reduzir o potencial carcinogênico para indivíduos que consomem esses produtos ou são expostos ao fumo secundário derivado desses produtos.

Também são apresentadas plantas transgênicas de *Nicotiana* que têm uma taxa
30 de conversão de nicotina em nornicotina de aproximadamente menos de 2%, sendo que as plantas abrangem um construto de ácido nucléico heterólogo que inclui um promotor capaz de funcionar em uma célula de planta, ligado operacionalmente a uma sequência de ácido nucléico, que tem a primeira sequência de nucleotídeos com um

fragmento de cerca de 100 a 400 nucleotídeos de polinucleotídeo nicotina demetilase de *Nicotiana*, e uma segunda sequência de nucleotídeos capaz de formar um RNA de dupla fita com a primeira sequência de nucleotídeos, sendo que as plantas transgênicas de *Nicotiana* são linhagens transgênicas conversoras de *Nicotiana*. Em
5 algumas concretizações, o polinucleotídeo de *Nicotiana* nicotina demetilase é um polinucleotídeo de nicotina demetilase do tabaco.

A presente invenção também apresenta um construto de ácido nucléico recombinante que abrange um promoter capaz de funcionar em uma célula de planta ligada operacionalmente a uma sequência de ácido nucléico que tem a primeira
10 sequência de nucleotídeos abrangendo um fragmento de cerca de 100 a 400 nucleotídeos de um polinucleotídeo de nicotina demetilase de tabaco e uma segunda sequência de nucleotídeos capaz de formar um RNA de dupla hélice com a primeira sequência de nucleotídeos.

Também são apresentados métodos para redução da conversão de nicotina em
15 normicotina em uma planta de *Nicotiana*, que abrange a transformação de uma planta de *Nicotiana* com um construto de ácido nucléico recombinante incluindo um promoter capaz de funcionar em uma célula de planta, ligado operacionalmente a uma sequência de ácido nucléico, que tem uma primeira sequência de nucleotídeos abrangendo um fragmento de cerca de 100 a 400 nucleotídeos de um polinucleotídeo
20 nicotina demetilase de *Nicotiana* e uma segunda sequência de nucleotídeos capaz de formar um RNA de dupla hélice com a primeira sequência de nucleotídeos; e regenerar uma planta transgênica de *Nicotiana*. Em algumas concretizações, o polinucleotídeo de *Nicotiana* nicotina demetilase é um polinucleotídeo de nicotina demetilase do tabaco.

A presente invenção também apresenta sementes obtidas a partir de plantas transgênicas de *Nicotiana* que têm uma taxa de conversão de nicotina em normicotina de cerca de menos 2%, em que as plantas abrangem um construto de ácido nucléico heterólogo que inclui um promoter capaz de funcionar em uma célula de planta,
25 ligado operacionalmente a uma sequência de ácido nucléico, que tem a primeira sequência de nucleotídeos com um fragmento de cerca de 100 a 400 nucleotídeos de polinucleotídeo nicotina demetilase de *Nicotiana*, e uma segunda sequência de nucleotídeos capaz de formar um RNA de dupla hélice com a primeira sequência de
30 nucleotídeos, sendo que as plantas transgênicas de *Nicotiana* são linhagens

transgênicas conversoras de *Nicotiana*. Em algumas concretizações, o polinucleotídeo de *Nicotiana* nicotina demetilase é um polinucleotídeo de nicotina demetilase do tabaco.

A presente invenção também apresenta células de planta transgênica que abrangem uma molécula de ácido nucléico que tem um promotor funcional em uma célula de planta e
5 uma sequência de ácido nucléico que codifica a nicotina demetilase tendo um resíduo de isoleucina na posição 274 e um resíduo de triptofano na posição 330.

A presente invenção também apresenta métodos de rastreamento para uma sequência de nicotina demetilase que abrange: a obtenção de uma sequência de ácido nucléico que possui mais de 90% de identidade da sequência com SEQ ID NO:5 e identificação de uma
10 sequência de códon que codifica um resíduo de triptofano na posição 330 do polipetídeo codificado.

Adicionalmente são apresentados métodos para rastreamento da nicotina demetilase que têm um isoleucina na posição 274 ou um triptofano na posição 330, abrangendo a obtenção de uma sequência de ácido nucléico que tem mais de 90% de identidade de
15 sequência com SEQ ID N°:5 e a identificação de uma primeira sequência de códon que codifica um resíduo de isoleucina na posição 274, ou uma segunda sequência de códon que codifica um resíduo de triptofano na posição 330 do polipetídeo codificado e ainda, um método para obtenção de uma planta transgênica de *Nicotiana* com uma taxa de conversão de nicotina em normicotina de aproximadamente menos de 2% compreendendo a transformação
20 de uma célula de planta com um construto de ácido nucléico heterólogo, seu cultivo e sua regeneração. E ainda, um método de obtenção de uma semente a partir da planta *Nicotiana* transformada.

BREVE DESCRIÇÃO DAS ILUSTRAÇÕES

25 A Figura 1 mostra as estruturas da nicotina, normicotina, e *N'*-nitrosonornicotina (NNN).

A Figura 2 mostra uma análise de Northern blot dos RNAs conversores e não-conversores usando 7D_A06 como probe hibridização. As bandas 1 e 2 mostram RNAs isolados a partir de folhas tratadas com bicarbonato de sódio de genótipos DH 98-325-5 (não-conversor) e DH 98-325-6 (conversor), respectivamente. As bandas 3 e 4 mostram RNAs
30 isolados a partir de folhas tratadas com etefon de genótipos DH 98-326-3 (não-conversor) e DH 98-326-1 (conversor), respectivamente. O tamanho estimado da banda hibridizada é indicado em kilobases (Kb).

A Figura 3A-3G mostra o alinhamento de uma sequência de nucleotídeos de
35 membros da família de gene 3D_C12. Os asteriscos denotam posições onde a

identidade da sequência é conservada entre todas as sequências comparadas. As posições onde se encontram diferenças são indicadas com hífens e os resíduos correspondentes estão sombreados em cinza. As sequências de nucleotídeos presentes no alinhamento incluem 3D_C12 (SEQ ID Nº:1), 3D_C12-10 (SEQ ID Nº:3);
 5 3D_C12-7 (SEQ ID Nº:5); 7D_A06 (SEQ ID Nº:7); 3D_C12-15 (SEQ ID Nº:9); e 131A_A02 (SEQ ID Nº:11). As entradas 3D_C12-15 e 131A_A02 são sequências de DNAc de comprimento parcial. A região 99 bp de 3D_C12 que foi usada para fazer o construto com base em RNAi está sublinhada.

A Figura 4 mostra um alinhamento das sequências de aminoácidos previstas
 10 para membros de comprimento inteiro da família 3D_C12 de genes P450. As sequências de aminoácidos presentes no alinhamento incluem 3D_C12 (ID SEQ Nº:2), 3D_C12-10 (ID SEQ Nº:4); 3D_C12-7 (ID SEQ Nº:6); e 7D_A06 (ID SEQ Nº:8). Os asteriscos denotam posições conservadas entre todas as quatro sequências. Os resíduos que diferem entre os membros estão sombreados em cinza.

15 A Figura 5 mostra uma análise de Northern blot de plantas transgênicas que possuem o construto 3D_C12/RNAi. (A) Hibridização da sonda 3D_C12-7 em RNAs isolados a partir de folhas de plantas transgênicas curadas e tratadas com etefon mostra fenótipos baixos de normicotina (3D_C12/RNAi-1, 3, e 4) e fenótipos altos de normicotina (3D_C12/RNAi-6, 7, e planta 11 de controle apenas por vetor). O
 20 tamanho estimado da banda hibridizada é indicado em kilobases (Kb). (B) Coloração com brometo de etídio da porção do gel usado em (A) que contém o RNA ribossômico 28S para mostrar a equivalência relativa da carga de RNA entre as bandas.

A Figura 6 mostra a análise de Northern blot de plantas transgênicas que
 25 possuem construtos de orientação senso de membros da família de gene 3D_C12. (A) Hibridização do probe 3D_C12-7 em RNAs isolados a partir de folhas não tratadas de linhagens transgênicas independentes que expressam os construtos 3D_C12-7, 3D_C12, e 7D_A06 e um controle apenas por vetor (controle 8). O tamanho estimado da banda hibridizada é indicado em kilobases (Kb). (B) Coloração com brometo de
 30 etídio da porção do gel usado em (A) que contém o RNA ribossômico 28S para mostrar a equivalência relativa da carga de RNA entre as bandas.

A Figura 7 mostra uma sequência genômica de um fragmento do gene 3D_C12-10 que possui um íntron. As sequências de íntrons estão indicadas em

negrito e itálico. As seqüências éxons são mostradas em tipo simples. As seqüências correspondentes aos primers de PCR (reação em cadeia de polimerase) usadas para amplificar o fragmento a partir do DNA genômico do tabaco estão sublinhadas.

A Figura 8 mostra um diagrama dos construtos de RNAi usados para silenciar a expressão dos membros da família do gene 3D_CD12.

DESCRIÇÃO DETALHADA DA INVENÇÃO

Antecedentes e definições

Antes de descrever a presente invenção detalhadamente, deve ser entendido que muitas modificações e outras concretizações das invenções aqui apresentadas virão à mente de uma pessoa habilitada na técnica a qual esta invenção se refere, tendo o benefício das instruções apresentadas nas descrições seguintes e nas ilustrações associadas. Portanto, deve ser entendido que a invenção não deve se limitar às concretizações específicas reveladas aqui e pretende-se que modificações e outras concretizações sejam incluídas dentro do escopo das reivindicações anexas. Preferivelmente, essas concretizações são apresentadas de modo que esta descoberta satisfaça aos requerimentos legais aplicáveis.

Embora aqui sejam empregados termos específicos, eles são usados apenas no sentido genérico e descritivo e não com o propósito de limitação. Números parecidos se referem aos elementos parecidos em todo o documento. Adicionalmente, o artigo “um” e “uma” são usados aqui se referindo a um ou mais que um (i.e., a no mínimo um) do objeto gramatical do artigo. A fim de exemplo, “um elemento” significa um ou mais elementos. Do início ao fim da especificação entende-se que a palavra “abranger” ou variações como “abrange” ou “abrangendo” implica a inclusão de um elemento declarado, inteiro ou fração, ou grupos de elementos, inteiros ou frações, mas não a exclusão de qualquer outro elemento, inteiro ou fração, ou grupo de elementos, inteiros ou frações.

A presente invenção é representada por composições e métodos para inibir a expressão ou função de polipeptídeos do citocromo P450, que estão envolvidos na conversão metabólica da nicotina em normicotina em uma planta, particularmente plantas do gênero *Nicotiana*, incluindo plantas de tabaco de diversas variedades comerciais. Conforme usado aqui, os termos “inibir”, “inibição”, e “inibindo” são definidos como qualquer método conhecido da técnica ou descrito aqui que reduza a

expressão ou função de um produto do gene de interesse (i.e., o produto do gene alvo) “Inibição” pode estar no contexto de uma comparação entre duas plantas, por exemplo, uma planta geneticamente alterada versus uma planta selvagem. Alternativamente, a inibição da expressão ou função do produto do gene alvo pode

5 estar dentro do contexto de uma comparação entre células, organelas, órgãos, tecidos, ou parte de plantas dentro da mesma planta ou entre plantas diferentes, e inclui comparações entre estágios do desenvolvimento ou temporários dentro da mesma planta, ou parte de planta, ou entre plantas ou partes de planta. “Inibição” inclui qualquer decréscimo relativo da função ou produção de um produto do gene de

10 interesse, até a (e incluindo a) eliminação completa da função ou produção daquele produto de gene. O termo “inibição” inclui qualquer método ou composição que regula para baixo a tradução e/ou transcrição do produto do gene alvo ou da atividade funcional do produto do gene alvo.

O termo “sequência inibidora” inclui qualquer polinucleotídeo ou sequência

15 de polipeptídeo capaz de inibir a expressão ou função de um polipeptídeo do citocromo P450 envolvido na conversão metabólica da nicotina em normicotina em uma planta, tal como um polinucleotídeo de comprimento inteiro ou sequências de polipeptídeos, polinucleotídeos ou sequências de polipeptídeos truncados, fragmentos de polinucleotídeos ou sequências de polipeptídeos, variantes de polinucleotídeos ou

20 sequências de polipeptídeos, sequências de nucleotídeos com orientação senso, sequências de nucleotídeos com orientação antisenso, o complemento de uma sequências de nucleotídeos com orientação senso ou antisenso, regiões invertidas de sequências de nucleotídeos, grampos de sequências de nucleotídeos, sequências de nucleotídeos de dupla fita, sequências de nucleotídeos de única fita, combinações

25 dessas, e semelhantes. O termo “sequência de polinucleotídeos” inclui sequências de RNA, DNA, ácidos nucléicos quimicamente modificados, análogos de ácidos nucléicos, combinações desses, e semelhantes.

Sequências inibidoras são designadas aqui pelo nome do produto do gene alvo. Portanto, uma “sequência inibidora do citocromo P450” se refere a uma sequência

30 inibidora capaz de inibir a expressão de um polipeptídeo de citocromo P450 que está envolvida na conversão metabólica da nicotina em normicotina em uma planta, por exemplo, no nível da transcrição e/ou tradução, ou que seja capaz de inibir a função de esse polipeptídeo do citocromo P450. Quando a frase “capaz de inibir” é usada no

contexto de uma sequência inibidora de polinucleotídeos, isso pretende significar que a sequência inibidora exerce em si o efeito inibidor; ou, que a sequência inibidora codifica uma molécula de nucleotídeo inibidora (por exemplo, um RNA em alça, RNAmi, ou polinucleotídeo de RNA de dupla hélice), ou codifica um polipeptídeo inibidor (i.e, um polipeptídeo que inibe a expressão ou função do produto de gene alvo), após sua transcrição (por exemplo, no caso de uma sequência inibidora que codifica um RNA em alça, RNAmi, ou polinucleotídeo de RNA de dupla hélice) ou sua transcrição e tradução (no caso de uma sequência inibidora que codifica um polipeptídeo inibidor), o produto transcrito ou traduzido, respectivamente, exerce o efeito inibidor sobre o produto do gene alvo (i.e., inibe a expressão ou função do produto do gene alvo).

“Célula hospedeira” significa uma célula de abrange uma sequência heteróloga de ácido nucléico da invenção. Embora as sequências de ácido nucléico da invenção, e fragmentos e variantes dessas, possam ser introduzidas dentro de qualquer célula de interesse, são de interesse particular as células de plantas, mais particularmente as células de espécies da planta *Nicotiana*, por exemplo, as espécies de planta de tabaco e variedades descritas abaixo.

O uso do termo “polinucleotídeo” não pretende limitar a presente invenção a polinucleotídeos que abrangem DNA. Pessoas com habilidade comum na técnica reconhecerão que polinucleotídeos podem abranger ribonucleotídeos e combinações de ribonucleotídeos e desoxirribonucleotídeos. Esses ribonucleotídeos e desoxirribonucleotídeos incluem tanto moléculas de ocorrência natural como análogos sintéticos. Os polinucleotídeos da invenção também incluem todas as formas de sequências, incluindo, mas não se limitando a, formas de única hélice, formas de dupla hélice, grampos, estruturas em grampo, e semelhantes.

O termo “variante” conforme usado aqui pretende significar uma sequência substancialmente similar, e o termo polinucleotídeo ou polipeptídeo “nativo” pretende significar uma sequência de nucleotídeos de ocorrência natural ou sequência de aminoácidos, respectivamente. Por “fragmento” entende-se uma porção de um polinucleotídeo ou uma porção de uma sequência de aminoácidos e então a proteína codificada por meio dela.

Conforme usado aqui, o termo “parte de planta” inclui células de plantas, protoplastos de plantas, culturas de tecidos de células de plantas, a partir das quais

uma planta inteira pode ser regenerada, calos de planta, grupos de plantas, e células de plantas que são intactas em plantas ou partes de plantas, como embriões, pólen, anteros, óvulos, sementes, folhas, flores, caules, ramos, frutos, raízes, extremidades de raízes e semelhantes. A progênie, as variantes e mutantes de plantas regeneradas
5 também estão incluídas dentro do escopo da invenção, desde que essas incluam sequências introduzidas de ácido nucléico da invenção.

Por “mudança fenotípica” entende-se uma mudança mensurável em uma ou mais funções celulares. Por exemplo, plantas que possuem uma modificação genética no locus gênico que codifica um polipeptídio do citocromo P450 da invenção podem
10 mostrar uma expressão ou atividade reduzida ou eliminada daquele polipeptídio do citocromo P450.

O termo “introduzir” significa apresentar à planta o polinucleotídeo ou polipeptídeo de modo que a sequência tenha acesso ao interior da célula da planta.

O termo “operacionalmente ligado” pretende significar uma ligação funcional
15 entre dois ou mais elementos. Por exemplo, uma ligação operacional entre um polinucleotídeo de interesse e uma sequência reguladora (i.e, um promotor) é uma ligação funcional que permite a expressão do polinucleotídeo de interesse. Elementos operacionalmente ligados podem ser contíguos ou não-contíguos. Quando usado para se referir à união de duas regiões que codificam proteína, por operacionalmente
20 ligadas entende-se que as regiões codificadoras estão no mesmo quadro de leitura.

O termo “heterólogo” de acordo com a presente invenção, quando usado em referência a uma sequência, pretende significar uma sequência que se origina a partir de uma espécie exótica, ou se a partir da mesma espécie, é substancialmente modificada a partir de sua forma nativa em composição e/ou locus genômico por
25 intervenção humana deliberada. O termo também é aplicável a construtos de ácido nucléico, também referido aqui como “construtos de polinucleotídeo” ou “construtos de nucleotídeo”. Dessa maneira, um construto de ácido nucléico “heterólogo” pretende significar um construto que se origina de espécies exóticas, ou, se a partir da mesma espécie, é substancialmente modificado a partir de sua forma nativa em
30 composição e/ou locus genômico por intervenção humana deliberada. Construtos de ácido nucléico heterólogo incluem, mas não se limitam a, construtos de nucleotídeos recombinantes que foram introduzidos dentro de uma planta ou parte da planta, por

exemplo, através de métodos de transformação ou cruzamento subsequentes de uma planta transgênica com outra planta de interesse.

Por exemplo, um promoter operacionalmente ligado a um polinucleotídeo heterólogo é de uma espécie diferente das espécies das quais o polinucleotídeo foi derivado, ou, se da mesma espécie ou análoga, uma ou ambas são substancialmente modificadas a partir de sua forma original e/ou locus genômico, ou o promoter não é o promoter nativo para o polinucleotídeo operacionalmente ligado. Além disso, conforme usado aqui o gene quimérico abrange uma sequência codificadora operacionalmente ligada a uma região de iniciação de transcrição que é heteróloga à sequência codificadora.

O termo "expressão" conforme usado aqui se refere à biossíntese de um produto de gene, incluindo a transcrição e/ou tradução desse produto do gene. Por exemplo, para a finalidade da presente invenção, um cassete de expressão, conforme descrito aqui, capaz de expressar um polinucleotídeo que inibe a expressão de no mínimo um polipeptídeo do citocromo P450 da invenção é um cassete de expressão capaz de produzir uma molécula de RNA que inibe a transcrição e/ou tradução de no mínimo um polipeptídeo do citocromo P450. A "expressão" ou "produção" de uma proteína ou polipeptídeo a partir de uma molécula de DNA se refere à transcrição e tradução da sequência codificadora para produzir a proteína ou polipeptídeo, enquanto que a "expressão" ou "produção" de uma proteína ou polipeptídeo a partir de uma molécula de RNA se refere à transcrição e tradução da sequência codificadora de RNA para produzir a proteína ou polipeptídeo.

Polinucleotídeos e polipeptídeos de citocromo P450, e variantes e fragmentos desses

As composições da presente invenção incluem polinucleotídeos e polipeptídeos isolados de citocromo P450 que estão envolvidos na conversão metabólica da nicotina em nornicotina em plantas, incluindo variedades comerciais de plantas de tabaco. Em particular, as composições da invenção incluem polipeptídeos isolados que abrangem sequências de aminoácidos como é mostrado nas SEQ ID N°s: 2, 4, 6, 8, 10, e 12, e polinucleotídeos isolados que abrangem sequências de nucleotídeos como é mostrado nas SEQ ID N°s: 1, 3, 5, 7, 9, e 11. Os polinucleotídeos da invenção têm uso na inibição da expressão desses polipeptídeos do citocromo P450

ou variantes desses, que estão envolvidos na conversão metabólica da nicotina em nornicotina em plantas, particularmente plantas de tabaco.

Dessa maneira, a invenção ainda apresenta cassetes de expressão que abrangem toda ou uma porção da sequência de polinucleotídeos mostrada nas SEQ ID

5 N°:1, 3, 5, 7, 9, ou 11, ou um complemento ou fragmento desses, ou uma sequência que tenha uma identidade de sequência substancial às SEQ ID N°:1, 3, 5, 7, 9, ou 11, ou um complemento ou fragmento dessas, operacionalmente ligadas a um promotor que é funcional em uma célula de planta para uso na expressão de um transcrito de RNA inibidor que interfere na expressão (i.e, transcrição e/ou tradução) de

10 polipeptídeos do citocromo P450 descritos aqui. Em algumas concretizações, os cassetes de expressão abrangem a sequência de nucleotídeos como é mostrado nas SEQ ID N°13, 14, 15, ou 16, ou um complemento um fragmento desses, ou uma sequência que tenha uma identidade de sequência substancial às SEQ ID N°:13, 14, 15, ou 16 ou complemento um fragmento dessas. A introdução desses cassetes de

15 expressão para dentro de uma planta de *Nicotiana* de interesse, especialmente uma planta de tabaco de variedades geralmente conhecidas como variedades de fumo ou brilhante, variedades Burley, variedades escuras, e variedade Oriental/Turca, resulta na produção de plantas de tabaco que têm quantidades reduzidas de nornicotina e nitrosamina, *N'*-nitrosonornicotina (NNN). A folha e o caule dessas plantas

20 transgênicas podem ser usados para produzir uma variedade de produtos de tabaco que possuem níveis reduzidos de nornicotina, e uma redução concomitante desse metabólito de nitrosamina carcinogênico.

Os polinucleotídeos do citocromo P450 e polipeptídeos codificados da presente invenção representam uma nova família de genes do citocromo P450,

25 designada de família de genes do citocromo P450 3D_C12, recentemente identificada por seu papel na conversão metabólica da nicotina em nornicotina em plantas de tabaco. A supressão da expressão de seus produtos de gene codificados nas plantas de tabaco transgênicas resulta em uma redução significativa da acumulação de nornicotina nas folhas dessas plantas transgênicas. Embora ainda não esteja vinculado

30 à teoria, o papel metabólico desses polipeptídeos pode ser direto, i.e., catalisando diretamente a reação de *N*-demetilação, ou indireto, i.e., na forma de produção de um produto que leva à regulação para cima da atividade de demetilase de nicotina da folha. Independentemente desse mecanismo, quaisquer meios pelos quais a expressão

e/ou função dos polipeptídios codificados por membros desta família de genes do citocromo P450 são direcionados para a inibição dentro da planta de *Nicotiana*, e serão efetivos na redução dos níveis de nornicotina, e dos níveis de seus metabólitos carcinogênicos, NNN, dentro de folhas e caules dessas plantas.

5 Os genes do citocromo P450 da invenção foram isolados de linhagens de tabaco de uma variedade do tipo Burley. O primeiro desses genes do citocromo P450, designado 3D_C12, codifica um transcrito de RNAm que corresponde aos nucleotídeos (nt) 1-1551 da sequência de DNAc apresentada na SEQ ID N°:1, que codifica o polipeptídio de comprimento inteiro de 517 resíduos apresentado na SEQ
10 ID N°:2. O segundo membro dessa nova família de citocromo P450, designada 3D_C12-10, codifica um transcrito de RNAm que corresponde aos nucleotídeos (NT) 1-1551 da sequência de DNAc apresentada na SEQ ID N°:3, que codifica o polipeptídeo de comprimento inteiro de 517 resíduos apresentado na SEQ ID N°:4. O terceiro desses genes de citocromo P450, designado 3D_C12-7, codifica um transcrito
15 de RNAm que corresponde aos nt 1-1551 da sequência de DNAc apresentada na SEQ ID N°:5, que codifica o polipeptídeo de comprimento inteiro de 517 resíduos apresentado na SEQ ID N°:6. O quarto membro dessa nova família de citocromo P45, designada 7D_A06, codifica um transcrito de RNAm que corresponde aos nt 1-1554 da sequência de DNAc apresentada na SEQ ID N°:7, que codifica o polipeptídio de
20 comprimento inteiro de 518 resíduos apresentado na SEQ ID N°:8.

Duas sequências de genes P450 de comprimento parcial que compartilham uma alta identidade de sequência para os membros de comprimento inteiro da família do gene 3D_C12 do citocromo P450 também foram isolados a partir dessas linhagens de tabaco Burley. O primeiro desses, designado 3D_C12-15, codifica um transcrito de
25 RNAm que corresponde à sequência de DNAc apresentada na SEQ ID N°:9, que codifica o polipeptídeo de comprimento parcial apresentado na SEQ ID N°:10. a segunda sequência de gene P450 de comprimento parcial, designada 131A_A02, codifica um transcrito de RNAm que corresponde à sequência de DNAc apresentada na SEQ ID N°:11, que codifica o polipeptídeo de comprimento parcial apresentado na
30 SEQ ID N°:12.

Sob um aspecto, os genes do citocromo P450 da presente invenção estão envolvidos na conversão de nicotina em nornicotina em uma planta. Sob um aspecto,

os genes do citocromo P450 da presente invenção têm atividade de nicotina demetilase.

Um alinhamento dos membros da família de genes do citocromo P450 3D_C12 é mostrado na Figura 3A-3G. As sequências de aminoácidos previstas para os clones de comprimento inteiro estão alinhadas na Figura 4. Essas sequências compartilham uma alta identidade de sequência uma com a outra (no mínimo 90% do nível de nucleotídeo e de aminoácido (ver Tabelas 2 e 3, Exemplo 4, mostrados mais adiante)

A invenção abrange composições de polinucleotídeos ou proteínas isoladas ou substancialmente purificadas. Um polinucleotídeo ou proteína "isolada" ou "purificada" ou uma porção biologicamente ativa desses, é substancialmente ou essencialmente livre de componentes que normalmente acompanham ou interagem com o polinucleotídeo ou proteína encontrados em sua ocorrência natural no ambiente. Portanto, um polinucleotídeo ou proteína isolada ou purificada é substancialmente livre de outro material celular, ou meio de cultura quando produzidos por técnicas recombinantes, ou substancialmente livre de precursores ou outros químicos quando quimicamente sintetizado. Da melhor forma, o polinucleotídeo "isolado" é livre de sequências (otimamente sequências codificadoras de proteína) que flanqueiam naturalmente o polinucleotídeo (i.e., sequências localizadas nas extremidades 5' e 3' do polinucleotídeo) no DNA genômico do organismo a partir do qual o polinucleotídeo é derivado. Por exemplo, em várias concretizações, o polinucleotídeo isolado pode conter menos de cerca de 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, ou 0.1 kb de sequência de nucleotídeos que flanqueiam naturalmente o polinucleotídeo no DNA genômico da célula a partir da qual o polinucleotídeo é derivado. Uma proteína que é substancialmente livre de material celular inclui preparações de proteína que têm menos de cerca de 30%, 20%, 10%, 5% ou 1% (por peso seco) de proteína contaminada. Quando a proteína da invenção ou uma porção biologicamente ativa dessa é produzida de maneira recombinante, o meio de cultura representa otimamente menos de cerca de 30%, 20%, 10%, 5%, ou 1% (por peso seco) de precursores químicos ou químicos que não constituem a proteína de interesse.

Fragmentos de polinucleotídeos do citocromo P450 aqui revelados e polipeptídios codificados desse modo também são abrangidos pela presente invenção.

Fragmentos de um polinucleotídeo podem codificar fragmentos de proteína que retêm a atividade biológica da proteína nativa e por isso estão envolvidos na conversão metabólica da nicotina em nornicotina em uma planta. De modo alternativo, fragmentos de um polinucleotídeo que são úteis como probes de hibridização ou primers PCR que usam métodos descritos abaixo, geralmente não codificam fragmentos de proteína que retêm atividade biológica. Além do mais, fragmentos das sequências de nucleotídeos reveladas aqui incluem aqueles que podem ser montados dentro de construtos recombinantes para uso no silenciamento do gene com qualquer método conhecido da técnica, incluindo, mas não se limitando a, supressão/cossupressão do senso, supressão antisense, interferência de RNA de dupla hélice (RNAds), interferência de RNA de alça e interferência de RNA de alça contendo íntron, ribozimas, e RNA ou microRNA de baixa interferência, conforme descrito abaixo. Portanto, fragmentos de uma sequência da polinucleotídeos do citocromo P450 podem variar de no mínimo cerca de 20 nucleotídeos, cerca de 50 nucleotídeos, cerca de 70 nucleotídeos, cerca de 100 nucleotídeos, cerca de 150 nucleotídeos, cerca de 200 nucleotídeos, cerca de 250 nucleotídeos, cerca de 300 nucleotídeos, cerca de 350 nucleotídeos, cerca de 400 nucleotídeos, e até o comprimento inteiro do polinucleotídeo que codifica as proteínas da invenção, dependendo do resultado desejado. Sob um aspecto, os fragmentos da sequência da polinucleotídeos do citocromo P450 podem ser um fragmento com comprimento entre de cerca de 50 e cerca de 400 nucleotídeos, entre cerca de 70 e cerca de 350 nucleotídeos, entre cerca de 90 e de 325 nucleotídeos, entre cerca de 90 e cerca de 300 nucleotídeos, entre cerca de 90 e de 275 nucleotídeos, entre cerca de 100 e cerca de 400 nucleotídeos, entre cerca de 100 e cerca de 350 nucleotídeos, entre cerca de 100 e de 325 nucleotídeos, entre cerca de 100 e cerca de 300 nucleotídeos, entre cerca de 125 e de 300 nucleotídeos, ou entre cerca de 125 e cerca de 275 nucleotídeos. Em certas concretizações, um fragmento de polinucleotídeo do citocromo P450 tem comprimento de cerca de 50, cerca de 60, cerca de 70, cerca de 80, cerca de 90, cerca de 100, de 125, cerca de 150, de 175, cerca de 200, cerca de 225, cerca de 250, cerca de 275, cerca de 300, cerca de 325, cerca de 350, cerca de 400 nucleotídeos, e outros valores como cerca de 70 e cerca de 400 nucleotídeos. Em outra concretização, um fragmento de um polinucleotídeo do citocromo P450 da invenção tem cerca de 90 bp a cerca de 110 bp de comprimento, incluindo 90, 91, 92, 93, 94, 95, 96, 97, 98, 99,

100, 101, 102, 103, 104, 105, 106, 107, 108, 109, e 110 bp de comprimento. Em outra concretização, um fragmento de um polinucleotídeo do citocromo P450 da invenção tem cerca de 290 bp a cerca de 310 bp de comprimento, incluindo 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, e 310
5 bp de comprimento.

Um fragmento de um polinucleotídeo do citocromo P450 da presente invenção que codifica uma porção biologicamente ativa de um polipeptídeo do citocromo P450 da presente invenção codificará no mínimo 15, 25, 30, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, ou 500 aminoácidos contíguos, ou até o número total de
10 aminoácidos presentes em um polipeptídeo do citocromo P450 de comprimento inteiro da invenção (por. ex., 517 para as ID SEQ N°s: 2, 4, e 6; e 518 para a ID SEQ N°:8), ou codificará no mínimo 15, 25, 30, 50, 75, 100, 125, 150, ou até o número total of aminoácidos presentes no polipeptídeo do citocromo P450 de comprimento parcial da invenção (por ex., 173 para ID SEQ N°:10; e 222 para a ID SEQ N°:12).
15 Sob um aspecto, um fragmento de um polinucleotídeo do citocromo P450 da presente invenção codifica uma polipeptídeo que abrange a posição 330 da sequência de polipeptídeo codificada. Sob outro aspecto, o fragmento de polinucleotídeo codifica um fragmento de um polipeptídeo do citocromo P450, sendo que o fragmento do polipeptídeo abrange os aminoácidos da posição 225 até o fim do aminoácido na
20 posição aproximada de 600 da ID SEQ N°:6. Em tal concretização, o fragmento de polinucleotídeo codifica um fragmento de polipeptídeo do citocromo P450, no qual o fragmento de polipeptídeo abrange aminoácidos da posição aproximada 239 até o fim do aminoácido na posição aproximada de 402 da ID SEQ N°:6. Uma porção biologicamente ativa do polipeptídeo do citocromo P450 pode ser preparada através
25 do isolamento de uma porção de um dos polinucleotídeos do citocromo P450 da presente invenção, que expressa uma porção codificada do polipeptídeo do citocromo P450 (por ex., por expressão recombinante *in vitro*), para avaliar a atividade da porção codificada do polipeptídeo do citocromo P450, i.e., a capacidade de promover a conversão da nicotina em nornicotina, usando ensaios conhecidos da técnica e aqueles mostrados mais adiante.
30

Os polinucleotídeos que são fragmentos de uma sequência de nucleotídeos do citocromo P450 da presente invenção abrangem no mínimo 16, 20, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 800, 900, 950, 1000, 1050,

1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, ou 1700 nucleotídeos contíguos, ou até o número de nucleotídeos presentes em um polinucleotídeo do citocromo P450 de comprimento inteiro conforme revelado aqui (por ex., 539 para a ID SEQ N°:9; 666 para a ID SEQ N°:11; 1733 para a ID SEQ N°s:1, 3, e 5; e 1727 para a ID SEQ N°:7). Os polinucleotídeos que são fragmentos de uma sequência de nucleotídeos do citocromo P450 da presente invenção abrangem fragmentos de cerca de 20 a cerca de 1700 nucleotídeos contíguos, de cerca de 50 a cerca de 1600 nucleotídeos contíguos, de cerca de 75 a cerca de 1500 nucleotídeos contíguos, de cerca de 100 a cerca de 1400 nucleotídeos, de cerca de 150 a cerca de 1300 nucleotídeos contíguos, de cerca de 150 a cerca de 1200 nucleotídeos contíguos, de cerca de 175 a cerca de 1100 nucleotídeos contíguos, ou de cerca de 200 a cerca de 1000 nucleotídeos contíguos do polinucleotídeo do citocromo P450, conforme revelado aqui.

Sob um aspecto, os fragmentos do polinucleotídeo do citocromo P450 abrangem uma sequência de polinucleotídeos que contém nucleotídeos da posição aproximada de 700 à posição aproximada de 1250 de uma sequência codificadora do citocromo P450. Sob outro aspecto, os fragmentos do polinucleotídeo do citocromo P450 abrangem uma sequência de polinucleotídeos que contém nucleotídeos da posição aproximada de 715 à posição aproximada de 1210, ou da posição aproximada de 717 à posição aproximada de 1207 de uma sequência codificadora do citocromo P450 revelada aqui.

Em outras concretizações da invenção, os fragmentos do polinucleotídeo do citocromo P450 abrangem uma sequência de polinucleotídeos que contém nucleotídeos da posição aproximada de 265 à posição aproximada de 625 de uma sequência codificadora do citocromo P450 revelada aqui, ou um complemento dessa. Em algumas concretizações, os fragmentos da sequência codificadora do citocromo P450 revelados aqui abrangem os nucleotídeos que correspondem à posição aproximada de 297 à posição aproximada de 594 da sequência codificadora do P450 apresentada na ID SEQ N°:3 ou ID SEQ N°:5, ou um complemento dessas.

Em outras concretizações da invenção, os fragmentos do polinucleotídeo do citocromo P450 abrangem uma sequência de polinucleotídeos que contém nucleotídeos da posição aproximada de 1420 à posição aproximada de 1580 de uma sequência codificadora do citocromo P450 revelada aqui, ou um complemento dessa.

Em algumas dessas concretizações, os fragmentos da sequência codificadora do citocromo P450 revelados aqui abrangem os nucleotídeos que correspondem à posição aproximada de 1453 à posição aproximada de 1551 da sequência codificadora do P450 apresentada na ID SEQ N°:1, ou um complementos dessa.

5 Variantes dos polinucleotídeos e polipeptídeos codificados desse modo também estão abrangidos pela presente invenção. Essas variantes de ocorrência natural incluem variantes que compartilham uma identidade de sequência substancial com os polinucleotídeos e polipeptídeos do citocromo P450 revelados aqui, conforme definido mais adiante. As composições e métodos da invenção podem ser usados para
10 direcionar a expressão ou função de qualquer citocromo P450 de ocorrência natural que compartilha uma identidade de sequência substancial com os polipeptídeos do citocromo P450 revelados aqui, e que possuem a atividade relevante do citocromo P450, i.e., envolvimento na conversão metabólica da nicotina em nornicotina em plantas. Tais variantes podem resultar de, por exemplo, polimorfismo genético ou da
15 manipulação humana, como ocorre com o cruzamento e seleção. Variantes biologicamente ativas de uma proteína do citocromo P450 da invenção, por exemplo, variantes do polipeptídeo apresentadas nas ID SEQ N°:2, 4, 6, 8, 10, ou 12, terão cerca de no mínimo 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais de identidade de
20 sequência com a sequência de aminoácidos da proteína nativa, conforme determinado pelo programas de alinhamento de sequência e parâmetros descritos aqui, e são caracterizadas por seu envolvimento funcional na conversão metabólica da nicotina em nornicotina em plantas. Uma variante biologicamente ativa de uma proteína pode diferir daquela proteína por apenas 1 a 15 resíduos de aminoácidos, apenas 10, apenas
25 9, apenas 8, apenas 7, apenas 6, apenas 5, apenas 4, apenas 3, apenas 2, ou apenas 1 resíduo de aminoácido.

 Variantes de um polinucleotídeo particular da presente invenção incluem aqueles polinucleotídeos de ocorrência natural que codificam um polipeptídeo do citocromo P450, o qual está envolvido na conversão metabólica da nicotina em
30 nornicotina em plantas. Tais variantes de polinucleotídeos podem abranger uma deleção e/ou adição de um ou mais nucleotídeos em um ou mais sítios dentro do polinucleotídeo nativo revelado aqui e/ou uma substituição de um ou mais nucleotídeos em um ou mais sítios no polinucleotídeo nativo. Por causa da

degeneração do código genético, variantes conservadoras para polinucleotídeos incluem aquelas sequências que codificam a sequência de aminoácidos para um dos polipeptídeos do citocromo P450 da invenção. Variantes de ocorrência natural como essas podem ser identificadas com o uso de técnicas bem conhecidas da biologia molecular, como por exemplo, uma reação da cadeia de polimerase (PCR) e técnicas de hibridização, conforme é descrito mais adiante. Polinucleotídeos variantes também incluem polinucleotídeos sinteticamente derivados, como aqueles gerados, por exemplo, pelo uso de mutagênese direcionada ao sítio, mas que ainda compartilham uma identidade de sequência substancial com as sequências de ocorrência natural reveladas aqui e, portanto, podem ser usados métodos da invenção para inibir a expressão ou função de um citocromo P450 envolvido na conversão metabólica de nicotina em nornicotina, incluindo os polipeptídeos do citocromo P450 apresentados nas ID SEQ N°s:2, 4, 6, e 8, e polipeptídeos que abrangem a sequência apresentada na ID SEQ N°:10 ou 12. Geralmente, variantes de um polinucleotídeo particular da invenção, por exemplo, a sequência apresentada nas ID SEQ N°:1, 3, 5, 7, 9, ou 11, terá cerca de no mínimo 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou maior identidade de sequência com aquele polinucleotídeo particular determinado por programas de alinhamento de sequência e parâmetros descritos aqui.

Variantes de um polinucleotídeo particular da presente invenção (também referido como o polinucleotídeo de referência) também podem ser avaliadas pela comparação da percentagem da identidade de sequência entre o polipeptídeo codificado pelo polinucleotídeo de referência, e o polipeptídeo codificado por um polinucleotídeo variante. Por exemplo, são revelados aqui polinucleotídeos isolados que codificam um polipeptídeo com uma percentagem de identidade de sequência com o polipeptídeo de comprimento inteiro das ID SEQ N°:2, 4, 6, ou 8, ou um polipeptídeo de comprimento parcial codificado pela ID SEQ N°:9 ou 11. Tais polinucleotídeos podem ser usados em métodos da presente invenção para direcionar a expressão de polipeptídeos do citocromo P450 envolvidos na conversão metabólica da nicotina em nornicotina em plantas, particularmente plantas de tabaco, inibindo desse modo a acumulação de nornicotina e seu metabólito *N'*-nitrosornicotina nos caules e folhas de uma planta geneticamente modificada. A percentagem de identidade de sequência entre dois polipeptídeos pode ser calculada com o uso de

programas de alinhamento de sequência e parâmetros descritos aqui. Quando qualquer par de polinucleotídeos da invenção é avaliado por comparação da percentagem da identidade de sequência compartilhada pelos dois polipeptídeos que eles codificam, a percentagem da identidade de sequência entre dois polipeptídeos codificados é no
5 mínimo cerca de 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou maior identidade de sequência. Sob um aspecto, um polipeptídeo variante da presente invenção inclui um polipeptídeo que tem um triptofano na posição 330 ou uma isoleucina na posição 274 do polipeptídeo do citocromo P450, ou ambos, i.e., triptofano na posição 330 e
10 isoleucina na posição 274.

Além disso, os polinucleotídeos da invenção podem ser usados para isolar as sequências de citocromo P450 a partir de outros organismos, particularmente outras plantas, mais particularmente outros membros do gênero *Nicotiana*. PCR, hibridização e outros métodos podem ser usados para identificar essas sequências com
15 base na sua homologia em relação às sequências apresentadas aqui. Sequências isoladas com base na identidade de sequência relativa às sequências de nucleotídeos apresentadas aqui ou às variantes e fragmentos dessas são incluídas pela presente invenção. Tais sequências incluem sequências que são ortólogas das sequências reveladas aqui.

De acordo com a presente invenção, “ortólogos” são genes derivados de um gene ancestral comum que são encontrados em diferentes espécies como resultado da especiação. Genes encontrados em diferentes espécies são considerados ortólogos quando sua sequências de nucleotídeos e/ou suas sequências de proteína codificada compartilham no mínimo 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%,
25 95%, 96%, 97%, 98%, 99%, ou uma maior identidade de sequência. As funções dos ortólogos geralmente são muito conservadas entre as espécies. Portanto, polinucleotídeos isolados que codificam um polipeptídeo do citocromo P450 envolvido na conversão metabólica da nicotina em nornicotina e que hibridizam sob condições adversas com as sequências do citocromo P450 reveladas aqui, ou com
30 variantes ou fragmentos dessas, são abrangidos pela presente invenção. Tais sequências podem ser usadas em métodos da presente invenção para inibir a expressão de polipeptídeos do citocromo P450 que estão envolvidos na conversão metabólica da nicotina em nornicotina em plantas.

Usando o PCR, os primers de oligonucleotídeos podem ser desenhados para uso em reações PCR para amplificar as sequências de DNA correspondentes a partir do DNAc ou DNA genômico extraído a partir de qualquer plantas de interesse. Métodos para desenhar primers de PCR e clonagem de PCR geralmente são conhecidos da técnica e são revelados em Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2º ed., Cold Spring Harbor Laboratory Press, Plainview, New York). Innis *et al.*, eds. (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York); Innis e Gelfand, eds. (1995) *PCR Strategies* (Academic Press, New York); e Innis e Gelfand, eds. (1999) *PCR Methods Manual* (Academic Press, New York). Métodos de PCR incluem, mas não se limitam a, métodos que usam primers pareados, primers aninhados, primers específicos simples, primers degenerados, primers gene-específicos, primers vetor-específicos, primers parcialmente mal emparelhados, e semelhantes.

Técnicas de hibridização envolvem o uso de todos ou parte de polinucleotídeos conhecidos como uma sonda que hibridiza seletivamente com outros polinucleotídeos correspondentes presentes em uma população de fragmentos de DNA genômico clonados ou fragmentos de DNAc (i.e. bibliotecas de DNAc ou genômicas) a partir de um organismo escolhido.

A hibridização pode ser realizada sob condições estringentes. Por “condições estringentes” ou “condições de hibridização estringentes” entende-se condições sob as quais uma sonda irá hibridizar com sua sequência alvo com um grau detectavelmente maior do que outras sequências (por ex., no mínimo duas vezes à sequência original). Condições estringentes são dependentes da sequência e serão diferentes sob circunstâncias diferentes. Através do controle da estringência da hibridização e/ou condições de lavagem, as sequências alvo que são 100% complementares a sonda podem ser identificadas (sonda homóloga). De modo alternativo, as condições estringentes podem ser ajustadas para permitir algum mau emparelhamento em sequências, de modo que graus mais baixos de similaridade sejam detectados (sonda heteróloga). Geralmente, uma sonda é menor do que 1000 nucleotídeos de comprimento, otimamente menor do que 500 nucleotídeos de comprimento.

Tipicamente, as condições estringentes serão aquelas nas quais a concentração de sal é menor do que cerca de 1,5 M de concentração de íon Na, tipicamente cerca de 0,01 a 1,0 M íon Na (ou outros sais) em pH 7,0 a 8,3 e temperatura de cerca de no

- mínimo 30°C. Condições estridentes também podem ser alcançadas pela adição de agentes desestabilizadores como a formamida. Condições exemplares de baixa estridência incluem a hibridização com uma solução tampão de 30 a 35% de formamida, 1 M NaCl, 1% SDS (dodecil sulfato de sódio) a 37°C, e uma lavagem em 5 1X a 2X SSC (20X SSC = 3,0 M NaCl/0,3 M de citrato de trisódio) à temperatura de 50 a 55°C. Condições exemplares de moderada estridência incluem a hibridização em 40 a 45% de formamida, 1,0 M NaCl, 1% SDS a 37°C, e uma lavagem em 0,5X a 1X SSC à temperatura de 55 a 60°C. Condições exemplares de alta estridência incluem a hibridização em 50% de formamida, 1,0 M NaCl, 1% SDS a 37°C, e uma 10 lavagem em 0,1X SSC à temperatura de 60 a 65°C. Otimamente, tampões de lavagem podem abranger cerca de 0,1% a cerca de 1% SDS. A duração da hibridização é de geralmente menos de 24 horas, geralmente cerca de 4 a cerca de 12 horas. A duração do tempo de lavagem será de no mínimo um período de tempo suficiente para alcançar o equilíbrio.
- 15 Em uma concretização específica, condições de estridência incluem a hibridização em uma solução contendo 5X SSC, 0,5% SDS, 5X de Denhardt, 0,45 µg/µl de Poli A RNA, 0,45 µg/µl de timo de bezerro e 50% de formamida a 42°C, e no mínimo uma lavagem pós-hibridização em uma solução que inclua cerca d 0,01X SSC a cerca de 1X SSC. A duração da hibridização é de aproximadamente 14 a 16 20 horas.

A especificidade é geralmente a função das lavagens pós-hibridização, sendo os fatores críticos a força iônica e temperatura da solução de lavagem final. Para híbridos DNA-DNA, a T_m pode ser aproximada a partir da equação de Meinkoth e Wahl (1984) *Anal. Biochem.* 138:267-284: $T_m = 81,5^\circ\text{C} + 16,6 (\log M) + 0,41 (\%GC) - 0,61 (\% \text{ form}) - 500/L$; em que M é a molaridade de cátions monovalentes, %GC é a percentagem de nucleotídeos de guanosina e citosina no DNA, % form é a percentagem de formamida na solução de hibridização, e L é o comprimento do híbrido em pares de bases. A T_m é a temperatura (na definição de força iônica e pH) na qual o 50% de uma sequência alvo complementar hibridiza com uma sonda

10 perfeitamente emparelhada. A T_m é reduzida cerca de 1°C para cada 1% de mau emparelhamento; portanto, T_m , hibridização, e/ou condições de lavagem podem ser ajustadas para hibridizar com sequências da identidade desejada. Por exemplo, se forem buscadas sequências com identidade $\geq 90\%$, a T_m pode ser diminuída 10°C . Geralmente, condições estridentes são selecionadas para serem cerca de menos de

15 5°C do ponto de fusão termal (T_m) para a sequência específica e seus complementos em uma força iônica e pH definidos. Contudo, condições severamente estridentes podem utilizar a hibridização e/ou lavagem a 1, 2, 3, ou 4°C abaixo do ponto de fusão termal (T_m); condições moderadamente estridentes podem utilizar a hibridização e/ou lavagem a 6, 7, 8, 9, ou 10°C abaixo do ponto de fusão termal (T_m); condições de

20 baixa estridentes podem utilizar uma hibridização e/ou lavagem a 11, 12, 13, 14, 15, ou 20°C abaixo do ponto de fusão termal (T_m). Usando a equação, as composições de hibridização e lavagem, e a T_m desejada, as pessoas habilitadas na técnica entenderão que as variações na estringência de hibridização e/ou soluções de lavagem são inerentemente descritas. Se o grau desejado de mau emparelhamento resultar em uma

25 a T_m menor do que 45°C (solução aquosa) ou 32% (solução de formamida), é mais adequado aumentar a concentração de SSC de modo que uma temperatura mais alta possa ser usada. Um guia amplo para a hibridização de ácidos nucléicos é encontrado em Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes*, Parte I, Capítulo 2 (Elsevier, New York); e

30 Ausubel *et al.*, eds. (1995) *Current Protocols in Molecular Biology*, Capítulo 2 (Greene Publishing and Wiley-Interscience, New York). Ver Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

As sondas de hibridização podem ser fragmentos do DNA genômico, fragmentos de DNAC, fragmentos de RNA, ou outros oligonucleotídeos, e podem ser rotulados com um grupo detectável como as ^{32}P , ou qualquer outro marcador detectável. Por exemplo, as sondas para hibridização podem ser feitas pela rotulagem de oligonucleotídeos sintéticos com base nas sequências de polinucleotídeos do citocromo P450 da presente invenção. Métodos para a preparação de sondas para hibridização e construção de DNAC e bibliotecas genômicas são geralmente conhecidos da técnica e são revelados em Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

Por exemplo, sequências de polinucleotídeos do citocromo P450 reveladas aqui, ou uma ou mais porções dessas, podem ser usadas como sondas capazes de hibridizar especificamente com polinucleotídeos do citocromo P450 correspondentes e com RNAs mensageiros. Para alcançar uma hibridização específica sob condições variadas, tais sondas incluem sequências que são únicas entre as sequências de polinucleotídeos do citocromo P450, incluindo regiões 5' à montante da sequência codificadora e regiões 3' à jusante da sequência codificadora, e é mais adequado terem um comprimento de cerca de no mínimo 10 nucleotídeos, e ainda mais adequado um comprimento de cerca de no mínimo 20 nucleotídeos. Tais sondas podem ser usadas para amplificar os polinucleotídeos do citocromo P450 correspondentes. Essa técnica pode ser usada para isolar sequências codificadoras adicionais a partir de uma planta desejada, ou como um ensaio diagnóstico para determinar a presença de sequências codificadoras em uma planta. Técnicas de hibridização incluem o rastreamento de hibridização de bibliotecas de DNA plaqueadas (sejam placas ou colônias; ver, por exemplo, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

Conforme usado aqui, a respeito dos relacionamentos de sequência entre dois ou mais polinucleotídeos ou polipeptídeos, o termo "sequência de referência" é uma sequência definida usada como base para comparação de sequências. Uma sequência de referência pode ser um subconjunto ou a totalidade de uma sequência especificada; por exemplo, como um segmento de um DNAC de comprimento inteiro ou sequência de genes, ou o DNAC completo ou sequência de genes.

Conforme usado aqui, o termo "janela de comparação" faz referência a um segmento contíguo e especificado de uma sequência de polinucleotídeos, em que a sequência de polinucleotídeos na janela de comparação pode abranger adições ou deleções (i.e., gaps) comparados com a sequência de referência (que não abrange adições ou deleções) para o alinhamento ótimo dos dois polinucleotídeos. Geralmente, a janela de comparação tem no mínimo 20 nucleotídeos contíguos de comprimento, e opcionalmente pode ter 30, 40, 50, 100, ou mais. As pessoas habilitadas na técnica entenderão que para evitar uma alta similaridade com uma sequência de referência, devido à inclusão de gaps na sequência de polinucleotídeos, geralmente é introduzida uma penalidade por gap, e é subtraída do número de emparelhamentos.

Métodos para o alinhamento de sequências para comparação são bem conhecidos da técnica. Portanto, a determinação da percentagem da identidade de sequência entre duas sequências quaisquer pode ser realizada com o uso de um algoritmo matemático. Exemplos não limitantes de tais algoritmos matemáticos são o algoritmo de Myers e Miller (1988) *CABIOS* 4:11-17; o algoritmo de alinhamento local de Smith *et al.* (1981) *Adv. Appl. Math.* 2:482 o algoritmo de alinhamento global de Needleman e Wunsch (1970) *J. Mol. Biol.* 48:443-453; o método de alinhamento de busca-do-local de Pearson e Lipman (1988) *Proc. Natl. Acad. Sci.* 85:2444-2448; o algoritmo de Karlin e Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264, modificado em Karlin e Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877.

Os programas BLAST de Altschul *et al.* (1990) *J. Mol. Biol.* 215:403 são baseados no algoritmo de Karlin e Altschul (1990) *supra*. As buscas de nucleotídeos do BLAST podem ser realizadas com o programa BLASTN, *escore* = 100, comprimento de palavra = 12, para obter sequências de nucleotídeos homólogos a uma sequência de nucleotídeos que codifica uma proteína da invenção. As buscas de proteína do BLAST podem ser realizadas com o programa BLASTX, *escore* = 50, comprimento de palavra = 3, para obter sequências de aminoácidos homólogos a uma proteína ou polipeptídeo da invenção. Para obter alinhamentos em gap para fins de comparação, o Gapped BLAST (em BLAST 2.0) pode ser utilizado conforme descrito em Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389. De modo alternativo, o PSI-BLAST (em BLAST 2.0) pode ser usado para realizar uma busca iterativa que detecta relacionamentos distantes entre moléculas. See Altschul *et al.* (1997) *supra*. Quando da utilização do BLAST, Gapped BLAST, e PSI-BLAST, os parâmetros padrão dos

respectivos programas (por ex., BLASTN para sequências de nucleotídeos, BLASTX para proteínas) podem ser usados (Ver www.ncbi.nlm.nih.gov). O alinhamento também pode ser realizado manualmente por inspeção.

Os valores de identidade/similaridade de sequência fornecidos aqui foram calculados com o uso do BLASTX (Altschul *et al.* (1997) *supra*), Clustal W (Higgins *et al.* (1994) *Nucleic Acids Res.* 22: 4673-4680), e os algoritmos GAP (pacote de software do Genetic Computing Group da Universidade de Wisconsin) com o uso de parâmetros padrão. A presente invenção também abrange o uso de qualquer programa equivalente para análise e comparação de sequências de ácido nucléico e proteína. Por “programa equivalente” entende-se qualquer programa de comparação de sequência que, para duas sequências quaisquer em questão, é gerado um alinhamento que tem emparelhamento idêntico de nucleotídeos ou resíduos de aminoácidos, e uma percentagem de identidade de sequência idêntica quando comparados ao alinhamento correspondente gerado pelo BLASTX, Clustal W, ou GAP.

Para fins da discussão precedente sobre sequências de nucleotídeos e polipeptídeos variantes abrangidos pela presente invenção, o termo “identidade de sequência” ou “identidade” no contexto de duas sequências de polinucleotídeos e polipeptídeos se refere aos resíduos das duas sequências, que são os mesmos quando alinhados para a correspondência máxima em uma janela de comparação especificada. Quando a percentagem de identidade de sequência é usada em relação a proteínas, reconhece-se que as posições de resíduo que não são geralmente idênticas diferem por meio da substituição de aminoácidos conservadores, sendo que os resíduos de aminoácidos são substituídos por outros resíduos de aminoácidos com propriedades químicas semelhantes (por ex., carga ou hidrofobicidade), e portanto, não modificam as propriedades funcionais da molécula. Quando as sequências diferem por substituições conservadoras, a percentagem de identidade de sequência pode ser ajustada para cima para corrigir a natureza conservadora da substituição. Diz-se que as sequências que diferem por tais substituições conservadoras têm “similaridade de sequência” ou “similaridade”. Meios para fazer esses ajustes são bem conhecidos por aqueles habilitados na técnica. Isso envolve tipicamente pontuar uma substituição conservadora como parcial em vez de um emparelhamento completo, aumentando dessa forma a percentagem da identidade de sequência. Portanto, por exemplo, quando se dá a um aminoácido idêntico um escore de 1 e à substituição conservadora

um escore de zero, dá-se à substituição conservadora um escore entre zero e 1. A contagem de substituições conservadoras é calculada, por ex., conforme implementado no programa PC/GENE (Intelligenetics, Mountain View, California).

O termo “percentagem de identidade de sequência” usado aqui significa o
5 valor determinado pela comparação entre duas sequências otimamente alinhadas sobre uma janela de comparação, em que a porção da sequência de polinucleotídeos na janela de comparação pode incluir adições ou deleções (i.e, gaps) quando comparada com a sequência de referência (que não inclui adições ou deleções) para o alinhamento ótimo das duas sequências. A percentagem é calculada através da
10 determinação do número de posições nas quais a base de ácido nucléico idêntico ou resíduo de aminoácido ocorre em ambas as sequências, para dar o número de posições emparelhadas, dividindo o número de posições emparelhadas pelo total do número de posições na janela de comparação, e multiplicando o resultado por 100 para obter a percentagem da identidade de sequência.

15 Quando dois polipeptídeos quaisquer estão otimamente alinhados para comparação, reconhece-se que os resíduos que aparecem opostos uns aos outros dentro do alinhamento ocupam posições dentro de seus respectivos polipeptídeos que correspondem uma a outra. Tais posições são chamadas aqui de “posições correspondentes” e os resíduos que residem nas posições correspondentes são
20 chamados de “resíduos correspondentes” ou resíduos que “correspondem” uns aos outros. Portanto, por exemplo, quando um polipeptídeo de interesse está otimamente alinhado a uma sequência de polipeptídeo de referência, que tenha por exemplo, 10 resíduos, o resíduo dentro do polipeptídeo de interesse que aparece oposto ao resíduo 5 da sequência de referência é chamado de “resíduo na posição correspondente ao
25 resíduo 5” da sequência de referência.

De modo semelhante, quando duas sequências de polinucleotídeos estão otimamente alinhadas para comparação, reconhece-se que os nucleotídeos que aparecem opostos uns aos outros dentro do alinhamento ocupam posições dentro de suas respectivas posições de polinucleotídeos que correspondem uma a outra. Tais
30 posições são chamadas aqui de “posições correspondentes” e os nucleotídeos que residem nas posições correspondentes são chamados de “nucleotídeos correspondentes” ou nucleotídeos que “correspondem” um ao outro. Portanto, por exemplo, quando um polinucleotídeo de interesse está otimamente alinhado a uma

sequência de polinucleotídeos de referência que tenha, por exemplo, 300 nucleotídeos, o nucleotídeo dentro do polinucleotídeo de interesse que aparece oposto ao nucleotídeo 275 da sequência de referência é chamado de “nucleotídeo na posição correspondente ao nucleotídeo 275” da sequência de referência .

- 5 Quando uma região de nucleotídeos está sendo comparada entre um polinucleotídeo de interesse e um polinucleotídeo de referência, diz-se que os nucleotídeos dentro dessas regiões “correspondem” um ao outro. Portanto, por exemplo, quando uma região de uma sequência de polinucleotídeos de referência, por exemplo, a sequência de polinucleotídeos apresentada na ID SEQ N°:5, reside na
- 10 posição de nucleotídeo 265 à posição de nucleotídeo 625 do polinucleotídeo de referência, e essa região de nucleotídeos está sendo comparada à região correspondente dos nucleotídeos dentro de uma sequência de polinucleotídeos de interesse otimamente alinhada, os nucleotídeos dentro da região correspondente do polinucleotídeo de interesse são chamados aqui de uma “região” do polinucleotídeo
- 15 de interesse que “corresponde à posição de nucleotídeo 265 à posição de nucleotídeo 625” da sequência de referência, nesse caso, a ID SEQ N°:5.

As sequências de polinucleotídeos e polipeptídeos do citocromo P450 podem ser identificadas usando as sequências apresentadas aqui. Tais métodos incluem a obtenção de uma sequência de polinucleotídeos ou polipeptídeos com no mínimo

20 80%, 85%, 90%, 95%, 98%, 99% de identidade de sequência com a sequência de polinucleotídeos das IDs SEQ N° 1, 3, 5, 7, 9, ou 11 ou um complemento ou fragmento dessas, ou uma sequência de polipeptídeos das ID SEQ N°:2, 4, 6, 7, 10, ou 12. Em uma concretização, a sequência identificada contém ou codifica um resíduo de triptofano na posição 330 e um resíduo de isoleucina na posição 274.

- 25 Dessa maneira, um aspecto da presente invenção é direcionado para um método de rastreamento de uma sequência de nicotina demetilase. Esse método abrange a obtenção de uma sequência de ácido nucléico que tem cerca de mais de 90% de identidade de sequência com ID SEQ N°:3 ou ID SEQ N°:5, e a identificação, dentro dessa sequência de ácido nucléico, de uma sequência de códon que codifica um
- 30 resíduo de triptofano na posição 330 do polipeptídeo codificado pela sequência de ácido nucléico. Em algumas concretizações, essa sequência de ácido nucléico codifica uma isoleucina na posição 274 do polipeptídeo codificado. Qualquer método adequado conhecido da técnica pode ser usado para identificar a sequência de códons

que codifica o resíduo de triptofano ou isoleucina. Em uma concretização, a sequência de códons é identificada por um método selecionado a partir de um grupo que consiste na identificação de um polimorfismo de nucleotídeo simples e RT-PCR. Em algumas concretizações desse aspecto da invenção, o método de rastreamento identifica uma
5 nicotina demetilase que converte a nicotina em nornicotina a uma taxa de cerca de no mínimo 5 vezes maior do que a taxa de conversão da nicotina demetilase codificada pela ID SEQ N°:3 (polipeptídeo apresentado na ID SEQ N°:4) ou ID SEQ N°:5 (polipeptídeo apresentado na ID SEQ N°:6). Em certas concretizações, a nicotina demetilase identificada com esse método de rastreamento tem uma taxa de conversão
10 que é cerca de no mínimo 8 vezes maior do que a taxa de conversão da nicotina demetilase codificada pela ID SEQ N°:3 ou ID SEQ N°:5.

Sob outro aspecto, a presente invenção apresenta um método para rastreamento de uma nicotina demetilase que tem uma isoleucina na posição 274 ou um triptofano na posição 329 do polipeptídeo. Esses métodos abrangem a obtenção de
15 uma sequência de ácido nucléico que tem cerca de mais de 90% de identidade de sequência com ID SEQ N°:3 ou ID SEQ N°:5; e a identificação de uma primeira sequência de códons para um resíduo de isoleucina na posição 274 do polipeptídeo codificado ou uma segunda sequência de códons que codifica um resíduo de triptofano na posição 330 do polipeptídeo codificado. Como foi observado acima,
20 qualquer método adequado conhecido de pessoas habilitadas na técnica pode ser usado para identificar esses códons. Em uma concretização, o primeiro ou segundo códon, ou ambos, são identificados por um método selecionado a partir de um grupo que consiste na identificação de um polimorfismo de nucleotídeo simples e RT-PCR. Em algumas concretizações desse aspecto da invenção, o método de rastreamento
25 identifica uma nicotina demetilase que converte a nicotina em nornicotina a uma taxa de cerca de no mínimo 5 vezes maior do que a taxa de conversão da nicotina demetilase codificada pela ID SEQ N°:3 (polipeptídeo apresentado na ID SEQ N°:4) ou ID SEQ N°:5 (polipeptídeo apresentado na ID SEQ N°:6). Em certas concretizações, a nicotina demetilase identificada com esse método de rastreamento
30 tem uma taxa de conversão que é cerca de no mínimo 8 vezes maior do que a taxa de conversão da nicotina demetilase codificada pela ID SEQ N°:3 ou ID SEQ N°:5.

A presente invenção também apresenta células de plantas, plantas e sementes transgênicas que abrangem uma molécula de ácido nucléico que tem um promotor

funcional em uma célula de planta, e uma sequência de ácido nucléico que codifica uma nicotina demetilase que tem um resíduo de isoleucina na posição 274 e um resíduo de triptofano na posição 330. Em algumas concretizações, a sequência de ácido nucléico que codifica essa nicotina demetilase é derivada de uma sequência selecionada do grupo que consiste das ID SEQ N°:3, ID SEQ N°:5, e ID SEQ N°:7. Essas células de plantas, plantas e sementes transgênicas incluem, mas não se limitam a, células de plantas de *Nicotiana*, plantas de *Nicotiana*, e sementes de plantas de *Nicotiana*. Em algumas concretizações, as células de plantas de *Nicotiana*, plantas de *Nicotiana*, e sementes de plantas de *Nicotiana* são de plantas de *Nicotiana* conversoras.

Cassetes de Expressão para Uso em Métodos da Invenção

Composições da presente invenção ainda incluem cassetes de expressão que abrangem sequências inibidoras capazes de inibir a expressão ou função de um polipeptídeo de um citocromo P450 envolvido na conversão da nicotina em nornicotina em uma planta de *Nicotiana* ou parte dessa planta, em que as sequências inibidoras estão operacionalmente ligadas a um promotor que é funcional em uma célula de planta. Dessa forma, são construídos cassetes de expressão que abrangem toda ou parte da sequência apresentada nas ID SEQ N°:1, 3, 5, 7, 9, ou 11, um complemento ou fragmento dessas, ou sequências que compartilham uma identidade de sequência substancial com as ID SEQ NO:1, 3, 5, 7, 9, 11, ou um complemento ou fragmento dessas, operacionalmente ligadas a um promotor que é funcional em uma célula de planta para uso em métodos de silenciamento de genes da presente invenção. Tais sequências são chamadas aqui de “sequências inibidoras” ou “sequências de polinucleotídeos inibidoras” conforme elas sejam capazes de serem expressas como uma molécula de RNA que inibe a expressão (i.e., transcrição e/ou tradução) do polipeptídeo alvo do citocromo P450, por exemplo, os polipeptídeos apresentados nas ID SEQ N°:2, 4, 6, ou 8 e variantes dessas, ou um polipeptídeo que inclui a sequência apresentada na ID SEQ N°:10 ou 12 e variantes dessas, em que os polipeptídeos variantes têm um identidade de sequência substancial com esses polipeptídeos do citocromo P450 revelados aqui, e estão envolvidos na conversão metabólica da nicotina em nornicotina em uma planta.

Como observado acima, essas sequências inibidoras incluem sequências de fragmentos dos polinucleotídeos alvo do citocromo P450. Por exemplo, uma sequência de fragmento pode incluir qualquer porção da sequência do citocromo P450, incluindo a sequência codificadora e não-codificadora (*por ex.*, as sequências 5' UTR, íntron, e 3' UTR), e pode incluir fragmentos entre cerca de 20 e cerca de 400 nucleotídeos, entre cerca de 50 e cerca de 400 nucleotídeos, entre cerca de 100 e cerca de 400 nucleotídeos, entre cerca de 125 e cerca de 325 nucleotídeos, entre cerca de 125 e cerca de 300 nucleotídeos, ou entre cerca de 125 e cerca de 275 nucleotídeos.

Dessa maneira, tais sequências inibidoras incluem, mas não se limitam a, um fragmento de uma sequência da polinucleotídeos do citocromo P450 que varia de no mínimo cerca de 20 nucleotídeos, cerca de 50 nucleotídeos, cerca de 70 nucleotídeos, cerca de 100 nucleotídeos, cerca de 150 nucleotídeos, cerca de 200 nucleotídeos, cerca de 250 nucleotídeos, cerca de 300 nucleotídeos, cerca de 350 nucleotídeos, cerca de 400 nucleotídeos, e até o comprimento inteiro do polinucleotídeo que codifica as proteínas da invenção, dependendo do resultado desejado. Sob um aspecto, as sequências inibidoras abrangem um fragmento da sequência da polinucleotídeo do citocromo P450 que possui entre cerca de 50 e cerca de 400 nucleotídeos, entre cerca de 70 e cerca de 350 nucleotídeos, entre cerca de 90 e cerca de 325 nucleotídeos, entre cerca de 90 e cerca de 300 nucleotídeos, entre cerca de 90 e cerca de 275 nucleotídeos, entre cerca de 100 e cerca de 400 nucleotídeos, entre cerca de 100 e cerca de 350 nucleotídeos, entre cerca de 100 e cerca de 325 nucleotídeos, entre cerca de 100 e cerca de 300 nucleotídeos, entre cerca de 125 e cerca de 300 nucleotídeos, ou entre cerca de 125 e cerca de 275 nucleotídeos de comprimento. Em certas concretizações, as sequências inibidoras abrangem um fragmento de um polinucleotídeo do citocromo P450 que tem cerca de 50, cerca de 60, cerca de 70, cerca de 80, cerca de 90, cerca de 100, cerca de 125, cerca de 150, cerca de 175, cerca de 200, cerca de 225, cerca de 250, cerca de 275, cerca de 300, cerca de 325, cerca de 350, cerca de 400 nucleotídeos de comprimento, e outros valores entre cerca de 50 e cerca de 400 nucleotídeos. Admite-se que a sequência inibidora também pode abranger uma sequência que é complementar a todo ou parte do fragmento da sequência de polinucleotídeo do citocromo P450.

Em uma concretização, as sequências inibidoras abrangem um fragmento de um polinucleotídeo de um citocromo P450 da invenção que tem cerca de 90 pb (pares

de bases) a cerca de 110 pb de comprimento, incluindo 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, e 110 pb de comprimento, e também podem abranger uma sequência que é complementar a toda ou parte da sequência do fragmento. Em outra concretização, um fragmento de um polinucleotídeo do citocromo P450 da invenção tem cerca de 290 pb a cerca de 310 pb de comprimento, incluindo 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, e 310 bp de comprimento, e também pode abranger uma sequência que é complementar a toda ou parte da sequência do fragmento.

5
10 Em outras concretizações da invenção, a sequência inibidora dentro de um cassete de expressão da invenção abrange uma sequência de polinucleotídeos que contém os nucleotídeos na posição aproximada de 265 à posição aproximada de 625 de uma sequência codificadora do citocromo P450 revelada aqui, e uma sequência que é completa ou parcialmente complementar a essa. Em algumas concretizações, a
15 sequência inibidora abrange os nucleotídeos correspondentes na posição aproximada de 297 à posição aproximada de 594 da sequência codificadora de P450 apresentada na ID SEQ N°:3 ou ID SEQ N°:5 e uma sequência que é completa ou parcialmente complementar a essa. Nas concretizações preferidas, essas sequências inibidoras são expressas como um RNA em alça conforme descrito mais adiante.

20 Ainda em outras concretizações da invenção, a sequência inibidora dentro de um cassete de expressão da invenção abrange uma sequência de polinucleotídeos que contém os nucleotídeos na posição aproximada de 1420 à posição aproximada de 1580 de uma sequência codificadora do citocromo P450 revelada aqui e uma sequência que é completa ou parcialmente complementar a essa. Em algumas
25 concretizações, a sequência inibidora abrange os nucleotídeos correspondentes na posição aproximada de 1453 à posição

aproximada de 1551 da sequência codificadora apresentada na ID SEQ Nº:1 e uma sequência que é completa ou parcialmente complementar a essa. Nas concretizações preferidas, essas sequências inibidoras são expressas como um RNA em alça conforme descrito mais adiante.

- 5 Admite-se que os cassetes de expressão da presente invenção incluem construtos nos quais a sequência de ácido nucléico desejada está operacionalmente ligada a um promotor que é funcional em uma célula de planta, particularmente na célula de uma planta de *Nicotiana*. Por “promotor” entende-se uma região reguladora de DNA que abrange geralmente uma caixa TATA capaz de conduzir a RNA
- 10 polimerase II a iniciar a síntese de RNA no sítio de iniciação de transcrição apropriado para uma sequência de nucleotídeo particular. Um promotor pode abranger adicionalmente outras sequências de reconhecimento posicionadas geralmente à montante ou 5’ para a caixa TATA, referidas como elementos de promotor à montante, que influenciam a taxa de iniciação de transcrição. Também se admite que
- 15 os cassetes de expressão da presente invenção incluem domínios adicionais que modulam o nível de expressão, tempo de expressão, ou tipo de tecido em que a expressão ocorre em (por ex., Patente Australiana Nº. AU-A-77751/94 e Patente dos EUA Nºs. 5.466.785 e 5.635.618). Por “funcional” entende-se que o promotor, quando operacionalmente ligado a uma sequência inibidora que codifica uma
- 20 molécula inibidora de nucleotídeo (por exemplo, um RNA em alça, RNA dupla fita, polinucleotídeo de RNAmi, e semelhantes), o promotor é capaz de iniciar a transcrição da sequência inibidora operacionalmente ligada de modo que a molécula do nucleotídeo inibidor seja expressa. Os promotores podem ser selecionados com base no resultado desejado. Os ácidos nucléicos podem ser combinados com
- 25 promotores constitutivos, preferivelmente de tecidos, ou outros para a expressão em plantas.

- Um cassete de expressão da presente invenção pode também conter no mínimo um gene adicional para ser cotransformado dentro da planta. De modo alternativo, o(s) gene(s) adicionais pode(m) ser fornecidos em cassetes de expressão
- 30 múltipla. Tal cassete de expressão é fornecido com uma pluralidade de sítios de restrição e/ou sítios de recombinação para que a inserção da sequência inibidora de polinucleotídeos do citocromo P450 ocorra sob a regulação transcricional das regiões

reguladoras. O cassete de expressão pode conter adicionalmente genes marcadores selecionáveis.

Desse modo, um cassete de expressão da presente invenção inclui uma região de iniciação transcricional e traducional (i.e., um promotor) na direção de transcrição 5'-3', uma sequência inibidora conforme descrito aqui, e uma região terminal transcricional e traducional (i.e., região terminal) funcional em uma célula de planta. As regiões reguladoras (i.e, promotores, regiões reguladoras transcricionais, e regiões reguladoras traducionais) e/ou a sequência inibidora da invenção podem ser nativas/análogas à célula hospedeira ou uma a outra. De modo alternativo, as regiões reguladoras e/ou a sequência inibidora da invenção podem ser heterólogas à célula hospedeira ou uma a outra. Apesar dos promotores heterólogos poderem ser usados para expressar sequências inibidoras da invenção, sequências de promotor nativas também podem ser usadas.

A região terminal pode ser nativa para a região de iniciação transcricional, pode ser nativa para a sequência inibidora operacionalmente ligada da invenção, pode ser nativa para a planta hospedeira, ou pode ser derivada de outra fonte (i.e., externa ou heteróloga ao promotor, à sequência inibidora de interesse, à planta hospedeira, ou qualquer combinação dessas). Regiões terminais convenientes estão disponíveis a partir do plasmídeo-Ti de *A. tumefaciens*, tais como regiões terminais de octopina sintase e nopalina sintase. Ver também Guerineau *et al.* (1991) *Mol. Gen. Genet.* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon *et al.* (1991) *Genes Dev.* 5:141-149; Mogen *et al.* (1990) *Plant Cell* 2:1261-1272; Munroe *et al.* (1990) *Gene* 91:151-158; Ballas *et al.* (1989) *Nucleic Acids Res.* 17:7891-7903; e Joshi *et al.* (1987) *Nucleic Acids Res.* 15:9627-9639.

Os cassetes de expressão da presente invenção podem conter adicionalmente sequências 5' líderes que podem atuar para intensificar a tradução. Líderes de tradução são conhecidos da técnica e incluem: líderes picornavírus, por exemplo, líder EMCV (região 5' não codificadora de Encefalomiocardite) (Elroy-Stein *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6126-6130); líderes potivírus, por exemplo, líder TEV (Tobacco Etch Vírus, em português, Vírus de Corrosão de Tabaco) (Gallie *et al.* (1995) *Gene* 165(2):233-238), líder MDMV (Maize Dwarf Mosaic Virus, em português, Vírus do Mosaico Anão do Milho) (*Virology* 154:9-20), e proteína de ligação de cadeia pesada da imunoglobulina (BiP) (Macejak *et al.* (1991) *Nature*

353:90-94); líder não traduzido da proteína de RNAm de cabra do vírus do mosaico da alfafa (AMV RNA 4) (Jobling *et al.* (1987) *Nature* 325:622-625); vírus líder do mosaico do tabaco (TMV, tobacco mosaic virus) (Gallie *et al.* (1989) in *Molecular Biology of RNA*, ed. Cech (Liss, New York), pp. 237-256); e vírus líder do milho mosqueado clorótico (MCMV, maize chlorotic mottle virus) (Lommel *et al.* (1991) *Virology* 81:382-385). Ver também, Della-Cioppa *et al.* (1987) *Plant Physiol.* 84:965-968. Outros métodos conhecidos para intensificar a tradução também podem ser utilizados.

Na preparação do cassete de expressão, fragmentos de DNA da invenção
 10 podem ser manipulados de modo a preparar as sequências de DNA na orientação adequada e, conforme apropriado, no quadro de leitura adequado. Adaptadores ou ligadores podem ser empregados para juntar os fragmentos de DNA ou podem estar envolvidas outras manipulações para preparar sítios de restrição convenientes, remoção de DNA supérfluo, remoção de sítios de restrição, ou semelhantes. Para esse
 15 fim, podem estar envolvidos a mutagênese, reparo de primer, restrição, recozimento, resubstituições, por ex. transições e transversões.

Os cassetes de expressão da presente invenção também podem abranger um gene marcador selecionável para seleção de células transformadas. Os genes marcadores selecionáveis são utilizados para expressão de células ou tecidos transformados. Genes
 20 marcadores incluem genes que codificam a resistência a antibióticos, como aqueles que codificam neomicina fosfotransferase II (NEO) e higromicina fosfotransferase (HPT), bem como genes que conferem resistência a compostos herbicidas, tais como o glufosinato de amônio, bromoxinil, imidazolinonas, e 2,4-diclorofenoxyacetato (2,4-D). Marcadores selecionáveis adicionais incluem marcadores fenotípicos como a β -
 25 galactosidase e proteínas fluorescentes como a proteína fluorescente verde (green fluorescent protein, GFP) (Rouwendaal *et al.* (1997) *Plant Mo. Biol.* 33:989-999; Su *et al.* (2004) *Biotechnol Bioeng* 85:610-9 e Fetter *et al.* (2004) *Plant Cell* 16:215-28), proteína fluorescente ciano (cyan florescent protein, CYP) (Bolte *et al.* (2004) *J. Cell Science* 117:943-54 e Kato *et al.* (2002) *Plant Physiol* 129:913-42), e proteína
 30 fluorescente amarela (yellow fluorescent protein, PhiYFP™) da Evrogen, ver, Bolte *et al.* (2004) *J. Cell Science* 117:943-954). Para marcadores selecionáveis adicionais, ver Yarranton (1992) *Curr. Opin. Biotech.* 3:506-511; Christopherson *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6314-6318; Yao *et al.* (1992) *Cell* 71:63-72; Reznikoff (1992)

- Mol. Microbiol.* 6:2419-2422; Barkley *et al.* (1980) em *The Operon*, pp. 177-220; Hu *et al.* (1987) *Cell* 48:555-566; Brown *et al.* (1987) *Cell* 49:603-612; Figge *et al.* (1988) *Cell* 52:713-722; Deuschle *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:5400-5404; Fuerst *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:2549-2553; Deuschle *et al.* (1990) *Science* 248:480-483; Gossen (1993) Tese de Ph.D., University of Heidelberg; Reines *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:1917-1921; Labow *et al.* (1990) *Mol. Cell. Biol.* 10:3343-3356; Zambretti *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3952-3956; Baim *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:5072-5076; Wyborski *et al.* (1991) *Nucleic Acids Res.* 19:4647-4653; Hillenand-Wissman (1989) *Topics Mol. Struc. Biol.* 10:143-162; Degenkolb *et al.* (1991) *Antimicrob. Agents Chemother.* 35:1591-1595; Kleinschmidt *et al.* (1988) *Biochemistry* 27:1094-1104; Bonin (1993) Tese de Ph.D., University of Heidelberg; Gossen *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Oliva *et al.* (1992) *Antimicrob. Agents Chemother.* 36:913-919; Hlavka *et al.* (1985) *Handbook of Experimental Pharmacology*, Vol. 78 (Springer-Verlag, Berlin); Gill *et al.* (1988) *Nature* 334:721-724. Tais descobertas estão incorporadas aqui por referência. A lista de genes marcadores selecionáveis acima não pretende ser limitante. Qualquer gene marcador selecionável pode ser usado na presente invenção.

Vários promotores podem ser usados na prática da invenção. De interesse particular são os promotores constitutivos, promotores induzíveis, particularmente ou promotores quimicamente induzíveis, e promotores que preferem tecidos, particularmente promotores que preferem folhas.

Os promotores quimicamente induzíveis podem ser usados para inibir a expressão de um citocromo P450 que está envolvido na conversão metabólica da nicotina à normicotina em uma planta através da aplicação de um regulador químico exógeno. Os promotores quimicamente induzíveis são conhecidos da técnica e incluem, mas não se limitam a, promotor PR-1a do tabaco, que é ativado pelo ácido salicílico. Outros promotores quimicamente induzíveis de interesse incluem os promotores responsivos a esteróides (ver, por exemplo, o promotor induzível por glucocorticóide em Schena *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:10421-10425 e McNellis *et al.* (1998) *Plant J.* 14(2):247-257) e promotores induzíveis por tetraciclina (ver, por exemplo, Gatz *et al.* (1991) *Mol. Gen. Genet.* 227:229-237, e patentes do EUA N^{os}. 5.814.618 e 5.789.156), incorporadas aqui por referência.

- Promotores constitutivos incluem, por exemplo, o promotor core do promotor Rsyn7 e outros promotores constitutivos revelados nas Patentes do EUA N° 6.072.050; o promotor core de CaMV 35S (Odell *et al.* (1985) *Nature* 313:810-812); ubiquitina (Christensen *et al.* (1989) *Plant Mol. Biol.* 12:619-632 e Christensen *et al.* (1992) *Plant Mol. Biol.* 18:675-689); pEMU (Last *et al.* (1991) *Theor. Appl. Genet.* 81:581-588); MAS (Velten *et al.* (1984) *EMBO J.* 3:2723-2730); promotor ALS (Patente dos EUA No. 5.659.026), e semelhantes. Outros promotores constitutivos incluem, por exemplo, Patentes dos EUA N°s 5.608.149; 5.608.144; 5.604.121; 5.569.597; 5.466.785; 5.399.680; 5.268.463; 5.608.142; e 6.177.611.
- 10 Promotores que preferem tecidos podem ser utilizados para direcionar a expressão de uma sequência inibidora de polinucleotídeo da presente invenção dentro de um tecido específico de planta. Promotores que preferem tecidos incluem Yamamoto *et al.* (1997) *Plant J.* 12(2):255-265; Kawamata *et al.* (1997) *Plant Cell Physiol.* 38(7):792-803; Hansen *et al.* (1997) *Mol. Gen Genet.* 254(3):337-343; Russell *et al.* (1997) *Transgenic Res.* 6(2):157-168; Rinehart *et al.* (1996) *Plant Physiol.* 112(3):1331-1341; Van Camp *et al.* (1996) *Plant Physiol.* 112(2):525-535; Canevascini *et al.* (1996) *Plant Physiol.* 112(2):513-524; Yamamoto *et al.* (1994) *Plant Cell Physiol.* 35(5):773-778; Lam (1994) *Results Probl. Cell Differ.* 20:181-196; Orozco *et al.* (1993) *Plant Mol Biol.* 23(6):1129-1138; Matsuoka *et al.* (1993) *Proc Natl. Acad. Sci. USA* 90(20):9586-9590; e Guevara-Garcia *et al.* (1993) *Plant J.* 4(3):495-505.

De interesse particular são os promotores que preferem folhas, que apresentam expressão predominantemente em tecidos de folhas. Ver, por exemplo, Yamamoto *et al.* (1997) *Plant J.* 12(2):255-265; Kwon *et al.* (1994) *Plant Physiol.* 105:357-67; Yamamoto *et al.* (1994) *Plant Cell Physiol.* 35(5):773-778; Gotor *et al.* (1993) *Plant J.* 3:509-18; Orozco *et al.* (1993) *Plant Mol. Biol.* 23(6):1129-1138; Baszczyński *et al.* (1988) *Nucl. Acid Res.* 16:4732; Mitra *et al.* (1994) *Plant Molecular Biology* 26:35-93; Kayaya *et al.* (1995) *Molecular and General Genetics* 248:668-674; e Matsuoka *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90(20):9586-9590. Também são usados genes regulados por senescência, como o SAM22 (Crowell *et al.* (1992) *Plant Mol. Biol.* 18:459-466); SAG12 (Lohman *et al.* (1994) *Physiol. Plant.* 92:322-328; Wingler *et al.* (1998) *Plant Physiol.* 116:329-335); SAG 13 (Gan e Amasino (1997) *Plant Physiol.* 113:313-319; SAG15 (Gan (1995) "Molecular Characterization and Genetic Manipulation of Plant Senescence," Tese de Ph.D., University of Wisconsin,

Madison); SEN1 (Oh *et al.* (1996) *Plant Mol. Biol.* 30:739-754; promotor de um gene específico de senescência para expressão de IPT (Gan e Amasino 1995) *Science* 270:1986-1988); e semelhantes (ver, por exemplo, Ou: *et al.* (1999) *Plant Cell* 11:1073-1080 e McCabe *et al.* (2001) *Plant Physiol.* 127:505-516).

Métodos para inibição da expressão ou função de um citocromo P450 envolvido na conversão de nicotina a nornicotina

Apresenta-se aqui métodos para a redução da concentração do conteúdo, e/ou atividade de um polipeptídeo de um citocromo P450 da presente invenção em uma planta de *Nicotiana* ou parte de planta, particularmente o tecido da folha. Muitos métodos podem ser usados, sozinhos ou em combinação, para reduzir ou eliminar a atividade de um polipeptídeo do citocromo P450 da presente invenção. Além disso, combinações de métodos podem ser empregadas para reduzir ou eliminar a atividade de dois ou mais polipeptídeos diferentes de citocromo P450.

De acordo com a presente invenção, a expressão de um polipeptídeo do citocromo P450 da presente invenção é inibido se o nível de proteína do polipeptídeo do citocromo P450 é estatisticamente mais baixo do que o nível de proteína do mesmo polipeptídeo do citocromo P450 em uma planta que não foi geneticamente modificada ou mutagenizada para inibir a expressão daquele polipeptídeo do citocromo P450, quando essas plantas foram cultivadas e colhidas com o uso dos mesmos protocolos. Em concretizações particulares da invenção, o nível de proteína do polipeptídeo do citocromo P450 em uma planta modificada de acordo com a invenção é menor que 95%, menor que 90%, menor que 80%, menor que 70%, menor que 60%, menor que 50%, menor que 40%, menor que 30%, menor que 20%, menor que 10%, menor que 5%, menor que 4%, menor que 3%, menor que 2%, ou menor que 1% do nível de proteína do mesmo polipeptídeo do citocromo P450 em uma planta que não é mutante ou que não tenha sido geneticamente modificada para inibir a expressão daquele polipeptídeo do citocromo P450 e que foi cultivada e colhida com o uso dos mesmos protocolos . O nível de expressão do polipeptídeo do citocromo P450 pode ser medido diretamente, por exemplo, por ensaio do nível do transcrito do citocromo P450 ou polipeptídeo do citocromo P450 expresso na planta ou parte de planta de *Nicotiana*, ou indiretamente, por exemplo, por mensuração da conversão de nicotina em nornicotina na planta ou parte de planta de *Nicotiana*. Métodos para monitorar o nível de expressão de uma proteína são conhecidos da técnica, e incluem, mas não se limitam a, análise de Northern blot, conforme discutido nos exemplos mostrados adiante. Métodos para determinação da atividade do polipeptídeo alvo do citocromo P450 na conversão de nicotina em nornicotina são descritos mais adiante, e incluem,

mas não se limitam a, análise de alcalóide com uso de cromatografia gasosa, como por exemplo, os procedimentos descritos nos exemplos mostrados abaixo.

Em outras concretizações da invenção, a atividade de um ou mais polipeptídeos do citocromo P450 é reduzida ou eliminada pela transformação de uma planta ou parte de planta com um cassete de expressão que abrange um polinucleotídeo que codifica um polipeptídeo que inibe a atividade de um ou mais polipeptídeos do citocromo P450 da presente invenção. A atividade do polipeptídeo do citocromo P450 na conversão de nicotina em nornicotina em uma planta ou parte de planta de *Nicotiana* é inibida de acordo com a presente invenção, se essa atividade de conversão é estatisticamente mais baixa do que atividade de conversão do mesmo do polipeptídeo do citocromo P450, em uma planta ou parte de planta de *Nicotiana* que não foi geneticamente modificada para inibir a atividade de conversão daquele polipeptídeo do citocromo P450, e que foi cultivada e colhida com o uso dos mesmos protocolos. Em concretizações particulares, a atividade do polipeptídeo do citocromo P450 na conversão de nicotina em nornicotina em uma planta ou parte de planta de *Nicotiana* de acordo com a invenção é inibida se a atividade é menor que 95%, menor que 90%, menor que 80%, menor que 70%, menor que 60%, menor que 50%, menor que 40%, menor que 30%, menor que 20%, menor que 10%, menor que 5%, menor que 4%, menor que 3%, menor que 2%, ou menor que 1% da atividade de conversão do mesmo polipeptídeo do citocromo P450 em uma planta *Nicotiana* que tenha sido geneticamente modificada para inibir a expressão daquele polipeptídeo do citocromo P450 e que foi cultivada e colhida com o uso dos mesmos protocolos. A atividade do polipeptídeo do citocromo P450 na conversão da nicotina em nornicotina em uma planta ou parte de planta de *Nicotiana* é eliminada de acordo com a invenção quando ela não é detectável pelos métodos de ensaio descritos aqui. Métodos de determinação da atividade de um polipeptídeo do citocromo P450 na conversão da nicotina em nornicotina em uma planta ou parte de planta de *Nicotiana* são descritos aqui, e incluem as análises de alcalóide com o uso de cromatografia gasosa reveladas nos exemplos abaixo.

Em concretizações específicas, uma sequência inibidora do citocromo p450 descrita aqui é introduzida em uma planta ou parte de planta de *Nicotiana*. Subsequentemente, a planta ou parte de planta de *Nicotiana* que tem a sequência inibidora de polinucleotídeo da invenção é selecionada usando métodos conhecidos

por aqueles habilitados na técnica como, mas não se limitando a, análise de Southern blot, sequenciamento de DNA, análise de PCR, ou análise fenotípica. Uma planta ou parte de planta alterada ou modificada através das concretizações anteriores é crescida sob condições de criação de plantas por tempo suficiente para modular a concentração e/ou atividade de polipeptídeos da presente invenção na planta. Condições de criação de plantas são bem conhecidas e discutidas brevemente mais adiante.

Em algumas concretizações, uma planta transformada de tabaco contendo uma sequência inibidora de polinucleotídeo de um citocromo P450 descrito aqui tem um nível reduzido de conversão de nicotina em nornicotina. Em concretizações particulares, a conversão de nicotina em nornicotina em uma planta ou parte de planta transformada de tabaco de acordo com a invenção é menor que 95%, menor que 90%, menor que 80%, menor que 70%, menor que 60%, menor que 50%, menor que 40%, menor que 30%, menor que 20%, menor que 10%, menor que 5%, menor que 4%, menor que 3%, menor que 2%, ou menor que 1% da conversão em uma planta de tabaco que não tenha sido geneticamente modificada para inibir a expressão daquele polipeptídeo do citocromo P450 e que foi cultivada e colhida com o uso dos mesmos protocolos. Em algumas concretizações, a planta transformada de tabaco é uma planta de tabaco conversora. Em algumas concretizações, a planta de tabaco transformada tem um taxa de conversão mais baixa do que a taxa observada em plantas de tabaco não-conversoras comerciais.

Também se admite que o nível e/ou atividade do polipeptídeo pode ser modulado pelo emprego de um polinucleotídeo que não é capaz de dirigir, em uma planta transformada, a expressão de uma proteína ou RNA. Por exemplo, os polinucleotídeos da invenção podem ser usados para desenhar construtos de polinucleotídeos que podem ser empregados em métodos para alterar o mutar uma sequência genômica em um organismo. Tais construtos de polinucleotídeos incluem, mas não se limitam a, vetores RNA:DNA, vetores mutacionais RNA:DNA, vetores de reparo RNA:DNA, oligonucleotídeos duplex mistos, oligonucleotídeos de RNA:DNA autocomplementares, e oligonucleases recombinogênicas. Tais construtos de nucleotídeos e métodos de uso são conhecidos da técnica. Ver Patente dos EUA nºs 5.565.350; 5.731.181; 5.756.325; 5.760.012; 5.795.972; e 5.871.984; todas as quais estão aqui incorporadas por referência. Ver também, WO 98/49350, WO 99/07865,

WO 99/25821, e Beetham *et al.* (1999) *Proc. Natl. Acad. Sci. USA* 96:8774-8778; incorporados aqui por referência.

Portanto admite-se que os métodos da presente invenção não dependem da incorporação do polinucleotídeo inibidor inteiro do citocromo P450 dentro do
5 genoma, somente que aquela planta ou parte de planta *Nicotiana* é alterada como resultado da introdução desse polinucleotídeo inibidor dentro de uma célula. Em uma concretização da invenção, o genoma pode ser alterado logo após a introdução do polinucleotídeo inibidor do citocromo P450 dentro de uma célula. Por exemplo, o polinucleotídeo inibidor, ou qualquer parte desse, pode se incorporar dentro do
10 genoma da planta. Alterações do genoma incluem, mas não se limitam a, adições, deleções, e substituições de nucleotídeos dentro do genoma. Embora os métodos da presente invenção não dependam de adições, deleções, e substituições de qualquer número específico de nucleotídeos, admite-se que essas adições, deleções, e substituições abrangem no mínimo um nucleotídeo.

Admite-se ainda que a redução do nível e/ou atividade de uma sequência do citocromo P450 da presente invenção pode ser realizada para extrair os efeitos da sequência somente durante certos estágios de desenvolvimento e para desligar o efeito em outros estágios em que a expressão não é mais desejável. O controle da expressão do citocromo P450 pode ser obtido através do uso de promotores com preferência por
15 tecidos ou induzíveis. De modo alternativo, o gene poderia ser invertido ou deletado usando recombinases sítio-específicas, transposons ou sistemas de recombinação, que também ligariam ou desligariam a expressão da sequência do citocromo P450.

De acordo com a presente invenção, mudanças nos níveis, proporções, atividade, ou distribuição de polipeptídeos do citocromo P450 da presente invenção,
25 ou mudanças no fenótipo da planta ou parte de planta *Nicotiana*, particularmente a acumulação reduzida de nicotina e seu metabólito carcinogênico, NNN, poderiam ser medidas por comparação de uma planta ou parte de planta-sujeito com uma planta ou parte de planta-controle, em que a planta ou parte de planta-sujeito e a planta ou parte de planta-controle tenham sido cultivadas e/ou colhidas com o uso dos mesmos
30 protocolos. Conforme usado aqui, uma planta ou parte de planta-sujeito é uma planta na qual a alteração genética, como uma transformação, foi influenciada quanto ao polipeptídeo de interesse, ou uma planta ou parte de planta *Nicotiana* que descende de uma planta ou parte de planta *Nicotiana* assim alterada e que abrange a alteração.

Uma planta ou parte de planta controle apresenta um ponto de referência para medir mudanças no fenótipo da planta ou parte da planta-sujeito.

A medição de mudanças no fenótipo pode ser feita em qualquer tempo em uma planta ou parte de planta, incluindo durante o desenvolvimento, senescência, ou
5 após a cura da planta. Em outras concretizações, a medição de mudanças no fenótipo pode ser medida em plantas crescidas sob quaisquer condições, incluindo plantas crescidas em uma câmara de crescimento, estufa, ou no campo. Em uma concretização, as mudanças no fenótipo podem ser medidas pela determinação da taxa de conversão de nicotina em nornicotina. Em uma concretização preferida, a
10 conversão pode ser medida através da divisão da percentagem de nicotina (como uma percentagem do peso total do tecido) pela soma da percentagem de nicotina em nornicotina (como percentagens do peso total do tecido) e multiplicando por 100.

De acordo com a presente invenção, uma planta ou parte de planta-controle pode abranger uma planta ou parte de planta selvagem de *Nicotiana*, i.e., do mesmo
15 genótipo como material de partida para a alteração genética que resultou na planta ou parte de planta-sujeito. Uma planta ou parte de planta-controle também pode abranger uma planta ou parte de planta de *Nicotiana* do mesmo genótipo como material de partida, mas que tenha sido transformada com um construto nulo (i.e, com um construto que tem um efeito conhecido sobre o traço de interesse, tal como um
20 construto que abrange um gene marcador selecionável). De modo alternativo, uma planta ou parte de planta-controle pode abranger uma planta ou parte de planta de *Nicotiana* que é uma segregante não transformada entre uma progênie de uma planta ou parte de planta-sujeito, ou uma planta ou parte de planta de *Nicotiana* geneticamente idêntica à planta ou parte de planta-sujeito, que não é exposta a
25 condições ou estímulos que induzam a supressão do gene do citocromo P450 de interesse. Por fim, uma planta ou parte de planta-controle pode abranger a planta ou parte de planta-sujeito em si sob condições nas quais a sequência inibidora do citocromo P450 não é expressa. Em todos os casos, a planta ou parte de planta-sujeito e a planta ou parte de planta-controle são cultivadas e colhidas usando os mesmos
30 protocolos.

Conforme descrito aqui, os métodos são apresentados para reduzir ou eliminar a atividade e/ou concentração do polipeptídeo do citocromo P450 da presente invenção pela introdução, dentro de uma planta ou parte de planta *Nicotiana*, de uma

sequência de polinucleotídeos inibidora do citocromo P450 que está envolvida na conversão metabólica da nicotina em normicotina. Em algumas concretizações, as sequências inibidoras são introduzidas pela transformação da planta ou parte de planta, tal como uma célula de planta, com um cassete de expressão que expressa um

5 polinucleotídeo que inibe a expressão do polipeptídeo do citocromo P450. O polinucleotídeo pode inibir diretamente a expressão de um polipeptídeo do citocromo P450, impedindo a tradução do RNA mensageiro do polipeptídeo do citocromo P450, ou indiretamente, pela codificação de um polipeptídeo que inibe a transcrição ou

10 tradução de um gene de polipeptídeo do citocromo P450 que codifica um polipeptídeo do citocromo P450. Métodos para inibição ou eliminação da expressão de um produto de gene em uma planta são bem conhecidos da técnica, e quaisquer desses métodos podem ser usados na presente invenção para inibir a expressão de polipeptídeos do citocromo P450.

Em outras concretizações, a atividade de um polipeptídeo do citocromo P450

15 da presente invenção pode ser reduzida ou eliminada através da disrupção do gene que codifica o polipeptídeo do citocromo P450. A invenção abrange plantas mutagenizadas que carregam mutações nos genes do citocromo P450, sendo que a mutações reduzem a expressão do gene do citocromo P450 ou inibem a atividade de um polipeptídeo do citocromo P450 codificado da presente invenção.

Em algumas concretizações da presente invenção, uma planta ou parte de planta de *Nicotiana* é transformada com um cassete de expressão capaz de expressar um polinucleotídeo que inibe a expressão de uma sequência de um citocromo P450. Tais métodos podem incluir o uso supressão/cossupressão senso, ou supressão

20 antisenso, interferência de RNA dupla fita (RNAds), interferência de RNA em alça e interferência de RNA em alça contendo íntron, interferência mediada por amplicon, ribozimas, e pequena interferência de RNA ou micro RNA.

Para a cossupressão, um cassete de expressão é desenhado para expressar uma molécula de RNA correspondente a todo ou parte de um RNA mensageiro que codifica um polipeptídeo do citocromo P450 de interesse (por exemplo, um

30 polipeptídeo do citocromo P450 que abrange a sequência apresentada nas ID SEQ N°:2, 4, 6, 8, 10, ou 12 ou uma sequência que tem uma identidade de sequência substancial às ID SEQ N°:2, 4, 6, 8, 10, ou 12) na orientação "senso". A expressão excessiva da molécula de RNA pode resultar na redução da expressão do gene nativo.

Linhagens múltiplas de plantas transformadas com o cassete de expressão de cossupressão são então rastreadas para identificar aquelas que mostram a maior inibição da expressão do polipeptídeo do citocromo P450.

5 O polinucleotídeo usado para a cossupressão pode corresponder a toda ou parte da sequência que codifica um polipeptídeo do citocromo P450 da presente invenção, toda ou parte da região 5' e/ou 3' não traduzida de um transcrito de polipeptídeo do citocromo P450, ou toda ou parte da sequência codificadora, bem como das regiões não traduzidas de um transcrito que codifica um polipeptídeo do citocromo P450. Em algumas concretizações em que o polinucleotídeo abrange toda
10 ou parte da região codificadora para um polipeptídeo do citocromo P450 da presente invenção, o cassete de expressão é desenhado para eliminar o códon de partida do polinucleotídeo, de modo que nenhum produto de proteína seja transcrito.

A cossupressão pode ser usada para inibir a expressão de genes de plantas para produzir plantas que têm níveis de proteína não detectáveis para as proteínas que
15 codificam através desses genes ou pode também ser usada para inibir a expressão de múltiplas proteínas na mesma planta (por ex., Broin *et al.* (2002) *Plant Cell* 14:1417-1432; Patente dos EUA Nº. 5.942.657). Métodos para usar a cossupressão para inibir a expressão de genes endógenos em plantas são descritos em Flavell *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3490-3496; Jorgensen *et al.* (1996) *Plant Mol. Biol.*
20 31:957-973; Johansen and Carrington (2001) *Plant Physiol.* 126:930-938; Broin *et al.* (2002) *Plant Cell* 14:1417-1432; Stoutjesdijk *et al.* (2002) *Plant Physiol.* 129:1723-1731; Yu *et al.* (2003) *Phytochemistry* 63:753-763; Patentes dos EUA Nºs. 5.034.323, 5.283.184, e 5.942.657; cada uma das quais está incorporada aqui por referência. A eficiência da cossupressão pode ser aumentada incluindo-se uma região
25 poli-dT no cassete de expressão na posição 3' para a sequência senso e 5' do sinal de poliadenilação. Ver a Publicação de Patente dos EUA Nº. 20020048814, incorporada aqui por referência. Tipicamente, tal sequência de nucleotídeos tem uma identidade de sequência substancial em relação à sequência do transcrito do gene endógeno, com identidade de sequência otimamente maior do que 65%, mais otimamente maior do
30 que 85%, e otimamente muito maior do que 95% (por ex., Patentes dos EUA Nºs 5.283.184 e 5.034.323), incorporadas aqui por referência.

Em algumas concretizações da invenção, a inibição da expressão do polipeptídeo do citocromo P450 da presente invenção pode ser obtida através da

supressão antisense. Para a supressão antisense, o cassete de expressão é desenhado para expressar uma molécula de RNA complementar a todo ou parte do RNA mensageiro que codifica o polipeptídeo do citocromo P450. A expressão excessiva da molécula de RNA antisense pode resultar na redução da expressão do gene nativo.

- 5 Consequentemente, linhagens múltiplas de plantas transformadas com o cassete de expressão de supressão antisense são rastreadas para identificar aquelas que mostram a maior inibição da expressão do polipeptídeo do citocromo P450.

O polinucleotídeo usado para supressão antisense pode corresponder a todo ou parte do complemento da sequência que codifica um polipeptídeo do citocromo P450,
10 toda ou parte da região 5' e/ou 3' não traduzida de um transcrito do polipeptídeo do citocromo P450, ou todo ou parte do complemento da sequência codificadora e das regiões não traduzidas de um transcrito que codifica um polipeptídeo do citocromo P450. Além disso, o polinucleotídeo antisense pode ser completamente complementar (i.e. 100% idêntico ao complemento da sequência alvo) ou
15 parcialmente complementar (i.e., menos de 100% idêntico ao complemento da sequência alvo) à sequência alvo. A supressão antisense também pode ser usada para inibir a expressão de múltiplas proteínas na mesma planta (por ex., Patente dos EUA Nº. 5.942.657). Além disso, porções dos nucleotídeos antisense podem ser usadas para a disrupção da expressão do gene alvo. Geralmente podem ser usadas sequências
20 de no mínimo 50 nucleotídeos, 100 nucleotídeos, 200 nucleotídeos, 300, 400, 450, 500, 550, ou maiores. Métodos de uso da supressão antisense para inibir a expressão de genes endógenos são descritos, por exemplo, em Liu *et al* (2002) *Plant Physiol.* 129:1732-1743 e Patentes do EUA Nºs. 5.759.829 e 5.942.657; cada uma das quais está incorporada aqui por referência. A eficiência da supressão pode ser aumentada
25 incluindo-se uma região poli-dT no cassete de expressão na posição 3' para a sequência senso e 5' do sinal de poliadenilação. Ver, Publicação de Patente dos EUA Nº. 20020048814, incorporada aqui por referência.

Para a interferência de RNAs, uma molécula de RNA senso como aquela descrita acima para a cossupressão e uma molécula de RNA antisense que é completa
30 ou parcialmente complementar à molécula de RNA senso são expressas na mesma célula, resultando na inibição da expressão do RNA mensageiro endógeno correspondente.

A expressão das moléculas senso e antisenso pode ser obtida através do desenho do cassete de expressão para abranger ambas as sequências, senso e a antisenso, para a sequência do citocromo P450 alvo. De modo alternativo, cassetes de expressão separados podem ser usados para as sequências senso e antisenso.

- 5 Linhagens múltiplas de plantas transformadas com o cassete de expressão de interferência de RNAs ou cassetes de expressão são então rastreadas para identificar linhagens de plantas que mostram a maior inibição da expressão do polipeptídeo do citocromo P450 alvo. Métodos para uso de interferência de RNAs para inibir a expressão de genes endógenos em plantas são descritos em Waterhouse *et al.* (1998)
- 10 *Proc. Natl. Acad. Sci. USA* 95:13959-13964, Liu *et al.* (2002) *Plant Physiol.* 129:1732-1743, e WO 99/49029, WO 99/53050, WO 99/61631, e WO 00/49035; cada um dos quais está incorporado aqui por referência.

- Cassetes de expressão de amplicon abrangem uma sequência derivada de vírus de planta que contém todo ou parte do gene alvo, mas geralmente não todos os genes
- 15 do vírus nativo. As sequências virais presentes no produto da transcrição do cassete de expressão permitem que o produto da transcrição dirija sua própria replicação. Os transcritos produzidos pelo amplicon podem ser senso ou antisenso em relação à sequência alvo (i.e., o RNA mensageiro para um polipeptídeo do citocromo P450 que está envolvido na conversão metabólica de nicotina em nornicotina). Métodos para o
- 20 uso de amplicons para inibir a expressão de genes endógenos de plantas são descritos, por exemplo, em Angell e Baulcombe (1997) *EMBO J.* 16:3675-3684, Angell e Baulcombe (1999) *Plant J.* 20:357-362, e Patentes do EUA Nº. 6.646.805, cada uma das quais está incorporada aqui por referência.

- Em concretizações adicionais da presente invenção, o polinucleotídeo
- 25 expresso pelo cassete de expressão da invenção é RNA catalítico ou tem uma atividade de ribozima específica para o RNA mensageiro de um polipeptídeo do citocromo P450 descrito aqui. Portanto, o polinucleotídeo causa a degradação do RNA mensageiro endógeno, resultando na expressão reduzida do polipeptídeo do citocromo P450. Esse método é descrito, por exemplo, na Patente dos EUA Nº
- 30 4.987.071, incorporada aqui por referência.

Em concretizações adicionais da invenção, a inibição da expressão de **um ou mais** polipeptídeos do citocromo P450 pode ser obtida pelo RNA de interferência (RNAi) por expressão de um gene que codifica um micro RNA (RNAmi). Os RNAmi

são agentes reguladores que consistem em cerca de 22 ribonucleotídeos. Os RNAmi são altamente eficientes na inibição de genes endógenos. Ver, por exemplo, Javier *et al.* (2003) *Nature* 425: 257-263), incorporados aqui por referência.

Para a interferência do RNAmi, o cassete de expressão é desenhado para
5 expressar uma molécula de RNA que é modelada sobre um gene de RNAmi endógeno. O gene do RNAmi codifica um RNA que forma uma estrutura em alça contendo 22 nucleotídeos que é complementar a outro gene endógeno (sequência alvo). Para a supressão da expressão do polipeptídeo do citocromo P450, a sequência
10 de 22 nucleotídeos é selecionada a partir de uma sequência de transcrito do polipeptídeo do citocromo P450 e contém 22 nucleotídeos de uma sequência antisenso correspondente que é complementar à sequência senso. Moléculas de RNAmi são altamente eficientes na inibição da expressão de genes endógenos, e a interferência de RNA que elas induzem é herdada pelas gerações de plantas subsequentes.

15 Ainda em outras concretizações da invenção, a inibição da expressão de um ou mais polipeptídeos do citocromo P450 pelo RNAi pode ser obtida pela interferência do RNA em alça (RNAhp, de “hairpin”) ou interferência de RNA em alça contendo íntron (RNAihp). Esses métodos são altamente eficientes na inibição da expressão de genes endógenos. Ver, Waterhouse and Helliwell (2003) *Nat. Rev. Genet.* 4:29-38 e
20 as referências citadas pelos autores.

Para a interferência de RNAhp, o cassete de expressão é desenhado para expressar uma molécula de RNA que hibridiza com ela mesma para formar uma estrutura em alça que abrange um região de loop de única fita e uma haste de bases
25 pareadas. A região de haste de bases pareadas abrange uma sequência senso correspondendo a todo ou parte do RNA mensageiro endógeno que codifica o produto do gene cuja expressão deve ser inibida, nesse caso, um polipeptídeo do citocromo P450 descrito aqui, e uma sequência antisenso que é completa ou parcialmente complementar à sequência senso. De modo alternativo, a região de haste de bases
30 pareadas pode corresponder a uma porção de uma sequência promotora que controla a expressão do gene que codifica o polipeptídeo do citocromo P450 a ser inibido. Portanto, a região de haste de bases pareadas da molécula geralmente determina a especificidade da interferência do RNA. As moléculas de RNAhp são altamente eficientes na inibição da expressão de genes endógenos, e a interferência que elas

induzem é herdada pelas gerações de plantas subsequentes. Ver, por exemplo, Chuang e Meyerowitz (2000) *Proc. Natl. Acad. Sci. USA* 97:4985-4990; Stoutjesdijk *et al.* (2002) *Plant Physiol.* 129:1723-1731; e Waterhouse e Helliwell (2003) *Nat. Rev. Genet.* 4:29-38. Métodos para uso na interferência de RNAhp para inibir ou

5 silenciar a expressão de genes são descritos, por exemplo, em Chuang e Meyerowitz (2000) *Proc. Natl. Acad. Sci. USA* 97:4985-4990; Stoutjesdijk *et al.* (2002) *Plant Physiol.* 129:1723-1731; Waterhouse e Helliwell (2003) *Nat. Rev. Genet.* 4:29-38; Pandolfini *et al.* *BMC Biotechnology* 3:7, e Publicação de Patente dos EUA Nº. 20030175965; cada uma das quais está incorporada aqui por referência. Um ensaio

10 transiente para a eficiência de construtos de RNAhp para silenciar a expressão de genes *in vivo* foi descrita por Panstruga *et al.* (2003) *Mol. Biol. Rep.* 30:135-140, incorporado aqui por referência.

Para o RNAihp, as moléculas de interferência têm a mesma estrutura geral do RNAhp, mas a molécula de RNA abrange adicionalmente um íntron que é capaz de

15 ser combinado na célula em que o RNAihp é expresso. O uso de um íntron minimiza o tamanho do loop na molécula de RNA em alça após a união, e isso aumenta a eficiência da interferência. Ver, por exemplo, Smith *et al.* (2000) *Nature* 407:319-320. Na verdade, Smith *et al.* mostram uma supressão de 100% da expressão de genes endógenos com o uso da interferência mediada por RNAihp. Métodos para uso

20 de interferência de RNAds para inibir a expressão de genes endógenos em plantas são descritos em Smith *et al.* (2000) *Nature* 407:319-320; Wesley *et al.* (2001) *Plant J.* 27:581-590; Wang e Waterhouse (2001) *Curr. Opin. Plant Biol.* 5:146-150; Waterhouse e Helliwell (2003) *Nat. Rev. Genet.* 4:29-38; Helliwell e Waterhouse (2003) *Methods* 30:289-295, e Publicação de Patente dos EUA Nº. 20030180945,

25 cada uma das quais está incorporada aqui por referência.

Em uma concretização, o RNAi é efetuado pela expressão de uma sequência inibidora que abrange uma primeira sequência de um polinucleotídeo do citocromo P450 da invenção que tem cerca de 90 pb (pares de bases) a cerca de 110 pb de comprimento, incluindo 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104,

30 105, 106, 107, 108, 109, e 110 pb de comprimento, uma sequência que é complementar a toda ou parte da primeira sequência. Em outra concretização, uma sequência inibidora que abrange uma primeira sequência de um polinucleotídeo do citocromo P450 da invenção que tem cerca de 290 pb a cerca de 310 pb de

comprimento, incluindo 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, e 310 bp de comprimento, e uma segunda sequência que é complementar a toda ou parte da sequência do fragmento.

Em outras concretizações da invenção, o RNAi é finalizado pela expressão de uma sequência inibidora que abrange uma sequência de polinucleotídeos que contém os nucleotídeos na posição aproximada de 265 à posição aproximada de 625 de uma sequência codificadora do citocromo P450 revelada aqui e uma segunda sequência que é completa ou parcialmente complementar a essa. Em algumas dessas concretizações, a sequência inibidora abrange como primeira sequência de nucleotídeos os nucleotídeos correspondentes à posição aproximada de 297 à posição aproximada de 594 da sequência codificadora do P450 apresentada na ID SEQ NO:3 ou ID SEQ NO:5 e a segunda sequência é o complemento (i.e, uma sequência antisenso) da primeira sequência. A sequência inibidora pode abranger opcionalmente uma sequência de íntrons ligada entre a primeira e a segunda sequências. Qualquer íntron conhecido por pessoas habilitadas na técnica pode ser usado desse forma. Em algumas concretizações, o íntron é o ácido graxo ômega-6 desaturase de soja (FAD, fatty acid desaturase) (ver GenBank Accession N°. DQ672337, e o Exemplo 7 adiante). Em tal concretização, o íntron abrange cerca de 151 nucleotídeos, o que abrange nucleotídeos de 100-247 do polinucleotídeo de ácido graxo ômega-6 desaturase de soja mostrado no GenBankk Accession N°. DQ672337. Exemplos de outros íntrons incluem, mas não se limitam às sequências de nucleotídeos de íntron de genes da enzima álcool desidrogenase (*adh1*). A expressão dessa sequência inibidora produz um RNA em alça contendo íntron que interfere fortemente na expressão dos polipeptídeos do citocromo P450 revelados aqui. Dessa forma, plantas de *Nicotiana* que normalmente são conversoras de nicotina em nornicotina que são transformadas com um cassete de expressão abrangendo essa sequência inibidora têm, vantajosamente, uma taxa de conversão de nicotina em nornicotina surpreendentemente ainda mais baixa do que a observada para plantas de *Nicotiana* que são não-conversoras de nicotina em nornicotina.

Ainda em outras concretizações da invenção, o RNAi é efetuado pela sequência inibidora que abrange uma primeira sequência de polinucleotídeos que contém os nucleotídeos na posição aproximada de 1420 à posição aproximada de 1580 de uma sequência codificadora do citocromo P450 revelada aqui e uma

sequência que é completa ou parcialmente complementar a essa. Em algumas dessas concretizações, a sequência inibidora abrange como primeira sequência de polinucleotídeos os nucleotídeos correspondentes na posição aproximada de 1453 à posição aproximada de 1551 da sequência codificadora do P450 apresentada na ID
 5 SEQ Nº: 1 e a segunda sequência é o complemento (i.e, uma sequência antisenso) da primeira sequência. A sequência inibidora pode abranger opcionalmente uma sequência de íntrons ligada entre a primeira e a segunda sequências. Qualquer íntron conhecido por pessoas habilitadas na técnica pode ser usado desse forma, como pode ser observado acima. A expressão dessa sequência inibidora produz um RNA em alça
 10 (ou RNA em alça contendo íntron quando o íntron está presente) que também interfere na expressão dos polipeptídeos do citocromo P450 revelados aqui.

O cassete de expressão para interferência do RNAhp também pode ser desenhado de modo que a sequência senso e a sequência antisenso não correspondam a um RNA endógeno. Nessa concretização, as sequências senso e antisenso
 15 flanqueiam uma sequência em loop que abrange uma sequência de nucleotídeos correspondentes a todo ou parte do RNA mensageiro endógeno do gene alvo. Portanto, é a região em loop que determina a especificidade da interferência do RNA. Ver, por exemplo, WO 02/00904, incorporado aqui por referência.

O silenciamento do gene transcricional (TGS, transcriptional gene silencing)
 20 pode ser realizado através do uso de construtos de RNAhp em que a repetição invertida em alça compartilha a identidade de sequência com uma região promotora do gene a ser silenciado. O processamento do RNAhp em RNAs curtos que podem interagir com a região promotora homóloga pode precipitar a degradação ou metilação para resultar no silenciamento (Aufsatz *et al.* (2002) *Proc. Natl. Acad. Sci.*
 25 99 (Suppl. 4):16499-16506; Mette *et al.* (2000) *EMBO J.* 19(19):5194-5201).

Em concretizações adicionais, um polinucleotídeo pode ser utilizado para codificar uma proteína do dedo de zinco que se liga a um gene que codifica o polipeptídeo do citocromo P450, resultando na expressão reduzida do gene. Em concretizações particulares, a proteína do dedo de zinco se liga a uma região
 30 reguladora do gene o polipeptídeo do citocromo P450. Em outras concretizações, a proteína do dedo de zinco se liga a um RNA mensageiro que codifica um polipeptídeo do citocromo P450 e impede sua tradução. Métodos de seleção de sítios para direcionamento ao alvo por parte de proteínas do dedo de zinco têm sido descritos,

por exemplo, na Patente dos EUA Nº. 6.453.242, e métodos para o uso de proteínas do dedo de zinco na inibição da expressão de genes em plantas são descritos, por exemplo, na Patente do EUA Nº. 20030037355; cada uma das quais está incorporada aqui por referência.

5 Em outras concretizações da invenção, o polipeptídeo codifica um anticorpo que se liga a no mínimo um polipeptídeo do citocromo P450, e reduz a atividade de um polipeptídeo do citocromo P450 da presente invenção. Em outra concretização, a ligação do anticorpo resulta na rotatividade do complexo anticorpo-polipeptídeo do citocromo P450 por mecanismos de controle de qualidade celular. A expressão de anticorpos em partes de plantas e a inibição de vias moleculares através de expressão e ligação de anticorpos a proteínas em partes de plantas são bem conhecidas da técnica. Ver, por exemplo, Conrad e Sonnewald (2003) *Nature Biotech.* 21:35-36, incorporados aqui por referência.

15 Em outras concretizações, a atividade de um polipeptídeo do citocromo P450 da presente invenção é reduzida ou eliminada através da disrupção do gene que codifica o polipeptídeo do citocromo P450. O gene codificador do polipeptídeo do citocromo P450 pode sofrer disrupção por qualquer método conhecido da técnica, por exemplo, por etiquetagem de transposon ou por mutagenização de plantas usando mutagênese aleatória ou direcionada ao alvo, e por seleção de plantas que têm atividade reduzida do citocromo P450.

20 A etiquetagem de transposon pode ser usada para reduzir ou eliminar a atividade de **um ou mais** polipeptídeos do citocromo P450 da presente invenção. A etiquetagem de transposon abrange a inserção de um transposon dentro do gene endógeno do citocromo P450 para reduzir ou eliminar a expressão do polipeptídeo do citocromo P450.

25 Nessa concretização, a expressão de **um ou mais** polipeptídeos do citocromo P450 é reduzida ou eliminada pela inserção de um transposon dentro da região reguladora ou codificadora do gene que codifica o polipeptídeo do citocromo P450. Um transposon que está dentro de um éxon, íntron ou sequência não traduzida 5' ou 3', promotor, ou de outra sequência reguladora de um gene do polipeptídeo do citocromo P450 pode ser usado para reduzir ou eliminar a expressão e/ou atividade do polipeptídeo do citocromo P450 codificado.

Métodos para a etiquetagem de transposon de genes específicos em plantas são bem conhecidos da técnica. Ver, por exemplo, Maes *et al.* (1999) *Trends Plant Sci.* 4:90-96; Dharmapuri e Sonti (1999) *FEMS Microbiol. Lett.* 179:53-59; Meissner *et al.* (2000) *Plant J.* 22:265-274; Phogat *et al.* (2000) *J. Biosci.* 25:57-63; Walbot (2000) *Curr. Opin. Plant Biol.* 2:103-107; Gai *et al.* (2000) *Nucleic Acids Res.* 28:94-96; Fitzmaurice *et al.* (1999) *Genetics* 153:1919-1928).

Métodos adicionais para diminuição ou eliminação da expressão de genes endógenos em plantas são bem conhecidos da técnica e podem ser aplicados de modo similar à presente invenção. Esses métodos incluem outras formas de mutagênese, tais como a mutagênese induzida por etil metanosulfonato, mutagênese de deleção, e mutagênese de deleção rápida de nêutrons usados em um senso genético reverso (com PCR) para identificar linhagens de plantas em que o gene endógeno foi deletado. Para exemplos desse método ver Ohshima *et al.* (1998) *Virology* 243:472-481; Okubara *et al.* (1994) *Genetics* 137:867-874; e Quesada *et al.* (2000) *Genetics* 154:421-436; cada um dos quais está incorporado aqui por referência. Além disso, um método rápido e automatizável para rastreamento de mutações quimicamente induzidas, o TILLING (Targeting Induced Local Lesions In Genomes, em português, Mira de Lesões Locais Induzidas em Genomas), que usa a desnaturação HPLC ou digestão seletiva de endonuclease de produtos de PCR selecionados também é aplicável à presente invenção. Ver McCallum *et al.* (2000) *Nat. Biotechnol.* 18:455-457, incorporados aqui por referência.

As mutações que têm impacto na expressão de genes ou que interferem na função da proteína codificada do citocromo P450 podem ser determinadas com o uso de métodos bem conhecidos da técnica. Mutações insercionais em éxons de genes geralmente resultam em mutantes nulos. Mutações em tecidos conservados podem ser particularmente efetivas na inibição da função metabólica da proteína codificada. Têm sido descritos resíduos conservados de polipeptídeos do citocromo P450 adequados para mutagênese com a meta de eliminar a atividade de um polipeptídeo do citocromo P450 na conversão da nicotina em nornicotina em uma planta ou parte de planta de *Nicotiana* (Ver, por exemplo, Figuras 3 e 4) Tais mutantes podem ser isolados de acordo com procedimentos bem conhecidos da técnica.

Em outra concretização desta invenção, mutantes dominantes podem ser usados para precipitar o silenciamento de RNA devido à inversão e recombinação de

genes de um locus de gene duplicado. Ver, por exemplo, Kusaba *et al.* (2003) *Plant Cell* 15:1455-1467.

Apesar de várias sequências serem reconhecidas na prática da invenção, em particular as ID SEQ N°:3 e ID SEQ N°:5 encontram uso específico. Embora não ligadas a quaisquer mecanismos de ação específicos, acredita-se que essas sequências codificam a nicotina demetilase que catalisa a N-demetilação oxidativa de nicotina em normicotina. Portanto, métodos para inibir especificamente essas sequências codificadoras e não outras sequências de P450 podem ser benéficas para a planta recombinante. Ou seja, estratégias que levariam à inibição da função do gene nesse locus individual podem se mostrar superiores àquelas que inibem a família inteira de genes. As enzimas P450 estão envolvidas em muitos mecanismos na planta, sendo que sua inibição pode ser mostrar deletéria ou prejudicial ao crescimento e desenvolvimento da planta, ou pode ter impacto negativo sobre fatores como a capacidade de defesa da planta contra doenças. Igualmente, devido à implicação das enzimas P450 da planta de *Nicotiana* em metabólitos como o fenilpropanóides, alcalóides, terpenóides, lipídeos, glicosídeos cianogênicos, glucosinolatos, e outras inúmeras entidades químicas, a disfunção da atividade de p450 pode alterar componentes envolvidos no sabor, textura, ou outras propriedades do tabaco que teriam impacto sobre a utilidade comercial da planta. Portanto, pode ser preferido o uso dos métodos discutidos acima para inibir a expressão de modo que tenha como alvo específico a sequência codificadora da ID SEQ N°:3 ou ID SEQ N°:5, incluindo estratégias mutacionais direcionadas ao alvo, como a quimeroplastia. Ver, por exemplo, Stewart *et al.* (2000) *Biotechniques* 29(4): 838-843; Graham *et al.* (2002) *Biochim Biophys Acta* 1587(1):1-6, incorporados aqui por referência.

A proteína codificada pelo DNAC designada 3D_C12-10 (ID SEQ N°:4) difere de 3D_C12-7 (ID SEQ N°:6) em apenas dois resíduos aminoácidos imediatamente após a metionina de partida. Os códons correspondentes a esses aminoácidos foram incluídos dentro do primer de PCR usado para gerar o DNAC 3D_C12-7. Portanto, o template de RNAm a partir do qual 3D_C12-7 foi amplificado pode ser o mesmo que o correspondente ao gene 3D_C12-10, com as sequências do primer de PCR mediando as mudanças observadas na segunda e terceira sequências de aminoácidos. Apesar disso, os produtos de proteína codificados funcionariam identicamente. O local dos dois aminoácidos que diferem entre as proteínas previstas está na sequência

de sinal N-terminal que serve meramente para ancorar a proteína na membrana do retículo endoplasmático e, portanto, não se deve esperar que isso influencie as propriedades catalíticas da enzima.

Em outra concretização da invenção, as composições da invenção têm uso nos métodos de rastreamento para identificar plantas não-conversoras para uso em programas de cruzamento. Dessa forma, as sequências de nucleotídeos da invenção podem ser usadas para rastrear germoplasmas nativos para plantas não-conversoras que têm uma mutação estável em um ou mais genes P450 identificados aqui. Essas plantas não-conversoras identificadas por métodos da invenção podem ser usadas para desenvolver linhagens de cruzamento.

Além disso, para as sequências de nucleotídeos que codificam sequências de codificação de P450, as composições da invenção incluem uma sequência de íntrons na sequência de 3D_C12-10 que pode ser usada em métodos de rastreamento. Embora não ligadas por qualquer mecanismo de ação o(s) gene(s) 3D_C12-7/3D_C12-10 podem representar o(s) único(s) membro(s) da família 3D_C12 envolvidos na conversão metabólica de nicotina em nornicotina (e conforme declarado previamente, há uma boa probabilidade de que os DNAs de 3D_C12-7 e 3D_C12-10 se originem a partir de um único locus genético simples. Para certas aplicações seria útil ter um meio de diferenciar diagnosticamente esse membro específico da família de genes 3D_C12 do resto das sequências intimamente relacionadas dentro dessa família. Por exemplo, é possível que dentro do germoplasma de tabaco existente naturalmente (ou em populações mutagenizadas), podem existir acessos nos quais esse gene é naturalmente disfuncional e pode, portanto, ser valioso como um recurso permanentemente não-conversor. Um método de ensaio específico para tais genótipos (por ex., mutantes de deleção, rearranjos, e semelhantes) poderia servir como uma ferramenta poderosa. Para obter tal ferramenta, o alinhamento de sequência mostrado na Figura 3A-3G foi usado para desenhar primers de PCR em regiões que possuem polimorfismos entre os membros. Uma combinação de primer (primer 5' mostrado na ID SEQ N°:25 e primer 3' mostrado na ID SEQ N°:26) usando sequências específicas para 3D_C12-10 produz dois resultados particularmente úteis. (1) todos os produtos de PCR

amplificados a partir do DNA genômico de tabaco resultaram no mesmo produto exclusivo (conforme determinado pela análise de sequência de DNA); e (2) foi revelado que a presença de um íntron de 992 pb está localizada entre as sequências do primer (Figura 7; íntron mostrado na ID SEQ Nº:24).

5 Quando um DNAC que corresponde a um membro da família de 3D_C12 é usado como uma sonda de hibridização em um ensaio de Southern Blot do DNA genômico do tabaco, um padrão complexo é observado. Isso é esperado, dado que existem múltiplos membros intimamente relacionados dessa família de genes. Por causa das regiões de íntrons de genes serem tipicamente menos conservadas do que as
10 de éxons, é previsto que o uso de uma sonda específica de íntron reduza essa complexidade e possibilite uma melhor distinção do(s) gene(s) que correspondem ao gene 3D_C12-7/3D_C12-10 de outros membros da família. De fato, a sonda correspondente à sequência mostrada na Figura 7 resultou em um padrão de Southern Blot com uma complexidade muito mais reduzida. O uso de uma sonda de 3D_C12-
15 10 íntron-específico, e/ou primers de PCR usados para gerar o fragmento mostrado na Figura 7 fornece, portanto, ferramentas poderosas em ensaios para determinar se plantas de tabaco de ocorrência natural ou mutagenizadas possuem deleções ou rearranjos que possam produzir o gene inativo. Tal planta pode ser então usada em programas de cruzamento para criar linhagens de tabaco que são incapazes de realizar
20 a conversão.

Plantas transformadas, parte de plantas e produtos que têm conteúdo reduzido de nornicotina e NNN

Os polinucleotídeos do citocromo P450 da invenção, e variantes e fragmentos
25 desses, podem ser usados em métodos da presente invenção para inibir a expressão ou função do citocromo P450 que está envolvida na conversão metabólica da nicotina à nornicotina em uma planta. Dessa forma, sequências inibidoras que a expressão ou função alvo de um polipeptídeo do citocromo P450 revelado aqui são introduzidas dentro de uma planta ou célula de planta de interesse. Em algumas concretizações os
30 cassetes de expressão descritos aqui são introduzidos dentro de uma planta de interesse, por exemplo, uma planta de *Nicotiana*, como observado adiante, usando quaisquer métodos adequados de transformação conhecidos da técnica, incluindo aqueles descritos aqui.

Os métodos da invenção não dependem de um método específico para introduzir uma sequência dentro de uma planta ou parte de planta, somente aquele em que a sequência desejada obtém acesso ao interior de no mínimo uma célula de planta ou parte de planta. Métodos para introdução de sequências de polinucleotídeos em plantas são conhecidos da técnica e incluem, mas não se limitam a, métodos de transformação estáveis, métodos de transformação transientes, e métodos mediados por vírus.

Protocolos de transformação bem como protocolos para introdução de sequências heterólogas de polinucleotídeos para dentro de plantas variam dependendo do tipo de planta ou célula de planta alvo para a transformação. Métodos adequados de introdução de polinucleotídeos para dentro de células de planta da presente invenção incluem microinjeção ((Crossway *et al.* (1986) *Biotechniques* 4:320-334), eletroporation (Shillito *et al.* (1987) *Meth. Enzymol.* 153:313-336; Riggs *et al.* (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-5606), transformação mediada por *Agrobacterium* (Patente dos EUA N.ºs. 5.104.310, 5.149.645, 5.177.010, 5.231.019, 5.463.174, 5.464.763, 5.469.976, 4.762.785, 5.004.863, 5.159.135, 5.563.055, e 5.981.840), transferência direta de gene (Paszowski *et al.* (1984) *EMBO J.* 3:2717-2722), e aceleração balística de partículas (ver, por exemplo, Patentes do EUA N.ºs. 4.945.050, 5.141.131, 5.886.244, 5.879.918, e 5.932.782; Tomes *et al.* (1995) in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips (Springer-Verlag, Berlin); McCabe *et al.* (1988) *Biotechnology* 6:923-926). Ver também Weissinger *et al.* (1988) *Ann. Rev. Genet.* 22:421-477; Christou *et al.* (1988) *Plant Physiol.* 87:671-674 (soja); McCabe *et al.* (1988) *Bio/Technology* 6:923-926 (soja); Finer e McMullen (1991) *In Vitro Cell Dev. Biol.* 27P:175-182 (soja); Singh *et al.* (1998) *Theor. Appl. Genet.* 96:319-324 (soja); De Wet *et al.* (1985) in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman *et al.* (Longman, New York), pp. 197-209 (pólen); Kaeppler *et al.* (1990) *Plant Cell Reports* 9:415-418 e Kaeppler *et al.* (1992) *Theor. Appl. Genet.* 84:560-566 (transformação mediada por whisker); D'Halluin *et al.* (1992) *Plant Cell* 4:1495-1505 (eletroporação); todos os quais estão incorporados aqui por referência.

Qualquer tecido de planta que possa ser posteriormente propagado com o uso de métodos de clonagem, seja por organogênese ou embriogênese, pode ser transformado com um construto recombinante abrangendo uma sequência inibidora

do citocromo P450, por exemplo, um cassete de expressão da presente invenção. Por “organogênese” entende-se o processo pelo qual brotos e raízes são desenvolvidos sequencialmente a partir de centros meristemáticos. Por “embriogênese” entende-se o processo pelo qual brotos e raízes desenvolvem-se juntamente de modo conjunto (não sequencialmente), seja a partir de células somáticas ou gametas. Tecidos exemplares que são adequados para vários protocolos de transformação descritos aqui incluem, mas não se limitam a, tecido do calo, tecido meristemático existente (por ex. meristemas apicais, brotos axilares, e meristemas de raiz) e tecido de meristema induzido (por ex. meristema do cotilédone e meristema do hipocótilo), hipocótilos, cotilédones, receptáculos de folha, pólen, embriões, e semelhantes.

Conforme usado aqui, o termo “transformação estável” pretende significar que o construto do nucleotídeo de interesse introduzido dentro de uma planta se integra dentro do genoma da planta e é capaz de ser herdado pela progênie da mesma. “Transformação transiente” pretende significar que uma sequência é introduzida dentro da planta e é expressa apenas temporariamente ou está presente apenas transitoriamente na planta.

Em concretizações específicas, as sequências inibidoras da invenção podem ser fornecidas a uma planta usando uma variedade de métodos de transformação transiente. As **sequência inibidoras da invenção** podem ser temporariamente transformadas dentro da planta com o uso de técnicas conhecidas. Tais técnicas incluem sistemas de vetores virais e a precipitação de polinucleotídeo de modo que evite a liberação subsequente de DNA. Portanto, a transcrição a partir de DNA ligado por partículas pode ocorrer, mas a frequência no qual ele é liberado para se tornar integrado dentro do genoma é muito reduzida. Tais métodos incluem o uso de partículas revestidas com polietilenoimina (PEI; Sigma #P3143).

Em outras concretizações, a sequência inibidora da invenção pode ser introduzida dentro de plantas colocando-as em contato com um vírus ou ácidos nucléicos virais. Geralmente, tais métodos envolvem a incorporação de um cassete de expressão da invenção dentro de uma molécula de DNA ou RNA viral. Admite-se que os promotores para uso em cassetes de expressão da invenção também abrangem promotores utilizados para transcrição por polimerases de RNA viral. Métodos para introdução de polinucleotídeos dentro de plantas e para expressão de uma proteína codificada dessas, envolvendo moléculas de DNA ou RNA viral, são conhecidos da

técnica. Ver, por exemplo, Patentes do EUA N°s 5.889.191, 5.889.190, 5.866.785, 5.589.367, 5.316.931, e Porta *et al.* (1996) *Molecular Biotechnology* 5:209-221; incorporadas aqui referência.

Células transformadas podem ser crescidas dentro de plantas de *Nicotiana* de acordo com métodos convencionais. Ver, por exemplo, métodos revelados aqui em Vasil e Hildebrandt (1965) *Science* 150:889; Negaard e Hoffman (1989) *Biotechniques* 7(8):808-812. Essas plantas podem então ser crescidas, e mesmo polinizadas com a mesma linhagem transformada ou linhagens diferentes, e a progênie resultante terá a expressão desejada da característica fenotípica identificada, i.e. expressão reduzida de um ou mais citocromos P450 que estão envolvidos na conversão metabólica de nicotina a normicotina, e portanto um conteúdo reduzido de normicotina, e um conteúdo concomitantemente reduzido de seu metabólito nitrosamina, NNN, na planta, particularmente em tecidos da folha. Duas ou mais gerações podem ser crescidas para assegurar que a expressão da característica fenotípica desejada seja mantida estável e herdada, e as sementes colhidas para assegurar que a expressão da característica fenotípica desejada foi alcançada. Dessa maneira, a presente invenção apresenta sementes transformadas (também referidas como "semente transgênica") que tem um polinucleotídeo da invenção, por exemplo, um cassete de expressão da invenção, estavelmente incorporado dentro de seu genoma.

As composições e métodos da invenção podem ser usados para reduzir o conteúdo de normicotina, particularmente nas folhas e caules, de qualquer planta do gênero *Nicotiana* incluindo, mas não se limitando às seguintes espécies: *acuminata*, *affinis*, *alata*, *attenuate*, *bigelovii*, *clevelandii*, *excelsior*, *forgetiana*, *glauca*, *glutinosa*, *langsдорffii*, *longiflora*, *obtusifolia*, *palmeri*, *paniculata*, *plumbaginifolia*, *quadrivalvis*, *repanda*, *rustica*, *suaveolens*, *sylvestris*, *tabacum*, *tomentosa*, *trigonophylla*, e *x sanderae*. A presente invenção também abrange a transformação de quaisquer variedades de uma planta do gênero *Nicotiana*, incluindo mas não se limitando a *Nicotiana acuminata multiflora*, *Nicotiana alata grandiflora*, *Nicotiana bigelovii quadrivalvis*, *Nicotiana bigelovii wallacei*, *Nicotiana obtusifolia obtusifolia*, *Nicotiana obtusifolia plameri*, *Nicotiana quadrivalvis bigelovii*, *Nicotiana quadrivalvis quadrivalvis*, *Nicotiana quadrivalvis wallacei*, and *Nicotiana trigonophylla palmeri*,

bem como variedades conhecidas como *variedades* de fumeiro ou brilhante, *variedades* Burley, *variedades* escuras, e *variedades* oriental/Turca.

As plantas transgênicas do gênero *Nicotiana* descritas aqui são adequadas para técnicas de crescimento e colheita convencionais, como o cultivo em solo rico em esterco ou sem esterco, com ensacamento das flores ou sem ensacamento, com cobertura ou sem cobertura. As folhas e caules colhidos podem ser usados em qualquer produto convencional do tabaco incluindo, mas não se limitando a, cachimbo, charuto, e cigarro de tabaco, tabaco para mascar, e em qualquer forma, incluindo tabaco em folha, tabaco triturado, ou tabaco cortado.

Portanto a presente invenção apresenta uma planta de *Nicotiana*, particularmente tecidos da folha dessas plantas, que abrangem um cassete de expressão da invenção e uma quantidade reduzida de nornicotina e *N'*-nitrosonornicotina. Conforme usado aqui, o termo “quantidade reduzida” ou “nível reduzido” pretende se referir a uma quantidade de nornicotina e/ou *N'*-nitrosonornicotina em uma planta tratada ou transgênica do gênero *Nicotiana* ou uma parte de planta ou produto de tabaco dessa planta, que é menor do que a quantidade que seria encontrada em uma planta ou parte de planta ou produto do tabaco a partir da mesma variedade de tabaco, processada (i.e. cultivada e colhida) da mesma forma, que não tenha sido tratada ou não transformada em transgênica para nornicotina e/ou *N'*-nitrosonornicotina reduzida. A quantidade de nornicotina pode ser reduzida em cerca de 10% a mais de cerca de 90%, incluindo mais do que cerca de 20%, cerca de 30%, cerca de 40%, cerca de 50%, cerca de 60%, cerca de 70%, e cerca de 80%.

O termo “produtos do tabaco”, conforme usado aqui incluem, mas não se limitam a, materiais do fumo (por ex., tabaco de cigarros, cachimbo, charuto), rapé, tabaco para mascar, goma, e tabletes. A presente invenção também abrange uma série de misturas de produtos do tabaco que podem ser feitas pela combinação do tabaco convencional com quantidades diferentes de baixa nornicotina e/ou *N'*-nitrosonornicotina descritas aqui. Em concretizações adicionais, a planta ou parte de planta do gênero *Nicotiana* como descrito acima é tabaco curado.

Em algumas concretizações da presente invenção, o produto do tabaco reduz o potencial carcinogênico da fumaça de tabaco que é inalada diretamente com o consumo de um produto de tabaco como charutos, cigarros, ou cachimbo, ou inalada por fumo secundário (i.e., por aqueles indivíduos que inalam a fumaça de tabaco

gerada por um indivíduo que consome um produto de tabaco como charutos, cigarros, ou cachimbo. O tabaco curado descrito aqui pode ser usado para preparar um produto de tabaco, particularmente um que passe por mudanças químicas devido ao aquecimento, abrangendo uma quantidade reduzida de nornicotina e/ou *N'*-nitrosonornicotina na fumaça inalada diretamente ou como fumo secundário. Da mesma forma, os produtos de tabaco da invenção podem ser úteis na preparação dos produtos de tabaco sem fumaça como tabaco para mascar, rapé, e semelhantes.

Os produtos de tabaco obtidos a partir de plantas de tabaco transgênicas da presente invenção encontram uso em métodos para reduzir o potencial carcinogênico desses produtos do tabaco, e reduzir a exposição de humanos à nitrosamina NNN carcinogênica, particularmente para indivíduos que são usuários desses produtos de tabaco.

Os seguintes exemplos são oferecidos como ilustração e não com o objetivo de limitação.

15

EXPERIMENTAL

Os seguintes materiais e protocolos foram utilizados nos experimentos descritos abaixo.

Materiais de plantas

Todos os materiais de plantas utilizados nesses experimentos foram fornecidos por Dr. Earl Wernsman, Department of Crop Science (Departamento de Ciências de Safra), North Carolina State University. DH 91-1307-46(NC) e DH91-1307-46(Con) são linhagens do tipo Burley duplo-haplóides quase isogênicas (não-conversoras e conversoras, respectivamente) recuperadas a partir da mesma planta haplóide materna. Linhagens do tipo Burley DH 98-326-3 (não-conversoras) e DH 98-326-1 (conversoras), e DH 98-325-5 (não-conversoras) e DH 98-325-6 (conversoras) representam dois pares adicionais de linhagens quase isogênicas. SC58 é uma variedade de tabaco curada em fumeiro, cujos indivíduos não-conversores são designados SC58($C_T C_T$). SC58($C_T C_T$) é uma linhagem conversora estável quase isogênica que é originada apesar da introgressão de um locus conversor dominante simples (C_T) encontrado nas espécies progenitoras de tabaco *N. tomentosiformis* dentro de SC58 (Mann *et al.* (1964) *Crop Sci.* 4:349-353. Após oito

retrocruzamentos adicionais para SC58, a linhagem SC58(C_TC_T) quase isogênica foi criada e posteriormente mantida via autofertilização.

5 Todas as plantas foram mantidas em câmaras de crescimento ou estufas usando envasamento com solo e fertilizante padrão. Para estudos de microarranjo, o metabolismo de nicotina em nornicotina foi acelerado pela excisão de folhas individuais e inserção dos pecíolos dentro de uma solução a 0,1% de etefon ou 1% de bicarbonato. As folhas foram então colocadas em uma câmara de crescimento (27°C) por 5 a 7 horas para facilitar a entrada do etefon ou bicarbonato de sódio através do
10 fluxo transpiracional. As folhas tratadas foram colocadas em pequenos sacos plásticos de armazenagem e foram levemente borrifadas com água (para manter a umidade alta) e curadas por três dias a 30°C no escuro. Para aumentar a conversão de nicotina em nornicotina nas plantas transgências geradas neste estudo, as folhas removidas foram imersas em uma solução de 0,2% de etefon, secas, e curadas em sacos plásticos de armazenagem por sete dias à temperatura ambiente.

15

Bibliotecas de DNAc e etiquetas de sequência expressa

O RNA celular total foi isolado a partir do tecido de folha de senescência de linhagens tipo Burley DH 91-1307-46(NC) e DH 91-1307-46(Con) usando o reagente TRIzol[®] de acordo com o protocolo do fabricante (Invitrogen). O RNA PolyA+ foi isolado a partir do RNA total usando o sistema MessageMaker (Invitrogen), e o DNAc foi subsequentemente sintetizado e clonado para dentro vetor lambda fago ZAP II usando a Síntese ZAP-DNAc e o Kit de clonagem GigaPack III Gold (Stratagene). Alíquotas das bibliotecas de fagos foram convertidas para as bibliotecas do plasmídeo com base em pBluescript seguindo o protocolo de excisão de massa descrito pela Stratagene.

Milhares de colônias de ambas as bibliotecas conversoras e não-conversoras foram crescidas em um meio sólido seletivo e colocadas numa placa de 384 poços contendo meio de cultura Luria (com ampicilina) em 10% de glicerol para facilitar a armazenagem dos clones em longo prazo a -80°C. Mais de 11.000 clones da cada biblioteca foram transferidos das placas de 384 poços para blocos de crescimento de 96 poços e crescidas em meio seletivo. Os plasmídeos foram isolados em formato de 96 poços usando o Kit de preparação R.E.A.L. (Qiagen) com a ajuda de uma Workstation BioRobot 3000 (Qiagen). Para gerar as ETS (etiquetas de sequência expressa, do inglês, "expressed sequence tags") os clones de plasmídeos foram sequenciados usando o Primer T3 (Qiagen) e o BigDye[®] Terminator system (Applied Biosystems) de acordo com o protocolo de sequenciamento de ciclo BigDye[®]. Placas de 96 poços Performa[®] DTR (Edge Biosystems) foram usadas para remover o corante não incorporado das reações de sequenciamento antes de carregar as amostras sobre um Sequenciador de DNA Automatizado de 96 capilares Perkin Elmer Prism 3700.

25

Preparação dos chips de DNA

Para obter DNAs adequados para o arranjo em lâminas de vidro, o M13 a seguir e os primers de sequenciamento reverso (Qiagen) foram usados como primers de PCR para amplificar os insertos de DNAc de plasmídeos contendo DNAs representados nas bases de dados de EST. Os clones de plasmídeos foram submetidos a PCR no formato de 96 poços usando um termociclador modelo 9700 Amp da Applied Biosystems Gene. Os produtos de PCR resultantes foram processados através de sistemas de purificação Millipore Multiscreen[™] PCR ou Montage[™]

30

PCR μ 96. Os produtos resultantes foram transferidos para dentro de placas de 384 poços contendo volumes iguais de DMSO. As concentrações finais de DNA foram estimadas como sendo iguais ou maiores do que 0,1 mg/ml. Os DNAs foram subsequentemente arranjados sobre uma lâmina cobertas com amino silano (Corning® GAPS II) usando uma impressora de arranjo Affymetrix GMS 417. Os DNAs foram imobilizados na superfície da lâmina através de ligação cruzada com UV (~120mJ/m²), seguida por cozedura a 75°C por duas horas.

Análise de hibridização de microarranjo

10 O método indireto com base em amino-alil dUTP de incorporação de corante descrito pelo "The Institute of Genome Research" (<http://pga.tigr.org/protocols.html>) foi usado para rotular RNAs conversores e não-conversores com corante Cy3 e Cy3 fluorescente (Amersham Biosciences). Brevemente, 20 µg de RNA total foi transcrito em reverso em um volume de 30 µl contendo 400 unidades de SuperScript II RT
15 (Invitrogen), 6 ng de primers random hexamer, 0,5 mM cada de dATP, dCTP, e dGTP, 0,3 mM dTTP, e 0,2 mM amino-alil dUTP (Sigma) no primeiro tampão de síntese de primeira fita (Invitrogen). As reações foram incubadas por 6 a 14 horas a 42°C, seguido por hidrólise do RNA com NaOH. As moléculas de DNAc de primeira fita resultantes foram purificadas (Qiagen) em coluna e lavadas com tampão de
20 fosfato. As reações de acoplamento dos corantes NHS-éster Cy3 ou Cy5 fluorescente para o DNAc ocorreram durante a incubação em 0,05 M de tampão de carbonato de sódio (pH 9) e 25% de DMSO à temperatura ambiente por 1,5 horas.

Lâminas de microensaio foram pré-hibridizadas em uma solução de 5X SSC, 0,1% de SDS, e 1% de BSA a 42°C por 45 minutos, enxaguadas gentilmente com
25 dH₂O e isopropanol, e secas por centrifugação em baixa velocidade. Os DNAs rotulados Cy3 e Cy5 foram purificados em coluna (Qiagen), combinados, e hibridizados em lâminas de DNA em uma solução contendo 5X SSC, 0,5% de SDS, 5X de Denhardt, 0,45 µg/µl Poli A RNA, 0,45 µg/µl de DNA de timo de bezerro, e 50% de formamida. As lâminas foram incubadas com a solução de hibridização por
30 14 a 16 horas a 42°C. As lavagens pós-hibridização consistiram em incubações sequenciais de 4 minutos com as seguintes soluções: 1X SSC, 0,2% de SDS; 0,1X SSC, 0,2% de SDS; 0,1X de SSC, em um enxágüe final de 10 segundos com 0,01X SSC.

Os microensaaios foram subsequentemente escaneados com o uso de ScanArray 2.1 (GSI Lumonics) ou ScanArray Express (PerkinElmer). O escaneamento sequencial para Cy5 e Cy3 fluorescente foi realizado em uma resolução máxima de 10 gm/pixel, e a força do laser e ganho de PMT ajustados para fornecer

5 forças de sinal confiáveis e equivalentes. As imagens de arranjo obtidas foram quantificadas para a intensidade de sinal com software de análise QuantArray™ (PerkinElmer), usando o método baseado em histograma. As intensidades totais foram usadas como campos de saída de quantificação, e os conjuntos de dados obtidos foram salvos como Unicode, arquivos de texto tab-delimitados. A importação dos arquivos

10 de texto para o Microsoft Excel possibilitou o cálculo subsequente das proporções de Cy5/Cy3 e Cy3/Cy5, a estatística empregada para a identificação de genes candidatos.

Clonagem de comprimento inteiro e membros adicionais da família de genes 3D C12

15 Para clonar a região de codificação inteira de 3D_C12 e 7D_A06, uma estratégia 5'-RACE modificada foi empregada com o uso de um primer avançado vetor-específico pBluescript II (BlueSK; 5'-CGCTCTAGAACTAGTGATC-3'; ID SEQ NO:17) e um conjunto de primers reversos gene-específicos. Dois primers reversos 3D_C12-específicos foram desenhados, um dos quais é complementar à

20 porção à jusante da região não traduzida (5'-TTTTTGGGACAATCAGTCAA-3'; ID SEQ N°:18) e o outro complementar à sequência dentro da região de codificação (5'-GTTAGATTTATCGTACTCGAATT-3'; ID SEQ N°:19). Para o primer anterior, as primeiras cinco Ts são complementares à cauda polyA do transcrito. Um primer reverso 7D_A06-específico (5'-TTCATTTCAAATTATTTTATGCACCA-3'; ID SEQ

25 N°:20) também foi desenhado, sendo complementar a um segmento na região de não codificada desse gene. As reações de PCR continham 10 ng de biblioteca de DNAc de folha de tabaco conversor (dentro do vetor pBluescript) como template, concentração de 2 µM de cada primer, 350 µM de cada dNTP, e 1,5 mM MgCl₂ em um volume de reação final de 50 µL. A amplificação foi iniciada pela adição de 2,5

30 unidades de mix de enzima UinPol, usando condições descritas pelo fabricante (Roche). Após uma etapa de desnaturação inicial a 94°C por 4 minutos, as amostras foram submetidas a 30 ciclos de desnaturação a 94°C por 15 segundo, enrijecimento a 57°C por 30 segundos, e extensão a 72°C por 90 segundos. Uma etapa de extensão

final a 72°C por 10 minutos foi incluída ao final dos 30 ciclos. Os amplicons foram ligados dentro do vetor pGEM Easy T/A (Promega), e 10 clones selecionados aleatoriamente de cada amplificação foram submetidos à análise de sequência de DNA. As sequências de ácido nucléico e proteína previstas para os vários membros da família de genes 3D_C12 foram analisadas e comparadas usando BLASTX (Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402), ClustalW (Higgins *et al.* (1994) *Nucleic Acids Res.* 4673-4680), e os algoritmos GAP (pacote de software do Genetic Computing Group da Universidade de Wisconsin).

A estratégia descrita acima foi efetiva na identificação da informação de sequência de comprimento inteiro para 3D_C12 e 7D_A06. Além disso, as amplificações de PCR usando o primer de PCR interno à região de codificação de 3D_C12 originou informação sobre a sequência parcial para o DNAc único 3D_C12-15. Na tentativa de obter informação sobre a sequência de comprimento inteiro para 3D_C12-15, um primer gene-específico complementar ao 5' terminal de sua região de codificação (5'-ATGGTTTTTCCCATAGAAGCC-3'; ID SEQ N°:21) foi usado em conjunto com um primer reverso pBluescript-específico (5'-TCGAGGTCGACGGTATC-3'; ID SEQ N°:22). Embora o DNAc 3D_C12-15 de comprimento inteiro não tenha sido recuperado, essa amplificação resultou no isolamento de 3D_C12-7, que provou ser outro membro único da família de genes 3D_C12.

Análise de plantas transgênicas

Os construtos silenciadores de gene baseados em RNAi foram montados numa versão do vetor de clonagem pKYL80 (Schardl *et al.* (1987) *Gene* 61:1-11) que foi engenheirado para conter um fragmento de 151-pb do íntron do gene FAD3 de soja entre os sítios de restrição XhoI e SacI do polylinker (pKYLX80I). Para criar um construto no qual o íntron FAD3 foi flanqueado por um fragmento senso e antisenso de 3D_C12, uma região de 99-pb localizada imediatamente à montante do códon de parada do 3D_C12 cDNA (Figura 3A-3G) foi clonada entre os sítios de restrição HindIII - XhoI e SacI - XbaI de pKYLX80I em sua orientação senso e antisenso, respectivamente. O fragmento HindIII - XbaI resultante contendo o braço senso 3D_C12, íntron FAD3, e braço antisenso 3D_C12 foi subclonado para dentro do vetor binário de expressão de planta pKYLX71 (Maiti *et al.* (1993) *Proc. Natl. Acad. Sci.*

USA 90:6110-6114) entre o promotor 35S CaMV e um terminador de pequena subunidade da rubisco.

Os construtos de expressão excessiva (sobre-expressão) foram criados para substituir a 3-glucuronidase do ORF (open reading frames, em português, quadro de
5 leitura aberta) do vetor binário de expressão da planta pBI121 (Clontech) com as regiões de codificação de comprimento inteiro 3D_C12, 7D_A06, e 3D_C12-7 cDNAs. Isso colocou o tabaco P450 sob o controle transcricional do promotor 35S do CaMV. Os construtos baseados em pBI121- e pKYLX71 foram transformados dentro de uma cepa de *Agrobacterium tumefaciens* LBA 4404 e introduzidos dentro dos
10 cultivares de tabaco Petite Havana e DH98-325-6 (conversor), respectivamente, usando protocolos estabelecidos (Horsch *et al* (1985) *Science* 227:1229-1231).

Análise de Northern Blot

Os RNAs celulares totais foram isolados a partir de folhas de tabaco usando o
15 método TRIZOL[®] conforme descrito pelo fabricante (Invitrogen). Cinco em dez microrganismos de RNA foram fracionados por tamanho sobre 1,2% de gel de agarose preparado em tampão TBE. A imobilização de RNA, rotulagem de sonda, e detecção de sinal foram realizados usando os kits DIG de detecção e rotulagem de ácido nucléico de acordo com as instruções do fabricante (Roche). De modo
20 alternativo, as sondas foram sintetizadas usando ³²P-dCTP de acordo com os protocolos que acompanham o kit Random Primed DNA Labeling (Roche).

Análise alcalóide

As folhas de tabaco foram colhidas e secas ao ar em um forno a 65°C por 2
25 dias. A uma amostra de 100 mg de folhas secas e esmagadas, foi adicionado 0,5 ml de NNaOH em um frasco de cintilação de 20 mL. A amostra foi misturada e deixou-se incubar por 15 minutos à temperatura ambiente. Alcalóides foram extraídos pela adição de 5 mL de solução de extração [0,04% de quinolina (peso/volume) dissolvida em metil-t-butil éter] e rotacionada gentilmente em um agitador linear por 3 horas.
30 Após a separação das fases, uma alíquota da fase orgânica foi transferida para um frasco de amostra. As amostras foram analisadas usando o cromatógrafo a gás Autosystem XL da PerkinElmer, equipado com um detector de ionização de chama, um liner de vidro de modo split/splitless de 4mm, uma coluna ID DB-5 de 30 m x

0.53 mm. As condições cromatográficas foram as seguintes: detector de temperatura: 250°C; temperatura do injetor: 250°C; taxa de fluxo de hélio da 120°C: 20 mL/min; volume de injeção: 2 µL; condições da coluna: 120°C, mantido por 1 minuto, taxa de alteração 120-280°C a 30°C /minuto, mantido a 280°C por 2 minutos. A composição
5 do alcalóide foi determinada pelo software TotalChrome Navigator usando uma curva de calibração.

Exemplo 1: Geração das bases de dados de EST

10 RNAs isolados a partir de folhas de senescência do genótipo conversor DH 91-1307-46(Con) e sua contraparte não-conversora quase isogênica DH 91-1307-46(NC) foram usados para gerar bibliotecas de DNAC. Sequenciadores automatizados de DNA de alto rendimento foram usados inicialmente para gerar informação de sequência de série única (ESTs) para 11.136 DNACs escolhidos aleatoriamente a partir da biblioteca conversora. A ferramenta de busca de alinhamento local (Altschul
15 *et al.* (1990) *J. Mol. Biol.* 215:403-410) foi usada para comparar a sequência de proteína prevista de cada DNAC de tabaco com a base de dados de proteínas não redundante controlada pelo National Center for Biotechnology Information da National Library of Medicine e National Institutes of Health (em português, Centro Nacional da Informação sobre Biotecnologia da Biblioteca Nacional de Medicina e
20 Institutos Nacionais de Saúde). Subsequentemente, uma base de dados de EST similar anotada foi gerada pela condução de séries de sequenciamento em 11.904 DNAs selecionados a partir da biblioteca não-conversora.

Exemplo 2: Análises de microarranjo da biblioteca de CDNA conversora

25 *Métodos*

Após a finalização da base de dados EST gerada a partir da biblioteca de DNAC conversora, os insertos a partir de 4992 clones foram amplificados por PCR e arranjados sobre lâminas de vidro. Dada a possibilidade da enzima nicotina demetilase poder ser catalisada por uma enzima da classe P450 de enzimas oxidativas,
30 atenção especial foi dada às entradas na biblioteca que foram previstas por análise BLASTX para codificar P450.

A partir de inspeção visual dos resultados da BLASTX, foi estimado que 31 genes P450 únicos foram representados na base de dados. Quando foram selecionadas placas de 96 poços específicas para serem incluídas no microensaio, tomou-se cuidado para assegurar que todos os genes P450 únicos fossem incluídos entre os 4992 selecionados.

O RNA isolado a partir dos genótipos do tipo Burley DH 98-326-3 quase isogênicos (não-conversores) e DH 98-326-1 (conversores), e DH 98-325-5 (não-conversores) e DH 98-325-6 (conversores) foram usados para gerar DNAs rotulados de Cy3 e Cy5. Para maximizar a conversão metabólica de nicotina em normicotina em genótipos conversores, as folhas retiradas foram tratadas com bicarbonato de sódio e etefon antes da cura, tratamentos que mostraram acelerar a produção de normicotina em plantas conversoras, tendo nenhum efeito em indivíduos não-conversores (Fannin e Bush (1992) *Med. Sci. Res.* 20:867-868; Shi *et al.* (2003) *J. Agric. Food Chem.* 51:7679-7683).

Para minimizar a variabilidade inerente aos experimentos de microarranjos, experimentos recíprocos foram conduzidos simultaneamente. Desse modo, o RNA DH 98-325-5 foi rotulado com Cy5 e o RNA DH 98-325-6 foi rotulado com Cy3, e então um experimento recíproco de RNA DH 98-325-5 foi rotulado com Cy3 e o RNA DH 98-325-6 RNA foi rotulado com Cy5 (coletivamente referido como Exp. 2.1). De modo similar, os RNAs DH 98-326-3 e DH 98-326-1 foram rotulados com Cy3 e Cy5, respectivamente, em um experimento, e então os mesmos RNAs foram rotulados com Cy5 e Cy3, respectivamente, em um experimento recíproco (coletivamente referido como Exp. 2.2).

Mesmo quando conduzidos reciprocamente, os resultados de quaisquer experimentos de microarranjos provavelmente incluem “falso positivos”, representando genes que são regulados diferencialmente entre um par genotípico específico e/ou unicamente em resposta a um tratamento específico, em oposição às diferenças diretamente associadas ao fenômeno de conversão. Para definir um conjunto de genes candidatos que mais provavelmente são regulados para cima devido ao processo de conversão, os DNAs foram identificados para satisfazer aos seguintes critérios: Para qualquer conjunto de experimentos recíprocos (i.e. Exp. 2.1, ou Exp. 2.2), a intensidade de hibridização de um dado DNA teve de ser no mínimo 2 vezes mais alta com a sonda conversora do que com a sonda não-conversora, em no

mínimo uma das hibridizações, não menos do que 1,5 vezes mais alta no experimento recíproco.

Experimento 2.1 – Folhas de linhagens quase isogênicas DH 98-325-5 e DH 98-325-6 foram tratadas com etefon e curadas por 3 dias a 30°C. A análise de
5 alcalóide revelou que virtualmente toda a nicotina tinha sido metabolizada em nornicotina na folha de DH 98-325-6 durante esse período enquanto o mínimo de nornicotina foi observado na folha de DH 98-325-5. RNAs da planta DH 98-325-5 não-conversora foram rotulados com corante fluorescente Cy3, e os RNAs extraídos da folha de DH 98-325-6 (conversora) foram rotulados com Cy5. Os DNAs
10 rotulados com Cy3 e Cy5 foram incubados juntamente no mesmo chip de DNA, deixando-os hibridizar por uma noite.

Experimento 2.2 – Uma análise de microarranjo similar a Exp. 2.1 foi conduzida com o uso de linhagens quase isogênicas DH 98-326-3 (não-conversoras) e DH 98-326-1 (conversoras). Nesses experimentos, as folhas de cada genótipo foram
15 tratadas com 1% de bicarbonato de sódio e curadas por 3 dias a 30°C. Ao final do período de tratamento, a nicotina foi o alcalóide predominante na folha de DH 98-326-3, enquanto que aproximadamente todo o alcalóide encontrado na folha de DH 98-326-1 foi a nornicotina. Conforme descrito no Exp. 2.1, esses experimentos foram conduzidos reciprocamente.

20

Resultados

Em ambos os Experimentos 2.1 e 2.1, a grande maioria dos 4992 DNAs arranjados sobre as lâminas de vidro mostraram diferenças substanciais em suas intensidades de hibridização em relação às sondas rotuladas com Cy3 e Cy5.

25 Dos 4992 DNAs arranjados sobre as lâminas de vidro, apenas cinco mostraram no mínimo uma expressão duas vezes mais alta em uma hibridização, e não menos do que 1,5 vezes na hibridização recíproca para ambos os Exp. 2.1, e Exp. 2.2. Essas entradas foram designadas 3D_C12, 7D_A06, 27C_C12, 33A_D06, e 34D_F06. A análise BLASTX da informação de sequência parcial para 3D_C12 e
30 7D_A06 encontrada em nossa base de dados EST previu que os DNAs codificam duas enzimas intimamente relacionadas à enzima P450. Previu-se que 27C_C12 e 33A_D06 codificam proteínas de parede celular rica em glicina, mostrando mais de 90% de identidade de sequência com pequenas proteínas ricas em glicina de tabaco

encontradas no GenBank (e.g., Nº de Acesso. AAK57546). Constatou-se que o clone 34D_F06 contém um duplo inserto de DNAC, um inserto que mostra homologia com as quinases das proteínas serina/treonina, e o outro mostra alta identidade de sequência com as mesmas proteínas de parede celular rica em glicina como os DNACs de 27C_C12 e 33A D06.

Exemplo 3: Análises de microarranjo da biblioteca de CDNA não-conversora

Após a finalização da base de dados de EST a partir da biblioteca não-conversora (gerada de folhas senescentes de genótipo DH 91-1307-46 (NC), outro conjunto de experimentos de microarranjo foi iniciado. Para a primeira geração de microarranjos, a meta foi produzir lâminas de vidro contendo um conjunto completo e não redundante de genes representados em ambas as bibliotecas.

Para obter uma estimativa do número de genes únicos que estão representados na base de dados, foi conduzida uma análise de agrupamento para identificar ESTs previstas para serem representadas múltiplas vezes na base de dados (contigs) versus aquelas previstas para serem representadas em apenas uma vez (singletons) (Huang and Madan (1999) *Genome Res.* 9:868-877). Devido à natureza dos algoritmos de agrupamento, as sequências que mostraram identidade de sequência alta, mas imperfeita, foram agrupadas dentro do mesmo contig. O conjunto total de genes únicos previstos, ou unigenes, dentro de uma base de dados é calculado como a soma dos contigs e singletons. A análise de agrupamento das bases de dados conversora e não-conversora combinadas previu 2246 contigs e 4717 singletons para um total de 6.963 unigenes. Insertos de todos os singletons e um individual para cada contig foram amplificados por PCR e arranjados em lâminas de vidro, resultando num chip de gene contendo o conjunto completo de 6.963 unigenes.

Além da criação de um novo chip de DNA, os materiais genéticos usados para gerar sondas de hibridização também diferiram daqueles usados no Exemplo 2. SC58 é uma variedade de tabaco curada por fumo, cujos indivíduos não-conversores são designados SC58(c_Tc_T). SC58(C_TC_T) é uma linhagem conversora estável quase isogênica que é originada apesar da introgressão de um locus conversor dominante simples (C_T) encontrado nas espécies progenitoras de tabaco *N. tomentosiformis* dentro de SC58 (Mann *et al.* (1964) *Crop Sci.* 4:349-353). Após oito

retrocruzamentos adicionais para SC58, a linhagem SC58($C_T C_T$) quase isogênica foi criada e posteriormente mantida via autofertilização. O fenótipo de conversão de plantas SC58($C_T C_T$) é único em relação às linhagens padrão de tabaco conversor em que o metabolismo de nicotina em nornicotina na folha não requer senescência ou curagem. Plantas que possuem o locus conversor C_T de *N. tomentosiformis* contém a nornicotina como o alcalóide predominante mesmo no tecido de folhas verdes (Wernsman e Matzinger (1968) *Tob. Sci.* 12:226-228)

Os RNAs isolados de tecido de folhas verdes SC58($c_T c_T$) e SC58($C_T C_T$) foram rotulados com Cy3 e Cy5, respectivamente, e simultaneamente hibridizados para um chip de DNA contendo o conjunto inteiro de 6.963 unigenes de DNACs. Os corantes fluorescentes foram revertidos para produzir as sondas para um experimento recíproco conforme descrito nos Exps. 2.1 e 2.2 do Exemplo 2.

Resultados

Os resultados foram avaliados usando os mesmos critérios do Exemplo 2, i.e., foram identificados os DNACs individuais que mostraram uma hibridização no mínimo duas vezes melhor em relação os DNACs rotulados de SC58($C_T C_T$) versus SC58($c_T c_T$) em um experimento, e no mínimo 1,5 vezes melhor no ensaio recíproco.

Os resultados foram comparados àqueles do Exp. 2.1, e Exp. 2.2 no Exemplo 2 acima. A hibridização melhorada dos RNAs conversores em membros codificadores de DNACs da mesma família de P450 intimamente relacionada foi o único resultado compartilhado por todos os três experimentos de microarranjo usando os critérios definidos. 131 A_A02 é o nome do cDNA que foi arranjado sobre o chip unigene de 6963 membros, representativo da família de P450 intimamente relacionada que inclui 3D_C12 e 7D_A06 (3D_C12 e 7D_A06 não foram arranjados sobre a lâmina unigene). Nenhum outro DNAC do arranjo do Exemplo 3, que incluiu representantes dos contigs contendo os DNACs codificadores de proteína rica em glicina 27C_C12 e 33A_D06 e 34D_F06, tiveram escores positivos pelo critérios definidos e também escore positivo no Exp. 2.1, ou Exp. 2.2 do Exemplo 2 acima, sem considerar se os resultados foram comparados individual ou coletivamente.

Tabela 1. Resultados de microarranjos de membros da família do gene 3D_C12

DNAc	Experimento 2.1			Experimento 2.1 (recíproco)		
	Leitura	Leitura	Razão	Leitura	Leitura	Razão
	Cy3	Cy5	Cy5/Cy3	Cy3	Cy5	Cy5/Cy3
3D_C12	15514,14	25928,95	1,67	19355,85	9507,87	2,04
7D_A06	15238,23	37196,19	2,44	13651,03	8121,04	1,68
	Experimento 2.2			Experimento 2.2 (recíproco)		
	Leitura	Leitura	Razão	Leitura	Leitura	Razão
	Cy3	Cy5	Cy5/Cy3	Cy3	Cy5	Cy5/Cy3
3D_C12	12756,43	28669,28	2,25	32198,81	16166,13	1,99
7D_A06	7571,06	19180,94	2,53	42408,85	18440,17	2,30
	Exemplo 3			Exemplo 3		
	Leitura	Leitura	Razão	Leitura	Leitura	Razão
	Cy3	Cy5	Cy5/Cy3	Cy3	Cy5	Cy5/Cy3
131A_A02	11138,96	19638,82	1,76	36963,45	10085,25	3,67

Resultados combinados

Os resultados combinados dos experimentos de microarranjo descritos acima definiram membros de uma família de genes do P450 intimamente relacionados, daqui em diante referida como a família 3D_C12, como sendo os melhores candidatos a desempenhar um papel direto na conversão metabólica da nicotina em nornicotina em plantas de tabaco conversoras. Os resultados da hibridização dos membros dessa família P450 em cada um dos três experimentos de microarranjos são mostrados na Tabela 1. Os resultados dos microarranjos foram confirmados independentemente com o uso de ensaios de Northern Blot. Conforme mostrado na Figura 2, um sinal aproximadamente duas vezes mais alto foi observado em folhas conversoras senescentes e curadas, em comparação as suas contrapartes não conversoras quando manchas (blots) de RNA foram incubadas com uma sonda de hibridização radiorotulada de 7D_A06.

Exemplo 4: Análise da sequência da família do gene 3D_C12

Uma vez que os experimentos de microarranjo definiram 3D_C12 e 7D_A06 como potencialmente envolvidos no processo de conversão, a obtenção de informação sobre a sequência completa de DNA para esses genes torna-se a próxima etapa para sua caracterização. Os clones originais de 3D_C12 e 7D_A06 que foram sequenciados quando da geração da base de dados de EST descrita aqui (e colocados sobre os microarranjos) não eram DNAs de comprimento inteiro. Para obter a sequência de comprimento inteiro, primers foram gerados para corresponder às regiões na região de flanqueamento 3' e no interior das regiões de codificação que foram suficientemente polimórficas para distinguir entre 7D_A06 e 3D_C12. Esses primeiros gene-específicos foram usados em combinação com primers específicos ao sítio de clonagem de pBluescript II para amplificar os DNAs a partir da biblioteca conversora de DNA numa tentativa de obter a sequência que incluísse os terminais 5' completos dos quadros de leitura de 3D_C12 e 7D_A06.

Essa estratégia levou à determinação da sequência de DNA correspondente às regiões de codificação completas de 3D_C12 (nt 1-1551 da ID SEQ Nº:1; sequência prevista de aminoácidos mostrada na ID SEQ Nº:2) e 7D_A06 (nt 1-1554 da ID SEQ Nº:7; sequência prevista de aminoácidos mostrada na ID SEQ Nº:8) (Figuras 3A-3G e 4). As análises GAP do DNA de 3D_C12 e 7D_A06 e as sequências de proteínas previstas mostraram que elas compartilham 93,4% da identidade de sequência e 92,3% da identidade em nível de proteína (Tabelas 2 e 3). A análise BLASTX inicial contra a base de dados GenBank não redundante revelou que 3D_C12 e 7D_A06 compartilham a maior homologia de sequência com CYP82E1, um gene P450 de tabaco de função desconhecida que é regulado para cima em resposta a elicitores fúngicos (Takemoto *et al.* (1999) *Plant Cell Physiol.* 40:1232-1242). A proteína CYP82E1 é 66,9% e 67,5% idêntica às sequências de aminoácidos previstas de 3D_C12 (ID SEQ Nº:2) e 7D_A06 (ID SEQ Nº:8), respectivamente, e a sequência de DNA de CYP82E1 é 72,1% e 73,5% idêntica às respectivas sequências de codificação de 3D_C12 (nt 1-1551 da ID SEQ Nº:1) e 7D_A06 (nt 1-1554 da ID SEQ Nº:7).

Tabela 2. Identities de sequência de nucleotídeo entre membros da família do gene 3D_C12

	3D_C12	7D_A06	3D_C12-7	3D_C12-10	3D_C12-15*
7D_A06	93,4**				
3D_C12-7	93,7	94,0			
3D_C12-10	93,7	94,4	99,7		
3D_C12-15*	95,5	92,6	93,1	92,8	
131A_A02*	98,0	94,0	93,4	93,1	93,1
*sequências parciais					
**Os números indicam as percentagens					

Tabela 3. Identities de sequência de aminoácidos previstos entre membros de comprimento inteiro da família do gene 3D_C12.

	3D_C12	7D_A06	3D_C12-7
7D_A06	92,3*		
3D_C12-7	92,8	94,8	
3D_C12-10	92,5	94,4	99,6
**os números indicam as percentagens			

- 5 Além de possibilitar a aquisição de informação sobre a sequência de comprimento inteiro para os DNAs de 3D_C12 e 7D_A06, as amplificações de PCR descritas acima produziram produtos adicionais que estavam intimamente relacionados a, ainda que claramente distintos, as sequências de DNA de 3D_C12 e 7D_A06. Usando um primer direcionado contra o interior da sequência de DNA de 3D_C12, em combinação com um primer específico ao pBluescript II, uma sequência

única designada 3D_C12-15 (Figura 3A-3G; ID SEQ Nº:9; a sequência de aminoácidos prevista mostrada em ID SEQ Nº:10) foi amplificada, além do produto esperado de 3D_C12. 3D_C12-15 é 95,5% idêntico à sequência de DNA correspondente de 3D_C12 e 92,6% idêntico a mesma região de 7D_A06 (Tabela 2).

5 Devido ao fragmento de 3D_C12-15 representar um membro distinto e adicional da família do gene 3D_C12, foi feita uma tentativa de obter a sequência de DNAC de comprimento inteiro desse gene. Um primer de PCR específico aos primeiros sete códons do quadro de leitura 3D_C12-15 foi usado em combinação com um primer pBluescript II-específico em uma reação de amplificação usando uma
10 biblioteca de DNAC conversora como template. A análise da sequência de vários produtos de amplificação independentes falhou em revelar o gene 3D_C12-15 de comprimento inteiro. Em vez disso, um novo membro dessa família foi recuperado, designado 3D_C12-7 (Figura 3A-3G; sequência codificadora apresentada como nt 1-1551 da ID SEQ Nº:5). Através de toda a sequência de nucleotídeos mostrada na ID
15 SEQ Nº:5, 3D_C12-7 compartilha 93,7% da identidade de sequência de nucleotídeos com 3D_C12 (através da ID SEQ Nº:1), 94,0% de identidade de sequência de nucleotídeo com 7D_A06 (através da ID SEQ Nº:7), e 93,1% de identidade com a região correspondente do fragmento 3D_C12-15 (ID SEQ Nº:9) (Tabela 2) A sequência de aminoácidos prevista de 3D_C12-7 (ID SEQ Nº:6) é 92,8% idêntica à
20 proteína de 3D_C12 (ID SEQ Nº:2), e 94,8% idêntica à proteína de 7D_A06 (ID SEQ Nº:8) (Tabela 3)

Dois membros adicionais da família 3D_C12 também foram identificados. Um gene designado 3D_C12-10 (Figura 3A-3G; sequência codificadora apresentada como nt 1-1551 da ID SEQ Nº:3; sequência codificadora apresentada como ID SEQ
25 Nº:4) foi recuperado a partir de uma reação de amplificação usando um primer de PCR complementar à sequência na região de flanqueamento 3' de 3D_C12 juntamente com um primer Bluescript II-específico (e a biblioteca conversora como template). 3D_C12-10 difere por apenas cinco posições de nucleotídeos da sequência de nucleotídeos de 3D_C12-7 (ID SEQ Nº:5) (Figura 3A-3G), e por apenas
30 duas posições de aminoácidos do produto de proteína de 3D_C12-7 previsto (ID SEQ Nº:6) (Figura 4).

Com a finalização da base de dados de EST não-conversora, outro membro da família de genes 3D_C12 foi revelado. A sequência parcial de DNA de 131A_A02

(ID SEQ N°:11; sequência de aminoácidos prevista apresentada na ID SEQ N°:12) que é encontrada nessa base de dados é 98,0% idêntica à sequência correspondente de 3D_C12, e 94,0% idêntica a mesma região de 7D_A06 (Figura 3A-3G e Tabela 2). Conforme descrito na seção anterior, 131A_A02 é membro da família de gene de 3D_C12 que foi representada no amplo chip unigene usado em ensaios de microarranjo conforme descrito aqui.

Exemplo 5: Análise de planta transgênica de membros da família de gene 3D_C12

Para determinar se os membros da família 3D_C12 de genes do citocromo P450 estão envolvidos na conversão metabólica da nicotina em nornicotina, foram geradas plantas transgênicas usando construtos desenhados seja para aumentar ou inibir a expressão do gene. Para testar os efeitos da atividade do gene relativa à regulação para cima, uma estratégia de RNA de interferência (RNAi) foi empregada. Uma região de 99-pb de 3D_C12 localizada imediatamente à montante do códon de parada (Figura 3A-3G), foi usada para criar um construto que formaria uma alça de RNAds dentro da célula da planta. Tais estruturas de RNAds são conhecidas como ativadoras de um complexo silenciador de RNAi que leva à degradação dos RNAs transgenes e dos RNAs endógenos que são idênticos ou altamente homólogos à sequência encontrada no RNAds (Wesley *et al.* (2001) *Plant J.* 27: 581-590; Waterhouse & Helliwell (2002) *Nat. Gen. Rev.* 4: 29-38).

Dado que cada membro do 3D_C12, caracterizado conforme descrito aqui, compartilha mais de 90% da identidade de sequência de DNA, espera-se que um construto de RNAi sintetizado a partir de um membro silencie a família de genes inteira. Especificamente, o construto de RNAi gerado a partir da sequência de 3D_C12 compartilha identidades de sequência de 90/99 e 91/99 com os DNAs de 7D_A06 e 3D_C12-7, respectivamente, sobre essa região (Figura 3A-3G). O construto 3D_C12/RNAi (também referido no Exemplo 7 como o construto 3D_C12Ri99) foi clonado à jusante do promotor 35S constitutivo do vírus do mosaico da couve-flor (CaMV, em inglês, cauliflower mosaic virus) e introduzido dentro de uma linhagem de tabaco Burley, DH 98-325-6 fortemente conversora usando a transformação mediada por *Agrobacterium*.

Uma característica distintiva de silenciamento mediado por RNAi é a redução marcada na acumulação do transcrito do gene em estado estacionário cuja atividade

tenha sido regulada para baixo. Para confirmar que o silenciamento de gene da família 3D_C12 ocorreu nas plantas que mostraram baixo fenótipo de nornicotina, uma análise de Northern blot foi conduzida usando RNAs isolados a partir de três das plantas transgênicas que possuíam construtos 3D_C12/RNAi e mostravam baixo fenótipo de nornicotina, dois indivíduos transformados com o construto 3D_C12/RNAi ainda mostrando altos níveis de nornicotina e uma das plantas-controle por vetor.

Para avaliar os efeitos da expressão excessiva da atividade do gene, os DNACs de três membros da família do gene 3D_C12, para os quais primeiro foi obtida informação da sequência de comprimento inteiro (3D_C12, 7D_A06, e 3D_C12-7) foram clonados na orientação senso à jusante do promotor de CaMV 35S. Esses construtos foram subsequentemente introduzidos para dentro de um cultivar de *N. tabacum* Petite Havana usando a transformação mediada por *Agrobacterium*. A linhagem Petite Havana é geralmente usada por pesquisadores por causa de sua baixa estatura e tempo abreviado de geração em relação aos cultivares comerciais de tabaco. A condição conversora/não-conversora do cultivar de Petite Havana é desconhecida, mas os ensaios de alcalóide do presente pedido de patente mostraram claramente que as plantas em nossa posse eram fortes conversoras.

Embora as plantas hospedeiras nesses experimentos fossem conversoras, a presente estratégia teve o objetivo de conduzir ensaios de alcalóides em tecidos verdes e não curados, em que uma quantidade mínima de nornicotina se acumula igualmente em planta conversoras e não-conversoras (e o promotor de CaMV 35S é muito ativo). Na verdade, uma linhagem não-conversora foi propositadamente escolhida por causa da cultura do tecido - conforme requerido quando se conduz a transformação mediada por *Agrobacterium* - a qual é conhecida por aumentar a frequência da conversão genética, e portanto complicaria potencialmente a interpretação dos resultados (por ex., avaliar se um novo fenótipo era atribuível somente ao transgene ou resultante da planta ter passado por conversão genética).

30 *Resultados*

Dado o alto grau de variabilidade tipicamente observado entre plantas transgênicas independentes com o mesmo construto de transgenes, 10 indivíduos transformados independentemente foram selecionados para avaliar os efeitos do

construto 3D_C12/RNAi sobre a conversão metabólica de nicotina em nornicotina. Folhas de cada um dos 10 indivíduos 3D_C12/RNAi, além de duas plantas-controle transformadas apenas pelo vetor pBI121, foram tratadas com etefon e curadas por sete dias. A análise de alcalóide desses materiais é mostrada na Tabela 4.

Tabela 4. Análise de alcalóide de plantas DH 98-325-6 transformadas independentemente com o construto 3D_C12/RNAi (e vetor controle pBI121). As folhas foram tratadas com etefon e curadas por sete dias.

Amostra	% Nicotina*	% Nornicotina*	% Anabasina*	% Anatabina*	% Conversão**
3D_C12 RNAi (1)	3,149	0,100	0,012	0,159	2,8
3D_C12 RNAi (2)	2,569	0,193	0,009	0,110	7,0
3D_C12 RNAi (3)	2,175	0,064	0,007	0,080	2,9
3D_C12 RNAi (4)	3,517	0,125	0,012	0,139	3,4
3D_C12 RNAi (5)	1,085	0,868	0,009	0,119	44,4
3D_C12 RNAi (6)	0,025	2,260	0,011	0,122	98,9
3D_C12 RNAi (7)	0,027	1,867	0,011	0,122	98,6
3D_C12 RNAi (8)	2,268	0,128	0,009	0,102	5,3
3D_C12 RNAi (9)	2,197	0,133	0,008	0,099	5,7
3D_C12 RNAi (10)	2,434	0,112	0,009	0,110	4,4
vetor controle (3)	1,811	1,1735	0,018	0,170	48,9
vetor controle (11)	0,290	2,090	0,013	0,143	87,8
°Percentagem de peso seco da folha.					
**[% de nornicotina/ (% de nicotina + nornicotina)] X 100					

Típico da linhagem DH 98-325-6, o tratamento e cura com etefon resultou na produção substancial de nornicotina em duas plantas-controle (48,9% e 87,8% de conversão de nicotina em nornicotina). Em um contraste surpreendente, sete de dez plantas transgênicas independentes que possuíam o construto 3D_C12/RNAi mostraram uma conversão de nicotina em nornicotina mínima, com as percentagens de conversão variando 2,8 a 7,0 por cento. As outras três linhagens de 3D_C12/RNAi mostraram conteúdos de alcalóides similares aos das plantas-controle apenas pelo vetor. Concentrações dos alcalóides menores anabasina e anatabina não parecem ser

significativamente influenciadas pela presença ou ausência do transgene 3D_C12/RNAi (Tabela 4).

Embora o inserto do DNAc do gene 3D_C12-7 tenha sido usado com sonda de hibridização específica, nas condições de hibridização e lavagem usadas nesse experimento, seria esperada a hibridização cruzada para a família inteira do gene 3D_C12. Conforme mostrado na Figura 5, um forte sinal de hibridização foi detectado em cada planta que mostrou um alto fenótipo de nornicotina, e uma hibridização mínima foi detectada nas plantas transformadas com o construto 3D_C12/RNAi que mostraram um baixo fenótipo de nornicotina. Portanto, conclui-se que o silenciamento efetivo da família do gene 3D_C12 inibe a conversão metabólica de nicotina em nornicotina no tabaco.

A análise de alcalóide das plantas transgênicas Petite Havana é mostrada na Tabela 5. Quatro plantas independentemente transformadas contendo os construtos 35S:3D_C12 e 35S:3D_C12-7 foram testadas juntamente com sete indivíduos 35S:7D_A06 independentes e três plantas independentemente transformadas com o vetor controle pBI121. Conforme esperado, as folhas verdes e não curadas das plantas-controle apenas por vetor continham quantidades mínimas de nornicotina. Da mesma forma, todas as plantas transformadas com os construtos 35S:3D_C12 e 35S:7D_A06 mostraram uma conversão metabólica mínima de nicotina em nornicotina. Contudo, um fenótipo muito diferente foi observado em plantas transformadas com 35S:3D_C12-7. Todas as quatro plantas independentemente transformadas com esse construto continham nornicotina como o alcalóide predominante na folha verde não tratada; as percentagens de conversão de nicotina em nornicotina variaram de 94,6 a 98,6.

Tabela 5. Análise de alcalóide de plantas individuais de Petite Havana transformadas com os construtos 3D_C12, 3D_C12-7, 7D_A06 ou com o vetor controle pBI121. As folhas verdes foram colhidas e analisadas sem tratamento ou cura.

Amostra	% Nicotina*	% Nornicotina*	% Anabasina*	% Anatabina*	% Conversão**
vetor controle (2)	0,673	0,018	0,006	0,018	2,6
vetor controle (8)	0,605	0,014	0,005	0,016	2,3
vetor controle (10)	0,694	0,017	0,004	0,018	2,4
35S:3D_C12 (1)	0,706	0,005	0,006	0,020	0,7
35S:3D_C12 (2)	0,814	0,022	0,007	0,017	2,6
35S:3D_C12 (3)	0,630	0,010	0,003	0,012	1,6
35S:3D_C12 (4)	0,647	0,010	0,004	0,011	1,5
35S:3D_C12-7 (1)	0,005	0,347	0,002	0,012	98,6
35S:3D_C12-7 (2)	0,006	0,255	0,002	0,009	97,4
35S:3D_C12-7 (3)	0,017	0,300	0,002	0,010	94,6
35S:3D_C12-7 (4)	0,010	0,384	0,002	0,015	97,5
35S:7D_A06 (1)	0,761	0,011	0,005	0,018	1,4
35S:7D_A06 (2)	0,507	0,009	0,003	0,007	1,7
35S:7D_A06 (4)	0,653	0,015	0,006	0,015	2,2
35S:7D_A06 (5)	0,643	0,013	0,004	0,018	2,0
35S:7D_A06 (6)	0,521	0,007	0,004	0,014	1,3
35S:7D_A06 (7)	0,716	0,015	0,005	0,020	2,1
35S:7D_A06 (8)	0,701	0,027	0,004	0,018	3,7
*Percentagem de peso seco da folha.					
**[% de nornicotina/ (% de nicotina + nornicotina)] X 100					

Exemplo 6: Cossupressão da família do gene 3D_C12

Além da conclusão principal de que o gene 3D_C12-7 foi capaz de mediar a conversão de nicotina em nornicotina, uma observação adicional sobressaiu-se nas análises de alcalóides de plantas transgênicas de Petite Havana. O resultados de
5 alcalóides relatados na Tabela 5 juntamente com os ensaios de alcalóides adicionais conduzidos independentemente (dados não mostrados) mostraram de modo consistente que uma das plantas transformadas com o construto 35S:3D_C12 (35S:3D_C12(1)) continha menos nicotina na folha verde não tratada do que qualquer outra planta desse estudo. Isso pode ser resultado da cossupressão da família do gene
10 3D_C12 nessa planta específica, um fenômeno observado frequentemente em plantas transgênicas mesmo quando um transgene é expresso na sua orientação senso (Fagard e Vaucheret (2000) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 51: 167-194).

Se a planta 35S:3D_C12 (1) estivesse mostrando verdadeiramente um fenótipo de cossupressão, seria esperada a manutenção desse fenótipo mesmo após o
15 tratamento com etefon e cura das folhas, similar aos fenótipos de baixa nornicotina conferidos pelo construto 3D_C12/RNAi no genótipo conversor DH 98-325-6 conforme mostrado acima. Para testar essa predição, os perfis de alcalóide foram determinados em folhas de 35S:3D_C12 (1) tratadas com etefon e curadas e duas plantas-controle apenas por vetor. Conforme mostrado da Tabela 6, o tratamento
20 com etefon e cura resultaram em mais de 97% de conversão de nicotina em nornicotina nas duas plantas-controle enquanto que folhas 35S:3D_C12 (1) tratadas de modo similar mostraram uma conversão negligente (0,6%). Folhas a partir de outras cinco plantas que expressaram os transgenes 35S:3D_C12 e 35S:7D_A06 também foram submetidas ao tratamento com etefon e cura. Em cada caso, um alto
25 fenótipo de nornicotina foi observado, similar ao de plantas-controle apenas por vetor (dados não mostrados).

Tabela 6. Análise de alcalóide de 35S:3D_C12 (1) e plantas de vetores controle pBI121. As folhas foram tratadas com etefon e curadas por sete dias.

Amostra	% Nicotina*	% Nornicotina*	% Anabasina*	% Anatabina*	% Conversão**
vetor controle (8)	0,009	0,425	n.d.	0,011	97,9
vetor controle (10)	0,008	0,560	n.d.	0,025	98,6
35S:3D_C12 (1)	1,185	0,007	n.d.	0,020	0,6
°Percentagem de peso seco da folha.					
**[% de nornicotina/ (% de nicotina + nornicotina)] X 100					
n.d., não detectado					

Finalmente, os ensaios de Northern blot foram conduzidos na seleção de plantas que representavam cada um dos genótipos transgênicos de Petite Havana (Figura 6). Usando um DNAc de 3D_C12-7 como sonda de hibridização, foi observado um sinal mínimo detectado com RNAs isolados a partir de folhas verdes não tratadas da planta-controle apenas por vetor. Em contraste, a hibridização foi facilmente detectada em amostras de RNA de todas as quatro plantas transgênicas independentes que possuíam o construto 35S:3D_C12-7. Um forte sinal de hibridização foi similarmente observado usando os RNAs de todas as outras plantas transgênicas testadas que foram transformadas com os construtos 35S:3D_C12 e 35S:7D_A06, com exceção da planta contendo baixa nornicotina 35S:3D_C12 (1).

Os resultados gerais dos ensaios de Northern blot mostram que o promotor CaMV 35S foi em geral efetivo na mediação de um alto nível de expressão gênica para cada um dos três membros da família do gene 3D_C12 testada nesse estudo. A falha na detecção de um sinal de hibridização em plantas 35S:3D_C12 (1) é consistente com a interpretação de que a família do gene 3D_C12 teria sido silenciada através da cossupressão nesse indivíduo.

Exemplo 7: Caracterização e supressão adicional de genes 3D_C12

Um Segundo construto de RNAi foi preparado usando sequências de polinucleotídeos a partir da sequência 3D_C12-7. A montagem do cassete de expressão 3D_C12-7/RNAi seguiu as mesmas etapas básicas daquelas descritas acima para 3D_C12/RNAi. Resumidamente, uma fita senso e antisenso de 298-pb do DNAc de 3D_C12-7 (ID SEQ N°:5) que corresponde à região entre as posições de nucleotídeos 297 e 594 da sequência de codificação (posições 1-1551 da ID SEQ N°:5) foram ligadas dentro do vetor pKYLX801 à jusante a à montante do íntron de ácido graxo ômega-6 desaturase de soja de 151-pb (ver GenBank Acesso N°. DQ672337), respectivamente. Os primers (E4SFwd e E4SRev) usados para o isolamento da região de 298-pb pelos braços senso e antisenso foram 5'-AAGCTTTGACGCCATTTTTTCCAATCG-3' (ID SEQ NO:27), e 5'-CTCGAGTTTTCCAGCGATCATCTTCAC-3' (ID SEQ NO:28), respectivamente. O cassete de RNAi foi removido de pKYL801 e colocado entre um promotor CaMV35S² forte e um terminador de pequena subunidade da rubisco do vetor binário de expressão da planta pKYLX71 (ver Figura 8). Nas discussões abaixo, o construto de RNAi é referido como construto 3D_C12-7-Ri298.

Plantas transgênicas de tabaco foram geradas através de transformação mediada por *Agrobacterium* seguindo os procedimentos apresentados acima. Resumidamente, as plantas de tabaco Burley transformadas foram regeneradas a partir de calos em um meio Murashige-Skoog (MS) suplementado com 100 mg/L de canamicina e hormônios vegetais em uma sala de crescimento a 25°C sob um ciclo claro/escuro de 16 horas/8 horas. Os calos foram transferidos para meios de seleção frescos a cada 2-3 semanas até aparecerem os ramos. Pequenos ramos foram transferidos para meio de enraizamento para permitir o desenvolvimento da raiz por 2 semanas. Plantas completamente regeneradas foram transferidas para uma estufa e crescidas sob condições padrão.

Química SYBR® Green I

O RNA total foi isolado a partir de folhas curadas de plantas de tabaco conversoras e não-conversoras usando o reagente TRIzol® (Invitrogen, Life Technologies, Carlsbad, CA) O RNA purificado foi tratado com DNase livre de

RNase (TURBO DNA-free™, Ambion, Austin, TX) A primeira fita de DNAc foi sintetizada usando 5 µg do RNA total e o StrataScript® First-Strand Synthesis System (Stratagene, Cedar Creek, TX). Foi empregada a RT-PCR para determinar a abundância do DNAc de 3D_C12-7 usando a química fluorescente SYBR® Green I, Morrison *et al.* (1998) *Biotechniques* 24:954-962.

Uma curva de calibração foi gerada com uma diluição serial do DNAc de 3D_C12-7 clonado para dentro do vetor pGEM®-T Easy (Promega Corporation, Madison, WI) A mistura RT-PCR continha 2,5 mM de MgCl₂, 125 µM de dNTP, 0,5 µM de cada primer, 0,5x SYBR® Green I, 0,5 µg de cDNA (ou 1 µl do plamídeo de referência), e 1,25 de polimerase U Platinum Taq (Invitrogen Life Technologies). As sequências dos primers 3D_C12-7 alelos-específicos (E4SyFwd e E4SyRev) foram 5'-ACGTGATCCTAAACTCTGGTCTG-3' (E4SyFwd (ID SEQ N°:29)) e 5'-GCCTGCACCTTCCTTCATG-3' (E4SyRev (ID SEQ N°:30)). A RT-PCR foi realizada em um termociclador BioRad (BioRad Laboratories, Hercules, CA) ajustado para o seguinte protocolo: 95°C por 2 min; 35 ciclos de 95°C por 30 seg, 55°C por 30 seg 72°C por 50 seg, seguido de uma extensão final de 72°C por 5 min. Um fragmento de 165-pb do gene α -tubulina foi usado como padrão interno. A concentração de DNAc de 3D_C12-7 foi determinada a partir da curva de calibração transcrito-específico e normalizada para o padrão interno. O número de indução foi calculado pela divisão dos valores de fluorescência normalizados do conversor pelas amostras não-conversoras. A análise da curva de fusão foi usada para confirmar a pureza de produtos da PCR conforme descrito em Ririe *et al.* (1997) *Anal. Biochem.* 245:154-160. Duas plantas foram amostradas por tratamento e as amplificações foram repetidas três vezes.

25

Química TaqMan®

O RNA total foi isolado de linhagens de tabaco usando o reagente TRIzol. O RNA purificado foi tratado com DNase livre de RNase (TURBO DNA-free™). A primeira fita DNAc foi sintetizada usando 10 µg de RNA total e o Kit High Capacity cDNA Archive (Applied Biosystems, Foster City, CA). A mistura RT-PCR continha 1X TaqMan® Universal PCR Master Mix (Applied Biosystems, Foster City, CA), 400 nM de cada primer (E4TmFwd e E4TmRev), 250 nM de ligador de depressões

menores TaqMan® (MGB) sonda (E4MGB), 2 ng de cDNA, e água livre de nuclease (Afonina *et al.* (2002) *Biotechniques* 32:940-949). As sequências do primer e da sonda foram 5'-CGGTAATCGGCCATCTTTTC-3' (E4TmFwd (ID SEQ N°:31)), 5'-CCGAGTTTTCGAGCTAATGGA-3' (E4TmRev(ID SEQ N°:32)), e sonda 5'-CAATGACGAACGGCGACAG-3' (MGB(ID SEQ N°:33)). A RT-PCR foi realizada em um ABI 7500 Real-Time System (Applied Biosystems, Foster City, CA) ajustado para o seguinte protocolo. 50°C por 2 min; 95°C por 10 min; 40 ciclos de 95°C por 15 seg, 60°C por 1 min. A Gliceraldeído-3-fosfato desidrogenase (G3PDH) foi usada como o controle endógeno para normalizar a quantidade do template de cDNA dentro das reações. O número de mudanças foi determinado pela divisão dos valores de fluorescência normalizados de cada amostra por aqueles obtidos a partir de uma amostra controle não-conversora ou não curada. Para cada tratamento, o RNA foi isolado a partir de três plantas independentes e as amplificações foram repetidas 3 vezes para cada amostra.

15

Análises de Northern e Southern Blot

O RNA total foi isolado a partir de folhas de tabaco curadas usando o reagente TRIzol de acordo com as instruções do fabricante (Invitrogen, Life Technologies). Amostras de RNA total foram separadas em de gel de agarose a 1,2% em TBE, e transferidas para membranas de nylon positivamente carregadas por eletroblotting com tampão de TBE 2X. As membranas passaram por ligação cruzada UV e lavadas 2X SSC por 5 min. A hibridização de Northern blot, lavagem e detecção foram realizadas usando o Sistema digoxigenina (DIG) conforme descrito pelo fabricante (Roche Diagnostics Corp., Indianapolis, IN). O OFR de comprimento inteiro de 1,8 kb no DNAC de 3D_C12-7 cDNA foi rotulado com DIG e usado como uma sonda.

O DNA genômico foi extraído com DNazol® (Invitrogen, Life Technologies) a partir de folhas de tabaco verdes de acordo com o protocolo do fabricante. Após a incubação com FcoRI ou enzimas de restrição NcoI por uma noite, 15 µg do DNA digerido foram separadas em gel de agarose a 0,7% em TBE, depurinadas com 0,25 M de HCl por 10 min, e desnaturadas com 0.5 N de NaOH por 30 minutos. O DNA foi marcado por uma noite por transferência capilar sobre membranas de nylon carregadas positivamente (Roche Diagnostics Corp.) e hibridizado a 65°C por uma noite com um fragmento do gene neomicina fosfotransferase II (NPT II) de 515-pb

rotulado por DIG. A hibridização, lavagem, e detecção foram realizadas de acordo com os protocolos fornecidos com o DIG System. Os primers usados para as amplificações das sondas de hibridização de Northern e Southern foram E4FIFwsd (5'-ATGGTTTTTCCCATAGAAGCC-3' (ID SEQ N°:34)), E4FIRev (5'-

5 TTTTGGGACAATCAGTCAAT-3' (ID SEQ N°:35)), KanFwd (5'-TGAATGAACTGCAGGACGAG-3' (ID SEQ N°:36)), e KanRev (5'-AATATCACGGGTAGCCAACG-3' (ID SEQ N°:37)).

Análise de alcalóide

10 As folhas de tabaco foram colhidas e secas ao ar em um forno a 50°C por 2 dias. A uma amostra de 100 mg de folhas secas e esmagadas foi adicionado 0,5 mL de 2 N de NaOH em um frasco de cintilação de 20 mL. A amostra foi misturada e deixou-se incubar por 15 minutos à temperatura ambiente. Alcalóides foram extraídos pela adição de 5 mL de solução de extração [0,04% de quinolina (peso/volume)

15 dissolvida em metil-t-butil éter] e rotacionada gentilmente em um agitador linear por 3 horas. Após a separação das fases, uma alíquota da fase orgânica foi transferida para um frasco de amostra. As amostras foram analisadas usando o cromatógrafo a gás Autosystem XL da PerkinElmer, equipado com um detector de ionização de chama, um liner de vidro de modo split/splitless de 4mm, uma coluna ID DB-5 de 30

20 m x 0.53 mm. As condições cromatográficas foram as seguintes: detector de temperatura: 250°C; temperatura do injetor: 250°C; fluxo de hélio da 120°C: 20 mL/min; volume de injeção: 2 µL; condições da coluna: 120°, mantido por 1 min, 20-280°C a taxa de alteração 30°C /min, mantido a 280°C por 2 min. A composição de alcalóide foi determinada pelo software TotalChrome Navigator usando uma curva de

25 calibração. As médias das medidas de alcalóides foram separadas de acordo com o Protected LSD de Fisher (PROC MIXED).

Plantas

30 Linhagens duplo-haplóides de tabaco Burley DH 98-325-5 (325-5; não-conversoras) e DH98-325-6 (325-6; conversoras) descritas acima foram usadas em todos os experimentos, exceto para os ensaios de RT-PCR fluorogênico da nuclease

5' (TaqMan®), em que foram usadas as linhas isogênicas DH 91-1307-46 (não-conversoras) e DH 91-1307-46 (conversoras). Todas as plantas foram crescidas em uma estufa de ambiente controlado equipada com iluminação suplementar que possibilitou um ciclo claro-escuro de 10h/14h.

5 Para a curagem, as folhas de tabaco foram coletadas a partir de plantas conversoras e não-conversoras cerca de um mês antes da floração e tratadas por imersão de cada folha duas vezes por 10 segundos em etefon a 2% e secas por duas horas. As folhas foram curadas por até duas semanas em sacos plásticos, em ambiente escuro, até se tornarem amarelas. As folhas curadas foram usadas para
10 análises de alcalóide e Northern. Amostras das folhas curadas submetidas à análise GC foram secas a 50°C por 2 dias. Para a análise de Southern, foram usadas folhas de tabaco verdes de plantas adultas. Para produzir a geração T₁ de plantas transgênicas, as transformantes primárias (T₀) foram autopolinizadas, e a semente de T₁ colhida foi abrigada por mudas de germinação em placas de ágar-MS contendo 100 mg/L de
15 canamicina por 6 semanas. As sobreviventes foram transplantadas para o solo e crescidas em uma estufa conforme descrito acima. As plantas foram fertilizadas com o Professional All Purpose Plant Food de Peter (20-20-20; Spectrum Brands Inc., Madison, WI) uma vez por semana.

20 Análise de RT-PCR da expressão de 3D_C12-7 em tabaco conversor e não-conversor

Ainda para caracterizar o papel de 3D_C12-7 na N-demetilação de nicotina, foram realizados experimentos para demonstrar que a regulação da expressão de 3D_C12-7 é consistente com os níveis de atividade de N-demetilação de nicotina observados em tabaco conversor versus não-conversor.

25 Para determinar a taxa de acumulação de RNAm de 3D_C12-7 em tabaco conversor e não-conversor, foi empregada uma estratégia de RT-PCR alelo-específica em tempo real. Devido ao RT-PCR envolver a detecção e medida dos produtos da amplificação de um template de PCR, o uso de primers alelo-específicos permite a quantificação de uma isoforma simples entre um grupo de sequências altamente
30 homólogas. Para a quantificação acurada do transcrito de 3D_C12-7, dois segmentos diferentes da região de codificação de 3D_C12-7 foram amplificados e as químicas SYBR® Green e TaqMan® foram usadas para gerar sinais fluorescentes. A análise RT-PCR que usou a química SYBR® Green I revelou um aumento de 80 vezes nos

níveis do transcrito de 3D_C12-7 nas folhas curadas de tabaco conversor versus não-conversor. As análises de curva de fusão com pico único e a eletroforese em gel dos amplicons confirmaram a homogeneidade dos produtos de PCR.

No experimento de RT-PCR baseado na química TaqMan®, os níveis do transcrito de 3D_C12-7 foram quantificados em folhas de tabaco conversoras e não-conversoras, tratadas e não tratadas com etefon, que foram curadas por 0, 1 ou 5 dias. Baixos níveis dos transcritos de 3D_C12-7 foram detectados nas folhas não curadas ou após um dia de período de cura, sem considerar o tipo de conversão ou tratamento com etefon. De modo similar, os níveis de linha de base da transcrição de 3D_C12-7 foram observados em folhas conversoras e não-conversoras que foram curadas por 5 dias sem tratamento com etefon. Em contraste, um aumento de 7,5 vezes na acumulação do transcrito de 3D_C12-7 foi detectado em folhas curadas de tabaco conversor versus não-conversor, e um aumento de 70 vezes foi observado em folhas não curadas versus curadas de uma variedade de tabaco conversor quando o tratamento com etefon precedeu o período de curagem de 5 dias. Embora não pretendam ser limitados por qualquer teoria particular, esses resultados sugerem que 3D_C12-7 é um contribuinte principal para a *N*-demetilação de nicotina, e é fortemente induzível por etileno em folhas senescentes de tabaco.

20 Supressão da conversão de nicotina em nornicotina pelos construtos 3D_C12-Ri99 e 3D_C12-7-Ri298

Para comparar a extensão na qual 3D_C12 e 3D_C12-7 mediam a supressão da produção de nornicotina, plantas de tabaco Burley conversoras e não-conversoras foram transformadas com dois vetores silenciadores de gene. Dez (10) plantas transgênicas foram regeneradas pelo construto RNAi. Cerca de 80% das plantas de tabaco que expressaram excessivamente a repetição invertida de 99-pb ou 298-pb mostraram níveis reduzidos de nornicotina em comparação aos controles de vetor vazio (Tabelas 7 e 8). No genótipo não-conversor, a expressão de 3D_C12-Ri99 e 3D_C12-7-Ri298 reduziu a conversão de nicotina em nornicotina em cerca de 1,8 vezes (2,0%) e 3,0 vezes (1,2%), respectivamente, em comparação com a taxa de conversão detectada nos controles de vetor (3,6%) (Tabela 7). Entre as plantas não-conversoras silenciadas, o nível de conversão mais baixo de 0,9% foi alcançado com o uso de construto 3D_C12-7-Ri298 (Tabela 7).

Tabela 7. Análise de alcalóide de plantas de tabaco Burley não-conversoras transformadas com o construto 3D_C12-Ri99 ou 3D_C12-7-Ri298.

Linhagem^c

3D-C12-Ri99	% de Nicotina^d	% de Nornicotina^d	% de Conversão^e
1	1,693	0,034	2,0
2	1,435	0,031	2,1
3	2,095	0,043	2,0
4	2,868	0,053	1,8
5	0,947	0,025	2,6
6	2,357	0,043	1,8
7	2,599	0,043	1,6
8	0,796	0,020	2,4
9	2,178	0,039	1,8
10	3,162	0,061	1,9
MÉDIA	2,013	0,039	2,0a
DP	0,748	0,012	0,3
3D_C12-7-Ri298			
3	1,806	0,020	1,1
4	1,948	0,207	1,4
5	2,061	0,020	1,0
6	2,704	0,040	1,5
8	2,652	0,023	0,9
9	1,074	0,015	1,3
MÉDIA	2,041	0,024	1,2b
DP	0,550	0,008	0,2
Vetor Controle^g			
1	1,206	0,052	4,2
2	1,265	0,038	2,9
3	1,752	0,058	3,2
4	1,230	0,072	5,6
5	1,777	0,060	3,3
6	1,536	0,044	2,8
MÉDIA	1,461	0,054	3,6c
DP	0,240	0,011	1,0

^aAs folhas de tabaco foram tratadas com etefon e curadas por 2 semanas a 25°C.

^bDas plantas transformadas com um construto de RNAi, apenas são mostrados os indivíduos silenciados. Os dados de alcalóide representam as médias de 2 mensurações.

^cOs números representam indivíduos independentemente transformados.

^dPorcentagem de peso seco da folha.

^e $[\% \text{ de nornicotina} / (\% \text{ de nicotina} + \% \text{ de nornicotina})] \times 100$; valores seguidos por diferentes letras são significativamente diferentes de acordo com o PSD Protegido de Fisher (0,05).

^fDP, desvio padrão

^gPlantas de tabaco transformadas apenas com o vetor pKYLX71 foram usadas como controles.

Em relação ao tabaco não-conversor, a acumulação de nornicotina foi suprimida ainda mais dramaticamente nos indivíduos silenciados das plantas fortemente conversoras (Tabela 8). Usando os construtos 3D_C12-Ri99, a conversão de nicotina foi reduzida a níveis tão baixos como 4,5% em plantas de tabaco 325-6 transformadas por 3D_C12-Ri 99, em contraste marcante com as plantas-controle 325-6 que exibiram taxas de conversão de cerca de 98%; Tabela 8). No entanto, com o uso do construto 3D_C12-7-Ri298 foram obtidas reduções ainda maiores na conversão e nicotina em nornicotina (Tabela 8). Quatro indivíduos transformados por 3D_C12-7-Ri298 converteram um valor tão baixo quanto 0,8% de sua nicotina em nornicotina, e a média aritmética das 9 transformantes silenciadas foi de 0,9% de conversão. Todas as plantas silenciadas foram morfológicamente indistinguíveis do vetor vazio e dos controles selvagens (dados não mostrados).

Tabela 8. Análise de alcalóide de plantas de tabaco Burley conversoras transformadas com o construto 3D_C12-Ri99 ou 3D_C12-7-Ri298.

Linhagem^c	% de Nicotina^d	% de Nornicotina^d	% de Conversão^e
3D_C12-Ri99			
1	3,419	0,100	2,8
2	2,569	0,193	7,0
3	2,175	0,064	2,9
4	3,517	0,125	3,4
8	2,268	0,128	5,3
9	2,197	0,133	5,7
10	2,434	0,112	4,4
MÉDIA	2,654	0,122	4,5a
DP^g	0,573	0,039	1,6
3D_C12-7-Ri298^b			
1	2,043	0,020	1,0
2	3,427	0,026	0,8
3	2,603	0,020	0,8
5	2,427	0,030	1,2
6	2,106	0,021	1,0
7	1,412	0,015	1,1
8	3,328	0,028	0,8
9	1,493	0,015	1,0
10	2,065	0,018	0,8
MÉDIA	2,323	0,021	0,9b
DP	0,669	0,005	0,1
Vetor Controle#1^e			
1	0,126	1,550	92,5
2	0,330	2,604	88,8
3	0,060	1,419	95,9
4	0,114	1,267	91,7
5	0,119	1,303	91,6
MÉDIA	0,150	1,628	92,1c
DP	0,093	0,498	2,3

^aAs folhas de tabaco foram tratadas com etefon e curadas por 2 semanas a 25°C.

^bDas plantas transformadas com um construto de RNAi, apenas são mostrados os indivíduos silenciados.

^cOs números representam indivíduos independentemente transformados.

^dPorcentagem de peso seco da folha.

^e $[\% \text{ de nornicotina} / (\% \text{ de nicotina} + \% \text{ de nornicotina})] \times 100$; valores seguidos por diferentes letras são significativamente diferentes de acordo com o PSD Protegido de Fisher (0,05).

^fDP, desvio padrão

^gPlantas de tabaco transformadas apenas com o vetor pKYLX71 foram usadas como controles.

Para testar a heretabilidade da supressão de normicotina nas plantas transformadas por 3D_C12-7-Ri298, um conjunto de linhagens conversoras e não-conversoras transformadas por 3D_C12-7-Ri298 que mostraram os níveis mais baixos de conversão de nicotina foram promovidas para a geração T₁ (Tabela 9). Devido à

5 segregação do(s) transgene(s) ocorrer na progênie T₁, indivíduos transgênicos foram identificados através da seleção de mudas capazes de crescer em meio contendo canamicina. Nove progênies resistentes à canamicina de cada transformante 3D_C12-7-Ri298 selecionado da geração T₀, e quadro indivíduos resistentes à canamicina de cada linhagem de controle de vetor selecionada foram analisados para

10 verificar o conteúdo de alcalóide. A taxa de nicotina não diferiu significativamente entre os transformantes 3D_C12-7-Ri298 primários e sua progênie T₁, indicando uma alta heritabilidade do traço de supressão de normicotina (ver Tabelas 7, 8 e 9). No entanto, a promoção da linhagem de controle por vetor “não-conversora” por uma única geração aumentou a taxa de conversão de nicotina em normicotina de 4,2% a um

15 valor médio de 11,6%, ilustrando o alto grau de instabilidade do locus de conversão em plantas transgênicas desprovidas do construto de RNAi 3D_C12-7-Ri298-específico (Tabela 7 e 9). Em geral, esses resultados mostram que o silenciamento mediado por RNAi da subfamília do gene 3D_C12 é um meio altamente efetivo de reduzir a produção de normicotina nas tanto nas plantas de tabaco não-conversoras

20 como nas fortemente conversoras.

Tabela 9. Análise de alcalóides de transformantes da geração T_1 de 3D_C12-7-Ri298

Linhagem	% de Nicotina^c	% de Nornicotina^c	% de Conversão^d
DH98-325-5 (não-conversor)			
3D_C12-7-Ri298#3			
Média	1,764	0,024	1,4a
DP	0,456	0,004	0,3
3D_C12-7-Ri298#5			
Média	1,500	0,020	1,3a
DP	0,306	0,006	0,3
3D_C12-7-Ri298#8			
Média	1,772	0,020	1,2a
DP	0,409	0,003	0,3
Vetor Controle#1^e			
Média	1,466	0,203	11,6b
DP	0,713	0,161	9,7
DH98-325-6 (conversor)			
3D_C12-7-Ri298#2			
Média	1,970	0,019	1,0a
DP	0,536	0,004	0,3
3D_C12-7-Ri298#8			
Média	1,623	0,022	1,3a
DP	0,300	0,002	0,2
3D_C12-7-Ri298#10			
Média	1,419	0,017	1,3a
DP	0,515	0,004	0,3
Vetor Controle#2^e			
Média	0,028	1,170	97,6c
DP	0,006	0,234	0,5

^aAs folhas de tabaco foram tratadas com etefon e curadas por 2 semanas a 25°C.

^bAs médias e desvios-padrão (DP) representam as progênies 9 e 4 T_1 do construto 3D_C12-7-Ri298 e linhagens vazias transformadas por vetor (vetor controle), respectivamente.

^cPercentagem de peso seco da folha.

^d $[\% \text{ de nornicotina} / (\% \text{ de nicotina} + \% \text{ de nornicotina})] \times 100$; valores seguidos por diferentes letras são significativamente diferentes de acordo com o PSD Protegido de Fisher (0,015).

^ePlantas de tabaco transformadas apenas com o vetor pKYLX71 foram usadas como controles.

Além do mais, a transformação de tabaco com o construto 3D_C12-7-298 conferiu uma redução de 3,6 vezes na conversão de nicotina em relação às plantas-controle não conversoras típicas, sem afetar o crescimento e desenvolvimento da planta.

- 5 Para demonstrar que a regulação para baixo da produção de nornicotina no tabaco transformado por 3D_C12-7-298 foi concomitante à redução dos transcritos da subfamília do gene 3D_C12, uma sonda de DNAC de 3D_C12-7 foi hibridizada com o RNA total isolado a partir de folhas curadas de plantas conversoras e não-conversoras. Um sinal fraco de hibridização foi gerado pelo RNA isolado a partir de
- 10 transformantes 3D_C12-7-Ri298 que mostraram baixo conteúdo de nornicotina em contraste com o forte sinal produzido pelo RNA extraído do plasmídeo de controle ou de plantas selvagens. Esses resultados indicam que a regulação para baixo de conversão de nicotina foi resultado do silenciamento de gene mediado por RNAi do(s) gene(s) de nicotina *N*-demetilase.

15

Determinação do número de cópias do transgene

- Para determinar se a integração de múltiplas cópias de 3D_C12-7-Ri298 era requerida para produzir transplantas que mostrassem uma atividade muito baixa de nicotina *N*-demetilase, foi realizada uma análise de Southern em indivíduos que
- 20 exibiram acumulação de nicotina menor do que 1,5%. O número de transgenes variou amplamente entre essas plantas incluindo indivíduos que continham 1 cópia (325-5, linhagens 5 e 8; 325-6, linhagens 2 e 8), 5 cópias (325-6, linhagem 10), e 6 cópias (325-5, linhagem 6) do construto 3D_C12-7-Ri298 com o uso da análise de Southern blot do DNA genômico digerido com a enzima de restrição *EcoRI*. O número de
- 25 cópias de transgene foi confirmado pelo uso de DNA digerido de *NcoI* (dados não mostrados). Esses resultados indicam que a integração de um único construto 3D_C12-7-Ri298 dentro do genoma de um tabaco fortemente conversor é suficiente para suprimir a produção de nornicotina a níveis muito baixos.

30 Conclusões gerais

As análises descritas nos Exemplos 1 a 6 acima resultaram na descoberta de uma família de genes P450 intimamente relacionados, designada família 3D_C12, cujos níveis de transcrito em estado estacionário coletivo foram significativamente

elevados nas plantas de tabaco conversoras que metabolizaram ativamente a nornicotina em comparação a suas contrapartes não-conversoras. A análise de plantas transgênicas demonstrou que a supressão da expressão gênica dessa família P450 em linhagens conversoras de tabaco inibiu o metabolismo de nicotina em nornicotina em níveis similares aos observados em plantas não-conversoras. Além disso, a expressão
5 senso de vários indivíduos dessa família de genes intimamente relacionados identificou um membro, designado 3D_C12-7, que desempenha um papel direto na conversão metabólica da nicotina em nornicotina. A expressão excessiva de 3D_C12-7 usando um forte promotor constitutivo causou um aumento dramático na produção
10 de nornicotina e acumulação em folhas verdes não curadas de plantas de tabaco transgênicas, um tecido onde a nicotina é o alcalóide predominante igualmente em plantas conversoras e não-conversoras. Dado que o membro da família do citocromo P450 designado 3D_C12-10 difere de 3D_C12-7 em apenas dois resíduos de aminoácidos imediatamente após a metionina de partida e dentro da sequência de
15 sinal N-terminal, prevê-se que esses produtos codificados funcionam de modo idêntico.

O contraste em fenótipos alcalóides entre a planta 35S:3D_C12 (1) e as plantas-controle apenas por vetor foi mais dramático em folhas que foram tratadas com etefon e curadas (0,6% de conversão versus mais de 97% de conversão; Tabela
20 6). Contudo, é digno de atenção que o conteúdo de nornicotina da planta 35S:3D_C12 (1) cosuprimida foi reduzido mesmo em folhas verdes não tratadas nas quais o alto fenótipo de nornicotina geralmente não se manifesta em linhagens de tabaco conversoras ou não-conversoras. As folhas verdes não tratadas da linhagem 35S:3D_C12 (1) mostraram apenas 0,7% de conversão de nicotina em nornicotina,
25 enquanto que cada outra planta nesse experimento mostrou percentagens de conversão variando de 1,3 a 3,7 (Tabela 5). Esse resultado sugere que a inibição da expressão do gene da família 3D_C12 pode provar ser efetivo na diminuição dos níveis de nornicotina mesmo em linhagens de tabaco nas quais a conversão genética geralmente não é um problema principal (como os tabacos curados por fumo) ou em indivíduos
30 não-conversores de linhagens que são propensas à conversão genética (como tabacos Burley).

Os ensaios de Southern Blot que usam membros da família de genes 3D_C12 como sondas de hibridização apresentam padrões de banda muito complexos,

sugerindo que podem existir mais membros dessa família de gene, mesmo além daqueles que foram identificados e caracterizados aqui (dados não mostrados). A hipótese de que a família de genes 3D_C12 é composta por membros adicionais ainda é apoiada pela publicação recente de 75 DNAs de tabacos P450 de comprimento
 5 inteiro com função desconhecida (Publicação de Pedido de Patente dos EUA 20040162420). Dentro dessa lista de P450s estão os DNAs adicionais que seriam, com base no trabalho descrito aqui, colocados dentro da família 3D_C12 em vista da identidade de sequência de aminoácidos de mais de 90% em relação às sequências de proteína mostradas na Figura 4.

10 A respeito da função molecular específica do gene 3D_C12-7 ou do membro aproximadamente idêntico da família 3D_C12-10, é possível que essa função seja codificar a enzima nicotina demetilase que catalisa a *N*-demetilação oxidativa de nicotina em nornicotina (Figura 1). De modo alternativo, a enzima codificada de 3D_C12-7 ou a enzima aproximadamente idêntica codificada de 3D_C12-10 pode
 15 produzir um produto que leva à regulação para cima da atividade da nicotina demetilase da folha, em oposição à catalisação direta da reação de *N*-demetilação.

Além disso, uma RT-PCR alelo-específica foi empregada para comparar a expressão de 3D-C12-7 entre as plantas conversoras e não-conversoras (Exemplo 7). Foi identificado um aumento de aproximadamente 80 vezes na expressão de 3D-C12-
 20 7 em plantas conversoras versus não-conversoras com o uso do ensaio SYBR® Green-chemistry RT-PCR. Uma regulação para cima de 7,5 vezes foi identificada pelo experimento RT-PCR baseado em química TaqMan®. Embora a variedade de tabaco DH 91-1307-46 usada no experimento de RT-PCR baseado em química TaqMan® exiba um nível leve a moderado de conversão de nicotina, as plantas não-
 25 conversoras DH98-325-5 usadas no ensaio RT-PCR baseado em SYBR® Green convertem de modo consistente uma percentagem muito baixa de nicotina em nornicotina. A expressão do gene 3D-C12-7 foi induzida no mínimo 7 vezes por etileno em folhas senescentes de plantas de tabaco conversoras.

Um construto de RNAi adicional, 3D_C12-7-Ri298, foi preparado com base
 30 em uma região do polinucleotídeo de 3D_C12-7 que corresponde às posições de nucleotídeos 297 a 594 da ID SEQ N°:5. A expressão desse construto de RNAi permite a supressão da produção de nornicotina em uma linhagem de tabaco fortemente conversora abaixo dos níveis normalmente encontrados em plantas não-

conversoras. O cassete de expressão do construto 3D_C12-7-Ri298 codificou um RNA em alça combinado por íntron no qual a região do caule foi engenheirada a partir do fragmento de 298-pb do DNAC de 3D_C12-7 inserido como uma repetição invertida. O loop da alça foi criado pela colocação de um íntron de 151-pb do gene *FAD* entre os dois lados das sequências palindrômicas. Um braço de comprimento de 298-pb foi usado para as repetições invertidas.

As plantas transformadas de 3D_C12-7-Ri298 acumularam menos nornicotina do que aquelas que hospedaram o construto 3D_C12-Ri99 (Tabelas 7 e 8). Nenhuma correlação foi encontrada entre o número de cópias do construto 3D_C12-7-Ri298 e a produção de nornicotina (Tabelas 7 e 8). O cassete de expressão 3D_C12-7-Ri298 possibilitou a produção de tabaco com uma taxa de conversão tão baixa quanto 0,8%, que está abaixo da taxa de 3 a 5% detectada em linhagens Burley usadas como produtoras de sementes. Essa redução surpreendente na produção de nornicotina pelo direcionamento a essa região particular do polinucleotídeo de 3D_C12-7 é um resultado inesperado. Também, a supressão da produção de nornicotina mostrou um alto grau de heritabilidade na progênie T₁ dos transformantes primários (Tabela 9). A supressão da produção de nornicotina nessas plantas transgênicas não produziu diferenças óbvias no crescimento e desenvolvimento quando comparadas às plantas de tipo selvagem.

Todas as publicações e pedidos de patentes mencionados na especificação são indicativos do nível daquelas pessoas habilitadas na técnica a qual essa invenção pertence. Todas as publicações e pedidos de patentes estão incorporados aqui por referência como se cada publicação individual ou pedido de patente fosse específica e individualmente indicado para ser incorporado por referência.

Embora a invenção precedente tenha sido descrita em detalhe através de ilustrações e exemplos para fins de clareza e compreensão, será óbvio que certas mudanças e modificações podem ser praticadas dentro do escopo das reivindicações anexas.

REIVINDICAÇÕES

1. Método para obtenção de uma planta transgênica de *Nicotiana* caracterizado por ter uma taxa de conversão de nicotina em nornicotina de aproximadamente menos de 2% e compreender:

a) transformar uma célula de planta com um construto de ácido nucléico heterólogo que inclui um promotor capaz de funcionar em uma célula de planta, ligado operacionalmente a uma sequência de ácido nucléico, que tem a primeira sequência de nucleotídeos com um fragmento de cerca de 100 a 400 nucleotídeos do polinucleotídeo de nicotina demetilase de *Nicotiana*, e uma segunda sequência de nucleotídeos capaz de formar um RNA de dupla fita com a primeira sequência de nucleotídeos, em que dito fragmento do polinucleotídeo de nicotina de demetilase de tabaco compreende nucleotídeos a partir de uma região do dito polinucleotídeo, e em que:

(i) dita região do polinucleotídeo de nicotina demetilase corresponde ao nucleotídeo da posição 265 ao nucleotídeo na posição 625 da SEQ ID NO: 3 ou SEQ ID NO:5; ou

(ii) dita região do dito polinucleotídeo de nicotina demetilase corresponde ao nucleotídeo da posição 297 ao nucleotídeo na posição 594 da SEQ ID NO: 3 ou SEQ ID NO:5;

b) cultivar a dita célula de planta sob condições de crescimento celular; e

c) regenerar a planta transformada de *Nicotiana*, em que dita planta transformada de *Nicotiana* é uma linhagem transgênica conversora de *Nicotiana*.

2. Método de acordo com a reivindicação 1, caracterizado pelo fato de que dita conversão é medida após a cura.

3. Método de acordo com a reivindicação 1, caracterizado por ter menos de aproximadamente 1% de conversão de nicotina em nornicotina.

4. Construto de ácido nucléico recombinante caracterizado por abranger um promotor capaz de funcionar em uma célula de planta ligada operacionalmente a uma sequência de ácido nucléico que tem uma primeira sequência de nucleotídeos abrangendo um fragmento de cerca de 100 nucleotídeos a cerca de 400 nucleotídeos de um tabaco de polinucleotídeo de nicotina demetilase e uma segunda sequência de nucleotídeo capaz de formar um RNA de

dupla fita com a primeira sequência de nucleotídeos, em que dito fragmento do dito polinucleotídeo de nicotina de demetilase de tabaco compreende nucleotídeos a partir de uma região do dito polinucleotídeo, e em que:

(i) dita região do polinucleotídeo de nicotina demetilase corresponde ao nucleotídeo na posição 265 ao nucleotídeo na posição 625 da SEQ ID NO: 3 ou SEQ ID NO:5; ou

(ii) dita região do dito polinucleotídeo de nicotina demetilase corresponde ao nucleotídeo na posição 297 ao nucleotídeo na posição 594 da SEQ ID NO: 3 ou SEQ ID NO:5.

5. Método de redução da conversão de nicotina em nornicotina em uma planta de *Nicotiana* caracterizado por abranger:

a) transformar uma planta de *Nicotiana* com um construto de ácido nucléico recombinante, que abrange um promotor capaz de funcionar em uma célula de planta ligada operacionalmente a uma sequência de ácido nucléico, que tem uma primeira sequência de nucleotídeos abrangendo um fragmento de cerca de 100 nucleotídeos a cerca de 400 nucleotídeos de um polinucleotídeo de nicotina demetilase de tabaco e uma segunda sequência de nucleotídeo capaz de formar um RNA de dupla fita com a primeira sequência de nucleotídeos, em que dito fragmento do dito polinucleotídeo de nicotina de demetilase de tabaco compreende nucleotídeos a partir de uma região do dito polinucleotídeo, e em que:

(i) dita região do dito polinucleotídeo de nicotina demetilase corresponde ao nucleotídeo na posição 265 ao nucleotídeo na posição 625 da SEQ ID NO: 3 ou SEQ ID NO:5; ou

(ii) dita região do dito polinucleotídeo de nicotina demetilase corresponde ao nucleotídeo na posição 297 ao nucleotídeo na posição 594 da SEQ ID NO: 3 ou SEQ ID NO:5.

b) regenerar uma planta transgênica de *Nicotiana*.

6. Método de acordo com a reivindicação 5, caracterizado pelo fato de que dita planta transgênica de *Nicotiana* tem menos de aproximadamente 2% de conversão de nicotina em nornicotina, preferencialmente, menor que 1%.

7. Constructo de acordo com a reivindicação 4, caracterizado pelo fato de que dito fragmento do polinucleotídeo de nicotina demetilase de tabaco abrange nucleotídeos da região do polinucleotídeo selecionada a partir de uma sequência codificadora e uma sequência de íntrons.

8. Constructo de acordo com a reivindicação 4, caracterizado pelo fato de que a primeira e segunda sequências de nucleotídeos são ligadas por uma sequência de íntrons.

9. Constructo de acordo com a reivindicação 8, caracterizado pelo fato de que dita sequência de íntrons é uma sequência de íntrons passível de ser combinada.

10. Constructo de acordo com a reivindicação 4, caracterizado pelo fato de que o polinucleotídeo de nicotina demetilase abrange a sequência apresentada na SEQ ID N°:3 ou SEQ ID N°:5, e na qual o dito fragmento abrange os nucleotídeos 265 a 625 da SEQ ID N°:3 ou SEQ ID N°:5.

11. Constructo de acordo com a reivindicação 4, caracterizado pelo fato de que o polinucleotídeo de nicotina demetilase abrange a sequência apresentada na SEQ ID N°:3 ou SEQ ID N°:5, e na qual o dito fragmento abrange os nucleotídeos 297 a 594 da SEQ ID N°:3 ou SEQ ID N°:5.

12. Constructo de acordo com a reivindicação 11, caracterizado pelo fato de que a dita segunda sequência de nucleotídeos abrange o complemento de nucleotídeos 297 a 594 da SEQ ID N°:3 ou SEQ ID N°:5.

13. Método de acordo com qualquer uma das reivindicações de 1 a 3 e 5, caracterizado pelo fato de que a planta de tabaco transgênica é crescida no campo.

14. Método de obtenção de uma semente caracterizado por esta ser obtida a partir da planta *Nicotiana* transformada como reivindicado de acordo com o método de qualquer uma das reivindicações de 1 – 3 e 5.

Figura 1

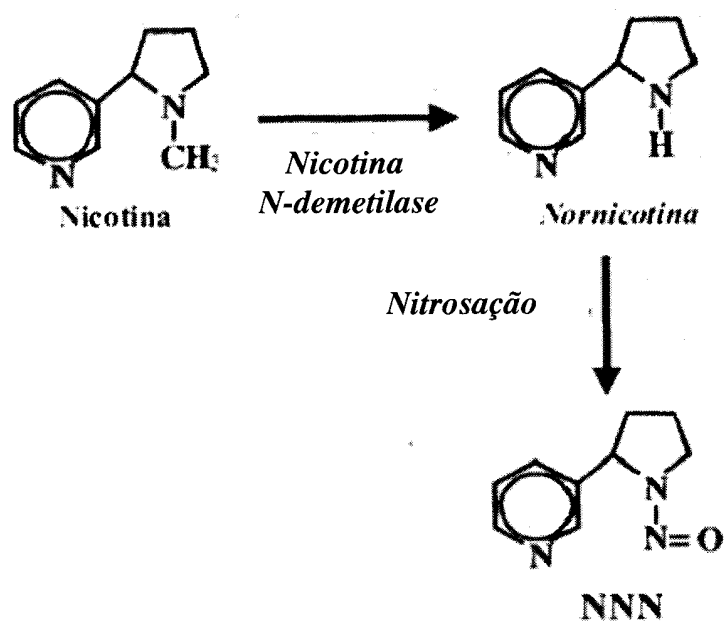


Figura 2

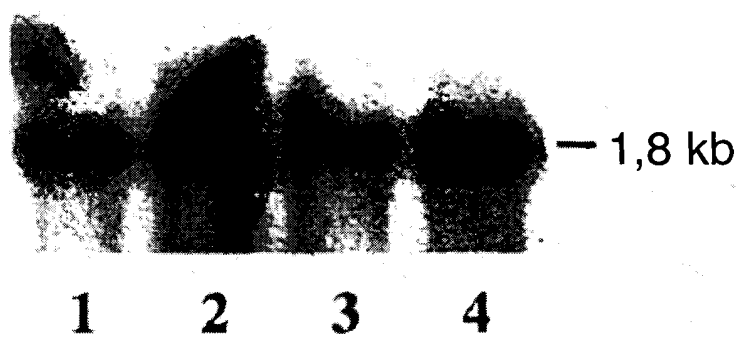


Figura 3A

3D_C12	ATGGTTTTTCCCATAGAGCC	TTGTAGGACTAGTAACCTT	---CACATTTCTCTTATAC	57		
131A_A02	ATGGTTTTTCCCATAGAGCCATTGTAGGA	AGTAACCTT	ACATTTCTCTTATAC	60		
3D_C12-15	ATGGTTTTTCCCATAGAGCC	TTGTAGGACTAGTAACCTT	---CACATTTCTCTTATAC	57		
3D_C12-10	ATG	TTT	TCCCATAGAGCCATTGTAGGACTAGTAACCTT	---CACATTTCTCTT	57	
3D_C12-7	ATGGTTTTTCCCATAGAGCCATTGTAGGACTAGTAACCTT	---CACATTTCTCTT	TT	57		
7D_A06	ATGGTTTTTCC	TAGAAGCCATTGTAGG	CTAGT	ACCTT	---CACATTTCTCTT	57

3D_C12	TTCCTATGGACAAAAAATCTCAAAAAC	T	CAAAACCCCTTACCACCGAAAATCCCCGGA	117		
131A_A02	TTCCTATG	ACAAAAAATCTCAAAAAC	TTCAAA	CCCTTACCA	CGAAAATCCCCGGA	120
3D_C12-15	TTCCTATGGACAAAAAATCTCAAAAAC	T	CAAAACCCCTTACCACCGAAAATCCCCGGA	117		
3D_C12-10	TTCCTATGGACAAAAAATCTCAAAAACCTTCAAAACCCCTTACCACCGAAAATCCCCGGA			117		
3D_C12-7	TTCCTATGGACAAAAAATCTCAAAAACCTTCAAAACCCCTTACCACCGAAAATCCCCGGA			117		
7D_A06	TTCCTATGGACAAAAAATCTCAAAAACCTTCAAAACCCCTTACCACCGAAAATCCCCGGA			117		

3D_C12	GGATGGCCGGTAATCGGCCATCTTTT	CAC	TTCAATAACGACGGCGACGACCGTCCATTA	177		
131A_A02	GGATGGCCGGTAATCGGCCATCTTTT	CCACTTCAATAACGACGGCGACGACCGTCCATT		180		
3D_C12-15	GGATGGCCGGTAATCGGCCATCTTTT	CAC	TTCAATAACGACGGCGACGACCGTCCATTA	177		
3D_C12-10	GGATGGCCGGTAATCGGCCATCTTTT	CCACTTCAAT	ACGACGGCGACGACCGTCCATTA	177		
3D_C12-7	GGATGGCCGGTAATCGGCCATCTTTT	CCACTTCAAT	ACGACGGCGACGACCGTCCATTA	177		
7D_A06	GGATGGCCGGTAATCGGCCATCTTTT	CCACTTCAAT	ACGACGGCGACGACCGTCCATTA	177		

3D_C12	GCTCGAAAACTCGGAGACTTAGCTGATAAAATACGGCCCCGTTTTT	CAC	TTTTTCGGCTAGGT	237		
131A_A02	GCTCGAAAACTCGGAGACTTAGCTGATAAAATACGGCCCCGTTTTT	CAC	TTTTTCGGCTAGGT	240		
3D_C12-15	GCTCGAAAACTCGGAGACTTAGCTGATAAAATACGGCCCCGTTTTT	CAC	TTTTTCGGCTAGGT	237		
3D_C12-10	GCTCGAAAACTCGGAGACTTAGCTGA	AAATACGGCCCCGTTTTT	CAC	TTTTTCGGCTAGG	237	
3D_C12-7	GCTCGAAAACTCGGAGACTTAGCTGA	AAATACGGCCCCGTTTTT	CAC	TTTTTCGGCTAGG	237	
7D_A06	GCTCGAAAACTCGGAGACTTAGCTGA	AAATACGG	CCCCGTTTTT	CAC	TTTTTCGGCTAGG	237

Figura 3B

3D_C12	CTTCCCCTTGTGCTAGTTGTAAGCAGTTACGAAGCTATAAAAGATTGCTTCTCTACAAAT	297
131A_A02	CTTCCCCTTGTGCTAGTTGTAAGCAGTTACGAAGCTATAAAAGATTGCTTCTCTACAAAT	300
3D_C12-15	CTTCCCCTTGTGCTAGTTGTAAGCAGTTACGAAGCTATAAAAGATTGCTTCTCTACAAAT	297
3D_C12-10	CTTCCCCTTGTGCTAGTTGTAAGCAGTTACGAAGCTATAAAAGATTGCTTCTCTACAAAT	297
3D_C12-7	CTTCCCCTTGTGCTAGTTGTAAGCAGTTACGAAGCTATAAAAGATTGCTTCTCTACAAAT	297
7D_A06	CTTCCCCTTGTGCTAGTTGTAAGCAGTTACGAAGCTATAAAAGATTGCTTCTCTACAAAT	297
*****--*****--*****--*****--*****--*****--*****--*****		
3D_C12	GACGCCATTTTCTCCAATCGTCCAGCTTTTCTTTACGGCGAATACCTTGGCTACAATAAT	357
131A_A02	GACGCCATTTTCTCCAATCGTCCAGCTTTTCTTTACGGCGAATACCTTGGCTACAATAAT	360
3D_C12-15	GAAGCCATTTTCTCCAATCGTCCAGCTTTTCTTTACGGCGAATACCTTGGCTACAATAAT	357
3D_C12-10	GACGCCATTTTCTCCAATCGTCCAGCTTTTCTTTACGGCGAATACCTTGGCTACAATAAT	357
3D_C12-7	GACGCCATTTTCTCCAATCGTCCAGCTTTTCTTTACGGCGAATACCTTGGCTACAATAAT	357
7D_A06	GACGCCATTTTCTCCAATCGTCCAGCTTTTCTTTACGGCGAATACCTTGGCTACAATAAT	357
--***--*****--*****--*****--*****--*****--*****		
3D_C12	GCATGCTTTTTCGCAAAATTACGGACCTTACTGGCGAAAAAATCGTAAATTAGTCATT	417
131A_A02	GCATGCTTTTTCGCAAAATTACGGACCTTACTGGCGAAAAAATCGTAAATTAGTCATT	420
3D_C12-15	GCCATGCTATTTTCGCAAAATTACGGACCTTACTGGCGAAAAAATCGTAAATTAGTCATT	416
3D_C12-10	GCCATGCTATTTTGGCAATTACGGACCTTACTGGCGAAAAAATCGTAAATTAGTCATT	417
3D_C12-7	GCCATGCTATTTTGGCAATTACGGACCTTACTGGCGAAAAAATCGTAAATTAGTCATT	417
7D_A06	GCCATGCTATTTTGGCAAAATTACGGACCTTACTGGCGAAAAAATCGTAAATTAGTCATT	417
--*****--*****--*****--*****--*****--*****--*****--*****		
3D_C12	CAGGAAGTTCTCTGCTAGTCGTCTCGAAAAATTCAAACAATGTGAGATTCCAGAATT	477
131A_A02	CAGGAAGTTCTCTGCTAGTCGTCTCGAAAAATTCAAACAATGTGAGATTCCAGAATT	480
3D_C12-15	CAGGAAGTTCTCTGCTAGTCGTCTCGAAAAATTCAAACACGTGAGATTGCAGAATT	476
3D_C12-10	CAGGAAGTTCTCTGCTAGTCGTCTCGAAAAATTCAAACACGTGAGATTGCAGAATT	477
3D_C12-7	CAGGAAGTTCTCTGCTAGTCGTCTCGAAAAATTCAAACACGTGAGATTGCAGAATT	477
7D_A06	CAGGAAGTTCTCTGCTAGTCGTCTCGAAAAATTCAAACACGTGAGATTGCAGAATT	477
*****--*****--*****--*****--*****--*****--*****--*****		

Figura 3D

3D_C12	GTGGATTTTCAAGG[CAT]TTAAGGC[ATGAAAAGGAC]TTTAA[GATATAGATTCTGTT	777
3D_C12-10	GTGGATTTTCAAGGGCATGTTAAGGCTATGAAAAGGACTTTTAAAGATATAGATTCTGTT	777
3D_C12-7	GTGGATTTTCAAGGGCATGTTAAGGCTATGAAAAGGACTTTTAAAGATATAGATTCTGTT	777
7D_A06	GTGGATTTTCAAGGGCATGTTAAGGCTATGAAAAGGACTTTTAAAGATATAGATTCTGTT	777
*****-*****-*****-*****-*****-*****		
3D_C12	TTTCAGAA[TCGTTAGAGGAACATATTAATAAAAGAGAAAAAAT]---GGAGGTTGGTGCA	834
3D_C12-10	TTTCAGAATTGGTTAGAGGAACATATTAATAAAAGAGAAAAAAT]---GGAGGTT[TCGCA	834
3D_C12-7	TTTCAGAATTGGTTAGAGGAACATATTAATAAAAGAGAAAAAAT]---GGAGGTT[TCGCA	834
7D_A06	TTTCAGAATTGGTTAGAGGAACATATTA[AAAAAGAGAAAAAAT][GGAGGTTGGT]CA	837
*****-*****-*****-*****-*****-*****		
3D_C12	GAAGGGAATGAACAAGATTTCAITGATGTGGTGCTTTCAAAA[TGAGTAA]GAATATCTT	894
3D_C12-10	GAAGGGAATGAACAAGATTTCAITGATGTGGTGCTTTCAAAAATGAGTAATGAATATCTT	894
3D_C12-7	GAAGGGAATGAACAAGATTTCAITGATGTGGTGCTTTCAAAAATGAGTAATGAATATCTT	894
7D_A06	GAAGGGAATGAACAAGATTT[AT]GATGTGGTGCTTTCAAAAATGAGTAATGAATATCTT	897
*****-*****-*****-*****-*****-*****		
3D_C12	G[GAAGGTTACTCTCGTGATACTGTCATTAAAGCAACAGT]TTTAGTTTGGTCTTGGAT	954
3D_C12-10	GGTGAAGGTTACTCTCGTGATACTGTCATTAAAGCAAC[GTGTTTAGTTTGGTCTTGGAT	954
3D_C12-7	GGTGAAGGTTACTCTCGTGATACTGTCATTAA[GCAAC]GTGTTTAGTTTGGTCTTGGAT	954
7D_A06	GG[GAAGGTTACTCTCG]GATACTGTCAT[AAAGCAACAGT]TTTAGTTTGGTCTTGGAT	957
*-----*****-*****-*****-*****-*****		
3D_C12	GCAGCAGACACAGTTGCTCTTCACATAAAATTGGGGGAATG[CATTATTGATAAACAATCAA	1014
3D_C12-10	GCAGCAGACACAGTTGCTCTTCACATAAAATTGGGGGAATGGCATTATTGATAAACAATCAA	1014
3D_C12-7	GCAGCAGACACAGTTGCTCTTCACATAAAATTGGGGGAATGGCATTATTGATAAACAATCAA	1014
7D_A06	GCAGCAGACACAGTTGCTCTTCACATAAAATG[GGAATGGCATTATTGATAAACAATCAA	1017
*****-*****-*****-*****-*****-*****		

Figura 3E

3D_C12	AATGCCTTGAAGAAAGCACAAGAAGAGATAGACACAAAAGTTGGTAAGGATAGATGGGTA	1074
3D_C12-10	AAAGCCTTGACGAAAGCACAAGAAGAGATAGACACAAAAGTTGGTAAGGATAGATGGGTA	1074
3D_C12-7	AAAGCCTTGACGAAAGCACAAGAAGAGATAGACACAAAAGTTGGTAAGGATAGATGGGTA	1074
7D_A06	AATGCCTTGAAGAAAGCACAAGAAGAGATAGACACAAAAGTTGGTAAGGATAGATGGGTA	1077
*****-*****-*****-*****-*****-*****-*****-*****-*****-*****		
3D_C12	GAAGAGAGTGATATTAAAGGATTGTATACCTCCAAGCTATTGTTAAAGAGTGTTACGA	1134
3D_C12-10	GAAGAGAGTGATATTAAAGGATTGGTATACCTCCAAGCTATTGTTAAAGAGTGTTACGA	1134
3D_C12-7	GAAGAGAGTGATATTAAAGGATTGGTATACCTCCAAGCTATTGTTAAAGAGTGTTACGA	1134
7D_A06	GAAGAGAGTGATATTAAAGGATTGGTATACCTCCAAGCTATTGTTAAAGAGTGTTACGA	1137
*****-*****-*****-*****-*****-*****-*****-*****-*****-*****		
3D_C12	TTATATCCACCAGGACCTTTGTTAGTACCACAGAAAAGTGAAGATTGTGTTGTTAGT	1194
3D_C12-10	TTATATCCACCAGGACCTTTGTTAGTACCACAGAAAAGTGAAGATTGTGTTGTTAGT	1194
3D_C12-7	TTATATCCACCAGGACCTTTGTTAGTACCACAGAAAAGTGAAGATTGTGTTGTTAGT	1194
7D_A06	TTATATCCACCAGGACCTTTGTTAGTACCACAGAAAAGTGAAGATTGTGTTGTTAGT	1197
*****-*****-*****-*****-*****-*****-*****-*****-*****-*****		
3D_C12	GGATATCACATTCCTAAAGGACTAGATTATTCGCAAAACGTCATGAAACTGCAACGAT	1254
3D_C12-10	GGATATCACATTCCTAAAGGGACAGATTATTCGCAAAACGTCATGAAACTGCAACGAT	1254
3D_C12-7	GGATATCACATTCCTAAAGGGACAGATTATTCGCAAAACGTCATGAAACTGCAACGAT	1254
7D_A06	GGATATCACATTCCTAAAGGACTAGATTATTCGCAAAACGTCATGAAACTGCAACGAT	1257
*****-*****-*****-*****-*****-*****-*****-*****-*****-*****		
3D_C12	CCTAAACCTCTGTCAAATCCTGATAAGTTCGATCCAGAGAGATTTCATCGCTGTGATATT	1314
3D_C12-10	CCTAAACCTCTGGTCATCCTGATAATTCGATCCAGAGAGATTTCATCGCTGTGATATT	1314
3D_C12-7	CCTAAACCTCTGGTCATCCTGATAATTCGATCCAGAGAGATTTCATCGCTGTGATATT	1314
7D_A06	CCTAAACCTCTGGTCATCCTGATAAGTTCGATCCAGAGAGATTTCATCGCTGTGATATT	1317
*****-*****-*****-*****-*****-*****-*****-*****-*****-*****		

Figura 3F

3D_C12	GACTTCGTCGTCAGTACTATGAGTATATCCCTTTGGTTCTGGAAGACGATCTTGTC	1374
3D_C12-10	GACTTTCGTGGTCAGTACTATAGTATATCCCGTTTGGTTCTGGAAGACGATCTTGTC	1374
3D_C12-7	GACTTTCGTGGTCAGTACTATAGTATATCCCGTTTGGTTCTGGAAGACGATCTTGTC	1374
7D_A06	GACTTTCGTCGTCAGTACTATGAGTATATCCCGTTTGGTTCTGGAAGACGATCTTGTC	1377
*****-*****-*****-*****-*****-*****-*****-*****-*****-*****		
3D_C12	GGGATGACTTATGCATTGCAAGTGGAAACACCTAACAAATGGCACATTTATCCAGGGTTTC	1434
3D_C12-10	GGGATGACTTATGCATTGCAAGTGGAAACACCTAACAAATGGCACATTTGATCCAGGTTTC	1434
3D_C12-7	GGGATGACTTATGCATTGCAAGTGGAAACACCTAACAAATGGCACATTTGATCCAGGTTTC	1434
7D_A06	GGGATGACTTATGCATTGCAAGTGGAAACACCTAACAAATGGCACATTTGATCCAGGGTTTC	1437
*****-*****-*****-*****-*****-*****-*****-*****-*****-*****		
3D_C12	AATTACAAGTCCAAATGACGAGCCTTGGATATGAAGGAAGGTGCAGGCATAACTATA	1494
3D_C12-10	AATTACAGAACTCCAAATGACGAGCCCTTGGATATGAAGGAAGGTGCAGGCATAACTATA	1494
3D_C12-7	AATTACAGAACTCCAAATGACGAGCCCTTGGATATGAAGGAAGGTGCAGGCATAACTATA	1494
7D_A06	AATTACAGAACTCCAAATGAGCCCTTGGATATGAAGGAAGGTGCAGGCATAACTATA	1497
*****-*****-*****-*****-*****-*****-*****-*****-*****-*****		
3D_C12	CGTAAGGTAAATCCGTGGAAATGATAATAACGCCTCGCTTGGCACCTGAGCTTTATTA	1554
3D_C12-10	CGTAAGGTAAATCCTGTGGAAGTGAATAACGCCTCGCTTGGCACCTGAGCTTTATTAA	1554
3D_C12-7	CGTAAGGTAAATCCTGTGGAAGTGAATAACGCCTCGCTTGGCACCTGAGCTTTATTAA	1554
7D_A06	CGTAAGGTAAATCCTGTGGAATGATAATAACGCCTCGCTTGGCACCTGAGCTTTATTAA	1557
*****-*****-*****-*****-*****-*****-*****-*****-*****-*****		
3D_C12	AACCTAAGATCTTTCATCTTGGTTGATCATTGTTTAATACTCCTAGATGGGTATTCATTT	1614
3D_C12-10	AACCTAAGATCTTTCATCTTGGTTGATCATTGTTTAATACTCCTAATGGTATTCATTT	1614
3D_C12-7	AACCTAAGATCTTTCATCTTGGTTGATCATTGTTTAATACTCCTAATGGTATTCATTT	1614
7D_A06	AACTAAGATCTTTCATCTTGGTTGATCATTGTTTAATACTCCTAGATGGGTATTCATTT	1617
~***-*****-*****-*****-*****-*****-*****-*****-*****-*****		

Figura 3G

3D_C12	ACCTTTTTCAATTAATTG	GTACGAG	TTTTTTAATTGGTA	AATTGTAA	AATAA	1674
3D_C12-10	ACCTTTT	TCAATTAATTGTCAGTACGAGTTTTT	TAATTGGTACATTGTAATAATAA			1674
3D_C12-7	ACCTTTT	TCAATTAATTGTCAGTACGAGTTTTT	TAATTGGTACATTGTAATAATAA			1674
7D_A06	ACCTTTTTCAATTA	TTGTC	GTACG	TTTTTTAATTGGTA	TTTGTAAATAATAA	1677
	*****-----*****					
3D_C12	GTAAAGAATGATTGTGCTAATATATAAAG	TTTGCAGAAGATAATTGACTGATTGTCCC				1733
3D_C12-10	GTAAAGAAT	AATTGTGCTAATATATAAAGGTTTG	AGAAGATAATTGACTGATTGTCCC			1733
3D_C12-7	GTAAAGAAT	AATTGTGCTAATATATAAAGGTTTG	AGAAGAT	AATTGACTGATTGTCCC		1733
7D_A06	GTAAAGAA	GATTGTGCTAATATATAA	GGT--GCA	AA	ATAATTGA	1727
	*****-----*****					

Figura 4

[illegible]

Figura 5

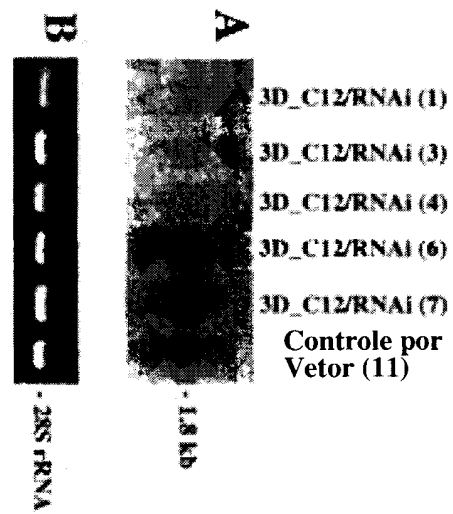


Figura 6

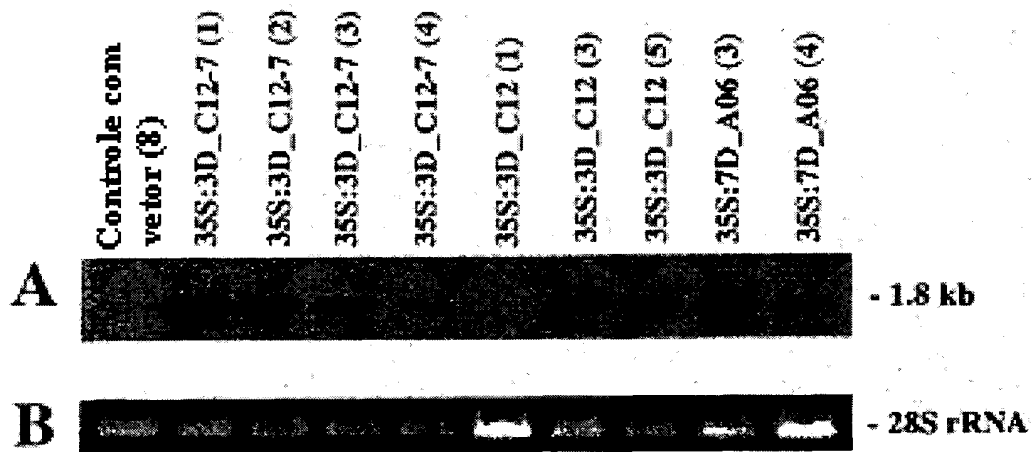


Figura 7

CTGTCATTAAAGCAACGGTGTGTTGTAAAGTTCACTGTCAATTTTCATTTATTCACTTTT
ATTTTGAGGAGCAGACATGTTAAATAATATTTGGAGCACTGTAAAGTTATCTATGTGT
ACAGGTTTCGAGCCTCAGGTGCAACCACTAATGCTTGTATTAGATTATGTTGTCTGCAT
CATACCCCTAATTTGGAGTGTGGCTCTTCCCGAACCTGCAATGGCTGGATGCTGGATGC
TTTATGTATCAGACTGACCTTTTTGTTAACTATCTAAATACAAAGGATGATTTAATAA
AAATATAGATGGTAACAGAAAGATGAGATTATTTTGGGGCTATATGGGATTCCG
CCGGGCTTTGGGAGGTAAACGGTATCTACCAATTGAGACTTTACTCCAGAACCTTAT
CTCGAGAGCTCTGAATAAAATGAAATAGTATTTACCACTCCAAATCTTTGATGGTAA
AAAGATGAGATATAACCTCTTATATTTGATTGAAACACGTTGATAGACTAAACTTCTT
TACTCCCATTCAGCATAGAAATGAACCAACGGGATTCCTTCTCTTTTTAGGGGG
AAATTCCTTATTTGCTTGTGAAATATAGATTCAATGTCGTTATTCTATTTTTAATATGAT
GAAATCAATATAGTCAAGTTAATACCTATGTCAATTTGGTTTGGGGACAAATATAT
GGAACTATATATACGTCTATATAGATAGTGATTATTTAGAGGATATACATTTTTTT
TGGATAAATATTTGATTTATTGGATTAAAAATAGAAATATACAGGTAAAGGTCTAACA
GTGTTTGTCTTTTACACTAAATAACTTGACCTCGTACAATTCTAAGAAATATTTGAA
TAAATGAATTAATTTATTTGTTATCAATTAATAAATCATAGTATAGATGAGATGTGTG
CATACTTGGCAATAACTATCTAATCAAAACAGGTATGTGAATATTTGATATTCCTTT
TTAATTAATCTTTTTTCCAGGTTTGGTCTTGGATGCAGCAGACACAGTTGCTCTTCA
CATAAATTGGCGAATGGCATTATTGATAAAACAATCAAAAGGGCTTGACCGAAAGCAC
AAG

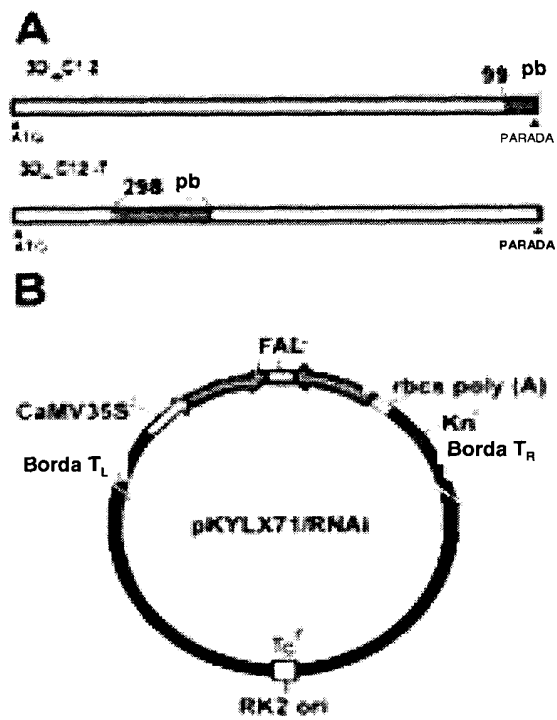


Figura 8