Alternation of Tobacco Alkaloid Content Through Modification of Specific Cytochrome P450 Genes

Ralph E. Dewey
Steven W. Bowen
Balazs Siminszky
University of Kentucky
Lily Gavilano
University of Kentucky

Follow this and additional works at: https://uknowledge.uky.edu/pss_patents

Part of the Plant Sciences Commons

Recommended Citation
Dewey, Ralph E.; Bowen, Steven W.; Siminszky, Balazs; and Gavilano, Lily, "Alternation of Tobacco Alkaloid Content Through Modification of Specific Cytochrome P450 Genes" (2008). Plant and Soil Sciences Faculty Patents. 1.
https://uknowledge.uky.edu/pss_patents/1

This Patent is brought to you for free and open access by the Plant and Soil Sciences at UKnowledge. It has been accepted for inclusion in Plant and Soil Sciences Faculty Patents by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
ALTERATION OF TOBACCO ALKALOID CONTENT THROUGH MODIFICATION OF SPECIFIC CYTOCHROME P450 GENES

Inventors: Ralph E. Dewey, Apex, NC (US);
Steven W. Bowen, Raleigh, NC (US);
Balazs Siminszky, Lexington, KY (US);
Lily Gavilano, Lexington, KY (US)

Assignees: North Carolina State University,
Raleigh, NC (US); University of Kentucky Research Foundation,
Lexington, KY (US)

Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 869 days.

Appl. No.: 11/580,765
Filed: Oct. 13, 2006

Prior Publication Data

Related U.S. Application Data
Continuation-in-part of application No. PCT/US2005/005665, filed on Feb. 23, 2005.

Int. Cl.
C12N 15/82 (2006.01)

U.S. Cl. ............................................. 800/285, 800/287

Field of Classification Search ....................... None
See application file for complete search history.

References Cited

U.S. PATENT DOCUMENTS
5,668,295 A 9/1997 Wahab et al.
2006/0037996 A1 2/2006 Xu

FOREIGN PATENT DOCUMENTS
WO WO 00/67558 A1 11/2000

OTHER PUBLICATIONS

* cited by examiner

Primary Examiner—Ashwin Mehta
(74) Attorney, Agent, or Firm—Alston & Bird LLP

ABSTRACT
Compositions and methods for reducing the level of nornicotine and N'-nitrosonornicotine (NNN) in Nicotiana plants and plant parts thereof are provided. The compositions comprise isolated polynucleotides and polypeptides for cytochrome P450s that are involved in the metabolic conversion of nicotine to nornicotine in these plants. Expression cassettes, vectors, plants, and plant parts thereof comprising inhibitory sequences that target expression or function of the disclosed cytochrome P450 polypeptides are also provided. Methods for the use of these novel sequences to inhibit expression or function of cytochrome P450 polypeptides involved in this metabolic conversion are also provided. The methods find use in the production of tobacco products that have reduced levels of nornicotine and its carcinogenic metabolite, NNN, and thus reduced carcinogenic potential for individuals consuming these tobacco products or exposed to secondary smoke derived from these products.

25 Claims, 14 Drawing Sheets
Figure 1

Nicotine $\xrightarrow{\text{nicotine } N$-demethylase} \text{Nornicotine}$

$\xrightarrow{\text{nitrosation}} \text{NNN}$
<table>
<thead>
<tr>
<th>Figure 3A</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D_C12</td>
</tr>
<tr>
<td>13IA_A02</td>
</tr>
<tr>
<td>3D_C12-15</td>
</tr>
<tr>
<td>3D_C12-10</td>
</tr>
<tr>
<td>3D_C12-17</td>
</tr>
<tr>
<td>7D_A06</td>
</tr>
</tbody>
</table>

| 3D_C12  |
| 13IA_A02 |
| 3D_C12-15 |
| 3D_C12-10 |
| 3D_C12-17 |
| 7D_A06  |

| 3D_C12  |
| 13IA_A02 |
| 3D_C12-15 |
| 3D_C12-10 |
| 3D_C12-17 |
| 7D_A06  |

| 3D_C12  |
| 13IA_A02 |
| 3D_C12-15 |
| 3D_C12-10 |
| 3D_C12-17 |
| 7D_A06  |
Figure 3C
### Figure 3E

<table>
<thead>
<tr>
<th>3D_C12</th>
<th>AATGCTTTGACGGAACCAAGAGAGATAAGACCACAAAAGTTGGATAGGATAGATGTTG</th>
<th>1074</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D_C12-10</td>
<td>AAGGCCCTGACGGAACCAAGAGAGATAAGACCACAAAAGTTGGATAGGATAGATGTTG</td>
<td>1074</td>
</tr>
<tr>
<td>3D_C12-7</td>
<td>AAGGCCCTGACGGAACCAAGAGAGATAAGACCACAAAAGTTGGATAGGATAGATGTTG</td>
<td>1074</td>
</tr>
<tr>
<td>7D_A06</td>
<td>AATGCTTTGACGGAACCAAGAGAGATAAGACCACAAAAGTTGGATAGGATAGATGTTG</td>
<td>1077</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3D_C12</td>
<td>GAAGAGATGTATTTAAAGGTATTCTATCCAGCTATTGTTAAAGAGAGTGTAGCA</td>
<td>1134</td>
</tr>
<tr>
<td>3D_C12-10</td>
<td>GAAGAGATGTATTTAAAGGTATTCTATCCAGCTATTGTTAAAGAGAGTGTAGCA</td>
<td>1134</td>
</tr>
<tr>
<td>3D_C12-7</td>
<td>GAAGAGATGTATTTAAAGGTATTCTATCCAGCTATTGTTAAAGAGAGTGTAGCA</td>
<td>1134</td>
</tr>
<tr>
<td>7D_A06</td>
<td>GAAGAGATGTATTTAAAGGTATTCTATCCAGCTATTGTTAAAGAGAGTGTAGCA</td>
<td>1137</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3D_C12</td>
<td>TTATATCCAGCTAGACCTTTTTTTCTTCTATACACTACGAAAATGTAGATGTATGTTAG</td>
<td>1194</td>
</tr>
<tr>
<td>3D_C12-10</td>
<td>TTATATCCAGCTAGACCTTTTTTTCTTCTATACACTACGAAAATGTAGATGTATGTTAG</td>
<td>1194</td>
</tr>
<tr>
<td>3D_C12-7</td>
<td>TTATATCCAGCTAGACCTTTTTTTCTTCTATACACTACGAAAATGTAGATGTATGTTAG</td>
<td>1194</td>
</tr>
<tr>
<td>7D_A06</td>
<td>TTATATCCAGCTAGACCTTTTTTTCTTCTATACACTACGAAAATGTAGATGTATGTTAG</td>
<td>1197</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3D_C12</td>
<td>GGATACTACATCTCTTAAAGCATCTAGATATTCCGAAAGGTCATGAAACTAGACGAT</td>
<td>1254</td>
</tr>
<tr>
<td>3D_C12-10</td>
<td>GGATACTACATCTCTTAAAGCATCTAGATATTCCGAAAGGTCATGAAACTAGACGAT</td>
<td>1254</td>
</tr>
<tr>
<td>3D_C12-7</td>
<td>GGATACTACATCTCTTAAAGCATCTAGATATTCCGAAAGGTCATGAAACTAGACGAT</td>
<td>1254</td>
</tr>
<tr>
<td>7D_A06</td>
<td>GGATACTACATCTCTTAAAGCATCTAGATATTCCGAAAGGTCATGAAACTAGACGAT</td>
<td>1257</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3D_C12</td>
<td>CCTAAACTCTGTGATAAATCTTGATCCAGAAGAGATTCATCGCTGCTGTATT</td>
<td>1314</td>
</tr>
<tr>
<td>3D_C12-10</td>
<td>CCTAAACTCTGTGATAAATCTTGATCCAGAAGAGATTCATCGCTGCTGCTGTATT</td>
<td>1314</td>
</tr>
<tr>
<td>3D_C12-7</td>
<td>CCTAAACTCTGTGATAAATCTTGATCCAGAAGAGATTCATCGCTGCTGCTGTATT</td>
<td>1314</td>
</tr>
<tr>
<td>7D_A06</td>
<td>CCTAAACTCTGTGATAAATCTTGATCCAGAAGAGATTCATCGCTGCTGCTGTATT</td>
<td>1317</td>
</tr>
</tbody>
</table>
Figure 5

A

- 1.8 kb

B

- 28S rRNA
Figure 7

CTGTCATTTAAAGCAACCGGTGTTTTTGAAGTCCATCTGTCAATTTTTCATTTATTACCT
TTATTTTGAGGAGCAGACATGTTAACAAATAATTGAGAACACTGTAATGAGATTTATCTATT
GTGTAACAGCTTCGAGCCTACGGAACCACTAAATGCTTGATATTGATTATTGTGT
CTCCATCATACCCTAATTGGAGTAGTGGCTTCTTCCGAAACCCCTGCAATGCGTGGATG
CTGGATGTCTTTATGTCAGACTCCCTTTTTGTAAACTATCCTAAATATACTAAGGAT
GATTAAATTAAAAATATAGAATGGAACAGAAAAAGATGAGATTATTTTGGGCT
ATATTGGAATCGCCGCGCTTTGAGCCAAAAACGATGTATCACCAGTGGACTTTA
CTCCAGAAGCTTATCTCGAGAGCTCTGAATAAAAATGAAATAGTATATTACACTCCA
AAATCTTTGTGATGTAATAAAGATGAGATATAACCTCTTTATAATTGATGGAAACCAGTT
GATAGAATAAAAGTTCTTTACCTCCATTCAAGCATAAGAAAAATGAAACCAAACCGGA
ATTCTCTCTTTTTTATGGGAAAATTCCCTTAATGCTTTGGAATATAGAATTACATGT
CGTTATTCATTTTTAATTATGAGAAAACTCAATATAGTCAAGTTTTAATCTTAAGAGT
CATTTGGTTTGCAGGCAAGTTATATTGGAACATATATATAATACGTCTATTATAGAATAG
TGATTTTATGAGATATATTTTTTGGAATAATATTGATTATTGAGTTAAAA
ATAGAATATAACGCTAAACGCTGTGTTTGCCTTTACACTAAATAKACTTG
ACCCTGTACAAATTCATAAACAAATAATTTGGAAATAATGAAATTATTATTTATTGTTAATCA
ATTAAAAATTCTCATATGAGATATGTTGACATCCTTGCGAACTAATGTATCTAA
CTAAAACAACGACTGTTGAATATTGATAATTCTTTTTTTAAATTATTCTTTTTTCGAG
GTGGTTGTCTTTGGATGACGACAGACACAGGTGCTACTTTGCAATGAAAAATTGGGGAATGGC
ATTATTGATAAAACAAATCAAAGGCGCTTTGACGAAAAGCACAAG
Figure 8

A

3D_C12

ATG

STOP

99 bp

3D_C12-7

ATG

STOP

298 bp

B

pKYLYX71/RNAi

CaMV35S

Tl-Border

rbcS poly (A)

Kn'

Tr-Border

FAD

Tc'

RK2 ori
ALTERATION OF TOBACCO ALKALOID CONTENT THROUGH MODIFICATION OF SPECIFIC CYTOCHROME P450 GENES

CROSS-REFERENCE TO RELATED APPLICATION

This application is a continuation-in-part of International Application No. PCT/US2005/005665, filed Feb. 23, 2005, the content of which is herein incorporated by reference in its entirety.

FIELD OF THE INVENTION

The invention relates to compositions and methods for reducing the level of nicotinic acid and its metabolite, N-nitrosornicotine, in a plant that is a member of the genus Nicotiana, particularly compositions and methods for inhibiting expression or function of a cytochrome P450 polypeptide involved in the metabolic conversion of nicotine to nor-nicotinie.

BACKGROUND OF THE INVENTION

The predominant alkaloid found in commercial tobacco varieties is nicotine, typically accounting for 90-95% of the total alkaloid pool. The remaining alkaloid function is comprised primarily of three additional pyridine alkaloids: nor-nicotinie, anabasine, and anatabine. Nicotinic acid is generated directly from nicotine through the activity of the enzyme nicotine N-demethylase (FIG. 1). Nicotinic acid usually represents less than 5% of the total pyridine alkaloid pool, but through a process termed “conversion,” tobacco plants that initially produce very low amounts of nicotine give rise to progeny that metabolically “convert” a large percentage of leaf nicotine to nor-nicotinie. In tobacco plants that have genetically converted (termed “converters”), the great majority of nor-nicotinie production occurs during the senescence and curing of the mature leaf (Wernsman and Matzinger (1968) Tob. Sci. 12:226-228). Burley tobaccos are particularly prone to genetic conversion, with rates as high as 20% per generation observed in some cultivars.

During the curing and processing of the tobacco leaf, a portion of the nor-nicotinie is metabolized to the compound N-nitrosornicotine (NNN; FIG. 1), a tobacco-specific nitrosoamine (TSNA) that has been shown to be carcinogenic in laboratory animals (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). In flue-cured tobaccos, TSNA were found to be predominantly formed through the reaction of alkaloids with the minute amounts of nitrogen oxides present in combustion gases formed by the direct-fired heating systems found in traditional curing barns (Peele and Gentry (1999) “Formation of Tobacco-specific Nitroamines in Flue-cured Tobacco,” CORESTA Meeting, Agro-Phyto Groups, Suzhou, China). Refrating these curing barns with heat-exchangers virtually eliminated the mixing of combustion gases with the curing air and dramatically reduced the formation of TSNA in tobaccos cured in this manner (Boyette and Ham (2001) Rec. Adv. Tob. Sci. 27:17-22). In contrast, in the air-cured Burley tobaccos, TSNA formation proceeds primarily through reaction of tobacco alkaloids with nitrile, a process catalyzed by leafborne microbes (Bush et al. (2001) Rec. Adv. Tob. Sci. 27:23-46). Thus far, attempts to reduce TNSAs through modification of curing conditions while maintaining acceptable quality standards have not proven to be successful for the air-cured tobaccos.

In Burley tobaccos, a positive correlation has been found between the nor-nicotinie content of the leaf and the amount of NNN that accumulates in the cured product (Bush et al. (2001) Rec. Adv. Tob. Sci. 27:23-46; Shi et al. (2000) Tob. Chem. Res. Conf. 54:Abstract 27). However, keeping nor-nicotinie levels at a minimum has been difficult because of the conversion phenomenon that results in a continual introduction of high nor-nicotinie-producing plants within commercially grown Burley populations. Minimizing the number of Burley plants that accumulate high levels of nor-nicotinie has traditionally been the responsibility of plant breeders and seed producers. Though the percentage of converter plants that are ultimately grown in farmers’ fields can be reduced through the roguing of converter plants during the propagation of seed stocks, this process is costly, time-consuming, and imperfect.

Previous studies have shown that once a plant has converted, the high nor-nicotinie trait is inherited as a single dominant gene (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). The nature of this gene, however, is currently unknown. In the most simple of scenarios, the conversion locus may represent a nonfunctional nicotine N-demethylase gene that regains its function in converter plants, possibly through the mobilization of a mutation-inducing transposable element. Alternatively, the converter locus may encode a protein that initiates a cascade of events that ultimately enables the plant to metabolize nicotine to nor-nicotinie, which would mean that multiple genes may be involved.

Regardless of whether there are one or many genes associated with the conversion process, it is clear that the gene(s) encoding polypeptides having nicotine demethylase activity play a pivotal role in this process. Although the inability to purify active nicotine N-demethylase from crude extracts has impeded the isolation and identification of this enzyme, there is some evidence that a member of the cytochrome P450 superfamily of monooxygenases may be involved (Hao and Yeoman (1996) Phytochem. 41:477-482; Hao and Yeoman (1996) Phytochem. 42:325-329; Chelvaraj et al. (1993) J. Agric. Food Chem. 41:858-862; Hao and Yeoman (1998) J. Plant Physiol. 152:420-426). However, these studies are not conclusive, as the classic P450 inhibitors carbon monoxide and tetcyclaisyl have failed to lower enzyme activity at rates comparable to other reported P450-mediated reactions (Chelvaraj et al. (1993) J. Agric. Food Chem. 41:858-862).

Furthermore, the cytochrome P450s are ubiquitous, transmembrane proteins that participate in the metabolism of a wide range of compounds (reviewed by Schuler (1996) Crit. Rev. Plant Sci. 15:235-284; Schuler and Werck-Reichhart (2003) Annu. Rev. Plant Biol. 54:629-667). Examples of biochemical reactions mediated by cytochrome P450s include hydroxylations, demethylations, and epoxidations. In plants, the cytochrome P450 gene families are very large. For example, total genome sequence examination has revealed 272 predicted cytochrome P450 genes in Arabidopsis and at least 455 unique cytochrome P450 genes in rice (see, for example, Nelson et al. (2004) Plant Physiol. 135(2):756-772). Even though cytochrome P450 has been implicated as having a role in the metabolic conversion of nicotine to nor-nicotinie, identification of key participating members of this protein family remains a challenge.

Aside from serving as a precursor for NNN, recent studies suggest that the nor-nicotinie found in tobacco products may have additional undesirable health consequences. Dickerson and Janda demonstrated that nor-nicotinie causes aberrant protein glycation within the cell (Dickerson and Janda (2002) Proc. Natl. Acad. Sci. USA 99:15084-15088). Concentrations
of nornicotine-modified proteins were found to be much higher in the plasma of smokers compared to nonsmokers. Furthermore, this same study showed that nornicotine can covalently modify commonly prescribed steroid drugs such as prednisone. Such modifications have the potential of altering both the efficacy and toxicity of these drugs.

In view of the difficulties associated with conversion and the undesirable health effects of nornicotine accumulation, improved methods for reducing the nornicotine content in tobacco varieties, particularly Burley tobacco, are therefore desirable. Such methods would not only help ameliorate the potential negative health consequences of the nornicotine per se as described above, but should also concomitantly reduce NNN levels.

SUMMARY OF THE INVENTION

Compositions and methods for reducing the nornicotine content in plants that are members of the genus Nicotiana are provided. Compositions include isolated cytochrome P450 polynucleotides and polypeptides that are involved in the metabolic conversion of nicotine to nornicotine in plants, particularly Nicotiana species. The isolated polynucleotides comprise the nucleotide sequence set forth in SEQ ID NO:1, 3, 5, 7, 9, or 11, a nucleotide sequence encoding a polypeptide as set forth in SEQ ID NO:2, 4, 6, 8, 10, or 12, and fragments and variants thereof. Isolated polypeptides of the invention comprise an amino acid sequence set forth in SEQ ID NO:2, 4, 6, 8, 10, or 12, an amino acid sequence encoded by the nucleotide sequence set forth in SEQ ID NO:1, 3, 5, 7, 9, or 11, and fragments and variants thereof.

The polynucleotides of the invention find use in suppressing expression of a cytochrome P450 that is involved in the metabolic conversion of nicotine to nornicotine in a plant, including the cytochrome P450s of the present invention. In this manner, compositions further include expression cassettes comprising an inhibitory sequence that is capable of inhibiting expression or function of a cytochrome P450 polypeptide of the invention, where the inhibitory sequence is operably linked to a promoter that is functional in a plant cell. In some embodiments, the inhibitory sequence comprises the sequence set forth in SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 14, 15, or 16, or a complement or fragment thereof. Compositions also include transformed plants and plant parts that comprise an expression cassette of the present invention, optionally stably incorporated into the genome of the plant. Further provided are tobacco products, including chewing tobacco, snuff, cigarettes, pipe tobacco, and cigars, having a reduced level of nornicotine, and its related nitrosamine, N'-nitrosornicotine.

The methods of the invention comprise inhibiting the expression or function of a cytochrome P450 polypeptide of the present invention. In some embodiments, an expression cassette comprising an inhibitory sequence that targets expression or function of a cytochrome P450 polypeptide of the present invention is introduced into the plant or plant part of interest, wherein expression of the inhibitory sequence produces a polynucleotide or polypeptide that inhibits expression or function of a cytochrome P450 polypeptide of the invention. In one such embodiment, the inhibitory sequence comprises a sequence set forth in SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 14, 15, or 16, or a complement or fragment thereof.

The methods of the invention find use in the production of Nicotiana plants that have decreased levels of nornicotine and its metabolite, the nitrosamine N'-nitrosornicotine, within the leaf and stem tissues. When harvested, the leaf and stem tissues of these plants can be utilized to produce tobacco products having reduced levels of nornicotine and this tobacco-specific nitrosamine, and thus reduced carcinogenic potential for individuals consuming these products or exposed to secondary smoke derived from these products.

Also provided are transgenic Nicotiana plants having a nicotine to nornicotine conversion rate of less than about 2%, where the plants comprise a heterologous nucleic acid construct comprising a promoter capable of functioning in a plant cell operably linked to a nucleic acid sequence having a first nucleotide sequence comprising a fragment of about 100 nucleotides and about 400 nucleotides of a Nicotiana nicotine demethylase polynucleotide and a second nucleotide sequence capable of forming a double-stranded RNA with the first nucleotide sequence, where the transgenic Nicotiana plants are transgenic converter lines of Nicotiana. In some embodiments, the Nicotiana nicotine demethylase polynucleotide is a tobacco nicotine demethylase polynucleotide.

The present invention also provides a recombinant nucleic acid construct comprising a promoter capable of functioning in a plant cell operably linked to a nucleic acid sequence having a first nucleotide sequence comprising a fragment of between about 100 nucleotides and about 400 nucleotides of a nicotine demethylase polynucleotide and a second nucleotide sequence capable of forming a double-stranded RNA with the first nucleotide sequence, and regenerating a transgenic Nicotiana plant. In some embodiments, the Nicotiana nicotine demethylase polynucleotide is a tobacco nicotine demethylase polynucleotide.

The present invention also provides seed obtained from the transgenic Nicotiana plant having a nicotine to nornicotine conversion rate of less than about 2%, where the seed comprises a heterologous nucleic acid construct comprising a promoter capable of functioning in a plant cell operably linked to a nucleic acid sequence having a first nucleotide sequence comprising a fragment of between about 100 nucleotides and about 400 nucleotides of a Nicotiana nicotine demethylase polynucleotide and a second nucleotide sequence capable of forming a double-stranded RNA with the first nucleotide sequence, where the transgenic Nicotiana plants are transgenic converter lines of Nicotiana. In some embodiments, the Nicotiana nicotine demethylase polynucleotide is a tobacco nicotine demethylase polynucleotide.

The present invention also provides transgenic plant cells comprising a nucleic acid molecule having a promoter functional in a plant cell and a nucleic acid sequence encoding a nicotine demethylase having an isoleucine residue at position 274 and a tryptophan residue at position 330.

The present invention also provides methods of screening for a nicotine demethylase sequence comprising: obtaining a nucleic acid sequence that has greater than about 90% sequence identity with SEQ ID NO:5 and identifying a codon sequence encoding for a tryptophan residue at position 330 of the encoded polypeptide.

Further provided are methods for screening for a nicotine demethylase having an isoleucine at position 274 or a tryptophan at position 330 comprising obtaining a nucleic acid
sequence that has greater than about 90% sequence identity with SEQ ID NO:5 and identifying a first codon sequence encoding for an isoleucine residue at position 274 or a second codon sequence encoding a tryptophan residue at position 330 of the encoded polypeptide.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows the structures of nicotine, nornicotine, and N'-nitrosornornicotine (NNN).

FIG. 2 shows Northern blot analysis of converter and non-converter RNAs using 7D_A06 as a hybridization probe. Lanes 1 and 2 show RNAs isolated from sodium bicarbonate-treated leaves of genotypes DH 98-325-5 (nonconverter) and DH 98-325-6 (converter), respectively. Lanes 3 and 4 show RNAs isolated from ethephon-treated leaves of genotypes DH 98-326-3 (nonconverter) and DH 98-326-1 (converter), respectively. Estimated size of the hybridizing band is indicated in kilobases (kb).

FIG. 3A-3G shows a nucleotide sequence alignment of members of the 3D_C12 gene family. Asterisks denote positions where sequence identity is conserved among all sequences compared. Positions where differences are found are indicated with dashes and the corresponding residues are shaded grey. The nucleotide sequences present in the alignment include 3D_C12 (SEQ ID NO:1); 3D_C12-10 (SEQ ID NO:3); 3D_C12-7 (SEQ ID NO:5); 7D_A06 (SEQ ID NO:7); 3D_C12-15 (SEQ ID NO:9); and 131A_A02 (SEQ ID NO:11). The 3D_C12-15 and 131A_A02 entries are partial-length cDNA sequences. The 99 bp region of 3D_C12 that was used to make the RNAi-based construct is underlined.

FIG. 4 shows an alignment of predicted amino acid sequences for full-length members of the 3D_C12 family of P450 genes. The amino acid sequences present in the alignment include 3D_C12 (SEQ ID NO:2); 3D_C12-10 (SEQ ID NO:4); 3D_C12-7 (SEQ ID NO:6); and 7D_A06 (SEQ ID NO: 8). Asterisks denote positions conserved among all four sequences. Residues that differ among the members are shaded in grey.

FIG. 5 shows a Northern blot analysis of transgenic plants possessing the 3D_C12/RNAi construct. (A) Hybridization of the 3D_C12-7 probe to RNAs isolated from ethephon-treated, cured leaves of transgenic plants displaying low or nornicotine phenotypes (3D_C12/RNAi-1, 3, and 4) and high nornicotine phenotypes (3D_C12/RNAi-6, 7, and vector-only control plant 11). Estimated size of hybridizing band is indicated in kilobases (kb). (B) Ethidium bromide staining of the portion of the gel used in (A) that contains the 28S ribosomal RNA to show the relative equivalence of RNA loading among the lanes.

FIG. 6 shows a Northern blot analysis of transgenic plants possessing sense-orientation constructs of members of the 3D_C12 gene family. (A) Hybridization of the 3D_C12-7 probe to RNAs isolated from non-treated leaves of independent transgenic lines expressing 3D_C12-7, 3D_C12, and 7D_A06 constructs and a vector-only control (control 8). Estimated size of hybridizing band is indicated in kilobases (kb). (B) Ethidium bromide staining of the portion of the gel used in (A) that contains the 28S ribosomal RNA to show the relative equivalence of RNA loading among the lanes.

FIG. 7 shows a genomic sequence of a fragment of the 3D_C12-10 gene (see SEQ ID NO: 23) possessing an intron. Intron sequences are indicated in bold, italicized type. Exon sequences are shown in plain type. The sequences corresponding to the PCR primers used to amplify the fragment from the tobacco genomic DNA are underlined.

FIG. 8A shows a diagram of the RNAi constructs used to silence expression of members of the 3D_C12 gene family; and FIG. 8B shows a diagram of the RNAi binary expression vector, pKYLX7/1/RNAi.

DETAILED DESCRIPTION OF THE INVENTION

Background and Definitions

Before describing the present invention in detail, it is to be understood that many modifications and other embodiments of the inventions set forth herein will come to mind to one skilled in the art to which this invention pertains having the benefit of the teachings presented in the foregoing descriptions and the associated drawings. Therefore, it is to be understood that the invention is not to be limited to the specific embodiments disclosed and that modifications and other embodiments are intended to be included within the scope of the appended claims. Rather, these embodiments are provided so that this disclosure will satisfy applicable legal requirements.

Although specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation. Like numbers refer to like elements throughout. Further, the article “a” and “an” are used herein to refer to one or more than one (i.e., to at least one) of the grammatical object of the article. By way of example, “an element” means one or more element. Throughout the specification the word “comprising,” or variations such as “comprises” or “comprising,” will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

The present invention is drawn to compositions and methods for inhibiting the expression or function of cytochrome P450 polypeptides that are involved in the metabolic conversion of nicotine to nornicotine in a plant, particularly plants of the Nicotiana genus, including tobacco plants of the various commercial varieties. As used herein, the terms “inhibit,” “inhibition,” and “inhibiting” are defined as any method known in the art or described herein that decreases the expression or function of a gene product of interest (i.e., the target gene product). “Inhibition” can be in the context of a comparison between two plants, for example, a genetically altered plant versus a wild-type plant. Alternatively, inhibition of expression or function of a target gene product can be in the context of a comparison between plant cells, organs, tissues, or plant parts within the same plant or between different plants, and includes comparisons between developmental or temporal stages within the same plant or plant part or between plants or plant parts. “Inhibition” includes any relative decrement of function or production of a gene product of interest, up to and including complete elimination of function or production of that gene product.

The term “inhibitory sequence” encompasses any polynucleotide or polypeptide sequence capable of inhibiting the expression or function of a cytochrome P450 polypeptide involved in the metabolic conversion of nicotine to nornicotine in a plant, such as full-length polynucleotide or polypeptide sequences, truncated polynucleotide or polypeptide sequences, fragments of polynucleotide or polypeptide sequences, variants of polynucleotide or polypeptide sequences, sense-oriented nucleotide sequences, antisense-
oriented nucleotide sequences, the complement of a sense- or antisense-oriented nucleotide sequence, inverted regions of nucleotide sequences, hairpins of nucleotide sequences, double-stranded nucleotide sequences, single-stranded nucleotide sequences, combinations thereof, and the like. The term "polynucleotide sequence" includes sequences of RNA, DNA, chemically modified nucleic acids, nucleic acid analogs, combinations thereof, and the like.

Inhibitory sequences are designated herein by the name of the target gene product. Thus, a "cytochrome P450 inhibitory sequence" refers to an inhibitory sequence that is capable of inhibiting the expression of a cytochrome P450 polypeptide that is involved in the metabolic conversion of nicotine to nornicotine in a plant, for example, at the level of transcription and/or translation, or which is capable of inhibiting the function of such a cytochrome P450 polypeptide. When the phrase "capable of inhibiting" is used in the context of a polynucleotide inhibitory sequence, it is intended to mean that the inhibitory sequence itself exerts the inhibitory effect; or, where the inhibitory sequence encodes an inhibitory nucleotide molecule (for example, hairpin RNA, miRNA, or double-stranded RNA polynucleotides), or encodes an inhibitory polypeptide (i.e., a polypeptide that inhibits expression or function of the target gene product), following its transcription (for example, in the case of an inhibitory sequence encoding a hairpin RNA, miRNA, or double-stranded RNA polynucleotide) or its transcription and translation (in the case of an inhibitory sequence encoding an inhibitory polypeptide), the transcribed or translated product, respectively, exerts the inhibitory effect on the target gene product (i.e., inhibits expression or function of the target gene product).

By "host cell" is meant a cell that comprises a heterologous nucleic acid sequence of the invention. Though the nucleic acid sequences of the invention, and fragments and variants thereof, can be introduced into any cell of interest, of particular interest are plant cells, more particularly cells of a Nicotiana plant species, for example, the tobacco plant species and varieties described herein below.

The use of the term "polynucleotide" is not intended to limit the present invention to polynucleotides comprising DNA. Those of ordinary skill in the art will recognize that polynucleotides can comprise ribonucleotides and combinations of ribonucleotides and deoxyribonucleotides. Such deoxyribonucleotides and ribonucleotides include both naturally occurring molecules and synthetic analogues. The polynucleotides of the invention also encompass all forms of sequences including, but not limited to, single-stranded forms, double-stranded forms, hairpins, stem-and-loop structures, and the like.

The term "variant" as used herein is intended to mean a substantially similar sequence, and the term "native" polynucleotide or polypeptide is intended to mean a naturally occurring nucleotide sequence or amino acid sequence, respectively. By "fragment" is intended a portion of a polynucleotide or a portion of an amino acid sequence and hence protein encoded thereby.

As used herein, the term "plant part" includes plant cells, plant protoplasts, plant cell tissue cultures from which a whole plant can be regenerated, plant calli, plant clumps, and plant cells that are intact in plants or parts of plants such as embryos, pollen, anthers, ovules, seeds, leaves, flowers, stems, branches, fruit, roots, root tips, and the like. Progeny, variants, and mutants of regenerated plants are also included within the scope of the invention, provided that these comprise the introduced nucleic acid sequences of the invention.

By "phenotypic change" is intended a measurable change in one or more cell functions. For example, plants having a genetic modification at the genomic locus encoding a cytochrome P450 polypeptide of the invention can show reduced or eliminated expression or activity of that cytochrome P450 polypeptide.

The term "introducing" is intended to mean presenting to the plant the polynucleotide or polypeptide in such a manner that the sequence gains access to the interior of a cell of the plant.

The term "openly linked" is intended to mean a functional linkage between two or more elements. For example, an operable linkage between a polynucleotide of interest and a regulatory sequence (i.e., a promoter) is a functional link that allows for expression of the polynucleotide of interest. Operably linked elements may be contiguous or non-contiguous. When used to refer to the joining of two protein coding regions, by openly linked is intended that the coding regions are in the same reading frame.

The term "heterologous" according to the present invention when used in reference to a sequence is intended to mean a sequence that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. The term also is applicable to nucleic acid constructs, also referred to herein as "polynucleotide constructs" or "nucleotide constructs." In this manner, a "heterologous" nucleic acid construct is intended to mean a construct that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. Heterologous nucleic acid constructs include, but are not limited to, recombinant nucleotide constructs that have been introduced into a plant or plant part thereof, for example, via transformation methods or subsequent breeding of a transgenic plant with another plant of interest.

For example, a promoter operably linked to a heterologous polynucleotide is from a species different from the species from which the polynucleotide was derived, or, if from the same/analogous species, one or both are substantially modified from their original form and/or genomic locus, or the promoter is not the native promoter for the operably linked polynucleotide. Furthermore, as used herein a chimeric gene comprises a coding sequence operably linked to a transcription initiation region that is heterologous to the coding sequence.

The term "expression" as used herein refers to the biosynthesis of a gene product, including the transcription and/or translation of said gene product. For example, for the purposes of the present invention, an expression cassette, as described elsewhere herein, capable of expressing a polynucleotide that inhibits the expression of at least one cytochrome P450 polypeptide of the invention is an expression cassette capable of producing an RNA molecule that inhibits the transcription and/or translation of at least one cytochrome P450 polypeptide. The "expression" or "production" of a protein or polypeptide from a DNA molecule refers to the transcription and translation of the coding sequence to produce the protein or polypeptide, while the "expression" or "production" of a protein or polypeptide from an RNA molecule refers to the translation of the RNA coding sequence to produce the protein or polypeptide.

Cytotoxic 450 Polynucleotides and Polypeptides, and Variants and Fragments Thereof

Compositions of the present invention include isolated cytochrome P450 polynucleotides and polypeptides that are involved in the metabolic conversion of nicotine to nornicotine in plants, including commercial varieties of tobacco.
US 7,884,263 B2

Inventors: Brian E. Burg, Rurik Lim, Mark K. Frerichs, Joseph S. Morari

Assignee: The Regents of the University of California

Filed: August 24, 2006

Inventors: Brian E. Burg, Rurik Lim, Mark K. Frerichs, Joseph S. Morari

Assignee: The Regents of the University of California

Filed: August 24, 2006

Claims:

1. A method of identifying a nicotine biosynthetic enzyme comprising:
   a. providing a nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotamide N-methyltransferase activity; and
   b. detecting a homology relationship between the nucleotide sequence and the nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity.

2. A method of identifying a nicotine biosynthetic enzyme comprising:
   a. providing a nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity; and
   b. detecting a homology relationship between the nucleotide sequence and the nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotamide N-methyltransferase activity.

3. A method of identifying a nicotine biosynthetic enzyme comprising:
   a. providing a nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity; and
   b. detecting a homology relationship between the nucleotide sequence and the nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity.

4. A method of identifying a nicotine biosynthetic enzyme comprising:
   a. providing a nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity; and
   b. detecting a homology relationship between the nucleotide sequence and the nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity.

5. A method of identifying a nicotine biosynthetic enzyme comprising:
   a. providing a nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity; and
   b. detecting a homology relationship between the nucleotide sequence and the nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity.

6. A method of identifying a nicotine biosynthetic enzyme comprising:
   a. providing a nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity; and
   b. detecting a homology relationship between the nucleotide sequence and the nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity.

7. A method of identifying a nicotine biosynthetic enzyme comprising:
   a. providing a nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity; and
   b. detecting a homology relationship between the nucleotide sequence and the nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity.

8. A method of identifying a nicotine biosynthetic enzyme comprising:
   a. providing a nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity; and
   b. detecting a homology relationship between the nucleotide sequence and the nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity.

9. A method of identifying a nicotine biosynthetic enzyme comprising:
   a. providing a nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity; and
   b. detecting a homology relationship between the nucleotide sequence and the nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity.

10. A method of identifying a nicotine biosynthetic enzyme comprising:
    a. providing a nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity; and
    b. detecting a homology relationship between the nucleotide sequence and the nucleotide sequence of a plant gene comprising an open reading frame that encodes a protein comprising a nicotine biosynthetic activity.
assembled within recombinant constructs for use in gene silencing with any method known in the art, including, but not limited to, sense suppression/cosuppression, antisense suppression, double-stranded RNA (dsRNA) interference, hairpin RNA interference and intron-containing hairpin RNA interference, silencers, trans-acting interfering ribozymes, and small interfering RNA or micro RNA, as described herein below. Thus, fragments of a cytochrome P450 polynucleotide sequence may range from at least about 20 nucleotides, about 50 nucleotides, about 70 nucleotides, about 100 nucleotides, about 150 nucleotides, about 200 nucleotides, about 250 nucleotides, about 300 nucleotides, about 350 nucleotides, about 400 nucleotides, and up to the full-length polynucleotide encoding the proteins of the invention, depending upon the desired outcome. In one aspect, the fragments of a cytochrome P450 polynucleotide sequence can be a fragment of both about 50 and about 400 nucleotides, between about 70 and about 350 nucleotides, between about 90 and about 325 nucleotides, between about 90 and about 300 nucleotides, between about 90 and about 275 nucleotides, between about 100 and about 400 nucleotides, between about 100 and about 350 nucleotides, between about 100 and about 325 nucleotides, between about 100 and about 300 nucleotides, between about 125 and about 300 nucleotides, or between about 125 and about 275 nucleotides in length. In some embodiments, a fragment of a cytochrome P450 polynucleotide is about 50, about 60, about 70, about 80, about 90, about 100, about 125, about 150, about 175, about 200, about 225, about 250, about 275, about 300, about 325, about 350, about 400 nucleotides in length, and other such values between about 70 and about 400 nucleotides. In one such embodiment, a fragment of a cytochrome P450 polynucleotide of the invention is about 90 bp to about 110 bp in length, including 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, and 110 bp in length. In another such embodiment, a fragment of a cytochrome P450 polynucleotide of the invention is about 290 to about 310 bp in length, including 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, and 310 bp in length.

A fragment of a cytochrome P450 polynucleotide of the present invention that encodes a biologically active portion of a cytochrome P450 polypeptide of the present invention will encode at least 15, 25, 30, 50, 75, 100, 125, 150, 175, 200, 250, 300, 400, 450, or 500 contiguous amino acids, or up to the total number of amino acids present in a full-length cytochrome P450 polypeptide of the invention (e.g., 517 for SEQ ID NO:2; 4, and 6; and 518 for SEQ ID NO:8), or will encode at least 15, 25, 30, 50, 75, 100, 125, 150, or up to the total number of amino acids present in a partial-length cytochrome P450 polypeptide of the invention (e.g., 173 for SEQ ID NO:10; and 222 for SEQ ID NO:12). In one aspect, a fragment of a cytochrome P450 polynucleotide of the invention encodes a polypeptide that comprises position 330 of the encoded polypeptide sequence. In another aspect, the polynucleotide fragment encodes a fragment of a cytochrome P450 polypeptide, where the polypeptide fragment comprises the amino acids from position 225 through the amino acid at about position 600 of SEQ ID NO:6. In one such embodiment, the polynucleotide fragment encodes a fragment of a cytochrome P450 polypeptide, where the polypeptide fragment comprises the amino acids from about position 239 through the amino acid at about position 402 of SEQ ID NO:6. A biologically active portion of a cytochrome P450 polypeptide can be prepared by isolating a portion of one of the cytochrome P450 polynucleotides of the present invention, expressing the encoded portion of the cytochrome P450 polypeptide (e.g., by recombinant expression in vitro), and assessing the activity of the encoded portion of the cytochrome P450 polypeptide, i.e., the ability to promote conversion of nicotine to nornicotine, using assays known in the art and those provided herein below.

Polynucleotides that are fragments of a cytochrome P450 nucleotide sequence of the present invention comprise at least 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, or 1700 contiguous nucleotides, or up to the number of nucleotides present in a full-length cytochrome P450 polynucleotide as disclosed herein (e.g., 539 for SEQ ID NO:9; 666 for SEQ ID NO:11; 1733 for SEQ ID NO:1; and 1727 for SEQ ID NO:7). Polynucleotides that are fragments of a cytochrome P450 nucleotide sequence of the present invention comprise fragments from about 20 to about 1700 contiguous nucleotides, from about 50 to about 1600 contiguous nucleotides, from about 75 to about 1500 contiguous nucleotides, from about 100 to about 1400 nucleotides, from about 150 to about 1300 contiguous nucleotides, from about 150 to about 1200 contiguous nucleotides, from about 175 to about 1100 contiguous nucleotides, or from about 200 to about 1000 contiguous nucleotides from a cytochrome P450 polynucleotide as disclosed herein.

In one aspect, fragments of a cytochrome P450 polynucleotide comprise a polynucleotide sequence containing the nucleotides from about position 700 to about position 1250 of a cytochrome P450 coding sequence. In another aspect, fragments of a cytochrome P450 polynucleotide comprise a polynucleotide sequence containing the nucleotides from about position 715 to about position 1210, or from about position 717 to about position 1207 of a cytochrome P450 coding sequence disclosed herein.

In other embodiments of the invention, fragments of a cytochrome P450 polynucleotide comprise a polynucleotide sequence containing the nucleotides from about position 265 to about position 625 of a cytochrome P450 coding sequence disclosed herein, or a complement thereof. In some of these embodiments, the fragments of a cytochrome P450 coding sequence disclosed herein comprise the nucleotides corresponding to about position 297 to about position 594 of the P450 coding sequence set forth in SEQ ID NO:3 or SEQ ID NO:5, or a complement thereof.

In yet other embodiments of the invention, fragments of a cytochrome P450 polynucleotide comprise a polynucleotide sequence containing the nucleotides from about position 1420 to about position 1580 of a cytochrome P450 coding sequence disclosed herein, or a complement thereof. In some of these embodiments, the fragments of a cytochrome P450 coding sequence disclosed herein comprise the nucleotides corresponding to about position 1453 to about position 1551 of the P450 coding sequence set forth in SEQ ID NO:1, or a complement thereof.

Variants of the disclosed polynucleotides and polypeptides encoded thereby are also encompassed by the present invention. Such naturally occurring variants include those variants that share substantial sequence identity to the disclosed cytochrome P450 polynucleotides and polypeptides disclosed herein as defined herein below. The compositions and methods of the invention can be used to target expression or function of any naturally occurring cytochrome P450 that shares substantial sequence identity to the disclosed cytochrome P450 polypeptides and which possesses the relevant cytochrome P450 activity, i.e., involvement in the metabolic conversion of nicotine to nornicotine in plants. Such variants may result from, for example, genetic polymorphism or from
human manipulation as occurs with breeding and selection. Biologically active variants of a cytochrome P450 protein of the invention, for example, variants of the polypeptide set forth in SEQ ID NO:2, 4, 6, 8, 10, or 12, will have at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to the amino acid sequence for the native protein as determined by sequence alignment programs and parameters described elsewhere herein. Where any given pair of polynucleotides of the invention is evaluated by comparison of the percent sequence identity shared by the two polynucleotides they encode, the percent sequence identity between the two encoded polypeptides is at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity. In one aspect, a variant polypeptide of the present invention includes a polypeptide having a tryptophan at position 330 or an isoleucine at position 274 of the cytochrome P450 polypeptide, or both a tryptophan at position 330 and an isoleucine at position 274.

Furthermore, the polynucleotides of the invention can be used to isolate corresponding cytochrome P450 sequences from other organisms, particularly other plants, more particularly other members of the Nicotiana genus. PCR, hybridization, and other like methods can be used to identify such sequences based on their sequence homology to the sequences set forth herein. Sequences isolated based on their sequence identity to the polynucleotide sequences set forth herein, or to variants and fragments thereof are encompassed by the present invention. Such sequences include sequences that are orthologs of the disclosed sequences.

According to the present invention, “orthologs” are genes derived from a common ancestral gene and which are found in different species as a result of speciation. Genes found in different species are considered orthologs when their nucleotide sequences and/or their encoded protein sequences share at least 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or greater sequence identity. Functions of orthologs are often highly conserved among species. Thus, isolated polynucleotides that encode for a cytochrome P450 polypeptide that is involved in the nicotine-to-nornicotine metabolic conversion and which hybridize under stringent conditions to the cytochrome P450 sequences disclosed herein, or to variants or fragments thereof, are encompassed by the present invention. Such sequences can be used in the methods of the present invention to inhibit expression of cytochrome P450 polypeptides that are involved in the metabolic conversion of nicotine to nornicotine in plants.


Hybridization techniques involve the use of all or part of a known polynucleotide as a probe that selectively hybridizes to other corresponding polynucleotides present in a population of cloned genomic DNA fragments or cDNA fragments (i.e., genomic or cDNA libraries) from a chosen organism. Hybridization may be carried out under stringent conditions. By “stringent conditions” or “stringent hybridization conditions” is intended conditions under which a probe will hybridize to its target sequence to a detectably greater degree than to other sequences (e.g., at least 2-fold over back-
Stringent conditions are sequence-dependent and will be different in different circumstances. By controlling the stringency of the hybridization and/or washing conditions, target sequences that are 100% complementary to the probe can be identified (homologous probing). Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of similarity are detected (heterologous probing). Generally, a probe is less than about 1000 nucleotides in length, optimally less than 500 nucleotides in length.

Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.1 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C. Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulfate) at 37°C, and a wash in 1× to 2×SSC (20×SSC=3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1.0 M NaCl, 1% SDS at 37°C, and a wash in 0.5× to 1×SSC at 55 to 60°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1× SSC at 60 to 65°C. Optionally, wash buffers may comprise about 0.1% to about 1% SDS. Duration of hybridization is generally less than about 24 hours, usually about 4 to 12 hours. The duration of the wash time will be at least a length of time sufficient to reach equilibrium.

In a specific embodiment, stringency conditions include hybridization in a solution containing 5×SSC, 0.5% SDS, 5×Denhardt’s, 0.45 µg/µl Poly A RNA, 0.45 µg/µl calf thymus DNA and 50% formamide at 42°C and at least one post-hybridization wash in a solution comprising from about 0.01× to about 1× SSC to about 1×SSC. The duration of hybridization is from about 14 to about 16 hours. Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the T_m can be approximated from the equation of Meinkoth and Wahl (1984) Anal. Biochem. 138:267-284: T_m=81.5°C +16.6 (log M)+0.41 (%GC)-0.61 (% form)-500/L; where M is the molality of monovalent cations, % GC is the percentage of guanosine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The T_m is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe. T_m is reduced by about 1°C for each 1% of mismatching; thus, T_m hybridization, and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with ≥90% identity are sought, the T_m can be decreased 11°C. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4°C lower than the thermal melting point (T_m); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 110°C lower than the thermal melting point (T_m); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20°C lower than the thermal melting point (T_m). Using the equation, hybridization and wash compositions, and desired T_m, those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a T_m of less than 45°C (aqueous solution) or 32°C (formamide solution), it is optimal to increase the SSC concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes, Part I, Chapter 2 (Elsevier, N.Y.); and Ausubel et al., eds. (1995) Current Protocols in Molecular Biology, Chapter 2 (Greene Publishing and Wiley-Interscience, New York). See Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y.).

Hybridization probes may be genomic DNA fragments, cDNA fragments, RNA fragments, or other oligonucleotides, and may be labeled with a detectable group such as 32P, or any other detectable marker. For example, probes for hybridization can be made by labeling synthetic oligonucleotides based on the cytochrome P450 polynucleotides sequences of the present invention. Methods for preparation of probes for hybridization and for construction of cDNA and genomic libraries are generally known in the art and are disclosed in Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y.).

For example, the cytochrome P450 polynucleotide sequences disclosed herein, or one or more portions thereof, may be used as probes capable of specifically hybridizing to corresponding cytochrome P450 polynucleotides and messenger RNAs. To achieve specific hybridization under a variety of conditions, such probes include sequences that are unique among cytochrome P450 polynucleotide sequences, including upstream regions 5' to the coding sequence and downstream regions 3' to the coding sequence, and are optimally at least about 20 nucleotides in length, and most optimally at least about 20 nucleotides in length. Such probes may be used to amplify corresponding cytochrome P450 polynucleotides. This technique may be used to isolate additional coding sequences from a desired plant or as a diagnostic assay to determine the presence of coding sequences in a plant. Hybridization techniques include hybridization screening of plated DNA libraries (either plaques or colonies; see, for example, Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y.).

As used herein, with respect to the sequence relationships between two or more polynucleotides or polypeptides, the term “reference sequence” is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset or the entirety of a specified sequence; for example, as a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence. As used herein, the term “comparison window” makes reference to a contiguous and specified segment of a polynucleotide sequence, wherein the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two polynucleotides. Generally, the comparison window is at least 20 contiguous nucleotides in length, and optionally can be 30, 40, 50, 100, or longer. Those of skill in the art understand that to avoid a high similarity to a reference sequence due to inclusion of gaps in the polynucleotide sequence a gap penalty is typically introduced and is subtracted from the number of matches.

Methods of alignment of sequences for comparison are well known in the art. Thus, the determination of percent

The BLAST programs of Altschul et al. (1990) J. Mol. Biol. 215:403 are based on the algorithm of Karlin and Altschul (1990) supra. BLAST nucleotide searches can be performed with the BLASTN program, score=100, wordlength=12, to obtain nucleotide sequences homologous to a nucleotide sequence encoding a protein of the invention. BLAST protein searches can be performed with the BLASTX program, score=50, wordlength=3, to obtain amino acid sequences homologous to a protein of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST (in BLAST 2.0) can be utilized as described in Altschul et al. (1997) Nucleic Acids Res. 25:3389. Alternatively, PSI-BLAST (in BLAST 2.0) can be used to perform an iterated search that detects distant relationships between molecules. See Altschul et al. (1997) supra. When utilizing BLAST, Gapped BLAST, PSI-BLAST, the default parameters of the respective programs (e.g., BLASTN for nucleotide sequences, BLASTX for proteins) can be used (See www.ncbi.nlm.nih.gov). Alignment may also be performed manually by inspection.

The sequence identity/similarity values provided herein were calculated using the BLASTX (Altschul et al. (1997) supra), Clustal W (Higgins et al. (1994) Nucleic Acids Res. 22: 4673-4680), and GAP (University of Wisconsin Genetic Computing Group software package) algorithms using default parameters. The present invention also encompasses the use of any equivalent program thereof for the analysis and comparison of nucleic acid and protein sequences. By “equivalent program” is intended any sequence comparison program that, for any two sequences in question, generates an alignment having identical nucleotide or amino acid residue matches and an identical percent sequence identity when compared to the corresponding alignment generated by BLASTX, Clustal W, or GAP.

For purposes of the foregoing discussion of variant nucleotide and polypeptide sequences encompassed by the present invention, the term “sequence identity” or “identity” in the context of two polynucleotides or polypeptide sequences makes reference to the residues in the two sequences that are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. When sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences that differ by such conservative substitutions are said to have “sequence similarity” or “similarity.” Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, Calif.).

The term “percentage of sequence identity” as used herein means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison, and multiplying the result by 100 to yield the percentage of sequence identity.

When any two polypeptide sequences are optimally aligned for comparison, it is recognized that residues appearing opposite of one another within the alignment occupy positions within their respective polypeptides that correspond to one another. Such positions are referred to herein as “corresponding positions” and the residues residing at corresponding positions are referred to as “corresponding residues” or residues that “correspond” to one another. Thus, for example, where a polypeptide of interest is optimally aligned to a reference polypeptide sequence having, for example, 10 residues, the residue within the polypeptide of interest appearing opposite residue 5 of the reference sequence is referred to as the “residue at the position corresponding to residue 5” of the reference sequence.

In like manner, when any two polynucleotide sequences are optimally aligned for comparison, it is recognized that the nucleotides appearing opposite of one another within the alignment occupy positions within their respective polynucleotide positions that correspond to one another. Such positions are referred to herein as “corresponding positions” and the nucleotides residing at corresponding positions are referred to as “corresponding nucleotides” or nucleotides that “correspond” to one another. Thus, for example, where a polynucleotide of interest is optimally aligned to a reference polynucleotide sequence having, for example, 300 nucleotides, the nucleotide within the polynucleotide of interest appearing opposite nucleotide 275 of the reference sequence is referred to as the “nucleotide at the position corresponding to nucleotide 275” of the reference sequence.

Where a region of nucleotides is being compared between a polynucleotide of interest and a reference polynucleotide, the nucleotides within these regions are said to “correspond” to one another. Thus, for example, where a region of a reference polynucleotide sequence, for example, the polynucleotide sequence set forth in SEQ ID NO:5, resides from nucleotide position 265 to nucleotide position 625 of the reference polynucleotide, and this region of nucleotides is being compared to the corresponding region of nucleotides within an optimally aligned polynucleotide sequence of interest, the nucleotides within the corresponding region of the polynucleotide of interest are referred to herein as “a region” of the polynucleotide of interest that “corresponds to nucleotide position 265 to nucleotide position 625” of the reference sequence, in this case, SEQ ID NO:5.

Cytochrome P450 polynucleotide and polypeptide sequences can be identified using the sequences provided.
herein. Such methods include obtaining a polynucleotide or polypeptide sequence at least 80%, 85%, 90%, 95%, 98%, 99% sequence identity with the polynucleotide sequence of SEQ ID NO:1, 3, 5, 7, 9, or 11 or a complement or fragment thereof, or a polypeptide sequence of SEQ ID NO:2, 4, 6, 7, 10, or 12. In one embodiment, the identified sequence contains or encodes for a tryptophan residue at position 330 and an isoleucine residue at position 274.

In this manner, one aspect of the present invention is directed to a method of screening for a nicotine demethylase sequence. This method comprises obtaining a nucleic acid sequence that has greater than about 90% sequence identity with SEQ ID NO:3 or SEQ ID NO:5; and identifying within this nucleic acid sequence a codon sequence encoding for a tryptophan residue at position 330 of the polypeptide encoded by the nucleic acid sequence. In some embodiments, this nucleic acid sequence also encodes an isoleucine at position 274 of the encoded polypeptide. Any suitable method known in the art can be used to identify the codon sequence encoding the tryptophan or isoleucine residue. In one embodiment, the codon sequence is identified by a method selected from the group consisting of identifying a single nucleotide polymorphism and RT-PCR. In some embodiments of this aspect of the invention, the screening method identifies a nicotine demethylase that converts nicotine to nornicotine at a rate that is at least about 5-fold greater than the conversion rate of the nicotine demethylase encoded by SEQ ID NO:3 (polypeptide set forth in SEQ ID NO:4) or SEQ ID NO:5 (polypeptide set forth in SEQ ID NO:6). In particular embodiments, the nicotine demethylase identified with this screening method has a conversion rate that is at least about 8-fold greater than the conversion rate of the nicotine demethylase encoded by SEQ ID NO:3 or SEQ ID NO:5.

In another aspect, the present invention provides a method for screening for a nicotine demethylase having an isoleucine at position 274 or a tryptophan at position 329 of the polypeptide. This method comprises obtaining a nucleic acid sequence that has greater than about 90% sequence identity with SEQ ID NO:3 or SEQ ID NO:5; and identifying a first codon sequence encoding for an isoleucine residue at position 274 of the encoded polypeptide or a second codon sequence encoding a tryptophan residue at position 330 of the encoded polypeptide. As noted above, any suitable method known to those of skill in the art can be used to identify these codons. In one embodiment, either or both of the first and second codons are identified by a method selected from the group consisting of identifying a single nucleotide polymorphism and RT-PCR. In some embodiments of this aspect of the invention, the screening method identifies a nicotine demethylase that converts nicotine to nornicotine at a rate that is at least about 5-fold greater than the conversion rate of the nicotine demethylase encoded by SEQ ID NO:3 (polypeptide set forth in SEQ ID NO:4) or SEQ ID NO:5 (polypeptide set forth in SEQ ID NO:6). In particular embodiments, the nicotine demethylase identified with this screening method has a conversion rate that is at least about 8-fold greater than the conversion rate of the nicotine demethylase encoded by SEQ ID NO:3 or SEQ ID NO:5.

The present invention also provides transgenic plant cells, plants, and seed comprising a nucleic acid molecule having a promoter functional in a plant cell and a nucleic acid sequence encoding a nicotine demethylase having an isoleucine residue at position 274 and a tryptophan residue at position 330. In some embodiments, the nucleic acid sequence encoding this nicotine demethylase is derived from a sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:5, and SEQ ID NO:7. These transgenic plant cells, plants, and seed include, but are not limited to, Nicotiana plant cells, Nicotiana plants, and seed of Nicotiana plants. In some embodiments, the Nicotiana plant cells, Nicotiana plants, and seed of Nicotiana plants are from converter Nicotiana plants.

Expression Cassettes for Use in the Methods of the Invention

Compositions of the present invention further include expression cassettes comprising inhibitory sequences capable of inhibiting expression or function of a cytochrome P450 polypeptide involved in the conversion of nicotine to nornicotine in a Nicotiana plant or plant part thereof, where the inhibitory sequences are operably linked to a promoter that is functional in a plant cell. In this manner, expression cassettes comprising all or part of the sequence set forth in SEQ ID NO:1, 3, 5, 7, 9, or 11, a complement or fragment thereof, or sequences sharing substantial sequence identity to SEQ ID NO:1, 3, 5, 7, 9, 11, or a complement or fragment thereof, operably linked to a promoter that is functional in a plant cell are constructed for use in the gene-silencing methods of the present invention described herein below. Such sequences are referred to herein as "inhibitory sequences" or "inhibitory polynucleotide sequences" as they are capable of being expressed as an RNA molecule that inhibits expression (i.e., transcription and/or translation) of the target cytochrome P450 polypeptide, for example, the polypeptide set forth in SEQ ID NO:2, 4, 6, or 8 and variants thereof, or a polypeptide comprising the sequence set forth in SEQ ID NO:10 or 12 and variants thereof, where the variant polypeptides have substantial sequence identity to these disclosed cytochrome P450 polypeptides and are involved in the metabolic conversion of nicotine to nornicotine in a plant.

As noted above, such inhibitory sequences include fragment sequences of the target cytochrome P450 polypeptides. For example, a fragment sequence can include any portion of the cytochrome P450 sequence, including coding and non-coding sequence (e.g., 5' UTR, intron, and 3' UTR sequences), and can include fragments of between about 20 and about 400 nucleotides, between about 50 and about 400 nucleotides, between about 100 and about 400 nucleotides, between about 125 and about 325 nucleotides, between about 125 and about 300 nucleotides, or between about 125 and about 275 nucleotides.

In this manner, such inhibitory sequences include, but are not limited, sequences that comprise a fragment of a cytochrome P450 polynucleotide sequence ranging from at least about 20 nucleotides, about 50 nucleotides, about 70 nucleotides, about 100 nucleotides, about 150 nucleotides, about 200 nucleotides, about 250 nucleotides, about 300 nucleotides, about 350 nucleotides, and up to the full-length polynucleotide encoding the proteins of the invention, depending upon the desired outcome. In one aspect, the inhibitory sequences comprise a fragment of a cytochrome P450 polynucleotide sequence that is between about 50 and about 400 nucleotides, between about 50 and about 350 nucleotides, between about 70 and about 350 nucleotides, between about 90 and about 325 nucleotides, between about 90 and about 300 nucleotides, between about 90 and about 275 nucleotides, between about 100 and about 400 nucleotides, between about 100 and about 350 nucleotides, between about 100 and about 325 nucleotides, between about 100 and about 300 nucleotides, between about 125 and about 300 nucleotides, or between about 125 and about 275 nucleotides in length. In some embodiments, the inhibitory sequences comprise a fragment of a cytochrome P450 polynucleotide sequence that is about 50, about 60, about 70, about 80, about 90, about 100, about 125, about 150, about 175, about 200, about 225, about 250, about 275, about 300, about 325, about
350, about 400 nucleotides in length, and other such values
between about 50 and about 400 nucleotides. It is recognized
that the inhibitory sequence can also comprise a sequence
that is complementary to all or a part of the fragment of the
cytochrome P450 polynucleotide sequence.
In one such embodiment, the inhibitory sequences com-
promise a fragment of a cytochrome P450 polynucleotide of
the invention that is about 90 bp to about 110 bp in length,
including 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102,
103, 104, 105, 107, 108, 109, and 110 bp in length, and
can also comprise a sequence that is complementary to all or
a part of the fragment sequence. In another such embodiment,
a fragment of a cytochrome P450 polynucleotide of the inven-
tion is about 290 to about 310 bp in length, including 290,
291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302,
303, 304, 305, 306, 307, 308, 309, and 310 bp in length, and
can also comprise a sequence that is complementary to all or
a part of the fragment sequence.
In other embodiments of the invention, the inhibitory
sequence within an expression cassette of the invention com-
promises a polynucleotide sequence containing the nucleo-
tides from about position 265 to about position 625 of a cyto-
chrome P450 coding sequence disclosed herein and sequence
that is fully or partially complementary thereto. In some of
these embodiments, the inhibitory sequence comprises the
nucleotides corresponding to about position 297 to about
position 594 of the P450 coding sequence set forth in SEQ ID
NO: 3 or SEQ ID NO: 5 and a sequence that is fully or partially
complementary thereto. In preferred embodiments, such inhibitory sequences are expressed as a hairpin RNA as
described herein below.
In yet other embodiments of the invention, the inhibitory
sequence within an expression cassette of the invention com-
promises a polynucleotide sequence containing the nucleo-
tides from about position 1420 to about position 1580 of a cyto-
chrome P450 coding sequence disclosed herein and a sequence
that is fully or partially complementary thereto. In some of
these embodiments, the inhibitory sequence comprises the
nucleotides corresponding to about position 1453 to about
position 1551 of the P450 coding sequence set forth in SEQ ID
NO: 1 and a sequence that is fully or partially complementary thereto. In preferred embodiments, such inhibitory sequences are expressed as a hairpin RNA as
described herein below.
It is recognized that expression cassettes of the present
invention encompass constructs in which a desired nucleic
acid sequence is operably linked to a promoter that is func-
tional in a plant cell, particularly in the cell of a Nicotiana
plant. By "promoter" is intended a regulatory region of DNA
usually comprising a TATA box capable of directing RNA
polymerase II to initiate RNA synthesis at the appropriate
transcription initiation site for a particular polynucleotide
sequence. A promoter may additionally comprise other rec-
ognition sequences generally positioned upstream or 5' to the
TATA box, referred to as upstream promoter elements, which
influence the transcription initiation rate. It is also recognized
that expression cassettes of the present invention encompass
additional domains that modulate the level of expression, the
developmental timing of expression, or tissue type that
expression occurs in (e.g., Australian Patent No. AU-A-
77751/94 and U.S. Pat. Nos. 5,466,785 and 5,635,618). By "func-
tional" is intended the promoter, when operably linked to
an inhibitory sequence encoding an inhibitory nucleotide
molecule (for example, a hairpin RNA, double-stranded
RNA, miRNA polynucleotide, and the like), the promoter is
capable of initiating transcription of the operably linked
inhibitory sequence such that the inhibitory nucleotide mol-
cule is expressed. The promoters can be selected based on
the desired outcome. The nucleic acids can be combined with
constitutive, tissue-preferred, or other promoters for expres-
sion in plants.
An expression cassette of the present invention may also
contain at least one additional gene to be cotransformed into
the plant. Alternatively, the additional gene(s) can be pro-
vided on multiple expression cassettes. Such an expression
cassette is provided with a plurality of restriction sites and/or
recombination sites for insertion of the cytochrome P450
inhibitory polynucleotide sequence to be under the transcrip-
tional regulation of the regulatory regions. The expression
cassette may additionally contain selectable marker genes.
In this manner, an expression cassette of the present inven-
tion includes a transcriptional and translational initiation
region (i.e., a promoter) in the 5'-3' direction of transcription,
an inhibitory sequence as described elsewhere herein, and a
transcriptional and translational termination region (i.e., termi-
nation region) functional in a plant cell. The regulatory
regions (i.e., promoters, transcriptional regulatory regions,
and translational termination regions) and/or the inhibitory
sequence of the invention may be native/analogous to the host
cell or to each other. Alternatively, the regulatory regions
and/or the inhibitory sequence of the invention may be het-
erologous to the host cell or to each other. While heterologous
promoters can be used to express the inhibitory sequences of
the invention, native promoter sequences may also be used.
The termination region may be native with the transcrip-
tional initiation region, may be native with the operably
linked inhibitory sequence of the invention, may be native
with the plant host, or may be derived from another source
(i.e., foreign or heterologous to the promoter, the inhibitory
sequence of interest, the plant host, or any combination thereof).
Convenient termination regions are available from the Ti-
plasmid of A. tumefaciens, such as the octopine synthase
and nopaline synthase termination regions. See also
Expression cassettes of the present invention may addition-
ally contain 5' leader sequences that can act to enhance transla-
tion. Translation leaders are known in the art and include:
picornavirus leaders, for example, EMCV leader (Eunomophy-
carcoidis 5' noncoding region) (Elroy-Stein et al. (1989)
Proc. Natl. Acad. Sci. USA 86:6126-6130); potyvirus leaders,
for example, TEV leader (Tomato Echte Virus) (Gallie et al.
Mosaic Virus) (Virology 154:9-20), and human immunoglo-
bulin heavy-chain binding protein (BiP) (Macejak et al.
(1991) Nature 353:90-94); untranslated leader from the coat
protein mRNA of alfalfa mosaic virus (AVM RNA 4)
(Witherington et al. (1997) Nature 325:622-625); tobacco mosaic
virus leader (TMV) (Gallie et al. (1989) in Molecular Biology of
RNA, ed. Cech (Liss, New York), pp. 237-256); and maize
chlorotic mottle virus leader (MCMV) (Lommel et al. (1991)
Virology 81:382-385). See also, Della-Cioppa et al. (1987)
Plant Physiol. 84:965-968. Other methods known to enhance
translation can also be utilized.
In preparing the expression cassette, DNA fragments of
the invention may be manipulated so as to provide for the DNA
sequences in the proper orientation and, as appropriate, in
the proper reading frame. Adapters or linkers may be employed
to join the DNA fragments or other manipulations may be
involved to provide for convenient restriction sites, removal
of superfluous DNA, removal of restriction sites, or the like. For this purpose, in vitro mutagenesis, primer repair, restriction, annealing, resubstitutions, e.g., transitions and transversions, may be involved.


A number of promoters can be used in the practice of the invention. Of particular interest are constitutive promoters, inducible promoters, particularly chemical-inducible promoters, and tissue-preferred promoters, particularly leaf-preferred promoters.

Chemical-inducible promoters can be used to inhibit the expression of a cytochrome P450 that is involved in the metabolic conversion of nicotine to nornicotine in a plant through the application of an exogenous chemical regulator. Chemical-inducible promoters are known in the art and include, but are not limited to, the tobacco PR-1a promoter, which is activated by salicylic acid. Other chemical-inducible promoters of interest include steroid-responsive promoters (see, for example, the glucocorticoid-inducible promoter in Schena et al. (1991) Proc. Natl. Acad. Sci. USA 88:10421-10425 and McNellis et al. (1998) Plant J. 14(2):247-257) and tetracycline-inducible promoters (see, for example, Gatz et al. (1991) Mol. Gen. Genet. 227:229-237, and U.S. Pat. Nos. 5,814,618 and 5,789,156, herein incorporated by reference.

Constitutive promoters include, for example, the core promoter of the Rsy7 promoter and other constitutive promoters disclosed in U.S. Pat. No. 6,072,050; the core CaMV 35S promoter (Odell et al. (1985) Nature 313:810-812); ubiquitin (Christensen et al. (1989) Plant Mol. Biol. 12:619-632 and 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); ALS promoter (U.S. Pat. Nos. 5,659,026, and the like. Other constitutive promoters include, for example, U.S. Pat. Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; and 6,177,611.


Methods for Inhibiting Expression or a Function of a Cytochrome P450 Involved in the Conversion of Nicotine to Nor-nicotine

Methods of reducing the concentration, content, and/or activity of a cytochrome P450 polyepitope of the present invention in a Nicotiana plant or plant part, particularly the leaf tissue, are provided. Many methods may be used, alone or in combination, to reduce or eliminate the activity of a cytochrome P450 polyepitope of the present invention. In addition, combinations of methods may be employed to reduce or eliminate the activity of two or more different cytochromes P450 polyepitopes.

In accordance with the present invention, the expression of a cytochrome P450 polyepitope of the present invention is
inhibited if the protein level of the cytochrome P450 polypeptide is statistically lower than the protein level of the same cytochrome P450 polypeptide in a plant that has not been genetically modified or mutagenized to inhibit the expression of that cytochrome P450 polypeptide, and where these plants have been cultured and harvested using the same protocols. In particular embodiments of the invention, the protein level of the cytochrome P450 polypeptide in a modified plant according to the invention is less than 95%, less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, less than 4%, less than 3%, less than 2%, or less than 1% of the protein level of the same cytochrome P450 polypeptide in a plant that is not a mutant or that has not been genetically modified to inhibit the expression of that cytochrome P450 polypeptide and which has been cultured and harvested using the same protocols. The expression level of the cytochrome P450 polypeptide may be measured directly, for example, by assaying for the level of the cytochrome P450 transcript or cytochrome P450 polypeptide expressed in the Nicotiana plant or plant part, or indirectly, for example, by measuring the conversion of nicotine to normicotine in the Nicotiana plant or plant part. Methods for monitoring the expression level of a protein are known in the art, and include, but are not limited to, Northern blot analysis as discussed in the examples herein below. Methods for determining the activity of the targeted cytochrome P450 polypeptide in converting nicotine to normicotine are described elsewhere herein below, and include, but are not limited to, alkaloid analysis using gas chromatography, for example the procedures described in the examples herein below.

In other embodiments of the invention, the activity of one or more cytochrome P450 polypeptides is reduced or eliminated by transforming a plant or plant part with an expression cassette comprising a polynucleotide encoding a polypeptide that inhibits the activity of one or more cytochrome P450 polypeptides of the present invention. The activity of a cytochrome P450 polypeptide in converting nicotine to normicotine in a Nicotiana plant or plant part is inhibited according to the present invention if this conversion activity is statistically lower than conversion activity of the same cytochrome P450 polypeptide in a Nicotiana plant or plant part that has not been genetically modified to inhibit the conversion activity of that cytochrome P450 polypeptide and which has been cultured and harvested using the same protocols. In particular embodiments, activity of a cytochrome P450 polypeptide in converting nicotine to normicotine in a modified Nicotiana plant or plant part according to the invention is inhibited if the activity is less than 95%, less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, less than 4%, less than 3%, less than 2%, or less than 1% of the conversion activity of the same cytochrome P450 polypeptide in a Nicotiana plant that has not been genetically modified to inhibit the expression of that cytochrome P450 polypeptide and which has been cultured and harvested using the same protocols. The activity of a cytochrome P450 polypeptide in converting nicotine to normicotine in a Nicotiana plant or plant part is determined by the assay methods described elsewhere herein. Methods of determining the activity of a cytochrome P450 polypeptide in converting nicotine to normicotine in a Nicotiana plant or plant part are described elsewhere herein, and include the alkaloid analyses using gas chromatography disclosed in the examples herein below.

In specific embodiments, a cytochrome P450 inhibitory polynucleotide sequence described herein is introduced into a Nicotiana plant or plant part. Subsequently, a Nicotiana plant or plant part having the introduced inhibitory polynucleotide sequence of the invention is selected using methods known to those of skill in the art such as, but not limited to, Southern blot analysis, DNA sequencing, PCR analysis, or phenotypic analysis. A plant or plant part altered or modified by the foregoing embodiments is grown under plant forming conditions for a time sufficient to modulate the concentration and/or activity of polypeptides of the present invention in the plant. Plant forming conditions are well known in the art and discussed briefly elsewhere herein.

In some embodiments, a transformed tobacco plant containing a cytochrome P450 inhibitory polynucleotide sequence described herein has a reduced level of conversion of nicotine to normicotine. In particular embodiments, conversion of nicotine to normicotine in a transformed tobacco plant or plant part according to the invention is less than 95%, less than 90%, less than 80% less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, less than 4%, less than 3%, less than 2%, or less than 1% of the conversion of a tobacco plant that has not been genetically modified to inhibit the expression of that cytochrome P450 polypeptide and which has been cultured and harvested using the same protocols. In some embodiments, the transformed tobacco plant is a converter tobacco plant. In some embodiments, the transformed tobacco plant has a conversion rate lower than the rate observed in commercial nonconverter tobacco plants.

It is also recognized that the level and/or activity of the polypeptide may be modulated by employing a polynucleotide that is not capable of directing, in a transformed plant, the expression of a protein or an RNA. For example, the polynucleotides of the invention may be used to design polynucleotide constructs that can be employed in methods for altering or mutating a genomic nucleotide sequence in an organism. Such polynucleotide constructs include, but are not limited to, RNA:DNA vectors, RNA:DNA mutational vectors, RNA:DNA repair vectors, mixed-duplex oligonucleotides, self-complementary RNA:DNA oligonucleotides, and recombinogenic oligonucleobases. Such nucleotide constructs and methods of use are known in the art. See, U.S. Pat. Nos. 5,565,350; 5,731,181; 5,756,325; 5,760,012; 5,795,972; and 5,871,984; all of which are herein incorporated by reference. See also, WO 98/49350, WO 99/07865, WO 99/25821, and Beetham et al. (1999) Proc. Natl. Acad. Sci. USA 96:8774-8778; herein incorporated by reference.

It is therefore recognized that methods of the present invention do not depend on the incorporation of the entire cytochrome P450 inhibitory polynucleotide into the genome, only that the Nicotiana plant or plant part thereof is altered as a result of the introduction of this inhibitory polynucleotide into a cell. In one embodiment of the invention, the genome may be altered following the introduction of the cytochrome P450 inhibitory polynucleotide into a cell. For example, the inhibitory polynucleotide, or any part thereof, may incorporate into the genome of the plant. Alterations to the genome include, but are not limited to, additions, deletions, and substitutions of nucleotides into the genome. While the methods of the present invention do not depend on additions, deletions, and substitutions of any particular number of nucleotides, it is recognized that such additions, deletions, or substitutions comprise at least one nucleotide.

It is further recognized that reducing the level and/or activity of a cytochrome P450 sequence of the present invention can be performed to elicit the effects of the sequence only during certain developmental stages and to switch the effect off in other stages where expression is no longer desirable.
Control of cytochrome P450 expression can be obtained via the use of inducible or tissue-preferred promoters. Alternatively, the gene could be inverted or deleted using site-specific recombinases, transposes or recombination systems, which would also turn on or off expression of the cytochrome P450 sequence.

According to the present invention, changes in levels, ratios, activity, or distribution of cytochrome P450 polypeptides of the present invention, or changes in Nicotiana plant or plant part phenotype, particularly reduced accumulation of nornicotine and its carcinogenic metabolite, NNN, could be measured by comparing a subject plant or plant part to a control plant or plant part, where the subject plant or plant part and the control plant or plant part have been cultured and/or harvested using the same protocols. As used herein, a subject plant or plant part is one in which genetic alteration, such as transformation, has been affected as to the cytochrome P450 polypeptide of interest, or is a Nicotiana plant or plant part that is descended from a Nicotiana plant or plant part so altered and which comprises the alteration. A control plant or plant part provides a reference point for measuring changes in phenotype of the subject plant or plant part.

The measurement of changes in phenotype can be measured at any time in a plant or plant part, including during plant development, senescence, or after curing. In other embodiments, the measurement of changes in phenotype can be measured in plants grown under any conditions, including from plants grown in a growth chamber, greenhouse, or in a field. In one embodiment, changes in phenotype can be measured by determining the nicotine to nornicotine conversion rate. In a preferred embodiment, conversion can be measured by dividing the percentage of nornicotine (as a percentage of the total tissue weight) by the sum of the percentage nicotine and nornicotine (as percentages of the total tissue weight) and multiplying by 100.

According to the present invention, a control plant or plant part may comprise a wild-type Nicotiana plant or plant part, i.e., of the same genotype as the starting material for the genetic alteration that resulted in the subject plant or plant part. A control plant or plant part may also comprise a Nicotiana plant or plant part of the same genotype as the starting material but that has been transformed with a null construct (i.e., with a construct that has no known effect on the trait of interest, such as a construct comprising a selectable marker gene). Alternatively, a control plant or plant part may comprise a Nicotiana plant or plant part that is a non-transformed segregant among progeny of a subject plant or plant part, or a Nicotiana plant or plant part genetically identical to the subject plant or plant part but that is not exposed to conditions or stimuli that would induce suppression of the cytochrome P450 gene of interest. Finally, a control plant or plant part may comprise the subject plant or plant part itself under conditions in which the cytochrome P450 inhibitory sequence is not expressed. In all such cases, the subject plant or plant part and the control plant or plant part are cultured and harvested using the same protocols.

As described elsewhere herein, methods are provided to reduce or eliminate the activity and/or concentration of a cytochrome P450 polypeptide of the present invention by introducing into a Nicotiana plant or plant part a cytochrome P450 inhibitory nucleotide sequence that is capable of inhibiting expression or function of a cytochrome P450 polypeptide that is involved in the metabolic conversion of nicotine to nornicotine. In some embodiments, the inhibitory sequence is introduced by transformation of the plant or plant part, such as a plant cell, with an expression cassette that expresses a nucleotide that inhibits the expression of the cytochrome P450 polypeptide. The polynucleotide may inhibit the expression of a cytochrome P450 polypeptide directly, by preventing translation of the cytochrome P450 polypeptide messenger RNA, or indirectly, by encoding a polypeptide that inhibits the transcription or translation of an cytochrome P450 polypeptide gene encoding a cytochrome P450 polypeptide. Methods for inhibiting or eliminating the expression of a gene product in a plant are well known in the art, and any such method may be used in the present invention to inhibit the expression of cytochrome P450 polypeptides.

In some embodiments, the activity of a cytochrome P450 polypeptide of the present invention may be reduced or eliminated by disrupting the gene encoding the cytochrome P450 polypeptide. The invention encompasses mutagenized plants that carry mutations in cytochrome P450 genes, where the mutations reduce expression of the cytochrome P450 gene or inhibit the activity of an encoded cytochrome P450 polypeptide of the present invention.

In some embodiments of the present invention, a Nicotiana plant or plant part is transformed with an expression cassette that is capable of expressing a polynucleotide that inhibits the expression of a cytochrome P450 sequence. Such methods may include the use of sense suppression/cosuppression, antisense suppression, double-stranded RNA (dsRNA) interference, hairpin RNA interference and intron-containing hairpin RNA interference, amplicon-mediated interference, ribozymes, and small interfering RNA or micro RNA.

For cosuppression, an expression cassette is designed to express an RNA molecule corresponding to all or part of a messenger RNA encoding a cytochrome P450 polypeptide of interest (for example, a cytochrome P450 polypeptide comprising the sequence set forth in SEQ ID NO:2, 4, 6, 8, 10, or 12 or a sequence having substantial sequence identity to SEQ ID NO:2, 4, 6, 8, 10, or 12) in the sense orientation. Overexpression of the RNA molecule can result in reduced expression of the native gene. Multiple plant lines transformed with the cosuppression expression cassette are then screened to identify those that show the greatest inhibition of cytochrome P450 polypeptide expression.

The polynucleotide used for cosuppression may correspond to all or part of the sequence encoding a cytochrome P450 polypeptide or the present invention, all or part of the 5' and/or 3' untranslated region of a cytochrome P450 polypeptide transcript, or all or part of both the coding sequence and the untranslated regions of a transcript encoding a cytochrome P450 polypeptide. In some embodiments where the polynucleotide comprises all or part of the coding region for a cytochrome P450 polypeptide of the present invention, the expression cassette is designed to eliminate the start codon of the polynucleotide so that no protein product will be transcribed.

Cosuppression may be used to inhibit the expression of plant genes to produce plants having undetectable protein levels for the proteins encoded by these genes or may also be used to inhibit the expression of multiple proteins in the same plant (e.g., Broin et al. (2002) Plant Cell 14:1417-1432; U.S. Pat. No. 5,942,657). Methods for using cosuppression to inhibit the expression of endogenous genes in plants are described in Flavell et al. (1994) Proc. Natl. Acad. Sci. USA 91:3490-3496; Jorgensen et al (1996) Plant Mol. Biol. 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; and U.S. Pat. Nos. 5,034,323, 5,283,184, and 5,942,657; each of which is herein incorporated by reference. The efficiency of cosuppression may be increased by including a poly-TR element in the expres-
29 sion cassette at a position 3' to the sense sequence and 5' of the polyadenylation signal. See, U.S. Patent Publication No. 20020048814, herein incorporated by reference. Typically, such a nucleotide sequence has substantial sequence identity to the sequence of the transcript of the endogenous gene, optimally greater than about 65% sequence identity, more optimally greater than about 85% sequence identity, most optimally greater than about 95% sequence identity (e.g., U.S. Pat. Nos. 5,283,184 and 5,034,323; herein incorporated by reference).

In some embodiments of the invention, inhibition of the expression of the cytochrome P450 polypeptide of the present invention may be obtained by antisense suppression. For antisense suppression, the expression cassette is designed to express an RNA molecule complementary to all or part of a messenger RNA encoding the cytochrome P450 polypeptide. Over expression of the antisense RNA molecule can result in reduced expression of the native gene. Accordingly, multiple plant lines transformed with the antisense suppression expression cassette are screened to identify those that show the greatest inhibition of cytochrome P450 polypeptide expression.

The polynucleotide for use in antisense suppression may correspond to all or part of the complement of the sequence encoding the cytochrome P450 polypeptide, all or part of the complement of the 5' and/or 3' untranslated region of the cytochrome P450 polypeptide transcript, or all or part of the complement of both the coding sequence and the untranslated regions of a transcript encoding the cytochrome P450 polypeptide. In addition, the antisense polynucleotide may be fully complementary (i.e., 100% identical to the complement of the target sequence) or partially complementary (i.e., less than 100% identical to the complement of the target sequence) to the target sequence. Antisense suppression may be used to inhibit the expression of multiple proteins in the same plant (e.g., U.S. Pat. No. 5,942,657). Furthermore, portions of the antisense nucleotides may be used to disrupt the expression of the target gene. Generally, sequences of at least 50 nucleotides, 100 nucleotides, 200 nucleotides, 300, 400, 450, 500, 550, or greater may be used. Methods for using antisense suppression to inhibit the expression of endogenous genes in plants are described, for example, in Liu et al. (2002) Plant Physiol. 129:1732-1743 and 5,795,829 and 5,942,657, each of which is herein incorporated by reference. Efficiency of antisense suppression may be increased by including a poly-dT region in the expression cassette at a position 3' to the antisense sequence and 5' of the polyadenylation signal. See, U.S. Patent Publication No. 20020048814, herein incorporated by reference.

For dsRNA interference, a sense RNA molecule that is described above for cosuppression and an antisense RNA molecule that is fully or partially complementary to the sense RNA molecule are expressed in the same cell, resulting in inhibition of the expression of the corresponding endogenous messenger RNA.

Expression of the sense and antisense molecules can be accomplished by designing the expression cassette to comprise both a sense sequence and an antisense sequence for the target cytochrome P450 sequence. Alternatively, separate expression cassettes may be used for the sense and antisense sequences. Multiple plant lines transformed with the dsRNA interference expression cassette or expression cassettes are then screened to identify plant lines that show the greatest inhibition of expression of the targeted cytochrome P450 polypeptide. Methods for using dsRNA interference to inhibit the expression of endogenous plant genes are described in Waterhouse et al. (1998) Proc. Natl. Acad. Sci. USA 95:13959-13964, Liu et al. (2002) Plant Physiol. 129:1732-1743, and WO 99/49029, WO 99/53050, WO 99/61631, and WO 00/49055; each of which is herein incorporated by reference.

Ambiplex expression cassettes comprise a plant virus-derived sequence that contains all or part of the target gene but generally not all of the genes of the native virus. The viral sequences present in the transcription product of the expression cassette allow the transcription product to direct its own replication. The transcripts produced by the ambiplex may be either sense or antisense relative to the target sequence (i.e., the messenger RNA for a cytochrome P450 polypeptide that is involved in the metabolic conversion of nicotine to nornicotine). Methods of using ambiplexes to inhibit the expression of endogenous plant genes are described, for example, in Angell and Baulcombe (1997) EMBO J. 16:3675-3684, Angell and Baulcombe (1999) Plant J 20:357-362, and U.S. Pat. No. 6,646,805, each of which is herein incorporated by reference.

In additional embodiments of the present invention, the polynucleotide expressed by the expression cassette of the invention is catalytic RNA or has ribozyme activity specific for the messenger RNA of a cytochrome P450 polypeptide described herein. Thus, the polynucleotide causes the degradation of the endogenous messenger RNA, resulting in reduced expression of the cytochrome P450 polypeptide. This method is described, for example, in U.S. Pat. No. 4,987,071, herein incorporated by reference.

In further embodiments of the invention, inhibition of the expression of one or more cytochrome P450 polypeptides may be obtained by RNA interference (RNAi) by expression of a gene encoding a micro RNA (miRNA). miRNAs are regulatory agents consisting of about 22 ribonucleotides. miRNAs are highly efficient at inhibiting the expression of endogenous genes. See, for example Javier et al. (2003) Nature 425: 257-263, herein incorporated by reference.

For miRNA interference, the expression cassette is designed to express an RNA molecule that is modeled on an endogenous miRNA gene. The miRNA gene encodes an RNA that forms a hairpin structure containing a 22-nucleotide sequence that is complementary to another endogenous gene (target sequence). For suppression of cytochrome P450 polypeptide expression, the 22-nucleotide sequence is selected from a cytochrome P450 polypeptide transcript sequence and contains 22 nucleotides encoding this cytochrome P450 polypeptide sequence in sense orientation and 21 nucleotides of a corresponding antisense sequence that is complementary to the sense sequence. miRNA molecules are highly efficient at inhibiting the expression of endogenous genes, and the RNA interference they induce is inherited by subsequent generations of plants.

In still other embodiments of the invention, inhibition of the expression of one or more cytochrome P450 polypeptides by RNAi may be obtained by hairpin RNA (hpRNA) interference or intron-containing hairpin RNA (ihpRNA) interference. These methods are highly efficient at inhibiting the expression of endogenous genes. See, Waterhouse and Hel-lowell (2003) Nat. Rev. Genet. 4:29-38 and the references cited therein.

For hpRNA interference, the expression cassette is designed to express an RNA molecule that hybridizes with itself to form a hairpin structure that comprises a single-stranded loop region and a base-paired stem. The base-paired stem region comprises a sense sequence corresponding to all or part of the endogenous messenger RNA encoding the gene product whose expression is to be inhibited, in this case, a cytochrome P450 polypeptide described herein, and an anti-


In one such embodiment, RNAi is accomplished by expressing an inhibitory sequence that comprises a first sequence of a cytochrome P450 polypeptide of the invention that is about 90 bp to about 110 bp in length, including 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, and 110 bp in length, and a second sequence that is complementary to all or a part of the first sequence. In another such embodiment, the inhibitory sequence comprises a first sequence of a cytochrome P450 polypeptide of the invention that is about 290 to about 310 bp in length, including 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, and 310 bp in length, and a second sequence that is complementary to all or a part of the fragment sequence.

In other embodiments of the invention, RNAi is accomplished by expressing an inhibitory sequence that comprises a first polynucleotide sequence containing the nucleotides from about position 265 to about position 625 of a cytochrome P450 coding sequence disclosed herein and a second sequence that is fully or partially complementary thereto. In some of these embodiments, the inhibitory sequence comprises as the first polynucleotide sequence the nucleotides corresponding to about position 297 to about position 594 of the P450 coding sequence set forth in SEQ ID NO:3 or SEQ ID NO:5 and the second sequence is the comple ment (i.e., antisen se sequence) of this first sequence. The inhibitory sequence can optionally comprise an intron sequence linked between the first and second sequences. Any intron known to those of skill in the art can be used in this manner. In some embodiments, the intron is from the soybean omega-6 fatty acid desaturase (FAD) (see GenBank Accession No. DQ672337, and Example 7 herein below). In one such embodiment, the intron comprises about 151 nucleotides that comprise nucleotides 100-247 of the soybean omega-6 fatty acid desaturase polynucleotide shown in GenBank Accession No. DQ672337. Examples of other introns include, but are not limited to, the intron nucleotide sequences of alcohol dehydrogenase (adh1) genes. Expression of this inhibitory sequence produces an intron-containing hairpin RNA that strongly interferes with expression of the cytochrome P450 polypeptides disclosed herein. In this manner, Nicotiana plants that are normally converters of nicotine to nornicotine that are transformed with an expression cassette comprising such an inhibitory sequence advantageously have a nicotine to nornicotine conversion rate that, surprisingly, is even lower than that observed for Nicotiana plants that are nonconverters of nicotine to nornicotine.

In yet other embodiments of the invention, RNAi is accomplished by expressing an inhibitory sequence that comprises a first polynucleotide sequence containing the nucleotides from about position 1420 to about position 1580 of a cytochrome P450 coding sequence disclosed herein and a sequence that is fully or partially complementary thereto. In some of these embodiments, the inhibitory sequence comprises as the first polynucleotide sequence the nucleotides corresponding to about position 1453 to about position 1551 of the P450 coding sequence set forth in SEQ ID NO:1 and the second sequence is the complement (i.e., antisense sequence) of this first sequence. The inhibitory sequence can optionally comprise an intron sequence linked between the first and second sequences. Any intron known to those of skill in the art can be used in this manner, as noted herein above. Expression of this inhibitory sequence produces a hairpin RNA (or intron-containing hairpin RNA when the intron is present) that also interferes with expression of the cytochrome P450 polypeptides disclosed herein.

The expression cassette for hairpin RNA interference may also be designed such that the sense sequence and the antisense sequence do not correspond to an endogenous RNA. In this embodiment, the sense and antisense sequence flank a loop sequence that comprises a nucleotide sequence corresponding to all or part of the endogenous messenger RNA of the target gene. Thus, it is the loop region that determines the specificity of the RNA interference. See, for example, WO 02/00904, herein incorporated by reference.

Transcriptional gene silencing (TGS) may be accomplished through use of hairpin RNA constructs wherein the inverted repeat of the hairpin shares sequence identity with the promoter region of a gene to be silenced. Processing of the hairpin RNA into short RNAs that can interact with the homologous promoter region may trigger degradation or methylation to result in silencing (Aufsatz et al. (2002) Proc. Natl. Acad. Sci. 99 (Suppl. 4):16499-16506; Mette et al. (2000) EMBO J 19(19):5194-5201).

In further embodiments, a polynucleotide may be utilized that encodes a zinc finger protein that binds to a gene encoding a cytochrome P450 polypeptide, resulting in reduced expression of the gene. In particular embodiments, the zinc finger protein binds to a regulatory region of a cytochrome P450 polypeptide gene. In other embodiments, the zinc finger protein binds to a messenger RNA encoding a cytochrome P450 polypeptide and prevents its translation. Methods of
selecting sites for targeting by zinc finger proteins have been described, for example, in U.S. Pat. No. 6,453,242, and methods for using zinc finger proteins to inhibit the expression of genes in plants are described, for example, in U.S. Patent Publication No. 20030037555; each of which is herein incorporated by reference.

In other embodiments of the invention, the polynucleotide encodes an antibody that binds to at least one cytochrome P450 polypeptide, and reduces the activity of a cytochrome P450 polypeptide of the present invention. In another embodiment, the binding of the antibody results in increased turnover of the antibody-cytochrome P450 polypeptide complex by cellular quality control mechanisms. The expression of antibodies in plant parts and the inhibition of molecular pathways by expression and binding of antibodies to proteins in plant parts are well known in the art. See, for example, Conrad and Sonnewald (2003) Nature Biotechnol. 21:35-36, incorporated herein by reference.

In other embodiments, the activity of a cytochrome P450 polypeptide of the present invention is reduced or eliminated by disrupting the gene encoding the cytochrome P450 polypeptide. The gene encoding the cytochrome P450 polypeptide may be disrupted by any method known in the art, for example, by transposon tagging or by mutagenizing plants using random or targeted mutagenesis and selecting for plants that have reduced cytochrome P450 activity.

Transposon tagging may be used to reduce or eliminate the activity of one or more cytochrome P450 polypeptides of the present invention. Transposon tagging comprises inserting a transposon within an endogenous cytochrome P450 gene to reduce or eliminate expression of the cytochrome P450 polypeptide.

In this embodiment, the expression of one or more cytochrome P450 polypeptides is reduced or eliminated by inserting a transposon within a regulatory region or coding region of the gene encoding the cytochrome P450 polypeptide. A transposon that is within an exon, intron, 5' or 3' untranslated sequence, a promoter, or any other regulatory sequence of a cytochrome P450 polypeptide gene may be used to reduce or eliminate the expression and/or activity of the encoded cytochrome P450 polypeptide.


Additional methods for decreasing or eliminating the expression of endogenous genes in plants are also known in the art and can be similarly applied to the instant invention. These methods include other forms of mutagenesis, such as ethyl methanesulfonate-induced mutagenesis, deletion mutagenesis, and fast neutron deletion mutagenesis used in a reverse genetics sense (with PCR) to identify plant lines in which the endogenous gene has been deleted. For examples of these methods see Ohshima et al. (1998) Virology 243:472-481; Okubara et al. (1994) Genetics 137:867-874; and Que-sada et al. (2000) Genetics 154:421-436; each of which is herein incorporated by reference. In addition, a fast and automatable method for screening for chemically induced mutations, TILLING (Targeting Induced Local Lesions In Genomes), using denaturing HPLC or selective endonuclease digestion of selected PCR products is also applicable to the instant invention. See McCallum et al. (2000) Nat. Biotechnol. 18:455-457, herein incorporated by reference.

Mutations that impact gene expression or that interfere with the function of the encoded cytochrome P450 protein can be determined using methods that are well known in the art. Insertional mutations in gene exons usually result in null-mutants. Mutations in conserved residues can be particularly effective in interfering the metabolic function of the encoded protein. Conserved residues of plant cytochrome P450 polypeptides suitable for mutagenesis with the goal to eliminate activity of a cytochrome P450 polypeptide in converting nicotine to nornicotine in a Nicotiana plant or plant part have been described (See, for example, Kusaba et al. (2003) Plant Cell 15:1455-1467.

While a number of sequences are recognized in the practice of the invention, in particular SEQ ID NO:3 and SEQ ID NO:5 find particular use. While not bound by any particular mechanisms of action, it is believed that these sequences encode a nicotine demethylation that catalyzes the oxidative N-demethylation of nicotine to nornicotine. Thus, methods to specifically inhibit these coding sequences and not other P450 sequences may be beneficial to the recombinant plant. That is, strategies that would lead to inhibition of gene function of this individual locus may prove to be superior to those that inhibit the entire gene family. The P450 enzymes are involved in many mechanisms in the plant, the inhibition of which may prove deleterious or detrimental to the growth and development of the plant or may negatively impact factors such as the disease defense capabilities of the plant. Likewise, because the Nicotiana plant P450 enzymes have been implicated in plant metabolites such as phenylpropanoids, alkaloids, terpenoids, lipids, cyanogenic glycosides, glucosinolates, and a host of other chemical entities, disruption of p450 activity may alter components involved in tobacco flavor, texture, or other properties that would impact the commercial usefulness of the plant. Therefore, the use of the methods discussed above to inhibit expression in a manner that specifically targets the coding sequence of SEQ ID NO:3 or SEQ ID NO:5 may be preferred, including targeted mutational strategies, such as chimera-plasty. See, for example, Stewart et al. (2000) Biotechniques 29(4): 838-843; Graham et al. (2002) Biochim Biophys Acta 1587(1): 1-6, herein incorporated by reference.

The protein encoded by the cDNA designated 3D_C12-10 (SEQ ID NO:4) differs from 3D_C12-7 (SEQ ID NO:6) at only two amino acid residues immediately following the start methionine. The codons corresponding to these amino acids were contained within the PCR primer used to generate the 3D_C12-7 cDNA. Thus, the original mRNA template from which 3D_C12-7 was amplified may be the same as that corresponding to the 3D_C12-10 gene, with the PCR primer sequences mediating the changes observed in the second and third amino acid sequence. Regardless, the encoded protein products would function identically. The location of the two amino acids that differ between the predicted proteins is in the N-terminal signal sequence that merely serves to anchor the protein to the endoplasmic reticulum membrane and therefore would not be expected to influence the catalytic properties of the enzyme.

In another embodiment of the invention, the compositions of the invention find use in screening methods to identify nonconverter plants for use in breeding programs. In this manner, the nucleotide sequences of the invention can be used to screen native germplasms for nonconverter plants having a stable mutation in one or more p450 genes identified herein.
These nonconverter plants identified by the methods of the invention can be used to develop breeding lines. In addition to the nucleotide sequences encoding P450 coding sequences, compositions of the invention include an intron sequence in the 3D C12-10 sequence that can be used in screening methods. While not bound by any mechanism of action, the 3D C12-7/3D C12-10 gene(s) may represent the only member(s) of the 3D C12 family involved in the metabolic conversion of nicotine to normicotine (and as stated previously there is a good likelihood that the 3D C12-7 and 3D C12-10 cDNAs originated from a single unique genetic locus). For certain applications it would be useful to have a means of diagnostically differentiating this specific member of the 3D C12 gene family from the rest of the closely related sequences within this family. For example, it is possible that within the naturally existing tobacco germplasm (or in mutagenized populations), accessions may exist in which this gene is naturally dysfunctional and may therefore be valuable as a permanently nonconverter resource. A method to specifically assay for such genotypes (e.g. deletion mutants, rearrangements, and the like) could serve as a powerful tool. To obtain such a tool, the sequence alignment shown in FIG. 3A-3G was used to design PCR primers in regions possessing polymorphisms among the members. One primer combination (5' primer shown in SEQ ID NO:25 and 3' primer shown in SEQ ID NO:26) using sequences specific to 3D C12-10 yields two particularly useful results: (1) all PCR products amplified from tobacco genomic DNA gave that same unique product (as determined by DNA sequence analysis); and (2) the presence of a 992 bp intron was revealed that is located between the primer sequences (FIG. 7; intron shown in SEQ ID NO:24).

When any cDNA corresponding to a member of the 3D C12 family is used as a hybridization probe in a Southern blotting assay of tobacco genomic DNA, a complex pattern is observed. This is expected, given that there are multiple, closely related members of this gene family. Because the intron regions of genes are typically less conserved than exons, it is predicted that the use of an intron-specific probe would reduce this complexity and better enable one to distinguish the gene(s) corresponding to the 3D C12-7/3D C12-10 gene from the other members of the family. Indeed, the probe corresponding to the sequence shown in FIG. 7 resulted in a Southern blotting pattern with greatly reduced complexity. The use of a 3D C12-10 intron-specific probe, and/or the PCR primers used to generate the fragment shown in FIG. 7, therefore provide powerful tools in assays to determine whether any naturally occurring, or mutagenized, tobacco plants possess deletions or rearrangements that may render the gene inactive. Such a plant can then be used in breeding programs to create tobacco lines that are incapable of converting.

Transformed Plants, Plant Parts, and Products Having Reduced Normicotine and NNN Content

The cytochrome P450 polynucleotides of the invention, and variants and fragments thereof, can be used in the methods of the present invention to inhibit expression or function of cytochrome P450s that are involved in the metabolic conversion of nicotine to normicotine in a plant. In this manner, inhibitory sequences that target expression or function of a cytochrome P450 polypeptide disclosed herein are introduced into a plant or plant cell of interest. In some embodiments the expression cassettes described herein are introduced into a plant or plant cell of interest, for example, a Nicotiana plant as noted herein below, using any suitable transformation methods known in the art including those described herein.

The methods of the invention do not depend on a particular method for introducing a sequence into a plant or plant part, only that the desired sequence gains access to the interior of at least one cell of the plant or plant part. Methods for introducing polynucleotide sequences into plants are known in the art and include, but are not limited to, stable transformation methods, transient transformation methods, and virus-mediated methods.


Any plant tissue that can be subsequently propagated using clonal methods, whether by organogenesis or embryogenesis, may be transformed with a recombinant construct comprising a cytochrome P450 inhibitory sequence, for example, an expression cassette of the present invention. By “organogenesis” is intended the process by which shoots and roots are developed sequentially from meristematic centers. By “embryogenesis” is intended the process by which shoots and roots develop together in a concerted fashion (not sequentially), whether from somatic cells or gametes. Exemplary tissues that are suitable for various transformation protocols described herein include, but are not limited to, callus tissue, existing meristematic tissue (e.g., apical meristems, axillary buds, and root meristems) and induced meristem tissue (e.g., cotyledon meristem and hypocotyl meristem), hypocotyls, cotyledons, leaf disks, pollen, embryos, and the like.

As used herein, the term “stable transformation” is intended to mean that the nucleotide construct of interest introduced into a plant integrates into the genome of the plant and is capable of being inherited by the progeny thereof.

“Transient transformation” is intended to mean that a sequence is introduced into the plant and is only temporarily expressed or is only transiently present in the plant.

In specific embodiments, the inhibitory sequences of the invention can be provided to a plant using a variety of transient transformation methods. The inhibitory sequences of the invention can be transiently transformed into the plant using techniques known in the art. Such techniques include
viral vector systems and the precipitation of the polynucleotide in a manner that precludes subsequent release of the DNA. Thus, the transcription from the particle-bound DNA can occur, but the frequency with which it is released to become integrated into the genome is greatly reduced. Such methods include the use of particles coated with polyethyleneimine (PEI; Sigma #P3143).

In other embodiments, the inhibitory sequence of the invention may be introduced into plants by contacting plants with a virus or viral nucleic acids. Generally, such methods involve incorporating an expression cassette of the invention within a viral DNA or RNA molecule. It is recognized that promoters for use in the expression cassettes of the invention also encompass promoters utilized for transcription by viral RNA polymerases. Methods for introducing polynucleotides into plants and expressing a protein encoded therein, involving viral DNA or RNA molecules, are known in the art. See, for example, U.S. Pat. Nos. 5,889,191, 5,889,190, 5,866,785, 5,589,367, 5,316,931, and Porta et al. (1996) Molecular Biotechnology 5:209-221; herein incorporated by reference.

Transformed cells may be grown into Nicotiana plants in accordance with conventional methods. See, for example, methods disclosed in Vasil and Hildebrandt (1965) Science 150:889; Nagara and Hoffman (1989) Biotechniques 7(8):808-812. These plants may then be grown, and either pollinated with the same transformed line or different lines, and the resulting progeny having expression of the desired phenotypic characteristic identified, i.e., reduced expression of one or more cytochrome P450s that are involved in the metabolic conversion of nicotine to nornicotine, and thus reduced content of nornicotine, and a concomitant reduced content of its nitrosamine metabolite, NNN, in the plant, particularly in the leaf tissues. Two or more generations may be grown to ensure that expression of the desired phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure expression of the desired phenotypic characteristic has been achieved. In this manner, the present invention provides transformed seed (also referred to as “transgenic seed”) having a polynucleotide of the invention, for example, an expression cassette of the invention, stably incorporated into its genome.

The compositions and methods of the invention can be used to reduce the nornicotine content, particularly in the leaves and stems, of any plant of the genus Nicotiana including, but not limited to, the following species: acuminata, affinis, alata, attenuata, bigelovii, clevelandii, excelsior, forgetiana, glauca, glutinosa, langsdorffi, longiflora, obtusifolia, palmeri, paniculata, plumbaginifolia, quadrivalvis, repanda, rustica, suaveolens, sylvester, tabacum, tomentosa, trigonophylla, and x. sanderae. The present invention also encompasses the transformation of any varieties of a plant of the genus Nicotiana, including but not limited to Nicotiana acuminata multiflora, Nicotiana alata grandiflora, Nicotiana bigelovii quadrivalvis, Nicotiana bigelovii wallacei, Nicotiana obtusifolia obtusifolia, Nicotiana obtusifolia plumeri, Nicotiana quadrivalvis bigelovii, Nicotiana quadrivalvis quadrivalvis, Nicotiana quadrivalvis wallacei, and Nicotiana trigonophylla palmeri, as well as varieties commonly known as flue or bright varieties, Burley varieties, dark varieties, and oriental/Turkish varieties.

The transgenic plants of the genus Nicotiana as described herein are suitable for conventional growing and harvesting techniques, such as cultivation in manure rich soil or without manure, bagging the flowers or no bagging, or topping or no topping. The harvested leaves and stems may be used in any traditional tobacco product including, but not limited to, pipe, cigar and cigarette tobacco, and chewing tobacco in any form including leaf tobacco, shredded tobacco, or cut tobacco.

Thus the present invention provides a Nicotiana plant, particularly leaf tissues of these plants, comprising an expression cassette of the invention and a reduced amount of nornicotine and N'-nitrosornicotine. As used herein, the term “a reduced amount” or “a reduced level” is intended to refer to an amount of nornicotine and/or N'-nitrosornicotine in a treated or transgenic plant of the genus Nicotiana or a plant part or tobacco product thereof that is less than what would be found in a plant of the genus Nicotiana or a plant part or tobacco product from the same variety of tobacco, processed (i.e., cultured and harvested) in the same manner, that has not been treated or was not made transgenic for reduced nornicotine and/or N'-nitrosornicotine. The amount of nornicotine may be reduced by about 10% to greater than about 90%, including greater than about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, and about 80%.

The term “tobacco products” as used herein include, but are not limited to, smoking materials (e.g., cigarettes, cigars, pipe tobacco), snuff, chewing tobacco, gum, and lozenges. The present invention also encompasses a range of tobacco products blends that can be made by combining conventional tobacco with differing amounts of the low nornicotine and/or N'-nitrosornicotine tobacco described herein. In further embodiments, the plant or plant part of the genus Nicotiana as described above is cured tobacco.

In some embodiments of the present invention, the tobacco product reduces the carcinogenic potential of tobacco smoke that is inhaled directly with consumption of a tobacco product such as cigars, cigarettes, or pipe tobacco, or inhaled as secondary smoke (i.e., by an individual that inhales the tobacco smoke generated by an individual consuming a tobacco product such as cigars, cigarettes, or pipe tobacco). The cured tobacco described herein can be used to prepare a tobacco product, particularly one that undergoes chemical changes due to heat, comprising a reduced amount of nornicotine and/or N'-nitrosornicotine in the smoke stream that is inhaled directly or inhaled as secondary smoke. In the same manner, the tobacco products of the invention may be useful in the preparation of smokeless tobacco products such as chewing tobacco, snuff, and the like.

The tobacco products obtained from the transgenic tobacco plants of the present invention thus find use in methods for reducing the carcinogenic potential of these tobacco products, and reducing the exposure of humans to the carcinogenic nitrosamine NNN, particularly for individuals that are users of these tobacco products.

The following examples are offered by way of illustration and not by way of limitation.

**EXPERIMENTAL**

The following materials and protocols were utilized in the experiments described herein below.

**Plant Materials**

All plant materials utilized in these experiments were provided by Dr. Earl Wernsman, Department of Crop Science, North Carolina State University. DH 91-1307-46(NC) and DH91-1307-46(Con) are near-isogenic doubled haploid Burley lines (nonconver and conver, respectively) recovered from the same maternal haploid plant. Burley lines DH 98-326-3 (nonconver) and DH 98-326-1 (conver), and DH 98-325-5 (nonconver) and DH 98-325-6 (conver) represent two additional pairs of near-isogenic lines. SC58 is a flue-cured tobacco variety, nonconver individuals of
which are designated SC58(C3) and SC58(C10) is a near-
isoisogenic stable convertor line that originated through the intro-
gression of the single dominant convertor locus (C7) found in
the tobacco progenitor species N. tomentosiformis into SC58
(Mann et al. (1964) Crop Sci. 4:340-353. After eight addi-
tional backcrosses to SC58, the near-isogenic SC58(C3) line
was created and subsequently maintained via self-fertil-
ization.

All plants were maintained in growth chambers or green-
houses using standard potting soil and fertilizer. For the micro-
array studies, the metabolism of nicotine to nornicotine
was accelerated by excising individual leaves and inserting
their petioles into a solution of 0.1% ethephon or 1% sodium
bicarbonate. The leaves were then placed in a growth cham-
ber (27°C) for 5 to 7 hours to facilitate the uptake of the
ethephon or sodium bicarbonate solutions throughout the
transpiration stream. The treated leaves were placed in
small plastic storage bags after being lightly sprayed with
water (to maintain high humidity) and cured for three days at
30°C in the dark. To enhance the nicotine to nornicotine
conversion in the transgenic plants generated in this study,
detached leaves were dipped into a solution of 0.2% etheph-
on, dried, and cured in plastic storage bags for seven days at
room temperature in the dark.

cDNA Libraries and Expressed Sequence Tags

Total cellular RNA was isolated from senescing leaf tissue
of Burley lines DH 91-1307-46 (NC) and DH 91-1307-46
(Cou) using the TRIzol® reagent according to the manufac-
turer’s protocol (Invitrogen). PolyA RNA was isolated from
total RNA using the MessageMaker system (Invitrogen), and
cDNA was subsequently synthesized and cloned into the lambda
ZAP II phage vector using the ZAP-cDNA Synthesis and
Gigapack III Gold Cloning Kit (Stratagene). Aliquots of the
phage libraries were converted to phleucis;se-based plasmid
libraries following the mass excision protocol outlined by
Stratagene.

Thousands of colonies from both the converter and non-
converter libraries were grown on selective solid media
and picked into 384-well plates containing Luria broth (with
ampicillin) in 10% glycerol to facilitate long term storage of
the clones at -80°C. Over 11,000 clones from each library
were transferred from the 384-well plates to 96-well growth
blocks and grown in selective media. Plasmids were isolated
in 96-well format using the F.E.A.L. Preparation Kit
(Qiagen) with the aid of a BioRobot 3000 Workstation
(Qiagen). To generate the ESTs, the plasmid clones were
sequenced using the T3 primer (Qiagen) and BigDye® Ter-
mator system (Applied Biosystems) according to the BigDye®
cycle sequencing protocol. Perform® DTR 96-well
plates (Edge Biosystems) were used to remove the unincor-
dated dye from the sequencing reactions prior to loading
the samples onto a Perkin Elmer Prism 3700 96-Capillary
Automated DNA Sequencer.

Preparation of DNA Chips

To obtain DNAs suitable for spotting onto glass slides, the
M13 forward and reverse sequencing primers (Qiagen) were
used as PCR primers to amplify cDNA inserts from the plas-
mids containing cDNAs represented in the EST databases.
Plasmid clones were subjected to PCR in 96-well format
using an Applied Biosystems Gene Amp 9700 model ther-
mo cycler. The resulting PCR products were processed
through Millipore Multiscreen TM PCR or Montage TM
PCR Purification systems. The resulting products were
transferred into 384-well plates containing equal volumes of
dDSMO. The final DNA concentrations were estimated to be
equal to or greater than 0.1 ng/μL. The DNAs were subse-
quently spotted onto amino silane-coated slides (Comin®
GAPS II) using an Affymetrix GMS 417 array printer. DNAs
were immobilized to the slide surface by UV crosslinking
(-120 mJ/cm²), followed by baking at 75°C for two hours.

Microarray Hybridization and Analysis

The amino allyl dUTP-based indirect method of dye incor-
poration described by “The Institute of Genome Research”
(http://pga.tigr.org/protocols.html) was used to label noncon-
verter and converter RNAs with Cy3 and Cy5 fluorescent
dyes (Amersham Biosciences). Briefly, 20 μg of total RNA
was reverse transcribed in a 30 μL volume containing 400 units
of SuperScript II RT (Invitrogen), 6 ng random hexamer
primers, 0.5 mM each of dATP, dCTP, and dGTP, 0.3 mM
dTTP, and 0.2 mM amino allyl dUTP (Sigma) in first strand
synthesis buffer (Invitrogen). Reactions were incubated for
6 to 14 hours at 42°C, followed by hydrolysis of the RNA
with NaOH. The resulting first strand cDNA molecules were
column purified (Qiagen) and washed with phosphate buffer.
Coupling reactions of the NHS-ester Cy3 or Cy5 fluorescent
dyes to the cDNA occurred during incubation in 0.05 M
sodium carbonate buffer (pH 9.0) and 25% DMSO at room
temperature for 1.5 hours.

Microarray slides were prehybridized in a solution of
5×SSC, 0.1% SDS, and 1% BSA at 42°C for 45 minutes,
rinsed gently with dH2O and isopropanol, and dried by low
speed centrifugation. The Cy3- and Cy5-labeled cDNAs were
column purified (Qiagen), combined, and hybridized to the
DNA slides in a solution containing 5×SSC, 0.5% SDS,
5×Denhardt’s, 0.45 μg/μl Poly A RNA, 0.45 μg/μl calf thymus
DNA, and 50% formamide. The slides were incubated with
the hybridization solution for 14 to 16 hours at 42°C. Post-
hybridization washes consisted of sequential 4-minute incubations
with the following solutions: 1×SSC, 0.2% SDS;
0.1×SSC, 0.2% SDS; 0.1×xSSC, and a final 10 second rinse
with 0.1×xSSC.

The microarrays were subsequently scanned using Scan-
Array 2.1 (GSI Lumonics) or ScanArray Express (Perki-
nElmer). Serial scanning for Cy5 and Cy3 fluorescence was
performed at a maximal resolution of 10 gm/pixel, and
laser power and PMT gain adjusted to provide reliable and
equivalent signal strengths. The acquired array images were
quantified for signal intensity with QuantaArray® analysis
software (PerkinElmer), using the histogram-based method.
Total intensities were used as quantification output fields, and
the acquired data sets were saved as Unicore, tab-delimited
text files. Importation of the text files into Microsoft Excel
enabled the subsequent calculation of Cy5/Cy3 and Cy3/Cy5
ratios, the statistic we employed for the identification of
candidate genes.

Cloning Full-Length and Additional Members of the 3D_C12
Gene Family

To clone the entire coding region of 3D_C12 and 7D_A06
a modified 5′-RACE strategy was employed using a pBlue-
script II vector-specific forward primer (BlueSK;
5′-GGCTCTAGACTATGTGATC-3′; SEQ ID NO:17) and a
set of gene-specific reverse primers. Two 3D_C12-specific
reverse primers were designed, one of which is complement-
tary to the downstream portion of the 3′ untranslated
region (5′-TTTTTGGGACATCAATGCA-3′; SEQ ID NO:18)
and the other complementary to a sequence within the coding
region (5′-GGTTAGATTTATCGTACCAT-3′; SEQ ID NO:19).
For the former primer, the first five Ts are comple-
mentary to the polyA tail of the transcript. A 7D_A06-specific
reverse primer (5′-TTCAATCTCAAAATTTTTTACACCA-3′;
SEQ ID NO:20) was also designed, and is complementary
to a segment in the 3′ untranslated region of this gene. PCR
41 reactions contained 10 ng of converter tobacco leaf cDNA library (within the pbPluescript vector) as template, 2 μM concentration of each primer, 350 μM of each dNTP, and 1.5 mM MgCl₂ in a final reaction volume of 50 μL. Amplification was initiated by the addition of 2.5 units of UniPol enzyme mix using conditions described by the manufacturer (Roche). After an initial denaturation step at 94°C for 4 minutes, the samples were subjected to 30 cycles of denaturation at 94°C for 15 seconds, annealing at 57°C for 30 seconds, and extension at 72°C for 90 seconds. A final extension step at 72°C for 10 minutes was included at the end of the 30 cycles. The amplicons were ligated into the pgEM Easy T/A vector (Promega), and 10 randomly selected clones from each amplification were subjected to DNA sequence analysis. Nucleic acid and predicted protein sequences of the various members of the 3D_C12 gene family were analyzed and compared using the BLASTX (Altschul et al. 1997) Nucleic Acids Res. 25:3389-3402, ClustalW (Higgins et al. 1994) Nucleic Acids Res. 22:4673-4680 and GAP (University of Wisconsin Genetic Computing Group software package) algorithms.

The above described strategy was effective in identifying full-length sequence information for 3D_C12 and 7D_A06. In addition, PCR amplifications using the PCR primer internal to the 3D_C12 coding region gave rise to partial-sequence information for the unique 3D_C12-15 cDNA. In an attempt to obtain full-length sequence information for 3D_C12-15, a gene-specific primer complementary to the 5’ terminus of its coding region (5’-ATGGTTTTTCTCCATGAGGCC-3’, SEQ ID NO:21) was used in conjunction with a pbPluescript-specific reverse primer (5’-TGAGGCTCGACGCGTACC-3’, SEQ ID NO:22). Although a full-length 3D_C12-15 cDNA was not recovered, this amplification resulted in the isolation of 3D_C12-7, which proved to be another unique member of the 3D_C12 gene family.

Transgenic Plant Analysis

The RNAi-based gene silencing constructs were assembled in a version of the pKYL80 cloning vector (Schardl et al. 1987) Gene 61:1-11) that was engineered to contain a 151-bp fragment of the soybean FAD3 gene intron between the Xhol and SacI restriction sites of the polylinker (pKYLX80I). To create a construct in which the FAD3 intron was flanked by a sense and antisense fragment of 3D_C12, a 99-bp region located immediately upstream of the stop codon of the 3D_C12 cDNA (Fig. 3A-3G) was cloned between the HindIII-Xhol and SacI-Xhol restriction sites of pKYLX80I in its sense and antisense orientation, respectively. The resulting HindIII-Xhol fragment containing the 3D_C12 sense arm, FAD3 intron, and 3D_C12 antisense arm was subcloned into the pKYLX71 plant binary expression vector (Maiti et al. 1993) Proc. Natl. Acad. Sci. USA 90:6110-6114) between the 3SS CaMV promoter and a rubisco small subunit terminator.

Overexpression constructs were created by replacing the 3-glucuronidase ORF of the plant binary expression vector pBI121 (Clontech) with the full-length coding regions of the 3D_C12, 7D_A06, and 3D_C12-7 cDNAs. This placed the tobacco P450s under the transcriptional control of the 3SS CaMV promoter. The pBI121- and pKYLX71-based constructs were transformed into Agrobacterium tumefaciens strain LBA4404 and introduced into tobacco cultivars Petite Havana and DH98-325-6 (converter), respectively, using established protocols (Horsch et al. 1985) Science 227:1229-1231).

Northern Blot Analysis

Total cellular RNAs were isolated from tobacco leaves using the TRIZOL® method as described by the manufacturer (Invitrogen). Five to ten micrograms of RNA were size fractionated on a 1.2% agarose gel prepared in TBE buffer. RNA immobilization, probe labeling, and signal detection were carried out using the DIG nucleic acid labeling and detection kits according to the manufacturer’s instructions (Roche). Alternatively, probes were synthesized using 32P-dCTP according to protocols accompanying the Random Primed DNA Labeling kit (Roche).

Alkaloid Analysis

Tobacco leaves were harvested and air dried in an oven at 65°C for 2 days. A 100 mg sample of crushed, dried leaf was added to 0.5 mL of 2 N NaOH in 20 mL scintillation vial. The sample was mixed and allowed to incubate for 15 minutes at room temperature. Alkaloids were extracted by the addition of 5 mL of extraction solution [0.04% quinoline (wt/vol) dissolved in methyl-t-butyl ether] and gently rotated on a linear shaker for 3 hours. Following phase separation, an aliquot of the organic phase was transferred to a sample vial. Samples were analyzed using a PerkinElmer Autosystem XL gas chromatograph equipped with a flame ionization detector, a 4 mm split/splitless glass liner, and a 30 m x 0.53 mm ID DB-5 column. Chromatographic conditions were as follows: detector temperature: 250°C; injector temperature: 250°C; helium flow rate at 120°C: 20 mL/minute; injection volume: 2 μL; column conditions: 120°C, hold 1 minute, 120-280°C at 30°C/minute ramping rate, hold at 280°C for 2 minutes. Alkaloid composition was determined by the TotalChrome Navigator software using a calibration curve.

Example 1

Generation of EST Databases

RNAs isolated from senescing leaves of the converter genotype DH 91-1307-46(Con) and its near-isogenic non-converter counterpart DH 91-1307-46(NC) were used to generate cDNA libraries. High-throughput automated DNA sequencers were initially used to generate single-run sequence information (ESTs) for 11,136 randomly chosen cDNAs from the converter library. The local alignment search tool BLASTX (Altschul et al. 1990) J. Mol. Biol. 215:403-410) was used to compare the predicted protein sequence of each tobacco cDNA with the nonredundant protein database curated by the National Center for Biotechnology Information of the National Library of Medicine and National Institutes of Health. Subsequently, a similar annotated EST database was generated by conducting sequencing runs on 11,904 cDNAs selected from the nonconverter library.

Example 2

Microarray Analyses of Converter CDNA Library

Methods

Upon completion of the EST database generated from the converter cDNA library, the inserts from 4992 clones were amplified by PCR and spotted onto glass slides. Given the possibility that the nicotine demethylation enzyme may be catalyzed by an enzyme of the P450 class of oxidative enzymes, special attention was given to library entries that were predicted by BLASTX analysis to encode P450s. From visual inspection of the BLASTX results, it was estimated that 31 unique P450 genes were represented in the database. When selecting specific 96-well plates to be
included on the microarray, care was taken to ensure that all unique P450 genes would be included among the 4992 cDNAs selected.

RNAs isolated from the near-isogenic Burley genotypes DH 98-326-3 (nonconverter) and DH 98-326-1 (converter), and DH 98-325-5 (nonconverter) and DH 98-325-6 (converter) were used to generate Cy3- and Cy5-labeled cDNAs. To maximize the metabolic conversion of nicotine to nornicotine in converter genotypes, detached leaves were treated with sodium bicarbonate or ethion in poultry or sheep prior to curing, treatments that have been shown to accelerate nornicotine production in converter plants while having no effect in nonconverter individuals (Faust and Bush (1992) Med. Sci. Res. 20:867-868; Shi et al. (2003) J. Agric. Food Chem. 51:7679-7683).

To minimize the variability inherent with microarray experiments, reciprocal experiments were conducted simultaneously. In this manner, DH 98-325-5 RNA was labeled with Cy5 and DH 98-325-6 RNA was labeled with Cy3, and then in a reciprocal experiment DH 98-325-5 RNA was labeled with Cy3 and DH 98-325-6 RNA was labeled with Cy5 (collectively referred to as Exp. 2.1). Similarly, DH 98-326-3 and DH 98-326-1 RNAs were labeled with Cy3 and Cy5, respectively, in one experiment, and then the same RNAs were labeled with Cy5 and Cy3, respectively, in a reciprocal experiment (collectively referred to as Exp. 2.2).

Even when conducted reciprocally, the results of any given microarray experiment are likely to include “false positives,” representing genes that are differentially regulated between a specific genotypic pair and or uniquely in response to a specific treatment, as opposed to differences directly associated with the conversion phenomenon. To define the set of candidate genes that are most likely to be upregulated due to the conversion process, cDNAs were identified that met the following criteria: for any set of reciprocal experiments (i.e., Exp. 2.1, or Exp. 2.2), the hybridization intensity of a given cDNA had to be at least 2-fold higher with the converter probe than nonconverter probe in at least one of the hybridizations, and not less than 1.5-fold higher in the reciprocal experiment.

Experiment 2.1.—Leaves from near-isogenic lines DH 98-325-5 and DH 98-325-6 were treated with ethion and cured for 3 days at 30°C. Alkaloid analysis revealed that virtually all of the nicotine had been metabolized to nornicotine in the DH 98-325-6 leaf during this period while minimal nornicotine was observed in the DH 98-325-5 leaf. RNAs from the DH 98-325-5 nonconverter plant were labeled with the Cy3 fluoroscent dye, and RNAs extracted from a DH 98-325-6 (converter) leaf were labeled with Cy5. The Cy3- and Cy5-labeled cDNAs were incubated together on the same DNA chip and allowed to hybridize overnight.

Experiment 2.2.—A microarray analysis similar to Exp. 2.1 was conducted using the DH 98-326-3 (nonconverter) and DH 98-326-1 (converter) near-isogenic lines. In these experiments, leaves from each genotype were treated with 1% sodium bicarbonate and cured for 3 days at 30°C. At the end of the treatment period, nicotine was the predominant alkaloid in the DH 98-326-3 leaf, while nearly all of the alkaloid in the DH 98-326-1 leaf was nornicotine. As described for Exp. 2.1, these experiments were reciprocally conducted.

Results
In both Experiment 2.1 and Experiment 2.2, the great majority of the 4992 cDNAs spotted on the glass slides showed no substantial differences in their hybridization intensities to the competing Cy3- and Cy5-labeled probes.

Of the 4992 cDNAs spotted on the glass slides, only five showed at least 2-fold higher expression in one hybridization and not less than 1.5-fold in the reciprocal hybridization for both Exp. 2.1 and Exp. 2.2. These entries were designated 3D_C12, 7D_A06, 27C_C12, 33A_D06, and 34D_F06. BLASTX analysis of the partial sequence information for 3D_C12 and 7D_A06 found in our EST database predicted that the cDNAs encode two closely related P450 enzymes. 27C_C12 and 33A_D06 were predicted to encode glycine-rich cell wall proteins, displaying over 90% sequence identity to small tobacco glycine-rich proteins found in GenBank (e.g., Accession No. AAK57546). Clone 34D_F06 was found to contain a double cDNA insert, one insert showing homology to serine/threonine protein kinases, and the other showing high sequence identity to the same glycine-rich cell wall proteins as the 27C_C12 and 33A_D06 cDNAs.

Example 3

Microarray Analysis of CDNA Non-Converter Library

Upon completion of the EST database from the nonconverter library (generated from senescing leaves of genotype DH191-1307-46 (NC)), another set of microarray experiments was initiated. For this next generation of microarrays, the goal was to produce glass slides containing the complete non-redundant set of genes represented in both libraries.

To obtain an estimate of the number of unique genes that are represented in the database, clustering analysis was conducted to identify ESTs predicted to be represented multiple times in the database (contigs) versus those predicted to be represented only once (singletons) (Huang and Madan (1999) Genome Res. 9:868-877). Due to the nature of the clustering algorithms, sequences showing high, but imperfect, sequence identities are clustered into the same contig. The total set of predicted unique genes, or unigenes, within a database is calculated as the sum of the contigs and singletons. Clustering analysis of the combined converter and nonconverter databases predicted 2246 contigs and 4717 singletons for a total of 6,963 unigenes. Inserts from all singletons and an individual from each contig were amplified by PCR and spotted onto glass slides, resulting in a gene chip containing the complete 6,963 unigene set.

In addition to creating a new DNA chip, the genetic materials used to generate hybridization probes also differed from those used in Example 2. SC58 is a flue-cured tobacco variety, nonconverter individuals of which are designated SC58 (C5-Y2), SC58(C5-Y2) is a near-isogenic stable converter line that originated through the introgression of the single dominant converter locus (C5) found in the tobacco progenitor species N. tomentosiformis into SC58 (Mann et al. (1964) Crop Sci. 4:349-353). After eight additional backcrosses to SC58, the near-isogenic SC58(C5-Y2) line was created and subsequently maintained via self-fertilization. The conversion phenotype of SC58(C5-Y2) plants is unique with respect to standard converter tobacco lines in that the metabolism of nicotine to nornicotine in the leaf does not require senescence or curing. Plants possessing the C5 converter locus from N. tomentosiformis contain nornicotine as the predominant alkaloid even in green leaf tissue (Wernsman and Matzinger (1968) Tob. Sci. 12:226-228).

RNAs isolated from green leaf tissue of SC58(C5-Y2) and SC58(C5-Y2) were labeled with Cy3 and Cy5, respectively, and simultaneously hybridized to a DNA chip containing the entire 6,963 unigene set of cDNAs. The fluorescent dyes were reversed to produce the probes for a reciprocal experiment as described in Exps. 2.1 and 2.2 of Example 2.
Results

Results were evaluated using the same criteria as in Example 2, i.e., individual cDNAs were identified that showed at least 2-fold enhanced hybridization to the labeled SC58(C, C) versus SC58(C, C) cDNAs in one experiment and at least 1.5-fold enhancement in the reciprocal assay.

Results were compared to those from Exp. 2.1 and Exp. 2.2 in Example 2 above. Enhanced hybridization of converter RNAs to cDNAs encoding members of the closely related P450 family was the only result shared by all three microarray experiments using the defined criteria. 131 A_A02 is the name of the cDNA that was spotted onto the 6963-member unigene chip that is representative of the closely-related P450 gene family that includes 3D_C12 and 7D_A06 (3D_C12 and 7D_A06 themselves were not spotted on the unigene slide). No other cDNAs on the array in Example 3, which included representatives of the contigs containing the glycine-rich protein-encoding 27C_C12 and 33A_A06 and 34D_F06 cDNAs, scored positive by the defined criteria and also scored positive in Exp. 2.1 or Exp. 2.2 of Example 2 above, regardless of whether the results were compared individually or collectively.

<table>
<thead>
<tr>
<th>cDNA</th>
<th>Cy3 reading</th>
<th>Cy5 reading</th>
<th>Cy5/Cy3 ratio</th>
<th>Cy3 reading</th>
<th>Cy5 reading</th>
<th>Cy5/Cy3 ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 2.1</td>
<td>15514.14</td>
<td>25928.95</td>
<td>1.67</td>
<td>19355.85</td>
<td>9507.87</td>
<td>2.04</td>
</tr>
<tr>
<td>Experiment 2.2</td>
<td>15238.23</td>
<td>37196.19</td>
<td>2.44</td>
<td>13651.03</td>
<td>8121.04</td>
<td>1.68</td>
</tr>
<tr>
<td>Experiment 3 (reciprocal)</td>
<td>12756.63</td>
<td>28669.28</td>
<td>2.25</td>
<td>32198.81</td>
<td>16166.13</td>
<td>1.99</td>
</tr>
<tr>
<td>Example 3</td>
<td>7571.06</td>
<td>19180.94</td>
<td>2.53</td>
<td>42408.85</td>
<td>18440.17</td>
<td>2.30</td>
</tr>
<tr>
<td>131A_A02</td>
<td>11138.96</td>
<td>19638.82</td>
<td>1.76</td>
<td>36963.45</td>
<td>10085.25</td>
<td>3.67</td>
</tr>
</tbody>
</table>

Combined Results

The combined results of microarray experiments described above identified members of a closely related P450 gene family, hereafter referred to as the 3D_C12 family, to be the best candidates for playing a direct role in the metabolic conversion of nicotine to normotonic in converter tobacco plants. The hybridization results of the members of this P450 family in each of the three microarray experiments are shown in Table 1. The results of the microarrays were independently confirmed using Northern blotting assays. As shown in FIG. 2, an approximately 2-fold higher signal was observed in senescing, cured converter leaves compared to their nonconverter counterparts when RNA blots were incubated with a radiolabeled 7D_A06 hybridization probe.

Example 4

Sequence Analysis of the 3D_C12 Gene Family

Once microarray experiments defined 3D_C12 and 7D_A06 as potentially being involved in the conversion process, obtaining complete DNA sequence information for these genes became the next step in their characterization. The original 3D_C12 and 7D_A06 clones that were sequenced when generating the EST database described elsewhere (and spotted onto the microarrays) were not full-length cDNAs. To obtain a full-length sequence, primers were designed corresponding to regions in the 5’ flanking region and in the interior of the coding regions that were sufficiently polymorphic to distinguish between 7D_A06 and 3D_C12. These gene-specific primers were used in combination with primers specific to the cloning site of pBluescript II to amplify cDNAs from the converter cDNA library in an attempt to obtain sequence that would include the complete 5’ ends of the 3D_C12 and 7D_A06 reading frames.

This strategy led to the determination of the DNA sequence corresponding to the complete coding regions of 3D_C12 (nt 1-1551 of SEQ ID NO:1; predicted amino acid sequence shown in SEQ ID NO:2) and 7D_A06 (nt 1-1554 of SEQ ID NO:7; predicted amino acid sequence shown in SEQ ID NO:8) (FIGS. 3A-3G and 4). GAP analyses of the 3D_C12 and 7D_A06 DNA and predicted protein sequences showed that they share 93.4% DNA sequence identity and 92.3% identity at the protein level (Tables 2 and 3). Initial BLASTX analysis against the nonredundant GenBank database revealed that 3D_C12 and 7D_A06 share greatest sequence homology to CYP82E1, a tobacco P450 gene of unknown function that is upregulated in response to fungal elicitors (Takimoto et al. (1999) Plant Cell Physiol. 40:1232-1242). The CYP82E1 protein is 66.9% and 67.5% identical to the predicted amino acid sequences of 3D_C12 (SEQ ID NO:2) and 7D_A06 (SEQ ID NO:8), respectively, and the CYP82E1 DNA sequence is 72.1% and 73.5% identical to the respective coding sequences for 3D_C12 (nt 1-1551 of SEQ ID NO:1) and 7D_A06 (nt 1-1554 of SEQ ID NO:7).

Table 2

| Nucleotide sequence identities between members of the 3D_C12 gene family |
|---------------------------|---------------------------|---------------------------|---------------------------|
| 3D_C12 | 7D_A06 | 3D_C12-7 | 3D_C12-10 | 3D_C12-15* |
| 7D_A06 | 93.4** | 93.7 | 94.0 | 93.4 | 93.1 | 93.1 |
| 3D_C12-7 | 93.7 | 94.4 | 90.7 | 92.8 |
| 3D_C12-10 | 93.5 | 92.6 | 93.1 | 93.1 |
| 3D_C12-15* | 98.0 | 94.0 | 93.4 | 93.1 |

*partial sequences
**numbers indicate percentages
<table>
<thead>
<tr>
<th>Predicted amino acid sequence identities between full-length members of the 3D_C12 gene family.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D_C12</td>
</tr>
<tr>
<td>7D_A06</td>
</tr>
</tbody>
</table>

*numbers indicate percentages

In addition to enabling the acquisition of full-length sequence information for the 3D_C12 and 7D_A06 cDNAs, the above described PCR amplifications yielded additional products that were closely related to, yet clearly distinct from, the 3D_C12 and 7D_A06 cDNA sequences. Using a primer directed against a sequence interior to the 3D_C12 cDNA, in combination with a primer specific to pluSscript II, a unique sequence designated 3D_C12-15 (FIG. 3A-3G; SEQ ID NO:3; predicted amino acid sequence shown in SEQ ID NO:10) was amplified in addition to the expected 3D_C12 product. 3D_C12-15 is 95.8% identical to the corresponding DNA sequence of 3D_C12 and 92.6% identical to the same region of 7D_A06 (Table 2).

Because the 3D_C12-15 fragment represented an additional, distinct member of the 3D_C12 gene family, an attempt was made to obtain a full-length cDNA sequence of this gene. A PCR primer specific to the first seven codons of the 3D_C12-15 reading frame was used in combination with a pluSscript II-specific primer in an amplification reaction using our converter cDNA library as template. Sequence analysis of several independent amplification products failed to reveal a full-length 3D_C12-15 gene. Instead, a new member of this family was recovered, designated 3D_C12-7 (FIG. 3A-3G; coding sequence set forth as nt 1-1551 of SEQ ID NO:5). Across the full-length nucleotide sequence shown in SEQ ID NO:5, 3D_C12-7 shares 93.7% nucleotide sequence identity with 3D_C12 (across SEQ ID NO:1), 94.0% nucleotide sequence identity with 7D_A06 (across SEQ ID NO:7), and 93.1% identity over the corresponding region of fragment 3D_C12-15 (SEQ ID NO:9) (Table 2). The predicted amino acid sequence of 3D_C12-7 (SEQ ID NO:6) is 92.8% identical to the 3D_C12 protein (SEQ ID NO:2), and 94.8% identical to the 7D_A06 protein (SEQ ID NO:8) (Table 3).

Two additional members of the 3D_C12 family were also identified. A gene designated 3D_C12-10 (FIG. 3A-3G; coding sequence set forth as nt 1-1551 of SEQ ID NO:3; predicted amino acid sequence set forth in SEQ ID NO:4) was recovered from an amplification reaction using a PCR primer complementary to a sequence in the 3' flanking region of 3D_C12 together with a Bluescript II-specific primer (and the converter library as template). 3D_C12-10 differs at only five nucleotide positions from the 3D_C12-7 nucleotide sequence (SEQ ID NO:5) (FIG. 3A-3G), and at only two amino acids positions from the predicted 3D_C12-7 protein sequence (SEQ ID NO:6) (FIG. 4).

With the completion of the nonconverter EST database, another member of the 3D_C12 gene family was revealed. The partial DNA sequence of 131 A_A02 (SEQ ID NO: 11; predicted amino acid sequence set forth in SEQ ID NO:12) that is found in this database is 98.0% identical to the corresponding sequence of 3D_C12, and 94.0% identical to the same region of 7D_A06 (FIG. 3A-3G and Table 2). As described in the previous section, 131A_A02 is a member of the 3D_C12 gene family that was represented on the comprehensive unigene chip used in microarray assays as described elsewhere herein.

Example 5

Transgenic Plant Analysis of Members of the 3D_C12 Gene Family

To determine whether members of the 3D_C12 family of cytochrome P450 genes are involved in the metabolic conversion of nicotine to nornicotine, transgenic plants were generated using constructs designed to either enhance or inhibit gene expression. To test the effects of down-regulating gene activity, an RNA interference (RNAi) strategy was employed. A 59-bp region of 3D_C12 located immediately upstream of the stop codon (FIG. 3A-3G), was used to create a construct that would form a dsRNA hairpin within the plant cell. Such dsRNA structures are known to activate an RNAi silencing complex that leads to the degradation of both transgene RNAs and endogenous RNAs that are identical or highly homologous to the sequence found in the dsRNA (Wesley et al. (2001) Plant J. 27: 581-590; Waterhouse & Hellwell (2002) Nat. Gen. Rev. 4: 29-38).

Given that each member of the 3D_C12 characterized as described herein shares over 90% DNA sequence identity, an RNAi construct synthesized against one member was expected to silence the entire gene family. Specifically, the RNAi construct generated against the 3D_C12 sequence shares sequence identities of 90.99 and 91.99 with the 7D_A06 and 3D_C12-7 cDNAs, respectively, over this region (FIG. 3A-3G). The 3D_C12/RNAi construct (also referred to in Example 7 as the 3D_C12/RNAi construct) was cloned downstream of the constitutive 3SS promoter of cauliflower mosaic virus (CaMV) and introduced into the strong converter Burley tobacco line DH 98-325-6 using Agrobacterium-mediated transformation.

A hallmark of RNAi-mediated silencing is the marked reduction in steady-state transcript accumulation of the gene whose activity has been down-regulated. To confirm that gene silencing of the 3D_C12 gene family had occurred in the plants showing low nornicotine phenotypes, a Northern blot analysis was conducted using RNAs isolated from three of the transgenic plants possessing 3D_C12/RNAi constructs and displaying low nornicotine phenotypes, two individuals transformed with the 3D_C12/RNAi construct yet still showing high levels of nornicotine, and one of the vector-only control plants.

To assess the effects of overexpression of gene activity, the cDNAs from the three members of the 3D_C12 gene family for which we first obtained full-length sequence information (3D_C12, 7D_A06, and 3D_C12-7) were cloned in their sense orientations downstream of the 35S CaMV promoter. These constructs were subsequently introduced into N. tabacum cultivar Petite Havana using Agrobacterium-mediated transformation. The Petite Havana line is commonly used by researchers because of his shorter stature and abbreviated generation time in relation to commercial tobacco cultivars. The converter/nonconverter status of the Petite Havana cultivar is unknown, but the alkaloid assays of the present application clearly showed that the plants in our possession were strong converters.

Although the host plants in these experiments were converters, the present strategy was to conduct alkaloid assays on green, non-cured tissue, where minimal nornicotine accumulates in converter and nonconverter plants alike (and the 35S CaMV promoter is very active). In fact, a nonconverter line
was purposely chosen because tissue culturing, as required when conducting Agrobacterium-mediated transformation, is known to enhance the frequency of genetic conversion and would thus potentially complicate interpretation of results (e.g., assessing whether a novel phenotype was solely attributable to the transgene as opposed to being the result of the plant having undergone genetic conversion).

Results

Given the high degree of variability typically observed among independent transgenic plants transformed with the same transgene construct, 10 independently transformed individuals were selected to assess the effects of the 3D_C12/ RNAi construct on the metabolic conversion of nicotine to nornicotine. Leaves from each of the 10 3D_C12/RNAi individuals, in addition to two control plants transformed with the pB121 vector alone, were treated with ethephon and cured for seven days. Alkaloid analysis of these materials is shown in Table 4.

### Table 4

<table>
<thead>
<tr>
<th>Sample</th>
<th>% Nicotine*</th>
<th>% Nornicotine*</th>
<th>% Anabasine*</th>
<th>% Anatabine*</th>
<th>% Conversion**</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D_C12 RNAi (1)</td>
<td>3.149</td>
<td>0.100</td>
<td>0.012</td>
<td>0.159</td>
<td>2.8</td>
</tr>
<tr>
<td>3D_C12 RNAi (2)</td>
<td>2.569</td>
<td>0.193</td>
<td>0.009</td>
<td>0.110</td>
<td>7.0</td>
</tr>
<tr>
<td>3D_C12 RNAi (3)</td>
<td>2.175</td>
<td>0.064</td>
<td>0.007</td>
<td>0.080</td>
<td>2.9</td>
</tr>
<tr>
<td>3D_C12 RNAi (4)</td>
<td>3.517</td>
<td>0.125</td>
<td>0.012</td>
<td>0.159</td>
<td>3.4</td>
</tr>
<tr>
<td>3D_C12 RNAi (5)</td>
<td>1.085</td>
<td>0.868</td>
<td>0.009</td>
<td>0.119</td>
<td>44.4</td>
</tr>
<tr>
<td>3D_C12 RNAi (6)</td>
<td>0.025</td>
<td>2.260</td>
<td>0.011</td>
<td>0.122</td>
<td>98.9</td>
</tr>
<tr>
<td>3D_C12 RNAi (7)</td>
<td>0.027</td>
<td>1.867</td>
<td>0.011</td>
<td>0.122</td>
<td>98.6</td>
</tr>
<tr>
<td>3D_C12 RNAi (8)</td>
<td>2.298</td>
<td>0.128</td>
<td>0.009</td>
<td>0.102</td>
<td>5.3</td>
</tr>
<tr>
<td>3D_C12 RNAi (9)</td>
<td>2.197</td>
<td>0.133</td>
<td>0.008</td>
<td>0.099</td>
<td>5.7</td>
</tr>
<tr>
<td>3D_C12 RNAi (10)</td>
<td>2.434</td>
<td>0.112</td>
<td>0.009</td>
<td>0.110</td>
<td>4.4</td>
</tr>
<tr>
<td>vector control (3)</td>
<td>1.811</td>
<td>1.735</td>
<td>0.018</td>
<td>0.170</td>
<td>48.9</td>
</tr>
<tr>
<td>vector control (11)</td>
<td>0.290</td>
<td>2.090</td>
<td>0.013</td>
<td>0.143</td>
<td>87.8</td>
</tr>
</tbody>
</table>

*percentage of leaf dry weight  
** [% nornicotine / (% nicotine + nornicotine)] × 100

Typical of line DH 98-325-6, ethephon treatment and curing resulted in substantial nornicotine production in the two control plants (48.9% and 87.8% conversion of nicotine to nornicotine). In dramatic contrast, seven of the ten independently transgenic plants possessing the 3D_C12/RNAi construct displayed minimal nicotine to nornicotine conversion, with conversion percentages ranging from 2.8 to 7.0 percent. The other three 3D_C12/RNAi lines displayed alkaloid contents similar to the vector-only control plants. Concentrations of the minor alkaloids anabasine and anatabine did not appear to be significantly influenced by the presence or absence of the 3D_C12/RNAi transgene (Table 4).

Although the cDNA insert of the 3D_C12-7 gene was used as the specific hybridization probe, at the hybridization and wash conditions used in this experiment, cross-hybridization to the entire 3D_C12 gene family would be expected. As shown in FIG. 5, a strong hybridization signal was detected in each plant showing a high nornicotine phenotype, and minimal hybridization was detected in the plants transformed with the 3D_C12/RNAi construct that showed a low nornicotine phenotype. We thus conclude that the effective silencing of the 3D_C12 gene family inhibits the metabolic conversion of nicotine to nornicotine in tobacco.

Alkaloid analysis of the Petite Havana transgenic plants is shown in Table 5. Four independently transformed plants containing the 35S:3D_C12 and 35S:3D_C12-7 constructs were tested along with seven independent 35S:7D_A06 individuals and three plants independently transformed with the pB121 control vector. As expected, the green, non-cured leaves of the three vector-only control plants contained minimal amounts of nornicotine. Likewise, all plants transformed with the 35S:3D_C12 and 35S:7D_A06 constructs showed minimal metabolic conversion of nicotine to nornicotine. A very different phenotype, however, was observed with plants transformed with 35S:3D_C12-7. All four plants independently transformed with this construct contained nornicotine as the predominant alkaloid in the green, non-treated leaf; nicotine to nornicotine conversion percentages ranged from 94.6 to 98.6.

### Table 5

<table>
<thead>
<tr>
<th>Sample</th>
<th>% Nicotine*</th>
<th>% Nornicotine*</th>
<th>% Anabasine*</th>
<th>% Anatabine*</th>
<th>% Conversion**</th>
</tr>
</thead>
<tbody>
<tr>
<td>vector control (2)</td>
<td>0.673</td>
<td>0.018</td>
<td>0.006</td>
<td>0.018</td>
<td>2.6</td>
</tr>
<tr>
<td>vector control (3)</td>
<td>0.695</td>
<td>0.017</td>
<td>0.005</td>
<td>0.016</td>
<td>2.3</td>
</tr>
<tr>
<td>vector control (10)</td>
<td>0.694</td>
<td>0.017</td>
<td>0.004</td>
<td>0.018</td>
<td>2.4</td>
</tr>
<tr>
<td>35S:3D_C12 (1)</td>
<td>0.706</td>
<td>0.005</td>
<td>0.006</td>
<td>0.020</td>
<td>0.7</td>
</tr>
<tr>
<td>35S:3D_C12 (2)</td>
<td>0.814</td>
<td>0.020</td>
<td>0.007</td>
<td>0.017</td>
<td>2.6</td>
</tr>
<tr>
<td>35S:3D_C12 (3)</td>
<td>0.650</td>
<td>0.010</td>
<td>0.003</td>
<td>0.012</td>
<td>1.6</td>
</tr>
<tr>
<td>35S:3D_C12 (4)</td>
<td>0.647</td>
<td>0.010</td>
<td>0.004</td>
<td>0.011</td>
<td>1.5</td>
</tr>
</tbody>
</table>
TABLE 5-continued

<table>
<thead>
<tr>
<th>Sample</th>
<th>% Nicotine*</th>
<th>% Nornicotine*</th>
<th>% Anabasine*</th>
<th>% Anatabine*</th>
<th>% Conversion**</th>
</tr>
</thead>
<tbody>
<tr>
<td>3SS:3D_C12-7 (1)</td>
<td>0.005</td>
<td>0.347</td>
<td>0.002</td>
<td>0.012</td>
<td>98.6</td>
</tr>
<tr>
<td>3SS:3D_C12-7 (2)</td>
<td>0.006</td>
<td>0.255</td>
<td>0.002</td>
<td>0.009</td>
<td>97.4</td>
</tr>
<tr>
<td>3SS:3D_C12-7 (3)</td>
<td>0.017</td>
<td>0.300</td>
<td>0.002</td>
<td>0.010</td>
<td>94.6</td>
</tr>
<tr>
<td>3SS:3D_C12-7 (4)</td>
<td>0.010</td>
<td>0.384</td>
<td>0.002</td>
<td>0.015</td>
<td>97.5</td>
</tr>
<tr>
<td>3SS:7D_A06 (1)</td>
<td>0.761</td>
<td>0.011</td>
<td>0.005</td>
<td>0.018</td>
<td>1.4</td>
</tr>
<tr>
<td>3SS:7D_A06 (2)</td>
<td>0.507</td>
<td>0.009</td>
<td>0.003</td>
<td>0.007</td>
<td>1.7</td>
</tr>
<tr>
<td>3SS:7D_A06 (4)</td>
<td>0.653</td>
<td>0.015</td>
<td>0.006</td>
<td>0.015</td>
<td>2.2</td>
</tr>
<tr>
<td>3SS:7D_A06 (5)</td>
<td>0.643</td>
<td>0.013</td>
<td>0.004</td>
<td>0.018</td>
<td>2.0</td>
</tr>
<tr>
<td>3SS:7D_A06 (6)</td>
<td>0.521</td>
<td>0.007</td>
<td>0.004</td>
<td>0.014</td>
<td>1.3</td>
</tr>
<tr>
<td>3SS:7D_A06 (7)</td>
<td>0.716</td>
<td>0.015</td>
<td>0.005</td>
<td>0.020</td>
<td>2.1</td>
</tr>
<tr>
<td>3SS:7D_A06 (8)</td>
<td>0.701</td>
<td>0.027</td>
<td>0.004</td>
<td>0.018</td>
<td>3.7</td>
</tr>
</tbody>
</table>

*Percentage of leaf dry weight
**% Nornicotine (% Nicotine + Nornicotine) × 100

Example 6

Cosuppression of the 3D_C12 Gene Family

In addition to the major conclusion that the 3D_C12-7 gene was capable of mediating nicotine to nornicotine conversion, one additional observation stood out in the alkaloid analyses of the Petite Havana transgenic plants. The alkaloid results reported in Table 5 together with additional alkaloid assays conducted independently (data not shown) consistently showed one of the plants transformed with the 3SS:3D_C12 construct (3SS:3D_C12(1)) as having less nornicotine in the green, nontreated leaf than any other plant in this study. This may be the result of cosuppression of the 3D_C12 gene family in this specific plant, a phenomenon frequently observed in transgenic plants even when a transgene is expressed in its sense orientation (Fagard and Vaucheret (2000) Annu. Rev. Plant Physiol. Plant Mol. Biol. 51: 167-194).

If plant 3SS:3D_C12 (1) was truly displaying a cosuppression phenotype, this phenotype would be expected to be maintained even upon ethephon treatment and curing of the leaves, similar to the low nornicotine phenotypes conferred by the 3D_C12 RNAi construct in the converter genotype DH 98-325-6 as described above. To test this prediction, alkaloid profiles were determined on ethephon treated, cured leaves of 3SS:3D_C12 (1) and two vector-only control plants. As shown in Table 6, ethephon treatment and curing resulted in over 97% nicotine to nornicotine conversion in the two control plants whereas similarly treated 3SS:3D_C12 (1) leaves displayed negligible conversion (0.6%). Leaves from five other plants expressing either 3SS:3D_C12 and 3SS:7D_A06 transgenes were also subjected to ethephon treatment and curing. In each case a high nornicotine phenotype was observed, similar to the vector-only control plants (data not shown).

Finally, Northern blot assays were conducted on select plants representing each of the Petite Havana transgenic genotypes (FIG. 6). Using a 3D_C12-7 cDNA as a hybridization probe, minimal signal was detected with RNAs isolated from green, nontreated leaves of the vector-only control plant. In contrast, hybridization was easily detected in RNA samples from all four independent transgenic plants possessing the 3SS:3D_C12-7 construct. A strong hybridization signal was similarly observed using RNAs from all other transgenic plants tested that were transformed with the 3SS:3D_C12 and 3SS:7D_A06 constructs, with the exception of the low nornicotine containing plant 3SS:3D_C12 (1).

Overall results of the Northern blotting assays show that the 3SS CaMV promoter was generally effective in mediating a high level of gene expression for each of the three members of the 3D_C12 gene family tested in this study. Failure to detect a hybridization signal in plant 3SS:3D_C12 (1) is con-

TABLE 6

<table>
<thead>
<tr>
<th>Sample</th>
<th>% Nicotine*</th>
<th>% Nornicotine*</th>
<th>% Anabasine*</th>
<th>% Anatabine*</th>
<th>% Conversion**</th>
</tr>
</thead>
<tbody>
<tr>
<td>vector control (8)</td>
<td>0.009</td>
<td>0.425</td>
<td>n.d.</td>
<td>0.011</td>
<td>97.9</td>
</tr>
<tr>
<td>vector control (10)</td>
<td>0.008</td>
<td>0.560</td>
<td>n.d.</td>
<td>0.025</td>
<td>98.6</td>
</tr>
<tr>
<td>3SS:3D_C12 (1)</td>
<td>1.185</td>
<td>0.007</td>
<td>n.d.</td>
<td>0.020</td>
<td>0.6</td>
</tr>
</tbody>
</table>

*Percentage of leaf dry weight
**% Nornicotine (% Nicotine + Nornicotine) × 100
n.d., not detected
sistent with the interpretation that the 3D_C12 gene family has been silenced via cosuppression in this individual.

Example 7

Additional Characterization and Suppression of Additional 3D_C12 Genes

A second RNAi construct was prepared using polynucleotide sequences from the 3D_C12-7 sequence. The assembly of the 3D_C12-7/RNAi expression cassette followed the same basic steps as those outlined for 3D_C12/RNAi above. Briefly, a 298-bp sense and antisense strand of the 3D_C12-7 cDNA (SEQ ID NO:5) corresponding to the region between nucleotide positions 297 and 594 of the coding sequence (positions 1-1551 of SEQ ID NO:5) were ligated into the pKLYX801 vector downstream and upstream of the 151-bp soybean omega-6 fatty acid desaturase intron (see GenBank Accession No. DQ672337), respectively. The primers (E4SFWd and E4SRev) used for the isolation of the 298-bp region by sense and antisense arms were 5'-AAGCTTGGCCGATTTCCCAAGCG-3' (SEQ ID NO:27), and 5'-CTCGAGGTTTCCAGCTATCATTAGA-3' (SEQ ID NO:28), respectively. The RNAi cassette was excised from pKLYX801 and placed between a strong CaMV35S promoter and a ribosuccinidoluburn terminator of the binary plant expression vector, pKLYX71 (see FIG. 8). In the discussions below, this RNAi construct is referred to as the 3D_C12-7-R29829 construct.

Transgenic tobacco plants were generated via Agrobacterium-mediated transformation following the procedures provided above. Briefly, transformed burley tobacco plants were regenerated from calli on Murashige-Skoog (MS) medium supplemented with 100 mg/L kanamycin and plant hormones in a growth room maintained at 25°C under a 16 hr/8 hr light/dark cycle. Calli were transferred to fresh selection media every 2-3 weeks until shoots appeared. Small shoots were transferred to rooting media to allow root development for 2 weeks. Fully regenerated plants were transferred to a greenhouse and grown under standard conditions.

SYBR® Green I Chemistry

Total RNA was isolated from cured leaves of convertor and nonconvertor burley tobacco plants using the TRIzol reagent (Invitrogen, Life Technologies, Carlsbad, Calif.). Purified RNA was treated with RNase-free DNase (TURBO DNA-free™, Ambion, Austin, Tex.). First strand cDNA was synthesized using 5 µg of total RNA and the StrataScript® First-Strand Synthesis System (Stratagene, Cedar Creek, Tex.). Relative quantitative RT-PCR was employed for determining the abundance of the 3D_C12-7 cDNA using SYBR® Green I fluorescence chemistry Morrison et al. (1998) Bio-techniques 24:954-962.

A calibration curve was generated with a serial dilution of the 3D_C12-7 cDNA cloned into the pGEM®-T Easy vector (Promega Corporation, Madison, Wis.). The RT-PCR mixture contained 2.5 mM MgCl₂, 125 µM each dNTP, 0.5 µM each primer, 0.5xSYBR® Green I, 0.5 µg cDNA (or 1 µl reference plasmid), and 1.25 U Platinum Taq polymerase (Invitrogen Life Technologies). The sequences of the allele-specific 3D_C12-7 primers (E4SFWd and E4SRrev) were 5'-ACGTGATCTCAAAACTCTGTCG-3' (E4SFWd (SEQ ID NO:29)) and 5'-GCCGACCCGTCCCTTCATG-3' (E4SRrev (SEQ ID NO:30)). RT-PCR was performed in a BioRad iCycler thermocycler (BioRad Laboratories, Hercules, Calif.) set to the following protocol: 95°C for 2 min; 35 cycles of 95°C for 30 sec, 55°C for 30 sec, and 72°C for 30 sec, followed by a final extension at 72°C for 5 min. A 165-bp fragment of the β-tubulin gene was used as an internal standard. 3D_C12-7 cDNA concentration was determined from the transcript-specific calibration curve and normalized to the internal standard. Fold-induction was calculated by dividing the normalized fluorescence values of the convertor by the nonconvertor samples. Melting-curve analysis was used to confirm the purity of PCR products as described in Ririe et al. (1997) Anal. Biochem. 245:154-160. Two plants were sampled for treatment and amplifications were repeated three times.

Taqman® Chemistry

Total RNA was isolated from tobacco leaves using TRIzol reagent. Purified RNA was treated with RNase-free DNase (TURBO DNA-free™). First strand cDNA was synthesized using 10 µg of total RNA and the High Capacity cDNA Archive Kit (Applied Biosystems, Foster City, Calif.). The RT-PCR mixture contained 1x TaqMan® Universal PCR Master Mix (Applied Biosystems, Foster City, Calif.), 400 nM of each primer (E4TmFWd and E4TmRev), 250 nM TaqMan® minor groove binder (MGB) probe (E4MGB), 2 ng of cDNA, and nuclelease-free water (Afonina et al. (2002) Bio-techniques 32:940-949). The primer and probe sequences were 5'-CCGTAATCGCCGATCTTTTC-3' (E4TmFWd (SEQ ID NO:31)), 5'-CCGAGTTTCCAGCTAATGGA-3' (E4TmRev (SEQ ID NO:32)), and 5'-CAATGACGAGCAGCACAG-3' (MGB probe (SEQ ID NO:33)). RT-PCR was performed in an ABI 7500 Real-Time System (Applied Biosystems, Foster City, Calif.) set to the following protocol: 50°C for 2 min; 95°C for 10 min; 40 cycles of 95°C for 15 sec, 60°C for 1 min. Glyceraldehyde-3-phosphate dehydrogenase (G3PDH) was used as the endogenous control to normalize the amount of cDNA template in the reactions. Fold-change was determined by dividing the normalized fluorescence values of each sample by those obtained from a nonconvertor or uncontrolled sample control. For each treatment, RNA was isolated from three independent plants and amplifications were repeated 3 times per RNA sample.

Northern and Southern Blot Analyses

Total RNA was isolated from cured tobacco leaves using the TRIzol reagent according to the manufacturer’s instructions (Invitrogen, Life Technologies). Total RNA samples were separated on 1.2% TBE agarose gel, and transferred to positively charged nylon membranes by electroblotting with 2xTBE buffer. Membranes were UV crosslinked and washed in 2xSCC for 5 min. Northern blot hybridization, washing, and detection were carried out using the digoxigenin (DIG) System as described by the manufacturer (Roche Diagnostics Corp., Indianapolis, Ind.). The 1.8 kb full-length ORF of the 3D_C12-7 cDNA was labeled with DIG and used as a probe.

Genomic DNA was extracted with DNAzol® (Invitrogen, Life Technologies) from green tobacco leaves according to manufacturer’s protocol. After incubation with FcIR or Ncol restriction enzymes overnight, 15 µg of the digested DNA was separated on 0.7% TBE agarose gel, depurinated with 0.25 M HCl for 10 min, and denatured with 0.5 N NaOH for 30 minutes. DNA was blotted overnight by capillary transfer onto positively charged nylon membranes (Roche Diagnostics Corp.) and hybridized at 65°C overnight with a 515-bp DIG-labeled fragment of the neomycin phosphotransferase II (NPT II) gene. Hybridization, washing, and detection were performed according to the protocols supplied with the DIG System. The primers used for the amplifications of the Northern and Southern hybridization probes were E4F1FWd (5'-TATGGTTTCCATCAAGGC-3' (SEQ ID NO:34)), E4F1Rev (5'-TTTGTGAGCACAATCATGGC-3' (SEQ ID NO:35)). KanFWd (5'-TGGATATGCAACTGCAAG-
Alkaloid Analysis

Tobacco leaves were harvested and air dried in an oven at 50° C. for 2 days. A 100 mg sample of crushed, dried leaf is added to 0.5 ml of 2 N NaOH in a 20 ml scintillation vial. The sample was mixed and allowed to incinate for 15 minutes at room temperature. Alkaloids were extracted by the addition of 5 ml of extraction solution [0.04% quinoline (wt/vol) dissolved in methyl-β-d-methyl ether] and gently rotated on a linear shaker for 3 hours. Following phase separation, an aliquot of the organic phase was transferred to a sample vial. Samples were analyzed using a Perkin-Elmer Autosystem XL (Perkin-Elmer, Boston, Mass.) gas chromatograph equipped with a flame ionization detector, a 4 mm split/splitless glass liner and a 30 m x 0.53 mm ID DB-5 column. Chromatographic conditions were as follows: detector temperature: 250° C.; injector temperature: 250° C.; helium flow at 120° C.: 20 ml/min; injection volume: 2 μl; column conditions: 120°, hold 1 min, 20-280° C. ramping rate, hold at 280° C., for 2 min. Alkaloid composition was determined by the Total-Chrome Navigator software using a calibration curve. Means of the alkaloid measurements were separated according to Fisher's Protected LSD (PROC MIXED).

Plants

Double haploid burley tobacco lines DH 98-325-5 (325-5; nonconverter) and DH98-325-6 (325-6; converter) described above were used in all experiments, except for the fluorogenic 5' nucleases (TaqMan®) chemistry-based RT-PCR assays where the isogenic DH 91-1307-46 (nonconverter) and DH 91-1307-46 (converter) lines were used. All plants were grown in a controlled environment greenhouse equipped with supplemental light-providing, a 14 hr/10 hr light/dark cycle.

For curing, tobacco leaves were collected from converter and nonconverter plants about 1 month before flowering and treated by dipping each leaf twice for 10 sec, into 2% ethephon and dried for 2 hours. Leaves were cured for up to 2 weeks in plastic bags, under dark conditions, until they turned yellow. Cured leaves were used for the Northern and alkaloid analyses. Samples of cured leaves subjected to GC analysis were dried at 50° C. for 2 days. For Southern analysis, green tobacco leaves of adult plants were used. To produce T₅ generation transgenic plants, primary transformants (T₀) were self-pollinated, and the harvested T₅ seed was screened by germinating seedlings on MSagar plates containing 100 mg/L kanamycin for 6 weeks. Survivors were transplanted to soil and grown in a greenhouse as described above. Plants were fertilized with Peter's Professional All Purpose Plant Food (20-20-20; Spectrum Brands Inc., Madison, Wis.) once a week.

RT-PCR Analysis of 3D_C12-7 Expression in Converter and Nonconverter Tobacco

To further characterize the role of 3D_C12-7 in nicotine N-demethylation, experiments were performed to demonstrate that the regulation of 3D_C12-7 expression is consistent with the levels of nicotine N-demethylation activity observed in converter versus nonconverter tobacco.

To determine the rate of 3D_C12-7 mRNA accumulation in converter and nonconverter tobacco, an allele-specific real-time RT-PCR strategy was employed. Because RT-PCR involves the detection and measurement of the amplification products of a PCR template, the use of allele-specific primers allows the quantification of a single isoform among a group of highly homologous sequences. For accurate quantification of the 3D_C12-7 transcript, two different segments of the 3D_C12-7 coding region were amplified and both SYBR® Green and TaqMan® chemistries were used to generate fluorescence signals. RT-PCR analysis using the SYBR® Green 1 chemistry revealed an 80-fold increase in the levels of the 3D_C12-7 transcript in the cured leaves of converter versus nonconverter tobacco. A single peak melting curve and gel electrophoretic analyses of the amplicons confirmed the homogeneity of the PCR products.

In the TaqMan® chemistry-based RT-PCR experiment, 3D_C12-7 transcript levels were quantified in untreated and ethephon-treated converter and nonconverter tobacco leaves that were cured for 0, 1 or 5 days. Low levels of 3D_C12-7 transcripts were detected in the uncured leaves or following a 1-day curing period regardless of conversion type or ethephon treatment. Similarly, base line levels of 3D_C12-7 transcription were observed in converter or nonconverter leaves that were cured for 5 days without ethephon treatment. In contrast, a 7.5-fold increase in 3D_C12-7 transcript accumulation was detected in the cured leaves of converter versus nonconverter tobacco, and a 70-fold increase was observed in the uncured versus cured leaves of a converter tobacco variety when ethephon treatment preceded the 5-day curing period. While not intending to be limited by any particular theory, these results suggest that 3D_C12-7 is a major contributor to nicotine N-demethylation and is strongly inducible by ethylene in senescing tobacco leaves.

Suppression of Nicotine to Nornicotine Conversion by the 3D_C12-R199 and 3D_C12-7-R1298 Constructs

To compare the extent to which 3D_C12 and 3D_C12-7 mediate the suppression of nornicotine production, converter and nonconverter burley tobacco plants were transformed with the two gene silencing vectors. Ten (10) transgenic plants were regenerated per RNAi construct. About 80% of tobacco plants overexpressing either the 99-bp or 298-bp inverted repeat showed reduced nicotine levels compared to the empty vector controls (Tables 7 and 8). In the nonconverter genotype, 3D_C12-R199 and 3D_C12-7-R1298 expression reduced nicotine to nornicotine conversion by about 1.8-fold (2.0%) and 3.0-fold (1.2%), respectively, in comparison to the rate of conversion detected in the vector controls (3.6%) (Table 7). Among the silenced nonconverter plants, the lowest conversion level of 0.9% was achieved using the 3D_C12-7-R1298 construct (Table 7).

<table>
<thead>
<tr>
<th>Line</th>
<th>% Nicotine</th>
<th>% Nornicotine</th>
<th>% Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D_C12-R199</td>
<td>1.693</td>
<td>0.034</td>
<td>2.0</td>
</tr>
<tr>
<td>2</td>
<td>1.435</td>
<td>0.031</td>
<td>2.1</td>
</tr>
<tr>
<td>3</td>
<td>2.095</td>
<td>0.043</td>
<td>2.0</td>
</tr>
<tr>
<td>4</td>
<td>2.868</td>
<td>0.053</td>
<td>1.8</td>
</tr>
<tr>
<td>5</td>
<td>0.947</td>
<td>0.025</td>
<td>2.6</td>
</tr>
<tr>
<td>6</td>
<td>2.357</td>
<td>0.043</td>
<td>1.8</td>
</tr>
<tr>
<td>7</td>
<td>2.599</td>
<td>0.043</td>
<td>1.6</td>
</tr>
<tr>
<td>8</td>
<td>0.796</td>
<td>0.020</td>
<td>2.4</td>
</tr>
<tr>
<td>9</td>
<td>2.178</td>
<td>0.039</td>
<td>1.8</td>
</tr>
<tr>
<td>10</td>
<td>3.162</td>
<td>0.061</td>
<td>1.9</td>
</tr>
<tr>
<td>MEAN</td>
<td>2.013</td>
<td>0.039</td>
<td>2.0a</td>
</tr>
<tr>
<td>STDEV</td>
<td>0.748</td>
<td>0.012</td>
<td>0.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Line</th>
<th>% Nicotine</th>
<th>% Nornicotine</th>
<th>% Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D_C12-7-R1298</td>
<td>1.806</td>
<td>0.020</td>
<td>1.1</td>
</tr>
<tr>
<td>4</td>
<td>1.948</td>
<td>0.207</td>
<td>1.4</td>
</tr>
</tbody>
</table>
### TABLE 7-continued

<table>
<thead>
<tr>
<th>Line</th>
<th>% Nicotine</th>
<th>% Nornicotine</th>
<th>% Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2.061</td>
<td>0.020</td>
<td>1.0</td>
</tr>
<tr>
<td>6</td>
<td>2.704</td>
<td>0.040</td>
<td>1.5</td>
</tr>
<tr>
<td>8</td>
<td>2.652</td>
<td>0.023</td>
<td>0.9</td>
</tr>
<tr>
<td>9</td>
<td>1.074</td>
<td>0.055</td>
<td>1.3</td>
</tr>
<tr>
<td>MEAN</td>
<td>2.041</td>
<td>0.024</td>
<td>1.2b</td>
</tr>
<tr>
<td>STE</td>
<td>0.530</td>
<td>0.008</td>
<td>0.2</td>
</tr>
</tbody>
</table>

**Vector Control**

<table>
<thead>
<tr>
<th>Line</th>
<th>% Nicotine</th>
<th>% Nornicotine</th>
<th>% Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.260</td>
<td>0.052</td>
<td>4.2</td>
</tr>
<tr>
<td>2</td>
<td>1.265</td>
<td>0.058</td>
<td>2.9</td>
</tr>
<tr>
<td>3</td>
<td>1.752</td>
<td>0.058</td>
<td>3.2</td>
</tr>
<tr>
<td>4</td>
<td>1.230</td>
<td>0.072</td>
<td>5.6</td>
</tr>
<tr>
<td>5</td>
<td>1.777</td>
<td>0.060</td>
<td>3.3</td>
</tr>
<tr>
<td>6</td>
<td>1.536</td>
<td>0.044</td>
<td>2.8</td>
</tr>
<tr>
<td>7</td>
<td>1.461</td>
<td>0.054</td>
<td>3.5c</td>
</tr>
<tr>
<td>8</td>
<td>0.240</td>
<td>0.011</td>
<td>1.0</td>
</tr>
</tbody>
</table>

**Table footnotes:**

1. Tobacco leaves were treated with ethylone and cured for 2 weeks at 25°C.
2. Of the plants transformed with an RNAi construct, only silenced individuals are shown. Alkaloid data represent the means of 2 measurements.
3. Numbers represent independent experiments.
4. Percentage of leaf dry weight.
5. Numbers represent independent experiments. Significant differences between lines were determined using Student's t-test.
6. Tobacco plants transformed with only pXYLX71 vector were used as controls.

Relative to nonconverter tobacco, nornicotine accumulation was suppressed even more dramatically in the silenced individuals of the strong converter plants (Table 8). Using 3D.C12-7-R199 constructs, nicotine conversion was reduced to levels as low as 4.5% in 3D.C12-R199-transformed 325-6 tobacco plants in sharp contrast to the 325-6 control plants exhibiting about 98% conversion rates (Table 8). However, using the 3D.C12-7-R298 construct even greater reductions in nicotine conversion were obtained (Table 8). Four 3D.C12-7-R298-transformed individuals converted as low as 0.8% of their nicotine to nornicotine, and the arithmetic mean across the 9 silenced transformants was 0.9% conversion. All silenced plants were morphologically indistinguishable from both the empty vector and wild-type controls (data not shown).

### TABLE 8

<table>
<thead>
<tr>
<th>Line</th>
<th>% Nicotine</th>
<th>% Nornicotine</th>
<th>% Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D.C12-7-R199</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3.419</td>
<td>0.100</td>
<td>2.8</td>
</tr>
<tr>
<td>2</td>
<td>2.569</td>
<td>0.193</td>
<td>7.0</td>
</tr>
<tr>
<td>3</td>
<td>2.175</td>
<td>0.064</td>
<td>2.9</td>
</tr>
<tr>
<td>4</td>
<td>3.517</td>
<td>0.125</td>
<td>3.4</td>
</tr>
<tr>
<td>5</td>
<td>2.268</td>
<td>0.128</td>
<td>5.3</td>
</tr>
<tr>
<td>6</td>
<td>2.197</td>
<td>0.133</td>
<td>5.7</td>
</tr>
<tr>
<td>7</td>
<td>2.434</td>
<td>0.112</td>
<td>4.4</td>
</tr>
<tr>
<td>8</td>
<td>2.197</td>
<td>0.133</td>
<td>5.7</td>
</tr>
<tr>
<td>MEAN</td>
<td>2.654</td>
<td>0.122</td>
<td>4.5a</td>
</tr>
<tr>
<td>STE</td>
<td>0.573</td>
<td>0.039</td>
<td>1.6</td>
</tr>
</tbody>
</table>

### TABLE 8-continued

<table>
<thead>
<tr>
<th>Line</th>
<th>% Nicotine</th>
<th>% Nornicotine</th>
<th>% Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1b98-325-5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.764</td>
<td>0.024</td>
<td>1.4a</td>
</tr>
<tr>
<td>STE</td>
<td>0.456</td>
<td>0.004</td>
<td>0.3</td>
</tr>
</tbody>
</table>

### Table footnotes:

1. Tobacco leaves were treated with ethylone and cured for 2 weeks at 25°C.
2. Of the plants transformed with an RNAi construct, only silenced individuals are shown. Numbers represent independent experiments.
3. Percentage of leaf dry weight.
4. Numbers represent independent experiments. Significant differences between lines were determined using Student’s t-test.
5. Tobacco plants transformed with only pXYLX71 vector were used as controls.

To test the heritability of nornicotine suppression in the 3D.C12-7-R298-transformed plants, a set of 3D.C12-7-R298-transformed converter and nonconverter lines that displayed the lowest levels of nicotine conversion were advanced to the T1 generation (Table 9). Because segregation of the transgene(s) occurs in the T1 progeny, transgenic individuals were identified by selecting seedlings capable of growing on kanamycin-containing media. Nine kanamycin-resistant progenies of each selected T1 generation 3D.C12-7-R298 transformant and four kanamycin-resistant individuals from each selected vector control line were analyzed for alkaloid content. The rate of nicotine conversion did not differ significantly between the primary 3D.C12-7-R298 transformants and their T1 progeny, indicating high heritability of the nornicotine suppression trait (see Tables 7, 8, and 9). However, advancing the “nonconverter” vector control line by a single generation increased the nicotine to nornicotine conversion rate from 4.2% to an average value of 11.6%, illustrating the high degree of instability of the conversion locus in transgenic plants lacking the 3D.C12-7-R298-specific RNAi construct (Tables 7 and 9). Overall, these results show that RNAi-mediated silencing of the 3D.C12 gene subfamily is a highly effective means of lowering nornicotine production in both nonconverter and strong converter tobacco plants.

### TABLE 9

<table>
<thead>
<tr>
<th>Line</th>
<th>% Nicotine</th>
<th>% Nornicotine</th>
<th>% Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D.C12-7-R298#3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.500</td>
<td>0.020</td>
<td>1.3a</td>
</tr>
<tr>
<td>STE</td>
<td>0.306</td>
<td>0.006</td>
<td>0.3</td>
</tr>
</tbody>
</table>

### Table footnotes:

1. Tobacco leaves were treated with ethylone and cured for 2 weeks at 25°C.
2. Of the plants transformed with an RNAi construct, only silenced individuals are shown. Numbers represent independent experiments.
3. Percentage of leaf dry weight.
4. Numbers represent independent experiments. Significant differences between lines were determined using Student’s t-test.
5. Tobacco plants transformed with only pXYLX71 vector were used as controls.
TABLE 9-continued

<table>
<thead>
<tr>
<th>Line</th>
<th>% Nicotine</th>
<th>% Nornicotine</th>
<th>% Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D_C12-T-7-R29888</td>
<td>1.772</td>
<td>0.020</td>
<td>1.2a</td>
</tr>
<tr>
<td>Mean</td>
<td>1.466</td>
<td>0.203</td>
<td>11.6b</td>
</tr>
<tr>
<td>Vector Control1*</td>
<td>0.409</td>
<td>0.003</td>
<td>0.3</td>
</tr>
<tr>
<td>3D_C12-T-7-R298882</td>
<td>1.623</td>
<td>0.022</td>
<td>1.3a</td>
</tr>
<tr>
<td>Mean</td>
<td>1.419</td>
<td>0.017</td>
<td>1.3a</td>
</tr>
<tr>
<td>Vector Control2*</td>
<td>0.515</td>
<td>0.004</td>
<td>0.3</td>
</tr>
<tr>
<td>3D_C12-T-7-R2988810</td>
<td>0.713</td>
<td>0.161</td>
<td>9.7</td>
</tr>
<tr>
<td>Mean</td>
<td>0.028</td>
<td>1.170</td>
<td>97.6c</td>
</tr>
<tr>
<td>Vector Control2*</td>
<td>0.006</td>
<td>0.234</td>
<td>0.5</td>
</tr>
</tbody>
</table>

1Tobacco leaves were treated with ethylene and cured for 2 weeks at 25°C.
2Means and standard errors (STE) (mean ± SD) of three replications (n = 3).
3Percentage of total dry weight.
4%Nornicotine (%nicotine + % nornicotine) is > 100; values followed by different letters are significantly different according to Fisher’s Protected LSD (0.05).

Furthermore, transforming tobacco with the 3D_C12-T-7-R298 construct conferred a 3.6-fold reduction in nicotine conversion relative to typical nonconverter control plants without affecting plant growth and development.

To demonstrate that the down-regulation of nornicotine production in 3D_C12-T-7-R298-transformed tobacco was concomitant with a reduction of the 3D_C12 gene subfamily transcripts, a 3D_C12-T cDNA probe was hybridized to the total RNA isolated from cured leaves of nonconverter and converter plants. A weak hybridization signal was generated by the RNA isolated from 3D_C12-T-7-R298 transformants displaying low nornicotine content in contrast to the strong signal produced by the RNA extracted from plasmid control or wild-type plants. These results indicate that the down-regulation of nicotine conversion was a result of RNA-mediated gene silencing of the nicotine N-demethylase gene(s).

Determination of Transgene Copy Number

To determine whether the integration of multiple 3D_C12-T-7-R298 copies were required for producing transplants displaying very low nicotine N-demethylase activity, Southern analysis was performed on selected individuals exhibiting <1.5% nornicotine accumulation. The number of transgenes varied widely among these plants including individuals containing 1 copy (325-5, lines 5 & 8; 325-6, lines 2 & 8), 5 copies (325-6, line 10), and 6 copies (325-5, line 6) of the 3D_C12-T-7-R298 construct using Southern blot analysis of genomic DNA digested with EcoRI restriction enzyme. Transgene copy number was confirmed using Neol digested DNA (data not shown). These results indicate that the integration of a single 3D_C12-T-7-R298 construct into the genome of a strong converter tobacco is sufficient for suppressing nornicotine production to very low levels.

The analyses outlined in Examples 1-6 above resulted in the discovery of a closely related P450 gene family, designated the 3D.C12 family, whose collective steady-state transcript levels were significantly elevated in converter tobacco plants that were actively metabolizing nicotine to nornicotine in comparison to their nonconverter counterparts. Transgenic plant analysis demonstrated that the suppression of gene expression of this P450 family in converter tobacco lines inhibited the metabolism of nicotine to nornicotine to levels similar to that observed in nonconverter plants. Furthermore, sense expression of several individuals of this closely related gene family identified one member, designated 3D.C12-7, as playing a direct role in the metabolic conversion of nicotine to nornicotine. Overexpression of 3D.C12-7 using a strong constitutive promoter caused a dramatic increase in nornicotine production and accumulation in non-cured green leaves of transgenic tobacco plants, a tissue where nicotine is normally the predominant alkaloid in converter and nonconverter plants alike. Given that the cytochrome P450 family member designated 3D.C12-10 differs from 3D.C12-7 at only two amino acid residues immediately following the start methionine and within the N-terminal signal sequence, it is predicted that these encoded products function identically.

The contrast in alkaloid phenotypes between the 35S: 3D.C12 (1) plant and vector-only control plants was most dramatic in leaves that had been ethylene treated and cured (0.6% conversion vs. 97% conversion; Table 6). However, it is noteworthy that the nornicotine content of the co-suppressed 35S:3D.C12 (1) plant was reduced even in green, non-treated leaves where the high nornicotine phenotype is typically not manifest in converter or nonconverter tobacco lines. The green, non-treated leaves of line 35S:3D.C12 (1) showed only 0.7% nicotine to nornicotine conversion, whereas every other plant in this experiment showed conversion percentages ranging from 1.3 to 3.7 (Table 5). This result suggests that the inhibition of gene expression of the 3D.C12 family may prove to be effective in the further lowering of nornicotine levels even in tobacco lines where genetic conversion isn’t typically a major problem (such as flue-cured tobaccos) or in the nonconverter individuals in lines that are prone to genetic conversion (such as Burley tobaccos).

Southern blotting assays using members of the 3D.C12 gene family as hybridization probes gives very complex banding patterns, suggesting that more members of this gene family may exist even beyond those that have been identified and characterized herein (data not shown). The hypothesis that the 3D.C12 gene family is comprised of additional members is further supported by the recent publication of 75 full-length tobacco P450 cDNAs of unknown function (US Patent Application Publication 20040162420). Within this list of P450s are additional cDNAs that would, based on the work described herein, be placed within the 3D.C12 family in view of their display of over 90% amino acid sequence identity to the protein sequences shown in FIG. 4.

With respect to the specific molecular function of the 3D.C12-7 gene or the nearly identical family member 3D.C12-10, it is possible that it encodes the actual nicotine demethylase enzyme which catalyzes the oxidative N-demethylation of nicotine to nornicotine (FIG. 1). Alternatively, the 3D.C12-7 encoded enzyme or nearly identical 3D.C12-10 encoded enzyme may produce a product that leads to the up-regulation of the nicotine demethylase activity of the leaf, as opposed to directly catalyzing the N-demethylation reaction.
In addition, an allele-specific RT-PCR was employed to compare 3D_C12-7 expression between converter and non-converter plants (Example 7). An approximately 80-fold increase in 3D_C12-7 expression in converter versus nonconverter plants was identified using the SYBR® Green-chemistry RT-PCR assay. A 7.5-fold up-regulation was identified by the TaqMan® chemistry-based RT-PCR experiment. While the DH 91-1307-46 tobacco variety used in the TaqMan® chemistry-based RT-PCR experiment exhibits low to moderate levels of nicotine conversion, the DH98-325-5 nonconverter plants used in the SYBR® Green-based RT-PCR assay consistently convert a very low percentage of their nicotine to nornicotine. Expression of the 3D_C12-7 gene was induced at least 7-fold by ethylene in senescing leaves of converter tobacco plants.

An additional RNAi construct, 3D_C12-7-Ri298, was prepared based on a region of the 3D_C12-7 polynucleotide that corresponds to nucleotide positions 297 through 594 of SEQ ID NO.5. Expression of this RNAi construct allows for the suppression of nornicotine production in a strong converter tobacco line below the levels normally found in nonconverter plants. The expression cassette of the 3D_C12-7-Ri298 construct encoded an intron-spliced hairpin RNA in which the stem region was engineered from this 298-bp fragment of the 3D_C12-7 cDNA inserted as an inverted repeat. The loop of the hairpin was created by placing a 151-bp intron of the FAD gene between the two sides of the palindromic sequences. An arm length of 298-bp was used for the inverted repeats.

3D_C12-7-Ri298-transformed plants accumulated less nornicotine than those harboring the 3D_C12-Ri99 construct (Tables 7 and 8). No correlation was found between the number of copies of the 3D_C12-7-Ri298 construct and nornicotine production (Tables 7 and 8). The 3D_C12-7-Ri298 expression cassette enabled the production of tobacco with a conversion rate as low as 0.8%, which is below the 3.5% rate detected in burley lines used by seed producers. Such dramatic reduction in nornicotine production by targeting this particular region of the 3D_C12-7 polynucleotide is an unexpected result. Also, suppression of nornicotine production showed a high degree of heritability in the T1 progeny of the primary transformants (Table 9). Suppression of nornicotine production in these transgenic plants yielded no obvious differences in growth and development when compared to wild-type plants.

All publications and patent applications mentioned in the specification are indicative of the level of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 37
<210> SEQ ID NO 1
<211> LENGTH: 1733
<212> TYPE: DNA
<213> ORGANISM: Nicotiana sp.
<220> FEATURE: .
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(1551)
<400> SEQUENCE: 1

atg gtt ttc ccc ata gaa gcc ttc gta gga cta cca acc ttc aca ttt
Met Val Phe Pro Ile Glu Ala Phe Val Gly Leu Val Thr Phe Thr Phe 1 5 10 15

ctc tta tac ttc cta tgg aca aaa aaa aag caa cca aaa ccc
Leu Leu Tyr Phe Leu Trp Thr Lys Ser Glu Leu Pro Lys Pro 20 25 30

cca cca cca gga gga tgt ccc cgg gta acc ggc cat ctt ttc
Leu Pro Lys Ile Pro Gly Gly Gly Trp Pro Val Ile Gly His Leu Phe 35 40 45

ccc ccc ccc gaa gag gcg agg cgg ctt cag ctc tgg cca gaa aag ctc gga
His Phe Aam Aam Aam Gly Aam Arg Arg Pro Leu Ala Arg Lys Leu Gly 50 55 60

agc tta gct gat aac tac ggg ctt ctt act ctt ccc ctt gct cca gaa aag ctc gga
Amp Leu Ala Aam Lys Tyr Gly Pro Val Phe Thr Phe Arg Leu Gly Leu 65 70 75 80

ccc ctt gtt ctt gta ggt aag ctt ttc gag ctt cca gaa aag gat tgc ttc
Pro Leu Val Leu Val Ser Ser Tyr Glu Ala Ile Lys Arg Cys Phe 85 90 95

tct ccc aat gac ggc att ttt tcc aat cct cca gct ttg ctt tac ggc
Ser Thr Aam Aam Ala Ile Phe Ser Ser Arg Pro Ala Phe Teur Gly 100 105 110

<215>
gaa tac ctt ggc tac aat aat aca atg ctt ttt cta gca aat tac gga
Glu Tyr Leu Gly Tyr Asn Asn Thr Met Leu Phe Leu Ala Asn Tyr Gly
115 120 125

cct tac tgg cga aaa aat cgt aaa tta gtc att cag gaa gtt ctc tct
Pro Tyr Trp Arg Lys Asn Arg Lys Leu Val Ile Gln Glu Val Leu Ser
130 135 140

gtt agt cgt ctc gaa aaa ttc aaa caa gta aga ttc acc aga att caa
Ala Ser Arg Leu Gly Lys Pro Arg Phe Thr Arg Ile Gln
145 150 155 160

acg acg att aag aat tta tac act cga att aat gga aat tcg agt acg
Thr Ser Ile Lys Asn Leu Tyr Thr Arg Ile Asn Gly Asn Ser Ser Thr
165 170 175

ata aat cta act gat tgg tta gaa gaa tgg aat ttt ggt ctc aag gtt
Ile Asn Leu Thr Asp Leu Glu Leu Glu Asp Asp Glu Ile Val
180 185 190

aaa atg atc gct ggg aaa aat tat gaa tcc ggt aaa gga gat gaa cca
lys Met Ala Gly Lys Asn Tyr Glu Ser Gly Lys Gly Asp Glu Gln
195 200 205

gtg gaa gaa att tgg aag aat gsg ttt aag gat ttt atc gtt tta tca atg
Val Glu Arg Phe Lys Ala Phe Asp Phe Met Val Leu Ser Met
210 215 220

gaa ttt gta tta tgg gat gca tta cca att cca tta ttt gaa tgg gtt
Glu Phe Val Leu Trp Asp Ala Phe Pro Ile Pro Leu Phe Lys Trp Val
225 230 235 240

gat ttt cca ggt cat att aag gca atg aaa agg cca ttt aag gat ata
Asp Phe Glu Gly His Ile Lys Ala Met Lys Arg Thr Phe Lys Asp Ile
245 250 255

gtt tct gtt ttt cag aac tgg tta gag gaa gat cat att aat aat aac aag
Asp Ser Val Phe Glu Asn Trp Leu Glu Glu His Ile Asn Leu Arg Glu
260 265 270

aaa atg gag gtt ggt gca gaa ggg aat gaa cca gat tcc att gat gtt
lys Met Glu Val Gly Ala Gly Asn Glu Gin Asp Phe Ile Asp Val
275 280 285

gtg ctt tca aaa ttt gtt gaa taa gat cct gat ggg ttt cct gat ctc tct
Val Leu Ser Leu Leu Ser Lys Tyr Leu Asp Gln Gly Tyr Ser Arg
290 295 300

gat act gtc att aaa gca aca gct ttt gtt gtt gtc tgg gat gca gca
Asp Thr Val Ile Lys Thr Val Phe Ser Leu Val Leu Asp Ala Ala
305 310 315 320

gac aca gtt gct ctc cac ata aat tgg ggs atg aca tta tgg atg ata aac
Asp Thr Val Ala Leu His Ile Asn Trp Gly Met Thr Leu Ile Ann
325 330 335

aat cca aat gac ttg atg aaaa gca cca gaa gag ata gac aca aac gtt
Asn Glu Asn Ala Leu Asp Gly Glu Glu Ile Asp Thr Tyr Val
340 345 350

ngt aag gat aga tgg gta gaa gag agt gat att aag gat tta gta tac
Gly Lys Asp Arg Trp Val Glu Glu Ser Asp Ile Lys Asp Leu Val Tyr
355 360 365

cct caa gct att gtt aaa aag gtt tta cga tta tat cca cca gga ctt
leu Glu Ala Ile Val Lys Val Leu Arg Leu Tyr Pro Pro Gly Pro
370 375 380

ttg tta gta cca cat gaa aat gta aag gat tgg gat ggg ttt gtt gtt
leu Leu Val Pro His Glu Asn Val Lys Asp Cys Val Ser Gly Tyr
385 390 395 400

cac att cct aaaa gag act aga tta ttc gca aac gtc atg aaaa ctc cag
His Ile Pro Lys Thr Arg Leu Phe Ala Asn Val Met Lys Leu Gln
405 410 415

cgc gat cct aat ctc ttc cta aat cct gat aag tcc gat cca gag aca
Arg Asp Pro Lys Leu Leu Ser Asn Pro Asp Lys Phe Asp Pro Glu Arg
420 425 430
ttc atc gct ggt gat att gac ttc ggt cac cac tat gag ttt atc
Phe Ile Ala Gly Asp Ile Asp Phe Arg Gly His His Tyr Glu Phe Ile
435 440 445

cca ttt ggt tot gga aga cga ctc tct cgc ggg att act tat gca tgg
Pro Phe Gly Ser Gly Arg Arg Ser Cys Pro Gly Met Thr Tyr Ala Leu
450 455 460

caa gtc gaa cac cta aca atg gca cat tta atc cag gtt ttc aat tac
Gln Val Glu His Leu Thr Met Ala His Leu Ile Glu Gly Phe Asn Tyr
465 470 475 480

aaa act cca aat gac gag gcc tgg gat atg aag gaa ggt gca gcc ata
Lys Thr Pro Asn Asp Glu Ala Leu Asp Met Lys Glu Gly Ala Gly Ile
485 490 495

acc ata cgt aag gta aat cca gtt gaa tgg ata ata acc cct cgc tgg
Thr Ile Arg Lys Val Asn Pro Val Glu Leu Ile Thr Pro Arg Leu
500 505 510

gca cct gag ctt tac taaaacctaa gatcttttcat cttgggtgat cattgtttaa
Ala Pro Glu Leu Tyr
615

tacttctaga tgggatattca tttacctttt ttcaatattg tgcagtttacy agcttctttsa
1651

atttggtata ttggtacac taagtaaaga argattgfgc tastatataa aagattgcaag
1711

agataatag actgattgtgc cc
1733

<210> SEQ ID NO 2
<211> LENGTH: 517
<212> TYPE: PRT
<213> ORGANISM: Nicotiana sp.

<400> SEQUENCE: 2

Met Val Phe Pro Ile Glu Ala Phe Val Gly Leu Val Thr Phe Thr Phe
1 5 10 15

Leu Leu Tyr Phe Leu Trp Thr Lys Ser Gln Lys Leu Pro Lys Pro
20 25

Leu Pro Pro Lys Ile Pro Gly Gly Trp Pro Val Ile Gly His Leu Phe
35 40 45

His Phe Asn Asn Asp Gly Asp Asp Arg Pro Leu Ala Arg Lys Leu Gly
50 55 60

Asp Leu Ala Asp Lys Tyr Gly Pro Val Phe Thr Phe Arg Leu Gly Leu
65 70 75 80

Pro Leu Val Leu Val Val Ser Ser Tyr Gly Ala Ile Lys Asp Cys Phe
85 90 95

Ser Thr Asn Asp Ala Ile Phe Ser Asn Arg Pro Ala Phe Leu Tyr Gly
100 105 110

Glu Tyr Leu Gly Tyr Asn Asn Thr Met Leu Phe Leu Ala Asn Tyr Gly
115 120 125

Pro Tyr Trp Arg Lys Asn Arg Lys Leu Val Ile Glu Glu Val Leu Ser
130 135 140

Ala Ser Arg Leu Glu Lys Phe Gly Glu Val Arg Phe Thr Arg Ile Glu
145 150 155 160

Thr Ser Ile Lys Asn Leu Tyr Thr Arg Ile Asn Gly Asn Ser Ser Thr
165 170 175

Ile Asn Leu Thr Asp Trp Leu Glu Leu Asn Phe Gly Leu Ile Val
180 185 190

Lys Met Ile Ala Gly Lys Asn Tyr Glu Ser Gly Lys Gly Asp Glu Glu
195 200 205

Val Glu Arg Phe Lys Asn Ala Phe Lys Asp Phe Met Val Leu Ser Met
Glu Phe Val Leu Trp Asp Ala Phe Pro Ile Pro Leu Phe Lys Trp Val
225 230 235 240
Asp Phe Gln Gly His Ile Lys Ala Met Lys Arg Thr Phe Lys Asp Ile
245 250 255
Asp Ser Val Phe Gln Asn Trp Leu Glu Glu His Ile Asn Lys Arg Glu
260 265 270
Lys Met Glu Val Gly Ala Gly Gly Asn Glu Gin Asp Phe Ile Asp Val
275 280 285
Val Leu Ser Lys Leu Ser Lys Glu Tyr Leu Asp Glu Gly Tyr Ser Arg
290 295 300
Asp Thr Val Ile Lys Ala Thr Val Phe Ser Leu Val Leu Asp Ala Ala
305 310 315 320
Asp Thr Val Ala His Ile Asn Trp Gly Met Thr Leu Leu Ile Asn
325 330 335
Asn Gln Asn Ala Leu Met Lys Ala Gln Glu Glu Ile Asp Thr Lys Val
340 345 350
Gly Lys Asp Arg Trp Val Glu Glu Ser Asp Ile Lys Asp Leu Val Tyr
355 360 365
Leu Gln Ala Ile Val Lys Val Leu Arg Leu Tyr Pro Pro Gly Pro
370 375 380
Leu Leu Val Pro His Glu Asn Val Lys Asp Cys Val Val Ser Gly Tyr
385 390 395 400
His Ile Pro Lys Glu Thr Arg Leu Phe Ala Asn Val Met Lys Leu Gln
405 410 415
Arg Asp Pro Lys Leu Leu Ser Asn Pro Asp Lys Phe Asp Pro Glu Arg
420 425 430
Phe Ile Ala Gly Asp Ile Asp Phe Arg Gly His Ile Thr Glu Phe Ile
435 440 445
Pro Phe Gly Ser Gly Arg Arg Ser Cys Pro Gly Thr Ser Thr Leu Ala Leu
450 455 460
Gln Val Glu His Leu Thr Met Ala His Leu Ile Gin Gly Phe Asn Tyr
465 470 475 480
Lys Thr Pro Asn Asp Glu Ala Leu Asp Met Lys Glu Gly Ala Gly Ile
485 490 495
Thr Ile Arg Lys Val Asn Pro Val Glu Leu Ile Ile Thr Pro Arg Leu
500 505 510
Ala Pro Glu Leu Tyr
515

<210> SEQ ID NO 3
<211> LENGTH: 1733
<212> TYPE: DNA
<213> ORGANISM: Nicotiana sp.
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(1551)

<400> SEQUENCE: 3
atg ctt ctc ccc ata gaa ggc att gta gga cta gta acc ttc acc ttt 48
Met Leu Ser Pro Ile Glu Ala Ile Val Gly Leu Val Thr Phe Thr Phe
1 5 10 15
cct ttc ttc ctc cta tgg aca aaa aaa tct caa aac cct cta aaa ccc 96
Leu Phe Phe Leu Thr Thr Lys Ser Gin Lys Pro Ser Lys Pro
20 25 30
tta cca cgc aaa atc ccc gga gga tgg cog gta atc gcc cat ctt ttc 144
Leu Pro Pro Lys Ile Pro Gly Gly Trp Pro Val Ile Gly His Leu Phe
35 40 45
ccc ttc aat gac gac ggc gac gct cca tta gct cga aaa ctc gga
His Phe Asn Amp Asp Gly Asp Arg Pro Leu Ala Arg Lys Leu Gly
50 55 60
192
gac tta gct gac aaa tac ggc ccc gtt ttc act ttt cgg cta ggc ctt
Amp Leu Ala Amp Lys Tyr Gly Pro Val Phe Thr Phe Arg Leu Gly Leu
65 70 75 80
240
ccc ctt gtc tta gtt gta aag aat gag tcc gaa aag aac gag tct ttc
Pro Leu Val Leu Val Ser Ser Tyr Gln Ala Val Lys Gly Cys Phe
85 90 95
288
tct cca aat gac gcc att ttt ccc aat cgt cca gct ttt ccc tac ggc
Ser Thr Asn Asp Ala Ile Phe Ser Asn Arg Pro Ala Phe Leu Tyr Gly
100 105 110
336
gat tcc att ggc tac aat aat goc atg cta ttt trg gcc aat ctc gga
Amp Tyr Leu Gly Tyr Asn Asn Met Leu Phe Leu Ala Asn Tyr Gly
115 120 125
384
cct tac tgg cga aac aat cga aaa tta gtt aat cag gaa gat ctt tcc
Pro Tyr Trp Arg Lys Asn Arg Arg Leu Val Ile Gln Glu Val Leu Ser
130 135 140
432
gtc aat gct ctc gaa aac ctc aaa cac gtg aga ttt gca aag att cac
Ala Ser Arg Leu Glu Lys Phe His Val Arg Phe Ala Arg Ile Gln
145 150 155 160
480
gcg aac att aag aat tta tat act cga att gat gca aat tgg aat tgc cgg
Ala Ser Ile Lys Asn Leu Tyr Thr Arg Ile Asp Gly Asn Ser Ser Thr
165 170 175
528
ata aat tta act gat tgg tta gaa ttt ttg aat ttt ggt aat tgt atg ctt
tgc Ile Asn Leu Thr Asp Trp Leu Glu Leu Phe Gly Leu Ile Val
180 185 190
576
aag atg atc gct gca aac aat tat gaa tcc gat ttt aac ggg gaa gat gaa
Lys Met Asn Ile Gly Lys Asn Ser Gly Thr Gly Lys Asp Arg Glu
195 200 205 210
624
gtg gag gaa ttt aag aag aac ggc tta aag gat ttt att tta tca att
gat Val Glu Arg Phe Lys Lys Asn Phe Asp Phe Met Ile Leu Ser Met
215 220
672
gag ttt gtc tta tgg gat gca ttt cca att cca tta ttt aac tgg gtt
Glu Phe Val Leu Trp Asp Ala Phe Pro Ile Pro Leu Phe Gly Lys Trp Val
225 230 235 240
720
gat ttt cca ggg cat gtt aag gct atg aac cag ttt aac gat atc
Amp Phe Gly Gln Gly His Val Tyr Ala Met Asp Ser Arg Thr Phe Lys Asp Ile
245 250 255
768
gat ctt gtt ttt cag aat tgg tta gag gaa cat att aat aac cag gaa
Amp Ser Phe Glu Asn Trp Leu Glu Glu His Ile Asn Lys Arg Glu
260 265 270
816
aaa atg gag gtt aat gca gaa ggg aat gaa cca gat tcc att cag gtt
lyc Met Glu Val Asn Ala Glu Gly Gly Asp Arg Phe Ile Asp Val
275 280 285
864
gtg ctt tca aac atg agt aat gaa ttt gtt gau gtt tac tct cgt
Val Leu Ser Lys Met Ser Asn Glu Tyr Leu Gly Gly Tyr Ser Arg
290 295 300
912
gat act gtc att aac gca agc gtt ttg tgt gtt gat gca gca
Amp Thr Val Glu Ala Thr Val Phe Ser Leu Val Leu Asp Ala Ala
305 310 315 320
960
gac aca gtt gct ctc aat tgg gga atg gca cta tgg ata aac
Amp Thr Val Ala Leu His Ile Asn Trp Gly Met Ala Leu Ile Asn
325 330 335
1008
aas caa aag goc ttg aag aac gca cca gaa gag ata gac aca aat
ggt Asn Glu Ala Leu Thr Lys Ala Glu Glu Ile Asp Thr Lys Val
340 345 350
1056
-continued

ggt aag gac aga tgg gta gaa gag agt att aag gat tgg gta tac
Gly Lys Arg Arg Trp Val Glu Ser Asp Ile Lys Aep Leu Val Tyr
355 360 365

tct caa gct att att gta aag gaa gtt tta cga tta tat cca cca gga cct
Leu Gln Ala Ile Val Lys Glu Val Arg Leu Tyr Pro Pro Gly Pro
370 375 380

ttg gta cca cac gaa aat gta aag gat tgt gtt gtt tgt gga tat
Leu Leu Val Pro His Arg Aep Glu Val Val Ser Gly Tyr
385 390 395 400

cac att cct aag gag cca aga tta tcc gca aac gtc atg aag ctc cca
His Ile Pro Lys Gly Thr Arg Ala Asn Val Met Lys Leu Gln
405 410 415

cgt gat cct aat ctc tgg tgt tgt gat cct gat act ttc gat cca gag aga
Arg Asp Pro Lys Thr Ser Asp Pro Asp Thr Phe Aep Gly Arg
420 425 430

ttc att gat act gat att gag ttt cgt ggt cag tac tat aag tat atc
Phe Ile Thr Aep Ile Aep Phe Arg Gly Gin Tyr Tyr Tyr Ile
435 440 445

cgg tgt tgt ggt gaa gga cta cct tgt cca ggg atg act tat gca tgt
Pro Phe Gly Ser Arg Arg Arg Ser Cys Pro Gly Met Thr Tyr Ala Leu
450 455 460

cag ggg gaa gca cta cca atg cca gaa ctt atg aag gaa ggt gcc gtc atc
Gln Val Glu His Leu Thr Met Ala His Leu Ile Gin Gly Phe Asn Tyr
465 470 475 480

<210> SEQ ID NO 4
<211> LENGTH: 517
<212> TYPE: PRT
<213> ORGANISM: Nicotiana sp.

<400> SEQUENCE: 4

Met Leu Ser Pro Ile Glu Ala Ile Val Gly Leu Val Thr Phe Thr Phe
1  5    10    15
Leu Phe Phe Leu Thr Thr Thr Thr Thr Thr Thr
20    25
Leu Pro Pro Asp Ile Lys Gly Trp Pro Val Ile Gly His Leu Phe
35    40    45
His Asp Asp Asp Lys Asp Pro Leu Ala Arg Gly Leu Gly
50    55    60
Asp Leu Ala Asp Lys Tyr Gly Pro Val Phe Thr Phe Arg Leu Gly
65    70    75    80
Pro Leu Val Leu Val Ser Ser Tyr Glu Ala Val Lys Asp Cys Phe
85    90    95
Ser Thr Asp Ala Ile Phe Ser Asn Arg Pro Ala Phe Leu Tyr Gly
100   105   110
Amp Tyr Leu Gly Tyr Asn Asn Ala Met Leu Phe Leu Ala Asn Tyr Gly
   115  120  125
Pro Tyr Trp Arg Lys Asn Arg Lys Leu Val Ile Gln Glu Val Leu Ser
   130  135   140
Ala Ser Arg Leu Glu Lys Phe Lys His Val Arg Phe Ala Arg Ile Gln
   145  150  155  160
Ala Ser Ile Lys Asn Leu Tyr Thr Arg Ile Asp Gly Asn Ser Ser Thr
   165  170  175
Ile Asn Leu Thr Asp Trp Leu Glu Leu Asn Phe Gly Leu Ile Val
   180  185  190
Lys Met Ile Ala Gly Lys Asn Tyr Glu Ser Gly Lys Gly Asp Glu Gln
   195  200  205
Val Glu Arg Phe Lys Ala Phe Lys Asp Phe Met Ile Leu Ser Met
   210  215  220
Glu Phe Val Leu Trp Asp Ala Phe Pro Ile Pro Leu Phe Lys Trp Val
   225  230  235  240
Asp Phe Gln Gly His Val Lys Ala Met Lys Arg Thr Phe Lys Asp Ile
   245  250  255
Asp Ser Val Phe Gln Asn Trp Leu Glu Glu His Ile Asn Lys Arg Glu
   260  265  270
Lys Met Glu Val Asn Ala Gly Asn Glu Gin Asp Phe Ile Asp Val
   275  280  285
Val Leu Ser Lys Met Ser Asn Glu Tyr Leu Gly Glu Gly Tyr Ser Arg
   290  295  300
Asp Thr Val Ile Lys Ala Thr Val Phe Ser Leu Val Leu Asp Ala Ala
   305  310  315  320
Asp Thr Val Ala Leu His Ile Asn Trp Gly Met Ala Leu Leu Ile Ann
   325  330  335
Asn Gin Lys Ala Leu Thr Lys Ala Gin Glu Glu Ile Asp Thr Lys Val
   340  345  350
Gly Lys Asp Arg Trp Val Glu Glu Ser Asp Ile Lys Asp Leu Val Tyr
   355  360  365
Leu Gin Ala Ile Val Lys Glu Val Leu Arg Leu Tyr Pro Pro Gly Pro
   370  375  380
Leu Leu Val Pro His Glu Asn Val Glu Asp Cys Val Val Ser Gly Tyr
   385  390  395  400
His Ile Pro Lys Gly Thr Arg Leu Phe Ala Asn Val Met Lys Leu Gin
   405  410  415
Arg Asp Pro Lys Leu Trp Ser Asp Pro Asp Thr Phe Asp Pro Glu Arg
   420  425  430
Phe Ile Ala Thr Asp Ile Phe Arg Gly Gin Tyr Tyr Lys Tyr Ile
   435  440  445
Pro Phe Gly Ser Gly Arg Ser Cys Pro Gly Met Thr Tyr Ala Leu
   450  455  460
Gln Val Glu His Leu Thr Met Ala His Leu Ile Glu Gly Phe Asn Tyr
   465  470  475  480
Arg Thr Pro Asn Asp Glu Pro Leu Met Lys Glu Gly Ala Gly Ile
   485  490  495
Thr Ile Arg Lys Val Asn Pro Val Glu Leu Ile Ala Pro Arg Leu
   500  505  510
Ala Pro Glu Leu Tyr
   515
<220> SEQUENCE: 5
<221> TYPE: DNA
<222> LOCATION: (1) . . . (1551)

atg gtt ttt ccc ata gaa gcc att gta gga cta gta acc ttc aca ttt 48
Met Val Phe Pro Ile Glu Ala Ile Val Gly Leu Val Thr Phe Thr Phe
1 5 10 15

cac ttc ttc ttc tca tgg aca aaa aaa ttt csa aaa cct taa aaa ccc 96
Leu Phe Phe Phe Leu Trp Thr Lys Ser Glu Lys Pro Ser Lys Pro
20 25 30

tta cca cgg aaa atc ccc gga gga tgg ccc gta atc ggc cat ttt 144
Leu Pro Pro Ile Pro Gly Gly Gly Trp Pro Val Ile Gly His Leu Phe
35 40 45

cac ttc aat gac gac ggc gac gac cgt cca tta gtt gga aat cca ctc gga 192
His Phe Asn Arg Asp Gly Asp Asp Pro Leu Ala Arg Lys Lys Gly
50 55 60

gac tta gct gac gaa taa gcc gtc aag ttt cga gta gca gtc ctt 240
Asp Leu Ala Arg Leu Tyr Gly Pro Val Phe Thr Arg Leu Gly Leu
65 70 75 80

cac ttc tgc tta gct gga aag aat tac gcc gtt tca act ttt cgg cta ggc ctt 288
Pro Leu Val Leu Val Val Ser Ser Tyr Glu Ala Val Lys Asp Cys Phe
85 90 95

tct aca aat gac gcc att ttt ccc aat cgt cca gct ttt cct tac ggc 336
Ser Thr Arg Asp Ala Ile Phe Ser Arg Asn Pro Ala Leu Tyr Gly
100 105 110

gat tac ctt gcc tac aat gcc atg cta ttt tgg gcc aat tac gga 384
Asp Tyr Leu Gly Tyr Asn Ala Met Leu Phe Leu Ala Asn Tyr Gly
115 120 125

cct tac tgg cga aaa aat cga aag tta gtt att cgg gaa gtt ctc tcc 432
Pro Tyr Trp Arg Lys Am Arg Lys Leu Val Ile Glu Glu Val Leu Ser
130 135 140

gat cgg aat cgt ttc gaa aat ctc aac cgg aat cgg tag gaa att taa aat cac 480
Ala Asp Asp Asn Asp Gly Arg Ala Ala Tyr Glu Phe Lys His Val Arg Asp Ile Glu
145 150 155 160

ggc aat ctc ctc gaa aca cgg gaa aat ctc atg cca gtt cca gtt cct tcc 528
Ser Ile Asp Leu Val Glu Thr Trp Gln Asn Ser Ser Thr Arg Trp Gly
170 175

aaa atg gaa gtt att gta gca gaa ggc cta gac aat gaa ggc aat cgg aat 576
Asp Ile Ala Ile Gly Lys Asp Tyr Glu Asp Glu Gly Glu Asp Glu
180 185 190

aaa atg gaa gtt atc gga aac gcc att gta gca gaa gtt taa aat gac gta 624
Asp Ile Ala Ile Gly Lys Asp Tyr Glu Asp Glu Gly Glu Asp Glu
195 200 205

gag gtg aag aat cac aat ttt ggc ttt act att tta cca aat ctt ttc 672
Glu Arg Glu Ala Leu Asp Phe Met Ile Leu Ser Met
210 215 220

gag gtg taa gta gag gta gca ttt ccc att cca tta ttt aat tgg gtg 720
Glu Phe Val Leu Trp Asp Ala Phe Pro Ile Pro Leu Phe Lys Trp Val
225 230 235 240

gct ttt cca ggg ctc att tgg aat gct gat aag ctt aat tgg cta gta 768
Asp Phe Glu Gly His Val Lys Ala Met Lys Arg Thr Phe Lys Asp Ile
245 250 255

gat tct gaa gaa gag gaa gat att aat aat agg aag aag aag act att 816
Asp Val Phe Glu Asn Trp Leu Glu Glu His Asn Lys Asp Arg
260 265 270

aaa atg gaa gtt aat gca gaa ggg aat gaa gca gat ttc att gat gtg 864
Lys Met Glu Val Asn Ala Glu Gly Aen Glu Gin Asp Phe Ile Asp Val
275  280  285

ttg ctt tca aas atg aat gat gaa gas tat cct ggt gaa gtt tac tct cgt 912
Val Leu Ser Lys Met Ser Asn Gly Thr Leu Gly Gly Tyr Ser Arg
290  295  300

gat act gtc att aag gca acg gtt ttt aat gtc tgg gat gca 960
Asp Thr Val Ile Lys Ala Thr Val Phe Ser Leu Val Leu Asp Ala Ala
305  310  315  320

gac cca gtt gct ctt cac ata aat tgg gga atg gca tta tgg ata aac 1008
Asp Thr Val Ala Leu His Ile Asn Trp Gly Met Ala Leu Leu Ile Asn
325  330  335

aat cca aag gcc tgt acg aas gca cac gas gag ats aac aca aas gtt 1056
Asn Gln Lys Ala Leu Thr Lys Ala Gln Glu Glu Ile Asp Thr Lys Val
340  345  350

ggt aag gac gcc tgt gta gaa gag act atg gat att aag gat tgg tga tac 1104
Gly Lys Aep Arg Trp Val Glu Ala Ser Asp Leu Arg Leu Val Tyr
355  360  365

cct cca gct att gtt aas gga gtt tta cga tta tat cca cca gga cct 1152
Leu Gln Ala Ile Val Lys Val Glu Leu Arg Leu Tyr Pro Gly Pro
370  375  380

ttg tta gta cca cac gas aat gta gaa gas gat tgg tgt gtt aat gga tat 1200
Leu Leu Pro His Glu Val Glu Arg Ser Val Val Ser Gly Tyr
385  390  395  400

cac att cct aas ggg aca aga tta ttc gca aac gtc act atg aas ctt cca 1248
His Ile Pro Lys Gly Thr Arg Leu Phe Ala Asn Val Met Lys Leu Gin
405  410  415

cgt gat cct aac ctc tgg tct gat oct gat act ttc gat cca gag aga 1296
Arg Asp Pro Lye Thr Ser Asp Pro Thr Phe Asp Pro Gly Arg
420  425  430

tcc att gct act atg att gag ttt cgt ggt cag tac tat aag tat atrc 1344
Phe Ala Thr Asp Ile Asp Phe Arg Gly Gin Tyr Tyr Lys Ile
435  440  445

cct ttt ggt tct gga aag cga ctc ttg cca ggg atg act tat gca ttg 1392
Pro Phe Gly Ser Gly Arg Arg Ser Cyu Pro Gly Met Thr Tyr Ala Leu
450  455  460

cac gtt gaa cac tta aca atg gca cat ttg act cac aat ggt ttc aat ttc 1440
Gln Val Glu His Leu Thr Met Ala His Leu Ile Gin Gly Phe Asn Tyr
465  470  475  480

aga act cca aat gac gag ccc ttg gac aag gaa ggt gca gcc ata 1488
Arg Thr Pro Asp Ala Glu Pro Leu Ala Met Lys Gly Ala Gly Ile
485  490  495

act ata cgt aag gta aat cct gtt gaa ctc ata ata gcc cct cgc tgc 1536
Thr Ile Arg Lys Val Asn Pro Val Leu Ile Ala Arg Pro Leu Arg
500  505  510

gca cgt gag ctt tat ttaaa cct gat ctttt cctc gttt gcc cttatc tatt tatt 1591
Ala Pro Gin Leu Tyr
515

tact tctt aag gaaacctttg gat ctttt ccttgag cagctttcttta 1651
att ttcttct aat taaa  aag gat cttttgcc aag ct tatt tatt 1711
aag tactttag gcttt tctttgc cc 1733

<210> SEQ ID NO 6
<211> LENGTH: 517
<212> TYPE: PRT
<213> ORGANISM: Nicotiana sp.
<400> SEQUENCE: 6

Met Val Phe Pro Ile Glu Ala Ala Val Gly Leu Val Thr Phe Thr Phe
1  5  10  15
Leu Phe Phe Phe Leu Trp Thr Lys Lys Ser Gln Lys Pro Ser Lys Pro
20 25 30
Leu Pro Pro Lys Ile Pro Gly Gly Trp Pro Val Ile Gly His Leu Phe
35 40 45
His Phe Asn Asp Gly Asp Asp Arg Pro Leu Ala Arg Lys Leu Gly
50 55 60
Asp Leu Ala Asp Lys Tyr Gly Pro Val Phe Thr Phe Arg Leu Gly Leu
65 70 75 80
Pro Leu Val Leu Val Val Ser Ser Tyr Glu Ala Val Lys Asp Cys Phe
85 90 95
Ser Thr Asn Asp Ala Ile Phe Ser Asn Arg Pro Ala Phe Leu Tyr Gly
100 105 110
Asp Tyr Leu Gly Tyr Asn Ala Met Leu Phe Leu Ala Asn Tyr Gly
115 120 125
Pro Tyr Trp Arg Lys Asn Arg Lys Leu Val Ile Gln Glu Val Leu Ser
130 135 140
Ala Ser Arg Leu Glu Phe Lys His Val Arg Phe Ala Arg Ile Gln
145 150 155 160
Ala Ser Ile Lys Asn Leu Thr Arg Ile Asp Gly Asn Ser Ser Thr
165 170 175
Ile Asn Leu Thr Asp Trp Leu Glu Leu Asn Phe Gly Leu Ile Val
180 185 190
Lys Met Ile Ala Gly Lys Asn Tyr Glu Ser Gly Lys Gly Asp Glu Gln
195 200 205
Val Glu Arg Phe Lys Ala Phe Lys Asp Phe Met Ile Leu Ser Met
210 215 220
Glu Phe Val Leu Trp Asp Ala Phe Pro Ile Pro Leu Phe Lys Trp Val
225 230 235 240
Asp Phe Gln Gly His Val Lys Ala Met Lys Arg Thr Phe Lys Asp Ile
245 250 255
Asp Ser Val Phe Gln Asn Trp Leu Glu Glu His Ile Asn Lys Arg Glu
260 265 270
Lys Met Glu Val Asn Ala Glu Gly Asn Glu Gin Asp Phe Ile Asp Val
275 280 285
Val Leu Ser Lys Met Ser Asn Glu Tyr Leu Glu Gly Tyr Ser Arg
290 295 300
Asp Thr Val Ile Lys Ala Thr Val Phe Ser Leu Val Asp Ala Ala
305 310 315 320
Asp Thr Val Ala Leu His Ile Asn Trp Gly Met Ala Leu Leu Ile Asn
325 330 335
Asn Gln Lys Ala Leu Thr Lys Ala Gin Glu Glu Ile Asp Thr Lys Val
340 345 350
Gly Lys Asp Arg Trp Val Glu Glu Ser Asp Ile Lys Asp Leu Val Tyr
355 360 365
Leu Gln Ala Ile Val Lys Glu Val Leu Arg Leu Tyr Pro Pro Gly Pro
370 375 380
Leu Leu Val Pro His Glu Asn Val Glu Asp Cys Val Val Ser Gly Tyr
385 390 395 400
His Ile Pro Lys Gly Thr Arg Leu Phe Ala Asn Val Met Lys Leu Gln
405 410 415
Arg Asp Pro Lys Leu Trp Ser Asp Pro Asp Thr Phe Asp Pro Glu Arg
420 425 430
<table>
<thead>
<tr>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro</td>
<td>Phe</td>
<td>Gly</td>
<td>Ser</td>
</tr>
<tr>
<td>435</td>
<td></td>
<td>440</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Arg</td>
<td>Asp</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>450</th>
<th>455</th>
<th>460</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg</td>
<td>Pro</td>
<td>Met</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>Leu</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>465</th>
<th>470</th>
<th>475</th>
<th>480</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr</td>
<td>Pro</td>
<td>Asp</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Met</td>
<td>Lys</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>485</th>
<th>490</th>
<th>495</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr</td>
<td>Arg</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>500</th>
<th>505</th>
<th>510</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>Pro</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>515</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 7
<212> LENGTH: 1727
<213> ORGANISM: Nicotiana sp.
<220> FEATURE: CDS
<222> LOCATION: (1) ... (1554)

<400> SEQUENCE:

```
ATG GTT TTT CCG GTA GAA GCC ATT GTA GGG CTA GTG ACC TTC ACA TTT
Met Val Phe Pro Val Glu Ala Val Gly Leu Val Thr Phe Thr Phe
1    5    10   15

CTC TTC TAC TTC CTA TGG ACA AAA AAA TCT CAA AAA CTC CAA AAC
Leu Phe Tyr Phe Leu Trp Thr Thr Lys Ser Gin Lys Pro Ser Lys Pro
20   25   30

TTT CCA CCC AAA ATC CCG GGA GGA TGG CCG GTA ATC GGC CAT CTT TTC
Leu Pro Pro Lys Ile Pro Gly Gly Trp Pro Val Ile Gly His Leu Phe
35   40   45

TAC TTC GAT GAC GAC GGC ACC TTA GCT CCA TTA GCT CAA AAA CTC GGA
Tyr Phe Asp Asp Asp Gly Asp Arg Pro Leu Ala Arg Leu Gly
50   55   60

GAC TTA GGT CAC AAA TAC GGA CCC GTT TTC ACT TTT CCG CTA GGC TTC
Asp Leu Ala Lys Asp Lys Gly Pro Val Thr Phe Arg Leu Gly Leu
65   70   75   80

CCC CTT GTG TTA GTT AGC AGT GAC ATT GAA GCT ATC AAA GAT TGT TTC
Pro Leu Val Leu Val Val Ser Ser Tyr Glu Ala Ile Lys Asp Cys Phe
85   90   95

tct aca aat gac gcc ctt ctc aat cgt cca gtt ttg ctt tac ggc
Ser Thr Asp Ala Ala Phe Ser Asp Pro Ala Phe Leu Tyr Gly
100  105  110

gaa tac ctt ggc tac aac gtt cta ttt ttg gca aat tac gga
Glu Tyr Leu Gly Tyr Lys Asp Ala Met Leu Phe Leu Ala Asn Tyr Gly
115  120  125

TCT TAC TGG CCA AAA AAT CGT AAT TTA ATT ATT CAG GAA GCT CTC TCT
Ser Thr Trp Arg Lys Asp Pro Asp Lys Leu Ile Gin Glu Val Leu Ser
130  135  140

gct ctt cct gcc ctt aac ttg aga tct gcc gcc aga act ctc
Ala Ser Arg Leu Gly Phe Leu Lys His Val Arg Phe Ala Arg Leu Gin
145  150  155  160

ACG AGC ATG AAA AAT TTA TAC ACT GCA ATG GGA AAT TGG ACT GTG
Ser Phe Ile Lys Arg Ala Tyr Thr Arg Lys Glu Asp Gln Ser Ser Thr
165  170  175

ATA AAT TTA ACT GTG TTA GAA GAA TTT GGT GTG GTG AAT TCT CTT
Ile Asn Leu Thr Arg Trp Leu Leu Leu Phe Leu Asp Arg Leu Ile Val
180  185  190

AAA ATG ATC GTG GGA AAT TAT GAA TCC GTT AAA GGA GAT GAA CAA
Leu Met Ile Ala Gly Lys Arg Thr Phe Ala Pro Asp Gly Glu Gin
195  200  205  210
```

<215> LOCATION:

<225> NAME/KEY:

<225> KEY:

<225> NAME:

<225> ORGANISM:

<table>
<thead>
<tr>
<th>81</th>
</tr>
</thead>
</table>

<210> SEQ ID NO 7
<212> LENGTH: 1727
<213> ORGANISM: Nicotiana sp.
<220> FEATURE: CDS
<222> LOCATION: (1) ... (1554)

<400> SEQUENCE:

```
ATG GTT TTT CCG GTA GAA GCC ATT GTA GGG CTA GTG ACC TTC ACA TTT
Met Val Phe Pro Val Glu Ala Val Gly Leu Val Thr Phe Thr Phe
1    5    10   15

CTC TTC TAC TTC CTA TGG ACA AAA AAA TCT CAA AAA CTC CAA AAC
Leu Phe Tyr Phe Leu Trp Thr Thr Lys Ser Gin Lys Pro Ser Lys Pro
20   25   30

TTT CCA CCC AAA ATC CCG GGA GGA TGG CCG GTA ATC GGC CAT CTT TTC
Leu Pro Pro Lys Ile Pro Gly Gly Trp Pro Val Ile Gly His Leu Phe
35   40   45

TAC TTC GAT GAC GAC GGC ACC TTA GCT CCA TTA GCT CAA AAA CTC GGA
Tyr Phe Asp Asp Asp Gly Asp Arg Pro Leu Ala Arg Leu Gly
50   55   60

GAC TTA GGT CAC AAA TAC GGA CCC GTT TTC ACT TTT CCG CTA GGC TTC
Asp Leu Ala Lys Asp Lys Gly Pro Val Thr Phe Arg Leu Gly Leu
65   70   75   80

CCC CTT GTG TTA GTT AGC AGT GAC ATT GAA GCT ATC AAA GAT TGT TTC
Pro Leu Val Leu Val Val Ser Ser Tyr Glu Ala Ile Lys Asp Cys Phe
85   90   95

tct aca aat gac gcc ctt ctc aat cgt cca gtt ttg ctt tac ggc
Ser Thr Asp Ala Ala Phe Ser Asp Pro Ala Phe Leu Tyr Gly
100  105  110

gaa tac ctt ggc tac aac gtt cta ttt ttg gca aat tac gga
Glu Tyr Leu Gly Tyr Lys Asp Ala Met Leu Phe Leu Ala Asn Tyr Gly
115  120  125

TCT TAC TGG CCA AAA AAT CGT AAT TTA ATT ATT CAG GAA GCT CTC TCT
Ser Thr Trp Arg Lys Asp Pro Asp Lys Leu Ile Gin Glu Val Leu Ser
130  135  140

gct ctt cct gcc ctt aac ttg aga tct gcc gcc aga act ctc
Ala Ser Arg Leu Gly Phe Leu Lys His Val Arg Phe Ala Arg Leu Gin
145  150  155  160

ACG AGC ATG AAA AAT TTA TAC ACT GCA ATG GGA AAT TGG ACT GTG
Ser Phe Ile Lys Arg Ala Tyr Thr Arg Lys Glu Asp Gln Ser Ser Thr
165  170  175

ATA AAT TTA ACT GTG TTA GAA GAA TTT GGT GTG GTG AAT TCT CTT
Ile Asn Leu Thr Arg Trp Leu Leu Leu Phe Leu Asp Arg Leu Ile Val
180  185  190

AAA ATG ATC GTG GGA AAT TAT GAA TCC GTT AAA GGA GAT GAA CAA
Leu Met Ile Ala Gly Lys Arg Thr Phe Ala Pro Asp Gly Glu Gin
195  200  205  210
```

<215> LOCATION:

<225> NAME/KEY:

<225> NAME:

<225> ORGANISM:
gtg gag aga ttt aag aat gaa ccc gcc ctc tac aga att tta tca aga
Val Glu Arg Phe Lys Lys Ala Phe Lys Arg Phe Met Ile Leu Ser Met
210 215 220

gag ttt gct gtt tta tgg gat gca ttt cca att cca tta ttt aag tgg gtg
Glu Phe Val Leu Trp Arg Ala Phe Pro Leu Phe Lys Trp Val
225 230 235 240

gat ttt cca gaa cat gtt aag gct atg aas aag act ttt aas gat ata
Amp Phe Gln Gly His Val Lys Ala Met Lys Arg Thr Phe Lys Asp Ile
245 250 255

gat tct gtt ttt cag aat tgg tta gag gaa cat att aag aas aga gaa
Amp Ser Val Phe Gln Asn Trp Leu Glu Glu His Ile Lys Arg Glu
260 265 270

aaa ata atg gag gct gcc aca gaa ggg aat gaa cca gat ttt ata gat
Lys Ile Met Glu Val Gly Thr Glu Gly Asp Glu Gln Arg Phe Ile Asp
275 280 285

gtc gtc ttt tca aaa atg aat gaa tat ctt ggc gaa ggt tac tct
Val Val Leu Ser Lys Met Ser Asn Glu Tyr Leu Glu Gly Tyr Ser
290 295 300

cgc gat act gtc ata aaa gca aca gta ttt atg tgg gct tgg gat gca
Arg Asp Thr Val Ile Ala Thr Val Phe Ser Leu Val Leu Asp Ala
305 310 315 320

gca gac aca gtt gtt ctt ccc aat aat tgg gaa atg gca tta tgg ata
Ala Asp Thr Val Ala Leu His Ile Asn Cys Gly Met Ala Leu Ile
325 330 335

aac aat cca aat gct tgg aag aag gca cca gaa gag ata gag aca aca
Aam Aam Glu Asn Ala Ala Lys Lys Ala Gln Glu Glu Ile Asp Thr Lys
340 345 350

gtt ggt aag gat aga tgg gta gaa gag agt gat att aag gat tgg gta
Val Gly Lys Arg Asp Thr Val Gly Thr Gln Val Glu Leu Asp Val
355 360 365

ttc ttc cca gtt act gtt aaa gaa gtt tca gaa tca ttt ctc aat gca
Tyr Leu Glu Ala Ile Val Glu Val Leu Arg Leu Tyr Pro Pro Gly
370 375 380

ttt tta gta cca cac gaa aat gta gaa gat tgg gtt gtt aag gca
Pro Leu Val Pro His Glu Asn Val Glu Arg Cys Val Val Ser Gly
385 390 395 400

tat ccc att cct aca gaa act aga cta ttc gca aat gta atg aca tta
Tyr Thr His Ile Pro Gly Thr Arg Leu Phe Ala Asn Val Met Lys Leu
405 410 415

cac gat cat ctt aca tcc ttc taa atg gat gaa ttc aat cca gag
Glu Arg Pro Lys Leu Trp Ser Asn Pro Asp Lys Phe Asn Pro Glu
420 425 430

aga ttc atc gct gtt gat att gcc ttt cat ggt cag cac tac tat gag tat
Arg Phe Ile Ala Arg Asp Ile Asp Phe His Gly Glu His Tyr Glu Tyr
435 440

atc gcc ttt gtt gtt gca aag gct tct tgg cgg gat gtc act ttt
ttc Pro Phe Gly Ser Gly Arg Ser Phe Pro Phe Pro Met Thr Tyr Ala
455 460

ttg ccc ggg gaa cca act tta gca cta gca atg gca ttc aat gaa gct
ttc Val Glu Val Glu His Thr Met Ala His Leu Ile Val Leu Phe Asn
465 470 475 480

tac gaa act cca act gat gac ccc ttc gat atg aas gaa ggt gca
Tyr Arg Thr Asp Pro Thr Glu Pro Leu Asp Met Lys Glu Ala Gly
485 490 495

ata act ata cgt aat gta aat cct gta aag gta aat atc arg cct cgc
Ile Thr Ile Arg Lys Val Asn Val Lys Val Ile Thr Pro Arg
500 505 510

ttg gca cct gag ctt taa atctaa gatgttctaa ctgtggtgtgat
520
Leu Ala Pro Glu Leu Tyr
515

<210> SEQ ID NO 8
<211> LENGTH: 518
<212> TYPE: PRT
<213> ORGANISM: Nicotiana sp.

<400> SEQUENCE: 8

Met Val Phe Pro Val Glu Ala Ile Val Gly Leu Val Thr Phe Thr Phe
1      5          10      15
Leu Phe Tyr Phe Leu Trp Thr Lys Ser Gin Lys Pro Ser Lys Pro
20     25             30
Leu Pro Pro Lys Ile Pro Gly Gly Trp Pro Val Ile Gly His Leu Phe
35    40          45
Tyr Phe Asp Asp Asp Gly Asp Asp Arg Pro Leu Ala Arg Lys Leu Gly
50    55          60
Asp Leu Ala Asp Lys Tyr Gly Pro Val Phe Thr Phe Arg Leu Gly Leu
65    70          75      80
Pro Leu Val Leu Val Val Ser Ser Tyr Glu Ala Ile Lys Asp Cys Phe
85    90          95
Ser Thr Asn Asp Ala Ile Phe Ser Asn Arg Pro Ala Phe Leu Tyr Gly
100   105           110
Glu Tyr Leu Gly Tyr Lys Asn Ala Met Leu Phe Leu Ala Asn Tyr Gly
115   120         125
Ser Tyr Trp Arg Lys Asn Arg Lys Leu Ile Ile Gin Glu Val Leu Ser
130   135         140
Ala Ser Arg Leu Glu Lys Phe Lys His Val Arg Phe Ala Arg Ile Gin
145   150         155      160
Thr Ser Ile Lys Asn Leu Tyr Thr Arg Ile Asp Gly Asn Ser Ser Thr
165   170         175
Ile Asn Leu Thr Asp Trp Leu Glu Leu Asn Phe Gly Leu Ile Val
180   185          190
Lys Met Ile Ala Gly Lys Asn Tyr Glu Ser Gly Lys Gly Asp Glu Gin
195   200          205
Val Glu Arg Phe Lys Ala Phe Lys Asp Phe Met Ile Leu Ser Met
210   215          220
Glu Phe Val Leu Trp Asp Ala Phe Pro Ile Pro Leu Phe Lys Trp Val
225   230          235      240
Asp Phe Gln Gly His Val Lys Ala Met Lys Arg Thr Phe Lys Asp Ile
245   250           255
Asp Ser Val Phe Gln Asn Trp Leu Glu Glu His Ile Lys Lys Arg Glu
260   265          270
Lys Ile Met Glu Val Gly Thr Glu Gly Asn Gin Asp Phe Ile Asp
275   280          285
Val Val Leu Ser Lys Met Ser Asn Glu Tyr Leu Gly Glu Gly Tyr Ser
290   295          300
Arg Asp Thr Val Ile Lys Ala Thr Val Phe Ser Leu Val Leu Asp Ala
305   310          315      320
Ala Asp Thr Val Ala Leu His Ile Asn Cys Gly Met Ala Leu Leu Ile
325   330          335
<table>
<thead>
<tr>
<th>Residue</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aasn</td>
<td>Asn Gln Asn Ala Leu Lys Lys Ala Gln Glu Glu Ile Leu Thr Lys</td>
</tr>
<tr>
<td>Gln</td>
<td>340</td>
</tr>
<tr>
<td>Asn</td>
<td>345</td>
</tr>
<tr>
<td>Ala</td>
<td>350</td>
</tr>
<tr>
<td>Leu</td>
<td>355</td>
</tr>
<tr>
<td>Lys</td>
<td>360</td>
</tr>
<tr>
<td>Lys</td>
<td>365</td>
</tr>
<tr>
<td>Arg</td>
<td>Tyr Leu Glu Ala Ile Val Lys Glu Val Leu Arg Leu Tyr Pro Pro Gly</td>
</tr>
<tr>
<td>Trp</td>
<td>370</td>
</tr>
<tr>
<td>Val</td>
<td>375</td>
</tr>
<tr>
<td>Glu</td>
<td>Pro Leu Leu Val Pro His Glu Asn Val Glu Asp Gly Cys Val Val Ser Gly</td>
</tr>
<tr>
<td>Ser</td>
<td>380</td>
</tr>
<tr>
<td>Leu</td>
<td>385</td>
</tr>
<tr>
<td>Tyr</td>
<td>390</td>
</tr>
<tr>
<td>His</td>
<td>395</td>
</tr>
<tr>
<td>Asn</td>
<td>400</td>
</tr>
<tr>
<td>Val</td>
<td>405</td>
</tr>
<tr>
<td>Asp</td>
<td>410</td>
</tr>
<tr>
<td>Tyr</td>
<td>415</td>
</tr>
<tr>
<td>Gln</td>
<td>Gln Arg Asp Pro Lys Leu Trp Ser Asn Pro Asp Lys Phe Asn Pro Glu</td>
</tr>
<tr>
<td>Arg</td>
<td>420</td>
</tr>
<tr>
<td>Phe</td>
<td>425</td>
</tr>
<tr>
<td>Ile</td>
<td>430</td>
</tr>
<tr>
<td>Pro</td>
<td>435</td>
</tr>
<tr>
<td>Phe</td>
<td>440</td>
</tr>
<tr>
<td>Gly</td>
<td>445</td>
</tr>
<tr>
<td>Ser</td>
<td>450</td>
</tr>
<tr>
<td>Gly</td>
<td>455</td>
</tr>
<tr>
<td>Arg</td>
<td>460</td>
</tr>
<tr>
<td>Asp</td>
<td>465</td>
</tr>
<tr>
<td>Ser</td>
<td>470</td>
</tr>
<tr>
<td>Cys</td>
<td>475</td>
</tr>
<tr>
<td>Pro</td>
<td>480</td>
</tr>
<tr>
<td>Gly</td>
<td>485</td>
</tr>
<tr>
<td>Met</td>
<td>490</td>
</tr>
<tr>
<td>Thr</td>
<td>495</td>
</tr>
<tr>
<td>Tyr</td>
<td>500</td>
</tr>
<tr>
<td>Ala</td>
<td>505</td>
</tr>
<tr>
<td>Pro</td>
<td>510</td>
</tr>
<tr>
<td>Glu</td>
<td>515</td>
</tr>
<tr>
<td>Leu</td>
<td>520</td>
</tr>
<tr>
<td>Tyr</td>
<td>525</td>
</tr>
</tbody>
</table>

**SEQ ID:** 9  
**LENGTH:** 539  
**TYPE:** DNA  
**ORGANISM:** Nicotiana sp.  
**FEATURE:** CDS  
**LOCATION:** (1) ... (539)  
**SEQUENCE:** 9  

```
<210> SEQ ID NO 9
<211> LENGTH: 539
<212> TYPE: DNA
<213> ORGANISM: Nicotiana sp.
<220> FEATURE: CDS
<222> LOCATION: (1)...(539)

<400> SEQUENCE: 9

atggtttcctataaggccttggataagcttaaccctctacaatttt
Met Val Phe Pro Ile Glu Ala Phe Val Gly Leu Val Thr Phe Thr Phen
  1    2    3    4    5    6

cctttacctctactgcaaaasaaatctgaaaccctctactaatttt
Leu Leu Tyr Phe Leu Thr Thr Lys Cys Ser Glu Lys Leu Pro Lys Pro
  7    8    9   10   11   12

ttcctaaaaagcccttaagtagatactactacctgaaaccctctactaatttt
Leu Pro Pro Lys Ile Pro Gly Gly Trp Pro Val Ile Gly His Leu Phen
  13   14   15   16   17

catacaacagcagggagcaggcttgacccatagtgaaaccctctactaatttt
His Phe Ann Asn Asp Gly Asp Asp Arg Pro Leu Ala Arg Lys Leu Gly
  18   19   20   21   22   23

gagcttagatagaaactggggctttgtcctgatccctactttgcttctcatgaaaccctctactaatttt
Aasp Leu Ala Asp Lys Tyr Gly Pro Val Phe Thr Phe Arg Leu Gly Leu
  24   25   26   27   28   29   30   31

cctttggatggagagcaattgtcctctcaattacctgaaaccctctactaatttt
Pro Leu Val Leu Val Ser Ser Tyr Glu Ala Lys Asp Cys Phe
  32   33   34   35   36   37   38

ttcctaaaaagcccttaagtagatactactacctgaaaccctctactaatttt
Ser Thr Ann Asn Ala Ile Phe Ser Ann Arg Pro Ala Phe Leu Tyr Gly
  39   40   41   42   43   44   45   46

gagcttagatagaaactggggctttgtcctgatccctactttgcttctcatgaaaccctctactaatttt
Glu Tyr Leu Gly Tyr Ser Asn Ala Met Leu Phe * Glu Ann Thr Asp
  47   48   49   50   51   52   53   54
```
ctt act ggc gaa aaa ata gaa aat tag tca ttc agg aag ttc tct gtg
Leu Thr Gly Glu Lys Ile Glu Asn * Ser Phe Arg Lys Phe Ser Val
130 135 140

ctt gtt gtc tct gaa aat tga agc acg tga gat tgg gtt gaa ttc aga
Leu Val Val Ser Lys Asn * Ser Thr * Asp Leu Val Lys Phe Arg
145 150 155

cga gca tta aga att tat cca ttc gaa tgg gaa att cga gta cga
Arg Ala Leu Arg Ile Tyr Thr Leu Glu Leu Met Glu Ile Arg Val Arg
160 165 170

taa atc taa ca
* Ile *
539

<210> SEQ ID NO 10
<211> LENGTH: 173
<212> TYPE: DNA
<213> ORGANISM: Nicotiana sp.

<400> SEQUENCE: 10
Met Val Phe Pro Ile Glu Ala Phe Val Gly Leu Val Thr Phe Thr Phe
1 5 10 15

Leu Leu Tyr Phe Leu Thr Tyr Thr Lys Ser Gln Lys Leu Pro Lys Pro
20 25 30

Leu Pro Pro Lys Ile Pro Gly Gly Trp Pro Val Ile Gly His Leu Phe
36 40 46

His Phe Asn Asn Asp Gly Asp Asp Arg Pro Leu Ala Arg Lys Leu Gly
50 55 60

Asp Leu Ala Asp Lys Tyr Gly Pro Val Phe Thr Phe Arg Leu Gly Leu
65 70 75 80

Pro Leu Val Leu Val Ser Ser Tyr Glu Ala Ile Lys Asp Cys Phe
85 90 95

Ser Thr Asn Asp Ala Ile Phe Ser Asn Arg Pro Ala Phe Leu Tyr Gly
100 105 110

Glu Tyr Leu Gly Tyr Ser Asn Ala Met Leu Phe Glu Asn Thr Asp Leu
115 120 125

Thr Gly Glu Lys Ile Glu Asn Ser Phe Arg Lys Phe Ser Val Leu Val
130 135 140

Val Ser Lys Asn Ser Thr Asp Leu Val Lys Phe Arg Arg Ala Leu Arg
145 150 155 160

Ile Tyr Thr Leu Glu Leu Met Glu Ile Arg Val Arg Ile
165 170

<210> SEQ ID NO 11
<211> LENGTH: 666
<212> TYPE: DNA
<213> ORGANISM: Nicotiana sp.

<400> SEQUENCE: 11
atg gtt ttt ccc ata gaa goc att gta gga gca gta acc cta att aca
Met Val Phe Pro Ile Glu Ala Ile Val Gly Ala Val Thr Leu Ile Thr
1 5 10 15

ttt ctt tta tac ttc cta tgt cca aca aca aca aca aca cta cca aag
Phe Leu Leu Tyr Phe Leu Cys Thr Lys Lys Ser Gln Lys His Ser Lys
20 25 30

ccc tta cca acg aaa atc ccc gga gga tgg ceg gta atc ggc cat ctt
Pro Leu Pro Thr Lys Ile Pro Gly Gly Trp Pro Val Ile Gly His Leu
35 40 45
<table>
<thead>
<tr>
<th>Phe</th>
<th>His</th>
<th>Phe</th>
<th>Asn</th>
<th>Asp</th>
<th>Gly</th>
<th>Asp</th>
<th>Asp</th>
<th>Arg</th>
<th>Pro</th>
<th>Phe</th>
<th>Ala</th>
<th>Arg</th>
<th>Lys</th>
<th>Leu</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>95</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gly</th>
<th>Asp</th>
<th>Leu</th>
<th>Ala</th>
<th>Asp</th>
<th>Lys</th>
<th>Tyr</th>
<th>Gly</th>
<th>Pro</th>
<th>Val</th>
<th>Phe</th>
<th>Thr</th>
<th>Phe</th>
<th>Arg</th>
<th>Leu</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ctt</th>
<th>Cct</th>
<th>Gtt</th>
<th>Ata</th>
<th>Gt</th>
<th>Cta</th>
<th>Gt</th>
<th>Aag</th>
<th>Tac</th>
<th>Tct</th>
<th>Gtt</th>
<th>Ctt</th>
<th>Gcc</th>
<th>Ctt</th>
<th>Gct</th>
<th>Ttt</th>
</tr>
</thead>
<tbody>
<tr>
<td>85</td>
<td>90</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ttc</th>
<th>Tct</th>
<th>Aca</th>
<th>Aat</th>
<th>Gcc</th>
<th>Att</th>
<th>Ttc</th>
<th>Ctt</th>
<th>Ccg</th>
<th>Cca</th>
<th>Gct</th>
<th>Ttt</th>
<th>Ctt</th>
<th>Tac</th>
<th>Phe</th>
<th>Ser</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>105</td>
<td>110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ggc</th>
<th>Gaa</th>
<th>Tac</th>
<th>Ctt</th>
<th>Ggc</th>
<th>Tac</th>
<th>Aat</th>
<th>Aat</th>
<th>Aca</th>
<th>Atg</th>
<th>Ctt</th>
<th>Tta</th>
<th>Gca</th>
<th>Aat</th>
<th>Tac</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>116</td>
<td>120</td>
<td>126</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gga</th>
<th>Cct</th>
<th>Tac</th>
<th>Tgg</th>
<th>Cga</th>
<th>Aaa</th>
<th>Aat</th>
<th>Cgt</th>
<th>Tta</th>
<th>Gtc</th>
<th>Att</th>
<th>Cag</th>
<th>Gaa</th>
<th>Gtt</th>
<th>Ctc</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>130</td>
<td>135</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tct</th>
<th>Gct</th>
<th>Agt</th>
<th>Ctt</th>
<th>Gas</th>
<th>Aaa</th>
<th>Caa</th>
<th>Csa</th>
<th>Ctg</th>
<th>Aaa</th>
<th>Tcc</th>
<th>Aaa</th>
<th>Aaa</th>
<th>Cgt</th>
<th>Aca</th>
<th>Ser</th>
</tr>
</thead>
<tbody>
<tr>
<td>145</td>
<td>150</td>
<td>155</td>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Caa</th>
<th>Aag</th>
<th>Aat</th>
<th>Tta</th>
<th>Tac</th>
<th>Act</th>
<th>Cga</th>
<th>Att</th>
<th>Aat</th>
<th>Gga</th>
<th>Aat</th>
<th>Tcg</th>
<th>Act</th>
<th>Gln</th>
<th>Thr</th>
<th>Ser</th>
</tr>
</thead>
<tbody>
<tr>
<td>166</td>
<td>170</td>
<td>176</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aag</th>
<th>Aat</th>
<th>Cta</th>
<th>Agt</th>
<th>Tgg</th>
<th>Tta</th>
<th>Gaa</th>
<th>Gaa</th>
<th>Ttg</th>
<th>Aat</th>
<th>Ttt</th>
<th>Ggt</th>
<th>Ctg</th>
<th>Act</th>
<th>Thr</th>
<th>Ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>180</td>
<td>185</td>
<td>190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gtg</th>
<th>Gaa</th>
<th>Aat</th>
<th>Cgg</th>
<th>Aaa</th>
<th>Tat</th>
<th>Gas</th>
<th>Tcc</th>
<th>Gtt</th>
<th>Aaa</th>
<th>Gga</th>
<th>Gat</th>
<th>Gaa</th>
<th>Val</th>
<th>Lys</th>
<th>Met</th>
</tr>
</thead>
<tbody>
<tr>
<td>190</td>
<td>200</td>
<td>205</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Caa</th>
<th>Gtg</th>
<th>Gaa</th>
<th>Att</th>
<th>Gcc</th>
<th>Ttt</th>
<th>Aag</th>
<th>Gat</th>
<th>Ttt</th>
<th>Atg</th>
<th>Gtt</th>
<th>Gln</th>
<th>Val</th>
<th>Glu</th>
<th>Arg</th>
<th>Phe</th>
</tr>
</thead>
<tbody>
<tr>
<td>210</td>
<td>215</td>
<td>220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

<210> SEQ ID NO 12
<211> LENGTH: 222
<212> TYPE: PRO
<213> ORGANISM: Nicotiana sp.

<400> SEQUENCE: 12

<table>
<thead>
<tr>
<th>Met</th>
<th>Val</th>
<th>Phe</th>
<th>Pro</th>
<th>Ile</th>
<th>Glu</th>
<th>Ala</th>
<th>Ile</th>
<th>Val</th>
<th>Gly</th>
<th>Ala</th>
<th>Val</th>
<th>Thr</th>
<th>Leu</th>
<th>Ile</th>
<th>Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phe</th>
<th>Leu</th>
<th>Leu</th>
<th>Tyr</th>
<th>Phe</th>
<th>Leu</th>
<th>Cys</th>
<th>Thr</th>
<th>Lys</th>
<th>Ser</th>
<th>Gln</th>
<th>Lys</th>
<th>His</th>
<th>Ser</th>
<th>Lys</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>25</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pro</th>
<th>Leu</th>
<th>Pro</th>
<th>Thr</th>
<th>Lys</th>
<th>Ile</th>
<th>Pro</th>
<th>Gly</th>
<th>Gly</th>
<th>Trp</th>
<th>Pro</th>
<th>Val</th>
<th>Ile</th>
<th>Gly</th>
<th>His</th>
<th>Leu</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>40</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phe</th>
<th>His</th>
<th>Phe</th>
<th>Asn</th>
<th>Asp</th>
<th>Gly</th>
<th>Asp</th>
<th>Asp</th>
<th>Arg</th>
<th>Pro</th>
<th>Phe</th>
<th>Ala</th>
<th>Arg</th>
<th>Lys</th>
<th>Leu</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gly</th>
<th>Asp</th>
<th>Leu</th>
<th>Ala</th>
<th>Asp</th>
<th>Lys</th>
<th>Tyr</th>
<th>Gly</th>
<th>Pro</th>
<th>Val</th>
<th>Phe</th>
<th>Thr</th>
<th>Phe</th>
<th>Arg</th>
<th>Leu</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Leu</th>
<th>Pro</th>
<th>Leu</th>
<th>Val</th>
<th>Leu</th>
<th>Val</th>
<th>Val</th>
<th>Ser</th>
<th>Ser</th>
<th>Tyr</th>
<th>Glu</th>
<th>Ala</th>
<th>Ile</th>
<th>Lys</th>
<th>Asp</th>
<th>Cys</th>
</tr>
</thead>
<tbody>
<tr>
<td>85</td>
<td>90</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phe</th>
<th>Ser</th>
<th>Thr</th>
<th>Asn</th>
<th>Asp</th>
<th>Ala</th>
<th>Ile</th>
<th>Phe</th>
<th>Ser</th>
<th>Asn</th>
<th>Arg</th>
<th>Pro</th>
<th>Ala</th>
<th>Phe</th>
<th>Leu</th>
<th>Tyr</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>105</td>
<td>110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gly</th>
<th>Glu</th>
<th>Tyr</th>
<th>Leu</th>
<th>Gly</th>
<th>Tyr</th>
<th>Asn</th>
<th>Thr</th>
<th>Met</th>
<th>Leu</th>
<th>Phe</th>
<th>Ala</th>
<th>Asn</th>
<th>Tyr</th>
</tr>
</thead>
<tbody>
<tr>
<td>115</td>
<td>120</td>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gly</th>
<th>Pro</th>
<th>Tyr</th>
<th>Trp</th>
<th>Arg</th>
<th>Lys</th>
<th>Asn</th>
<th>Arg</th>
<th>Lys</th>
<th>Leu</th>
<th>Val</th>
<th>Ile</th>
<th>Gln</th>
<th>Glu</th>
<th>Val</th>
<th>Leu</th>
</tr>
</thead>
<tbody>
<tr>
<td>130</td>
<td>135</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Ser Ala Ser Arg Leu Glu Lys Phe Lys Glu Val Arg Phe Thr Arg Ile 145 150 155 160
Gln Thr Ser Ile Lys Asn Leu Tyr Thr Arg Ile Asn Gly Asn Ser Ser 165 170 175
Thr Ile Asn Leu Ser Asp Trp Leu Glu Leu Asn Phe Gly Leu Ile 180 185 190
Val Lys Met Ile Ala Gly Lys Tyr Glu Ser Gly Lys Gly Asp Glu 195 200 205
Gln Val Glu Arg Phe Lys Ala Phe Lys Asp Phe Met Val 210 215 220

<210> SEQ ID NO 13
<211> LENGTH: 99
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Partial coding sequence selected from SEQ ID NO:1

<400> SEQUENCE: 13
gacagagctc tgatgatga ggaaggtgca gcataaca ca tgaaggtt aataccagtg 60
gaatgataaa tacgctctcg cttggcacct gactttac 99

<210> SEQ ID NO 14
<211> LENGTH: 99
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Partial coding sequence selected from SEQ ID NO:3

<400> SEQUENCE: 14
gacagagctc tgatgatga ggaaggtgca gcataaca ta cgaaggtt aataccagtg 60
gaatgataaa tacgctctcg cttggcacct gactttat 99

<210> SEQ ID NO 15
<211> LENGTH: 99
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Partial coding sequence selected from SEQ ID NO:5

<400> SEQUENCE: 15
gacagagctc tgacatgaa ggaaggtgca gcataacta ta cgaaggtt aataccgttg 60
gaatgataaa tacgctctcg cttggcacct gactttat 99

<210> SEQ ID NO 16
<211> LENGTH: 99
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Partial coding sequence selected from SEQ ID NO:7

<400> SEQUENCE: 16
gatgagcctc tgatgatga ggaaggtgca gcataacta ta cgaaggtt aataccgttg 60
gaatgataaa tacgctctcg cttggcacct gactttat 99

<210> SEQ ID NO 17
<211> LENGTH: 19
<210> SEQ ID NO 18  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Primer  
<400> SEQUENCE: 18  

ctctctagaa ctggtgactc  

<210> SEQ ID NO 19  
<211> LENGTH: 19  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Primer  
<400> SEQUENCE: 19  

ttttgggac astcagtcaa  

<210> SEQ ID NO 20  
<211> LENGTH: 26  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Primer  
<400> SEQUENCE: 20  

tttcattcga attatatttat gcacca  

<210> SEQ ID NO 21  
<211> LENGTH: 21  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Primer  
<400> SEQUENCE: 21  

atgtttttt ccattgaagc c  

<210> SEQ ID NO 22  
<211> LENGTH: 17  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Primer  
<400> SEQUENCE: 22  

tcgagtcgca cggtatc  

<210> SEQ ID NO 23  
<211> LENGTH: 118  
<212> TYPE: DNA  
<213> ORGANISM: Nicotiana sp.  
<400> SEQUENCE: 23  

cgtctcattaa agcaacggggt tttgtaagtt catcgctcat ttttctattta ttctactttta
ttttgaggacagactgtgtaaataaattttggacgactgtgtaattatatagtgcggctccttcttctac}
aggtcagcctctccggcaaacattagctatagtaggtagtctgccatatataacgtagattttttg
ccttaatacggtgccgaactcgagcggtaagtgaggtctgttttatggagttagtcttttttttgc
ttatgcagctcctctcttttagggattgcagtagaaattagggattaggattgat
agaattgcagttggggtgtaattttttgaattttttgccctttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
<210> SEQ ID NO 25
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR Primer

<400> SEQUENCE: 25
ctgtcattaa agcaacggtg

<210> SEQ ID NO 26
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer

<400> SEQUENCE: 26
ggccttgacg aagcacaag

<210> SEQ ID NO 27
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer

<400> SEQUENCE: 27
aagcctttgac gcattttttt ccatcg

<210> SEQ ID NO 28
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer

<400> SEQUENCE: 28
tctgagtttt tcaagctgataa tcttcac

<210> SEQ ID NO 29
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer

<400> SEQUENCE: 29
acgtgatct aaactctgtgt ctg

<210> SEQ ID NO 30
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer

<400> SEQUENCE: 30
gcctgcacct tctttctcag

<210> SEQ ID NO 31
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer
<400> SEQUENCE: 31

cggtaatcg gcactttttc

<210> SEQ ID NO 32
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer

<400> SEQUENCE: 32
cggagttttc gagctaatgg a

<210> SEQ ID NO 33
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer

<400> SEQUENCE: 33
catgacgac cgccgacag

<210> SEQ ID NO 34
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer

<400> SEQUENCE: 34
atggtttttc catagaagc c

<210> SEQ ID NO 35
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer

<400> SEQUENCE: 35
tttttgggac atcagtccaa t

<210> SEQ ID NO 36
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer

<400> SEQUENCE: 36
tgaatgaact gcaggacgag

<210> SEQ ID NO 37
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer

<400> SEQUENCE: 37
aatcaccg gtgcaccacg
What is claimed is:

1. A method for reducing the level of nicotine or N’-nictrosonornicotine in a plant of the genus *Nicotiana* or a plant part thereof, said method comprising growing a plant of the genus *Nicotiana*, wherein said plant, or a plant part thereof, comprises an introduced expression cassette that comprises an inhibitory sequence that is operably linked to a promoter that is functional in a plant cell, wherein said inhibitory sequence comprises a first nucleotide sequence, wherein said first nucleotide sequence is a fragment of a polynucleotide having at least 90% sequence identity to the sequence set forth in SEQ ID NO:3, wherein said fragment is from a region of said polynucleotide selected from the group consisting of:

   a) a region that corresponds to nucleotide position 265 to nucleotide position 625 of SEQ ID NO:3, or a complement thereof; and

   b) a region that corresponds to nucleotide position 1420 to nucleotide position 1580 of SEQ ID NO:3, or a complement thereof;

   the size of said fragment being from about 100 nucleotides to the full length of said region.

2. The method of claim 1, wherein said inhibitory sequence is capable of being transcribed as an inhibitory polynucleotide selected from the group consisting of a single-stranded RNA polynucleotide, a double-stranded RNA polynucleotide, and a combination thereof.

3. The method of claim 1, wherein said promoter is selected from the group consisting of a constitutive promoter, an inducible promoter, and a tissue-preferred promoter.

4. The method of claim 3, wherein said tissue-preferred promoter is a leaf-preferred promoter.

5. The method of claim 1, wherein said expression cassette is incorporated into a vector.


7. A method for reducing the level of nicotine or N’-nictrosonornicotine in a tobacco product, said method comprising preparing said tobacco product from a tobacco plant, or a plant part thereof, that has been genetically modified to inhibit the expression of a cytochrome P450 that is involved in the metabolic conversion of nicotine to nornicotine in said tobacco plant, or said plant part thereof, wherein expression of said cytochrome P450 in said tobacco plant, or said plant part thereof, is inhibited by introduction of an expression cassette into said tobacco plant, or a plant part thereof, wherein said expression cassette comprises an inhibitory sequence that is operably linked to a promoter that is functional in a plant cell, wherein said inhibitory sequence comprises a first nucleotide sequence, wherein said first nucleotide sequence is a fragment of a polynucleotide having at least 90% sequence identity to the sequence set forth in SEQ ID NO:3, wherein said fragment is from a region of said polynucleotide selected from the group consisting of:

   a) a region that corresponds to nucleotide position 265 to nucleotide position 625 of SEQ ID NO:3, or a complement thereof; and

   b) a region that corresponds to nucleotide position 1420 to nucleotide position 1580 of SEQ ID NO:3, or a complement thereof;

   the size of said fragment being from about 100 nucleotides to the full length of said region.

8. The method of claim 7, wherein said inhibitory sequence is capable of being transcribed as an inhibitory polynucleotide selected from the group consisting of a single-stranded RNA polynucleotide, a double-stranded RNA polynucleotide, and a combination thereof.

9. The method of claim 7, wherein said promoter is selected from the group consisting of a constitutive promoter, an inducible promoter, or a tissue-preferred promoter.

10. The method of claim 9, wherein said tissue-preferred promoter is a leaf-preferred promoter.

11. The method of claim 7, wherein said expression cassette is incorporated into a vector.

12. The method of claim 7, wherein said tobacco product is selected from the group consisting of chewing tobacco, snuff, cigarettes, pipe tobacco, and cigars.

13. The method of claim 7, wherein said tobacco plant is a Burley tobacco variety.

14. The method of claim 7, wherein said tobacco product has a reduced carcinogenic potential.

15. The method of claim 1, wherein said inhibitory sequence comprises a second nucleotide sequence that is fully or partially complementary to said first nucleotide sequence.

16. The method of claim 15, wherein said polynucleotide comprises the sequence set forth in SEQ ID NO:3, and wherein said fragment consists of nucleotides 265 to 625 of SEQ ID NO:3.

17. The method of claim 15, wherein said fragment of said polynucleotide consists of nucleotides from a region of said polynucleotide that corresponds to nucleotide position 297 to nucleotide position 594 of SEQ ID NO:3.

18. The method of claim 17, wherein said polynucleotide comprises the sequence set forth in SEQ ID NO:3, and wherein said fragment consists of nucleotides 297 to 594 of SEQ ID NO:3.

19. The method of claim 18, wherein said second nucleotide sequence comprises the complement of nucleotides 297 to 594 of SEQ ID NO:3.

20. The method of claim 15, wherein said first and said second nucleotide sequences are linked by an intron sequence.

21. The method of claim 20, wherein said intron sequence is a spliceable intron sequence.

22. The method of claim 20, wherein said polynucleotide comprises the sequence set forth in SEQ ID NO:3, and wherein said fragment consists of nucleotides 265 to 625 of SEQ ID NO:3.

23. The method of claim 20, wherein said fragment of said polynucleotide consists of nucleotides from a region of said polynucleotide that corresponds to nucleotide position 297 to nucleotide position 594 of SEQ ID NO:3.

24. The method of claim 23, wherein said polynucleotide comprises the sequence set forth in SEQ ID NO:3, and wherein said fragment consists of nucleotides 297 to 594 of SEQ ID NO:3.

25. The method of claim 24, wherein said second nucleotide sequence comprises the complement of nucleotides 297 to 594 of SEQ ID NO:3.