(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 July 2011 (21.07.2011)

PCT

(10) International Publication Number
WO 2011/088180 A1

(51) International Patent Classification:
C12N 15/82 (2006.01) *A01H 5/12* (2006.01)
C12N 15/53 (2006.01)

(21) International Application Number:
PCT/US2011/021088

(22) International Filing Date:
13 January 2011 (13.01.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/295,671 15 January 2010 (15.01.2010) US

(71) Applicant *(for all designated States except US)*: **NORTH CAROLINA STATE UNIVERSITY** [US/US]; Campus Box 8210, Raleigh, NC 27695 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **DEWEY, Ralph, E.** [US/US]; 2403 Colony Woods Drive, Apex, NC 27523 (US). **LEWIS, Ramsey, S.** [US/US]; 8640 Glade Hill Drive, Apex, NC 27539 (US).

(74) Agents: **HENRY, Leslie, T.** et al.; ALSTON & BIRD LLP, Bank Of America Plaza, 101 South Tryon Street, Suite 4000, Charlotte, NC 28280-4000 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: COMPOSITIONS AND METHODS FOR MINIMIZING NORNICOTINE SYNTHESIS IN TOBACCO

(57) Abstract: Compositions and methods for reducing the level of nornicotine and *N'*-nitrosonornicotine (NNN) in tobacco plants and plant parts thereof are provided. The compositions comprise isolated polynucleotides and polypeptides for a root-specific nicotine demethylases, CYP82E10, and variants thereof, that are involved in the metabolic conversion of nicotine to nornicotine in these plants. Compositions of the invention also include tobacco plants, or plant parts thereof, comprising a mutation in a gene encoding a CYP82E10 nicotine demethylase, wherein the mutation results in reduced expression or function of the CYP82E10 nicotine demethylase. Seed of these tobacco plants, or progeny thereof, and tobacco products prepared from the tobacco plants of the invention, or from plant parts or progeny thereof, are also provided. Methods for reducing the level of nornicotine, or reducing the rate of conversion of nicotine to nornicotine, in a tobacco plant, or plant part thereof are also provided. The methods comprise introducing into the genome of a tobacco plant a mutation within at least one allele of each of at least three nicotine demethylase genes, wherein the mutation reduces expression of the nicotine demethylase gene, and wherein a first of these nicotine demethylase genes encodes a root- specific nicotine demethylase involved in the metabolic conversion of nicotine to nornicotine in a tobacco plant or a plant part thereof. 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.

WO 2011/088180 A1

COMPOSITIONS AND METHODS FOR MINIMIZING NORNICOTINE SYNTHESIS IN TOBACCO

INCORPORATION OF SEQUENCE LISTING

An official copy of the Sequence Listing is submitted electronically via EFS-Web as an ASCII formatted Sequence Listing with a file named "400712SequenceListing.txt," created on January 12, 2011, having a size of 149 KB and is filed concurrently with the specification. The Sequence Listing contained in this ASCII formatted document is part of the specification and is herein incorporated by reference in its entirety.

FIELD OF THE INVENTION

The invention relates to compositions and methods for minimizing nornicotine synthesis, and hence its metabolite N'-nitrosonornicotine, in tobacco plants and plant parts thereof, particularly compositions and methods for inhibiting expression or function of a root-specific nicotine demethylase in combination with a green leaf and a senescence-induced nicotine demethylase.

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 fraction is comprised primarily of three additional pyridine alkaloids: nornicotine, anabasine, and anatabine. Nornicotine is generated directly from nicotine through the activity of the enzyme nicotine N-demethylase. Nornicotine 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 nornicotine give rise to progeny that metabolically "convert" a large percentage of leaf nicotine to nornicotine. In tobacco plants that have genetically converted (termed "converters"), the great majority of nornicotine 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 nornicotine is metabolized to the compound N-nitrosonornicotine (NNN), a tobacco-specific nitrosamine (TSNA) that has been asserted 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, TSNAs are 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 Nitrosamines in Flue-cured Tobacco," CORESTA Meeting, Agro-Phyto Groups, Suzhou, China). Retrofitting these curing barns with heat-exchangers virtually eliminated the mixing of combustion gases with the curing air and dramatically reduced the formation of TSNAs in tobaccos cured in this manner (Boyette and Hamm (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 nitrite, a process catalyzed by leaf-borne microbes (Bush *et al.* (2001) *Rec. Adv. Tob. Sci.* 27:23-46). Thus far, attempts to reduce TSNAs through modification of curing conditions while maintaining acceptable quality standards have not proven to be successful for the air-cured tobaccos.

Aside from serving as a precursor for NNN, recent studies suggest that the nornicotine found in tobacco products may have additional undesirable health consequences. Dickerson and Janda (2002) *Proc. Natl. Acad. Sci. USA* 99: 15084-15088 demonstrated that nornicotine causes aberrant protein glycation within the cell. Concentrations of nornicotine-modified proteins were found to be much higher in the plasma of smokers compared to nonsmokers. This same study also 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. Furthermore, studies have been reported linking the nornicotine found in tobacco products with age-related macular degeneration, birth defects, and periodontal disease (Brogan et al. (2005) *Proc. Natl. Acad. Sci. USA* 102: 10433-10438; Katz et al. (2005) *J. Periodontol.* 76: 1171-1174).

In Burley tobaccos, a positive correlation has been found between the nornicotine 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). Therefore, strategies that could effectively reduce the nornicotine content of the leaf 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. This correlation was further solidified in the recent study by Lewis et al. (2008) *Plant Biotech. J.* 6: 346-354 who demonstrated that lowering nornicotine levels using an RNAi transgene construct directed against the *CYP82E4v2* gene, which encodes a senescence-induced nicotine demethylase, lead to concomitant reductions in the NNN content of the cured leaf. Although this study demonstrated that transgenic technologies can be used to greatly reduce the nornicotine and NNN content of tobacco, a combination of public perception and intellectual property issues make it very difficult for commercialization of products derived from transgenic plants.

Therefore a great need exists for a means to effectively minimize nornicotine accumulation in tobacco that does not rely on the use of transgenics.

SUMMARY OF THE INVENTION

Compositions and methods for minimizing the nornicotine content in tobacco plants and plant parts thereof are provided. Compositions include an isolated root-specific cytochrome P450 polynucleotide designated the *CYP82E10* polynucleotide, as set forth in SEQ ID NO:1, and CYP82E10 nicotine demethylase polypeptide encoded thereby, as set forth in SEQ ID NO:2, and variants and fragments thereof, including, but not limited to, polypeptides comprising the sequence set forth in SEQ ID NO:5, 6, 7, 8, 9, 10, 11, 12, or 13, as well as polynucleotides encoding the polypeptide set forth in SEQ ID NO:5, 6, 7, 8, 9, 10, 11, 12, or 13. The CYP82E10 polypeptide of the invention is a nicotine demethylase that is involved in the metabolic conversion of nicotine to nornicotine in the roots of tobacco plants. Isolated polynucleotides of the invention also include a polynucleotide comprising the sequence set forth in SEQ ID NO:3 or 4, and variants and fragments thereof. Compositions of the invention also include tobacco plants, or plant parts thereof, comprising a mutation in a gene encoding a CYP82E10 nicotine demethylase, wherein the mutation results in reduced expression or function of the CYP82E10 nicotine demethylase. In some embodiments, the tobacco plants of the invention further comprise a mutation in a gene encoding a CYP82E4 nicotine demethylase and/or a mutation in a gene encoding a CYP82E5 nicotine demethylase, wherein the mutation within these genes results in reduced expression or function of the CYP82E4 or CYP82E5 nicotine demethylase. Seed of these tobacco plants, or progeny thereof, and tobacco products prepared from the tobacco plants of the invention, or from plant parts or progeny thereof, are also provided.

Methods for reducing the level of nornicotine, or reducing the rate of conversion of nicotine to nornicotine, in a tobacco plant, or plant part thereof are also provided. The methods comprise introducing into the genome of a tobacco plant a mutation within at least one allele of each of at least three nicotine demethylase genes, wherein the mutation reduces expression of the nicotine demethylase gene, and wherein a first of these nicotine demethylase genes encodes a root-specific nicotine demethylase involved in the metabolic conversion of nicotine to nornicotine in a tobacco plant or a plant part thereof. In some embodiments, the root-specific nicotine demethylase is CYP82E10 or variant thereof. In other embodiments, these methods comprise introducing into the genome of a tobacco plant a mutation within at least one allele of a nicotine demethylase gene encoding CYP82E10 or variant thereof, and a mutation within at least one allele of a nicotine demethylase encoding CYP82E4 or variant thereof, and/or a nicotine demetylase encoding CYP82E5 or variant thereof. Methods for identifying a tobacco plant with low levels of nornicotine are also provided, wherein the plant or plant part thereof is screened for the presence of a mutation in a gene encoding CYP82E10 or variant thereof, alone or in combination with screening for the presence of a mutation in a gene encoding CYP82E4 or variant thereof, and/or the presence of a mutation in a gene encoding CYP82E5 or variant thereof.

The following embodiments are encompassed by the present invention.

1. A tobacco plant, or plant part thereof, comprising a mutation in a gene encoding a CYP82E10 nicotine demethylase, wherein said mutation results in reduced expression or function of said CYP82E10 nicotine demethylase.

2. The tobacco plant, or plant part thereof, according to embodiment 1, wherein said CYP82E10 nicotine demethylase is selected from the group consisting of the sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, and 9.

3. The tobacco plant, or plant part thereof, according to embodiment 1 or 2, wherein said mutation results in a modification of said CYP82E10 nicotine demethylase occurring at a position selected from the group consisting of amino acid residues 79, 107, 382, 419, and any combination thereof, wherein said numbering is according to SEQ ID NO:2.

4. The tobacco plant, or plant part thereof, according to embodiment 3, wherein said mutation is selected from the group consisting of:

a) a serine substitution for the glycine residue at position 79;

b) a serine substitution for the proline residue at position 107;

c) a serine substitution for the proline residue at position 382;

d) a serine substitution for the proline residue at position 419; and

e) any combination thereof.

5. The tobacco plant, or plant part thereof, according to any of embodiments 1-4, further comprising a mutation in a gene encoding a CYP82E4 nicotine demethylase, wherein said mutation results in reduced expression or function of said CYP82E4 nicotine demethylase.

6. The tobacco plant, or plant part thereof, according to embodiment 5, wherein said CYP82E4 nicotine demethylase is selected from the sequence set forth in SEQ ID NO:14, 15, 16, 17, 18, 19, and 20.

7. The tobacco plant, or plant part thereof, according to embodiment 5 or 6, wherein said mutation results in a modification of said CYP82E4 nicotine demethylase occurring at a position selected from the group consisting of amino acid residues 329, 364, 376, 382, and 458, wherein said numbering is according to SEQ ID NO:14.

8. The tobacco plant, or plant part thereof, according to embodiment 7, wherein said mutation is selected from the group consisting of:

a) a stop codon substitution for the tryptophan residue at position 329;

b) an asparagine substitution for the lysine residue at position 364;

c) a methionine substitution for the valine residue at position 376;

d) a serine substitution for the proline residue at position 382;

d) a serine substitution for the proline residue at position 458; and

e) any combination thereof.

9. The tobacco plant, or plant part thereof, according to any of embodiments 1-8, further comprising a mutation in a gene encoding a CYP82E5 nicotine demethylase, wherein said mutation results in reduced expression or function of said CYP82E5 nicotine demethylase.

10. The tobacco plant, or plant part thereof, according to embodiment 9, wherein said CYP82E5 nicotine demethylase is selected from the sequence set forth in SEQ ID NO:26, 27, 28, 29, 30, 31, and 32.

11. The tobacco plant, or plant part thereof, according to embodiment 9 or 10, wherein said mutation results in a modification of said CYP82E5 nicotine demethylase occuring at a position selected from the group consisting of amino acid residues 422 and 449, wherein said numbering is according to SEQ ID NO:26.

12. The tobacco plant, or plant part thereof, according to embodiment 11, wherein said mutation is selected from the group consisting of:

a) a stop codon substituted for the tryptophan residue at position 422;

b) a leucine substituted for the proline residue at position 449; and

c) any combination thereof.

13. The tobacco plant, or plant part thereof, according to any of embodiments 9-12, comprising a mutation in said CYP82E10 nicotine demethylase gene and said CYP82E4 nicotine demethylase gene.

14. The tobacco plant, or plant part thereof, according to any of embodiments 1-13, wherein said tobacco plant, or plant part thereof, is homozygous for said mutation.

15. The tobacco plant, or plant part thereof, according to embodiment 14, wherein said CYP82E10 nicotine demethylase comprises a mutation at position 382, said CYP82E4 nicotine demethylase comprises a mutation at position 329, and said CYP82E5 nicotine demethylase comprises a mutation at position 422, wherein said numbering is according to SEQ ID NO:2, 14, and 26, respectively.

16. The tobacco plant, or plant part thereof, according to embodiment 15, wherein said mutation is selected from the group consisting of:

a) a serine substitution for the proline residue at position 382;

b) a stop codon substitution for the tryptophan residue at position 329;

c) a stop codon substitution for the tryptophan residue at position 422; and

d) any combination thereof.

17. The tobacco plant, or plant part thereof, according to any of embodiments 13-16, wherein said plant or plant part thereof has less than 1.5% conversion of nicotine to nornicotine.

18. The tobacco plant, or plant part thereof, according to embodiment 17, wherein said plant or plant part thereof has no more than 0.5% conversion of nicotine to nornicotine.

19. Seed of the tobacco plant according to any of embodiments 1-18, or progeny thereof.

20. A tobacco product prepared from a tobacco plant, or plant part or progeny thereof, according to any of embodiments 1-19.

21. A method for reducing a carcinogenic potential of a tobacco product, said method comprising preparing said tobacco product from a tobacco plant, or plant part or progeny thereof, according to any of embodiments 1-18.

22. A method for reducing the level of nornicotine, or reducing the rate of conversion of nicotine to nornicotine, in a tobacco plant, or a plant part thereof, said method comprising introducing into the genome of said plant a mutation within at least one allele of each of at least three nicotine demethylase genes, wherein said mutation reduces expression of said nicotine demethylase gene, and wherein a first of said nicotine demethylase genes encodes a root-specific nicotine demethylase involved in the metabolic conversion of nicotine to nornicotine in a tobacco plant or a plant part thereof.

23. The method of embodiment 22, wherein said root-specific nicotine demethylase is a CYP82E10 nicotine demethylase comprising an amino acid sequence selected from the group consisting of:

a) the amino acid sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, 9, or 10; and

b) an amino acid sequence having at least 98% sequence identity to the amino acid sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, 9, or 10.

24. The method of embodiment 23, wherein said amino acid sequence for said CYP82E10 nicotine demethylase has a substitution at an amino acid residue in a position selected from the group consisting of residues 79, 107, 382, 419, and any combination thereof, where the numbering is according to SEQ ID NO:2.

25. The method of embodiment 24, wherein said substitution at position 79, 107, 382, or 419 is a serine residue.

26. The method of any one of embodiments 22-25, wherein a second of said nicotine demethylase genes encodes a CYP82E4 nicotine demethylase.

27. The method of embodiment 26, wherein said CYP82E4 nicotine demethylase comprises an amino acid sequence selected from the group consisting of:

a) the amino acid sequence set forth in SEQ ID NO:14, 15, 16, 17, 18, 19, 20, or 21; and

b) an amino acid sequence having at least 98% sequence identity to the sequence set forth in SEQ ID NO:14, 15, 16, 17, 18, 19, 20, or 21.

28. The method of embodiment 27, wherein said amino acid sequence for said CYP82E4 nicotine demethylase has a substitution at an amino acid residue in a position selected from the group consisting of residues 329, 364, 382, 458, and any combination thereof, where the numbering is according to SEQ ID NO:14.

29. The method of embodiment 28, wherein said substitution at position 329 is a stop codon, said substitution at position 364 is an asparagine residue, said substitution at position 382 is a serine residue, said substitution at position 458 is a serine residue, or any combination thereof.

30. The method of any one of embodiments 22-29, wherein a third of said nicotine demethylase genes encodes a CYP82E5 nicotine demethylase.

31. The method of embodiment 30, wherein said CYP82E5 nicotine demethylase comprises an amino acid sequence selected from the group consisting of:

a) the amino acid sequence set forth in SEQ ID NO:26, 27, 28, 29, 30, 31, or 32; and

b) an amino acid sequence having at least 98% sequence identity to the sequence set forth in SEQ ID NO: 26, 27, 28, 29, 30, 31, or 32.

32. The method of embodiment 31, wherein said amino acid sequence for said CYP82E5 nicotine demethylase has a substitution at an amino acid residue in a position selected from the group consisting of residues 422 and 449, and any combination thereof, where the numbering is according to SEQ ID NO:26.

33. The method of embodiment 32, wherein said substitution at position 422 is a stop codon, said substitution at position 449 is a leucine residue, or any combination thereof.

34. The method of any one of embodiments 22-33, wherein said plant or plant part thereof is homozygous for said mutation.

35. The method of any one of embodiments 22-34, wherein said introducing comprises a breeding protocol.

36. The method of any one of embodiments 22-35, wherein said plant is a Burley, Virginia, flue-cured, air-cured, fire-cured, Oriental, or a dark tobacco plant.

37. The tobacco plant, or plant part thereof, according to any of embodiments 1-18, wherein said tobacco plant is a Burley, Virginia, flue-cured, air-cured, fire-cured, Oriental, or a dark tobacco plant.

38. A method for identifying a tobacco plant with low levels of nornicotine, said method comprising screening a DNA sample from a tobacco plant of interest for the presence of a mutation in SEQ ID NO:1 or 3.

39. The method according to embodiment 38, wherein said tobacco plant is a nonconverter.

40. The method according to embodiment 38 or 39, wherein said screening is carried out using a sequence selected from the group consisting of SEQ ID NOS:1, 3, 35, 36, 37, and 38.

41. The method according to any one of embodiments 38-40, further comprising screening said DNA sample, or another DNA sample from said tobacco plant of interest, for the presence of a mutation in SEQ ID NO:14, the presence of a mutation in SEQ ID NO:26, or the presence of a mutation in SEQ ID NO:14 and SEQ ID NO:26.

42. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

a) a nucleotide sequence comprising SEQ ID NO:1, 3, or 4;

b) a nucleotide sequence comprising a fragment of at least 20 consecutive nucleotides of SEQ ID NO:1, 3, or 4;

c) a nucleotide sequence having at least 97% sequence identity to the entirety of the sequence set forth in SEQ ID NO:1, wherein said polynucleotide encodes a polypeptide involved in the metabolic conversion of nicotine to nornicotine in a plant;

d) a nucleotide sequence encoding a polypeptide selected from the group consisting of SEQ ID NOS:2 and 5-13, or a fragment thereof comprising at least 115 contiguous residues;

e) a nucleotide sequence encoding a polypeptide having at least 98% sequence identity to the sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, 9, 10, 11, 12, or 13; and

f) a nucleotide sequence that is complementary to the sequence according to any of preceding items (a) through (e).

43. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

a) an amino acid sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, 9, 10, 11, 12, or 13;

b) an amino acid sequence that is at least 98% identical to an amino acid sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, 9, 10, 11, 12, or 13; and

c) an amino acid sequence that is a fragment of the amino acid sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, 9, 10, 11, 12, or 13, wherein said fragment comprises at least 115 contiguous residues of the amino acid sequence of SEQ ID NO:2, 5, 6, 7, 8, 9, 10, 11, 12, or 13.

44. A tobacco plant, or plant part thereof that is homozygous for a mutation in a gene encoding a CYP82E10 nicotine demethylase, a gene encoding a CYP82E4 nicotine demethylase, and a gene encoding a CYP82E5 nicotine demethylase, wherein said mutation results in reduced expression or function of said CYP82E10, CYP82E4, and CYP82E5 nicotine demethylase, wherein said CYP82E10 nicotine demethylase comprises a mutation at position 382, said CYP82E4 nicotine demethylase comprises a mutation at position 329, and said CYP82E5 nicotine demethylase comprises a mutation at position 422, wherein said numbering is according to SEQ ID NO:2, 14, and 26, respectively.

45. A mutation in a gene encoding a CYP82E10 nicotine demethylase, wherein said mutation results in reduced expression or function of said CYPe2E10 nicotine demethylase.

46. A plant having a mutation in a *CYP82E10* gene that inhibits nicotine demethylase activity in roots, a mutation in a *CYP82E4v2* gene that inhibits nicotine demethylase activity in senescent leaves, and a mutation in a *CYP83E5* gene that inhibits nicotine demethylase activity in green leaves.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1A-C shows the DNA (SEQ ID NO:4) and predicted protein sequences of the *CYP82E10* nicotine demethylase gene. The protein coding sequences are in uppercase, and 5' and 3' flanking sequences are in lowercase. The intron sequence (SEQ ID NO:3) is lowercase italicized. Numbers for the nucleotide sequence are shown on the left and numbers for the protein sequence are labeled on the right. Nucleotide sequences corresponding to the PCR primers used to specifically amplify exon 1 for mutation screening are underlined (not shown in bold), whereas underlined sequences in bold denote the exon 2-specific primer sites. Individual nucleotide and amino acid residues that were found to be altered in the mutation screen (Table 2) are underlined and in bold.

Figure 2A-C shows an alignment of genomic sequences for *CYP82E10* (SEQ ID NO:4), *CYP82E5v2* (SEQ ID NO:38), and *CYP82E4v2* (SEQ ID NO:37). Protein-encoding sequences are in upper case type; 5' and 3' untranslated regions are indicated in lower case type; and intron sequences are shown in lower case italicized type. Positions of shared sequence identity are box shaded.

Figure 3 shows thin layer chromatographic data of nicotine demethylase activities of microsomal membranes from yeast cells expressing CYP82E10, and CYP82E10 possessing the Ser382Pro (S382P) mutation from plant 1041. CPM, counts per minute.

Figures 4A and 4B show mean percent nicotine conversion for burley tobacco plants with varying mutant combinations at *CYP82E4v2*, *CYP82E5v2*, and *CYP82E10* loci. Means with different letters are significantly different at the $P < 0.05$ level.

DESCRIPTION OF THE SEQUENCES OF THE SEQUENCE LISTING

The following listing sets forth the sequence information for the Sequence Listing. Standard notation for amino acid substitutions is used. Thus, for example, CYP82E10 P419S indicates the variant protein has a serine substitution for the proline residue at position 419, where the numbering is with respect to the wild-type sequence, in this case, the CYP82E10 sequence set forth in SEQ ID NO:2. As another example, CYP82E4 P38L indicates the variant protein has a leucine substitution for the proline residue at position 38, where the numbering is with respect to the wild-type sequence, in this case, the CYP82E4 sequence set forth in SEQ ID NO:14. As yet another example, CYP82E5 P72L indicates the variant protein has a leucine substitution for the proline residue at position 72, where the numbering is with respect to the wild-type sequence, in this case, the CYP82E5 sequence set forth in SEQ ID NO:26.

SEQ ID NO:1 sets forth a coding sequence for CYP82E10.

SEQ ID NO:2 sets forth the amino acid sequence for CYP82E10.

SEQ ID NO:3 sets forth the nucleotide sequence of an intron of the *CYP82E10* gene.

SEQ ID NO:4 sets forth the genomic sequence for *CYP82E10*.

SEQ ID NO:5 sets forth the amino acid sequence for CYP82E10 L148F.

SEQ ID NO:6 sets forth the amino acid sequence for CYP82E10 G172R.

SEQ ID NO:7 sets forth the amino acid sequence for CYP82E10 A344T.

SEQ ID NO:8 sets forth the amino acid sequence for CYP82E10 A410T.

SEQ ID NO:9 sets forth the amino acid sequence for CYP82E10 R417H.

SEQ ID NO:10 sets forth the amino acid sequence for CYP82E10 P419S.

SEQ ID NO:11 sets forth the amino acid sequence for CYP82E10 G79S.

SEQ ID NO:12 sets forth the amino acid sequence for CYP82E10 P107S.

SEQ ID NO:13 sets forth the amino acid sequence for CYP82E10 P382S.

SEQ ID NO:14 sets forth the amino acid sequence for CYP82E4.

SEQ ID NO:15 sets forth the amino acid sequence for CYP82E4 P38L.

SEQ ID NO:16 sets forth the amino acid sequence for CYP82E4 D171N.

SEQ ID NO:17 sets forth the amino acid sequence for CYP82E4 E201K.

SEQ ID NO:18 sets forth the amino acid sequence for CYP82E4 R169Q.

SEQ ID NO:19 sets forth the amino acid sequence for CYP82E4 G459R.

SEQ ID NO:20 sets forth the amino acid sequence for CYP82E4 T427I.

SEQ ID NO:21 sets forth the amino acid sequence for CYP82E4 V376M.

SEQ ID NO:22 sets forth the amino acid sequence for CYP82E4 W329Stop.

SEQ ID NO:23 sets forth the amino acid sequence for CYP82E4 K364N.

SEQ ID NO:24 sets forth the amino acid sequence for CYP82E4 P382S.

SEQ ID NO:25 sets forth the amino acid sequence for CYP82E4 P458S.

SEQ ID NO:26 sets forth the amino acid sequence for CYP82E5.

SEQ ID NO:27 sets forth the amino acid sequence for CYP82E5 P72L.

SEQ ID NO:28 sets forth the amino acid sequence for CYP82E5 L143F.

SEQ ID NO:29 sets forth the amino acid sequence for CYP82E5 S174L.

SEQ ID NO:30 sets forth the amino acid sequence for CYP82E5 M224I.

SEQ ID NO:31 sets forth the amino acid sequence for CYP82E5 P235S.

SEQ ID NO:32 sets forth the amino acid sequence for CYP82E5 A410V.

SEQ ID NO:33 sets forth the amino acid sequence for CYP82E5 W422Stop.

SEQ ID NO:34 sets forth the amino acid sequence for CYP82E5 P449L.

SEQ ID NO:35 sets forth the forward primer sequence for exon 1 of *CYP82E10*.

SEQ ID NO:36 sets forth the reverse primer sequence for exon 1 of *CYP82E10*.

SEQ ID NO:37 sets forth the forward primer sequence for exon 2 of *CYP82E10*.

SEQ ID NO:37 sets forth the reverse primer sequence for exon 2 of *CYP82E10*.

SEQ ID NO:38 sets forth the genomic sequence for *CYP82E4v2*.

SEQ ID NO:39 sets forth the genomic sequence for *CYP82E5v2*.

DEFINITIONS

The present invention includes compositions and methods for inhibiting expression or function of root-specific nicotine demethylase polypeptides that are involved in the metabolic conversion of nicotine to nornicotine in the roots of a plant, particularly plants of the *Nicotiana* genus, including tobacco plants of various commercial varieties.

As used herein, "inhibit," "inhibition" and "inhibiting" are defined as any method known in the art or described herein, which decreases the expression or function of a gene product of interest (*i.e.*, the target gene product), in this case a nicotine demethylase, such as a root-specific nicotine demethylase of the invention. It is recognized that nicotine demethylase polypeptides can be inhibited by any suitable method known in the art, including sense and antisense suppression, RNAi suppression, knock out approaches such as mutagenesis, and the like. Of particular interest are methods that knock out, or knock down, expression and/or function of these root-specific nicotine demethylases, particularly mutagenic approaches that allow for selection of favorable mutations in the *CYP82E10* nicotine demethylase gene.

By "favorable mutation" is intended a mutation that results in a substitution, insertion, deletion, or truncation of the CYP82E10 polypeptide such that its nicotine demethylase activity is inhibited. In some embodiments, the nicotine demethylase activity is inhibited by at least 25%, 30%, 35, 40%, 45, 50%, 55%, or 60% when compared to the activity of the wild-type CYP82E10 polypeptide under the same test conditions. In other embodiments, the nicotine demethylase activity is inhibited by at least 65%, 70%, 75%, 80%, 85%, 90%, or 95%. In preferred embodiments, the favorable mutation provides for complete inhibition (i.e., 100% inhibition), and the nicotine demethylase activity is knocked out (i.e., its activity cannot be measured).

"Inhibiting" can be in the context of a comparison between two plants, for example, a genetically altered plant versus a wild-type plant. The comparison can be between plants, for example, a wild-type plant and one of which lacks a DNA sequence capable of producing a root-specific nicotine demethylase that converts nicotine to nornicotine. Inhibition of expression or function of a target gene product also can be in the context of a comparison between plant cells, organelles, 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.

"Inhibiting" can include any relative decrement of function or production of a gene product of interest, in this case, a root-specific nicotine demethylase, up to and including complete elimination of function or production of that gene product. When levels of a gene product are compared, such a comparison is preferably carried out between organisms with a similar genetic background. Preferably, a similar genetic background is a background where the organisms being compared share 50% or greater, more preferably 75% or greater, and, even more preferably 90% or greater sequence identity of nuclear genetic material. A similar genetic background is a background where the organisms being compared are plants, and the plants are isogenic except for any genetic material originally introduced using plant transformation techniques or a mutation generated by human intervention. Measurement of the level or amount of a gene product may be carried out by any suitable method, non-limiting examples of which include, but are not limited to, comparison of mRNA transcript levels, protein or peptide levels, and/or phenotype, especially the conversion of nicotine to nornicotine. As used herein, mRNA transcripts can include processed and non-processed mRNA transcripts, and polypeptides or peptides can include polypeptides or peptides with or without any post-translational modification.

As used herein, "variant" means a substantially similar sequence. A variant can have different function or a substantially similar function as a wild-type polypeptide of interest. For a nicotine demethylase, a substantially similar function is at least 99%, 98%, 97%, 95%, 90%, 85%, 80%, 75%, 60%, 50%, 25% or 15% of wild-type enzyme function of converting nicotine to nornicotine under the same conditions or in a near-isogenic line. A wild-type CYP82E10 is set forth in SEQ ID NO:2. A wild-type CYP82E4 is set forth in SEQ ID NO:14. A wild-type CYP82E5 is set forth in SEQ ID NO:26. Exemplary variants of the wild-type CYP82E10 of the present invention include polypeptides

comprising the sequence set forth in SEQ ID NO:5, 6, 7, 8, 9, 10, 11, 12, or 13. The variant set forth in SEQ ID NO:10 (CYP82E10 P419S) advantageously has a favorable mutation that results in the enzyme having only about 25% of the nicotine demethylase activity of the wild-type CYP82E10 polypeptide. The variants set forth in SEQ ID NOs: 11 (CYP82E10 G79S), 12 (CYP82E10 with P107S), and 13 (CYP82E10 with P382S) advantageously have favorable mutations that result in their nicotine demethylase activity being knocked out (i.e., 100% inhibition, and thus a nonfunctional polypeptide). In like manner, exemplary variants of the wild-type CYP82E4 include polypeptides comprising the sequence set forth in SEQ ID NO:15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25. The variant set forth in SEQ ID NO:21 (CYP82E4 V376M) advantageously has a favorable mutation that results in the enzyme having only about 50% of the nicotine demethylase activity of the wild-type CYP82E4 polypeptide. The variants set forth in SEQ ID NOs: 22 (CYP82E4 W329Stop), 23 (CYP82E4 K364N), 24 (CYP82E4 P382S), and 25 (CYP82E4 P458S) advantageously have favorable mutations that result in their nicotine demethylase activity being knocked out (i.e., 100% inhibition). Similarly, exemplary variants of the wild-type CYP82E4 include polypeptides comprising the sequence set forth in SEQ ID NO: 27, 28, 29, 30, 31, 32, 33, or 34. The variant set forth in SEQ ID NO:34 (CYP82E5 P449L) advantageously has a favorable mutation that results in inhibition of its nicotine demethylase activity, and the variant set forth in SEQ ID NO:33 advantageously has a favorable mutation that results in its nicotine demethylase activity being knocked out (i.e., 100% inhibition).

As used herein, a "variant polynucleotide" or "variant polypeptide" means a nucleic acid or amino acid sequence that is not wild-type.

A variant can have one addition, deletion or substitution; two or less additions, deletions or substitutions; three or less additions, deletions or substitutions; four or less additions, deletions or substitutions; or five or less additions, deletions or substitutions. A mutation includes additions, deletions, and substitutions. Such deletions or additions can be at the C-terminus, N-terminus or both the C- and N-termini. Fusion polypeptides or epitope-tagged polypeptides are also included in the present invention. "Silent" nucleotide mutations do not change the encoded amino acid at a given position. Amino acid substitutions can be conservative. A conservative substitution is a change in the amino acid where the change is to an amino acid within the same family of amino acids as the original amino acid. The family is defined by the side chain of the individual amino acids. A family of amino acids can have basic, acidic, uncharged polar or nonpolar side chains.

See, Alberts *et al.*, (1994) *Molecular biology of the cell* (3rd ed., pages 56-57, Garland Publishing Inc., New York, New York), incorporated herein by reference as if set forth in its entirety. A deletion, substitution or addition can be to the amino acid of another CYP82E family member in that same position. As used herein, a "fragment" means a portion of a polynucleotide or a portion of a polypeptide and hence protein encoded thereby.

As used herein, "plant part" means 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 present invention, provided that they comprise the introduced polynucleotides of the invention. As used herein, "tobacco plant material" means any portion of a plant part or any combination of plant parts.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to a novel nicotine demethylase gene, *CYP82E10* (genomic sequence set forth in SEQ ID NO:4), and its encoded CYP82E10 nicotine demethylase (SEQ ID NO:2), that is involved in root-specific conversion of nicotine to nornicotine in roots of tobacco plants and its use in reducing or minimizing nicotine to nornicotine conversion and thus reducing levels of nornicotine in tobacco plants and plant parts thereof. By "root-specific" is intended it is preferentially expressed within the roots of tobacco plants, as opposed to other plant organs such as leaves or seeds. By introducing selected favorable mutations into this root-specific nicotine demethylase or variants thereof having nicotine demethylase activity, in combination with one or more selected favorable mutations within a gene encoding a green-leaf nicotine demethylase (for example, CYP82E5 set forth in SEQ ID NO:26) or variant thereof having nicotine demethylase activity, and further in combination with one or more selected favorable mutations within a gene encoding a senescence-induced nicotine demethylase (for example, CYP82E4 set forth in SEQ ID NO:14) or variant thereof having nicotine demethylase activity, it is possible to produce nontransgenic tobacco plants having minimal nicotine to nornicotine conversion, where the conversion rate is less than about 1.5%, preferably less than about 1%.

Lowering nornicotine levels in tobacco is highly desirable because this alkaloid serves as a precursor to the well-documented carcinogen N'-nitrosonornicotine (NNN). Two genes encoding proteins having nicotine demethylase activity in tobacco have been previously identified and designated as *CYP82E4v2* and *CYP82E5v2*. The CYP82E4 polypeptide (SEQ ID NO:14) is a senescence-induced nicotine demethylase. The *CYP82E4v2* gene (including the coding and intron regions), its role in nornicotine production in tobacco plants, and methods for inhibiting its expression and function are described in U.S. Patent Application No. 11/580,765, which published as U.S. Patent Application Publication No. 2008/0202541 A1. The CYP82E5 polypeptide (SEQ ID NO:26) is a green-leaf nicotine demethylase (i.e., its predominant expression is in green leaves). The *CYP82E4* gene (including the coding and intron regions), its role in nornicotine production in tobacco plants, and methods for inhibiting its expression and function are described in U.S. Patent Application No. 12/269,531, which published as U.S. Patent Application Publication No. 2009/0205072 A1. The contents of these two U.S. patent applications and their respective publications are herein incorporated by reference in their entirety.

Plants homozygous for favorable mutant *cyp82e4v2* and *cyp82e5v2* alleles (i.e., mutant alleles that knock down, or knock out, expression of these respective nicotine demethylase genes), however, can still metabolize more than 2% of their nicotine to nornicotine, which represent nornicotine levels that can still lead to substantial NNN formation. The discovery of the *CYP82E10* nicotine demethylase gene provides a further avenue for minimizing the nicotine to nornicotine conversion rate in tobacco plants, and thus further reducing the levels of nornicotine and thus NNN in tobacco plants and plant materials derived therefrom. Combining favorable mutant *cyp82e10* alleles with favorable mutant *cyp82e4v2* and *cyp82e5v2* alleles provides for tobacco plants possessing more than a 3-fold reduction in nornicotine when compared to that observed for tobacco plants having the *cyp82e4v2* mutation alone, or the *cyp82e5v2* mutations together. In one embodiment, the present invention provides a homozygous triple mutant combination of nicotine demethylase genes *cyp82e4v2*, *cyp82e5v2*, and *cyp82e10*) that results in nontransgenic tobacco plants that produce very low levels of nornicotine comparable to that only previously achieved via transgenic gene suppression approaches, such as those described in U.S. Patent Application Publication Nos. 2008/0202541 A1 and 2009/0205072 A1.

Nicotine Demethylase Polynucleotides and Polypeptides, and Variants and Fragments Thereof

Compositions of the present invention include the CYP82E10 polypeptide and variants and fragments thereof. Such nicotine demethylase polynucleotides and polypeptides are involved in the metabolic conversion of nicotine to nornicotine in plants, including commercial varieties of tobacco plants. In particular, compositions of the invention include isolated polypeptides comprising the amino acid sequences as shown in SEQ ID NOs:2, and 5-13, isolated polynucleotides comprising the nucleotide sequences as shown in SEQ ID NOs:1, 3, and 4, and isolated polynucleotides encoding the amino acid sequences of SEQ ID NOs:2 and 5-13. The polynucleotides of the present invention can find use in inhibiting expression of nicotine demethylase polypeptides or variants thereof that are involved in the metabolic conversion of nicotine to nornicotine in plants, particularly tobacco plants. Some of the polynucleotides of the invention have mutations which result in inhibiting the nicotine demethylase activity of the wild-type nicotine demethylase. The inhibition of polypeptides of the present invention is effective in lowering nornicotine levels in tobacco lines where genetic conversion occurs in less than 30%, 50%, 70%, 90% of the population, such as flue-cured tobaccos. The inhibition of polypeptides of the present invention is effective in lowering nornicotine levels in tobacco populations where genetic conversion occurs in at least 90%, 80%, 70%, 60%, 50% of a plant population. A population preferably contains greater than about 25, 50, 100, 500, 1,000, 5,000, or 25,000 plants where, more preferably at least about 10%, 25%, 50%, 75%, 95% or 100% of the plants comprise a polypeptide of the present invention.

The nicotine demethylase polynucleotides and encoded polypeptides of the present invention include a novel cytochrome P450 gene, designated the *CYP82E10* nicotine demethylase gene, that is newly identified as having a role in the metabolic conversion of nicotine to nornicotine in roots of tobacco plants. Transgenic approaches such as sense, antisense, and RNAi suppression may be used to knock down expression of this nicotine demethylase, in a manner similar to that described for the CYP82E4 and CYP82E5 nicotine demethylases, as described in U.S. Patent Application Publication Nos. 2008/0202541 A1 and 2009/0205072 A1, the disclosures of which are herein incorporated by reference in their entirety. The preferred approach is one that introduces one or more favorable mutations into this gene, as this approach advantageously provides nontransgenic tobacco plants having reduced nicotine to nornicotine conversion rates, and

thus reduced levels of nornicotine and NNN. Such approaches include, but are not limited to, mutagenesis, and the like, as described elsewhere herein below.

The invention encompasses isolated or substantially purified polynucleotide or protein compositions of the present invention. An "isolated" or "purified" polynucleotide or protein, or biologically active portion thereof, is substantially or essentially free from components that normally accompany or interact with the polynucleotide or protein as found in its naturally occurring environment. Thus, an isolated or purified polynucleotide or protein is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. Optimally, an "isolated" polynucleotide is free of sequences (optimally protein encoding sequences) that naturally flank the polynucleotide (i.e., sequences located at the 5' and 3' ends of the polynucleotide) in the genomic DNA of the organism from which the polynucleotide is derived. For example, in various embodiments, the isolated polynucleotide can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequence that naturally flank the polynucleotide in genomic DNA of the cell from which the polynucleotide is derived. A protein that is substantially free of cellular material includes preparations of protein having less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of contaminating protein. When the protein of the invention or biologically active portion thereof is recombinantly produced, optimally culture medium represents less than about 30% 20%, 10%, 5%, or 1% (by dry weight) of chemical precursors or non-protein-of-interest chemicals.

Fragments of the disclosed polynucleotides and polypeptides encoded thereby are also encompassed by the present invention. Fragments of a polynucleotide may encode protein fragments that retain the biological activity of the native protein and hence are involved in the metabolic conversion of nicotine to nornicotine in a plant. Alternatively, fragments of a polynucleotide that are useful as hybridization probes or PCR primers generally do not encode fragment proteins retaining biological activity. Furthermore, fragments of the disclosed nucleotide sequences include those that can be 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, amplicon-mediated interference, ribozymes, and small interfering RNA or micro RNA, as described in the art and herein below. Thus, fragments of a nucleotide sequence may range from at least about 20 nucleotides, about 50

nucleotides, about 70 nucleotides, about 100 nucleotides about 150 nucleotides, about 200 nucleotides, 250 nucleotides, 300 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 nucleotide sequence can be a fragment between 100 and about 350 nucleotides, between 100 and about 325 nucleotides, between 100 and about 300 nucleotides, between about 125 and about 300 nucleotides, between about 125 and about 275 nucleotides in length, between about 200 to about 320 contiguous nucleotides, between about 200 and about 420 contiguous nucleotides in length between about 250 and about 450 contiguous nucleotides in length. Another embodiment includes a recombinant nucleic acid molecule having between about 300 and about 450 contiguous nucleotides in length.

A fragment of a nicotine demethylase polynucleotide of the present invention that encodes a biologically active portion of a CYP82E10 polypeptide of the present invention will encode at least 15, 25, 30, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, or 500 contiguous amino acids, or up to the total number of amino acids present in a full-length nicotine demethylase polypeptide of the invention (e.g., 517 amino acids for SEQ ID NOs:2 and 5-13). A biologically active portion of a nicotine demethylase polypeptide can be prepared by isolating a portion of one of the CYP82E10 polynucleotides of the present invention, expressing the encoded portion of the CYP82E10 polypeptide (e.g., by recombinant expression *in vitro*), and assessing the activity of the encoded portion of the CYP82E10 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 CYP82E10 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 CYP82E10 polynucleotide as disclosed herein (e.g., 1551 for SEQ ID NO: 1; 2636 for SEQ ID NO:4). Polynucleotides that are fragments of a CYP82E10 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, about 200 to about 1000 contiguous nucleotides, about 225 to about 900 contiguous

nucleotides, about 500 to about 1600 contiguous nucleotides, about 775 to about 1700 contiguous nucleotides, about 1000 to about 1700 contiguous nucleotides, or from about 300 to about 800 contiguous nucleotides from a CYP82E10 polynucleotide as disclosed herein. In one aspect, fragment polynucleotides comprise a polynucleotide sequence containing the polynucleotide sequence from the nucleotide at about position 700 to about position 1250 of a CYP82E10 coding sequence, at about position 700 to about position 1250 of a CYP82E10 genomic sequence, at about position 10 to about position 900 of a CYP82E10 intron sequence, or at about position 100 to about position 800 of a CYP82E10 intron sequence.

Variants of the disclosed polynucleotides and polypeptides encoded thereby are also encompassed by the present invention. Naturally occurring variants include those variants that share substantial sequence identity to the CYP82E10 polynucleotides and polypeptides disclosed herein as defined herein below. In another embodiment, naturally occurring variants also share substantial functional identity to the CYP82E10 polynucleotides disclosed herein. The compositions and methods of the invention can be used to target expression or function of any naturally occurring CYP82E10 that shares substantial sequence identity to the disclosed CYP82E10 polypeptides. Such CYP82E10 polypeptides can possess the relevant nicotine demethylase activity, i.e., involvement in the metabolic conversion of nicotine to nornicotine in plants, or not. Such variants may result from, for example, genetic polymorphism or from human manipulation as occurs with breeding and selection, including mutagenesis approaches. Biologically active variants of a CYP82E10 protein of the invention, for example, variants of the polypeptide set forth in SEQ ID NO:2 and 5-13, 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 wild-type protein as determined by sequence alignment programs and parameters described elsewhere herein, and can be characterized by their functional involvement in the metabolic conversion of nicotine to nornicotine in plants, or lack thereof. A biologically active variant of a protein of the invention may differ from that protein by as few as 1-15 amino acid residues, as few as 10, as few as 9, as few as 8, as few as 7, as few as 6, as few as 5, as few as 4, as few as 3, as few as 2, or as few as 1 amino acid residue. A biologically inactive variant of a protein of the invention may differ from that protein by as few as 1-15 amino acid residues, as few as 10, as few as 9, as few as 8, as few as 7, as few as 6, as few as 5, as few as 4, as few as 3, as few as 2, or as few as 1 amino acid residue.

Variants of a particular polynucleotide of the present invention include those naturally occurring polynucleotides that encode a CYP82E10 polypeptide that is involved in the metabolic conversion of nicotine to nornicotine in the roots of plants. Such polynucleotide variants can comprise a deletion and/or addition of one or more nucleotides at one or more sites within the native polynucleotide disclosed herein and/or a substitution of one or more nucleotides at one or more sites in the native polynucleotide. Because of the degeneracy of the genetic code, conservative variants for polynucleotides include those sequences that encode the amino acid sequence of one of the CYP82E10 polypeptides of the invention. Naturally occurring variants such as these can be identified with the use of well-known molecular biology techniques, as, for example, with polymerase chain reaction (PCR) and hybridization techniques as are known in the art and disclosed herein. Variant polynucleotides also include synthetically derived polynucleotides, such as those generated, for example, by using site-directed mutagenesis but which still share substantial sequence identity to the naturally occurring sequences disclosed herein, and thus can be used in the methods of the invention to inhibit the expression or function of a nicotine demethylase that is involved in the metabolic conversion of nicotine to nornicotine, including the nicotine demethylase polypeptides set forth in SEQ ID NOS:2, 5, 6, 7, 8, 9, and 10. Generally, variants of a particular polynucleotide of the invention, for example, the polynucleotide sequence of SEQ ID NO:3 or the polynucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO:2, and 5-13, 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 that particular polynucleotide as determined by sequence alignment programs and parameters described elsewhere herein.

Variants of a particular polynucleotide of the present invention (also referred to as the reference polynucleotide) can also be evaluated by comparison of the percent sequence identity between the polypeptide encoded by the reference polynucleotide and the polypeptide encoded by a variant polynucleotide. Percent sequence identity between any two polypeptides can be calculated using 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 polypeptides 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.

Furthermore, the polynucleotides of the invention can be used to isolate corresponding root-specific nicotine demethylase sequences, particularly CYP82E10 sequences, from 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 nucleotide 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 nicotine demethylase polypeptide that is involved in the nicotine-to-nornicotine metabolic conversion and which hybridize under stringent conditions to the CYP82E10 sequence 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 nicotine demethylase polypeptides that are involved in the metabolic conversion of nicotine to nornicotine in plants.

Using PCR, oligonucleotide primers can be designed for use in PCR reactions to amplify corresponding DNA sequences from cDNA or genomic DNA extracted from any plant of interest. Methods for designing PCR primers and PCR cloning 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, New York). Innis *et al.*, eds. (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York); Innis and Gelfand, eds. (1995) *PCR Strategies* (Academic Press, New York); and Innis and Gelfand, eds. (1999) *PCR Methods Manual* (Academic Press, New York). Known methods of PCR include, but are not limited to, methods using paired primers, nested primers, single specific primers, degenerate primers, gene-specific primers, vector-specific primers, partially mismatched primers, and the like.

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 background). 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 0.01 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 sulphate) at 37°C, and a wash in 1X to 2X SSC (20X 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 0.5X to 1X 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 0.1X 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 about 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 5X SSC, 0.5% SDS, 5X Denhardt's, 0.45 ug/ul Poly A RNA, 0.45 ug/ul calf thymus DNA and 50% formamide at 42°C, and at least one post-hybridization wash in a solution comprising from about 0.01X SSC to about l X 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 molarity 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 l°C for each l% 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 10°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 10°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, New York); 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, New York).

Hybridization probes may be genomic DNA fragments, cDNA fragments, RNA fragments, or other oligonucleotides, and may be labeled with a detectable group such as ^{32}P, or any other delectable marker. For example, probes for hybridization can be made by

labeling synthetic oligonucleotides based on the CYP82E10 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, New York).

For example, the CYP82E10 polynucleotide sequences disclosed herein, or one or more portions thereof, may be used as probes capable of specifically hybridizing to corresponding root-specific nicotine demethylase polynucleotides and messenger RNAs. To achieve specific hybridization under a variety of conditions, such probes include sequences that are unique among the CYP82E10 polynucleotide sequences or unique to one of the CYP82E10 polynucleotide sequences, including upstream regions 5' to the coding sequence and downstream regions 3' to the coding sequence and an intron region (for example, SEQ ID NO:3), and are optimally at least about 10 contiguous nucleotides in length, more optimally at least about 20 contiguous nucleotides in length, more optimally at least about 50 contiguous nucleotides in length, more optimally at least about 75 contiguous nucleotides in length, and more optimally at least about 100 contiguous nucleotides in length. Such probes may be used to amplify corresponding CYP82E10 polynucleotides. This technique may be used to isolate additional coding sequences or mutations 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, New York).

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, where the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) compared to the deference 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 sequence identity between any two sequences can be accomplished using a mathematical algorithm. Non-limiting examples of such mathematical algorithms are the algorithm of Myers and Miller (1988) *CABIOS* 4:11-17; the local alignment algorithm of Smith *et al.* (1981) *Adv. Appl. Math.* 2:482; the global alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443-453; the search-for-local alignment method of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci.* 85:2444-2448; the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 872264, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877.

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 or polypeptide 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.nlna.nih.gov). Alignment may also be performed manually by inspection.

In some embodiments, the sequence identity/similarity values provided herein are 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 malting 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, California).

The term "percentage of sequence identity" as used herein means the value determined by comparing two optimally aligned sequences over a comparison window, where 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.

Thus, CYP82E10 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, or 4 or a complement or fragment thereof, or a polypeptide sequence of SEQ ID NO: 2, or 5-13. A preferred embodiment includes a polypeptide corresponding to SEQ ID NO:2 that has a serine at position 79, 107, or 382 of the CYP82E10 polypeptide, where the numbering corresponds to SEQ ID NO:2.

Methods for Inhibiting Expression or Function of a Nicotine Demethylase

Methods of reducing the concentration, content, and/or activity of a CYP82E10 polypeptide of the present invention in a *Nicotiana* plant or plant part, particularly the root tissue, are provided. Many methods may be used, alone or in combination, to reduce or eliminate the activity of the CYP82E10 polypeptide of the present invention (SEQ ID NO:2), and variants thereof that retain nicotine demethylases activity (for example, SEQ ID NOs:7, 8, 9, and 10). In addition, combinations of methods may be employed to reduce or eliminate the activity of two or more different nicotine demethylases, more particularly the root-specific CYP82E10 nicotine demethylase and one or both of the green-leaf CYP82E5 and senescence-induced CYP82E4 nicotine demethylases. In a particular embodiment, the CYP82E5 is a polypeptide with at least one amino acid mutation in the sequence of SEQ ID NO: 26 that negatively affects conversion in green leaves and the CYP82E4 has the sequence set forth in SEQ ID NO:14 with at least one amino acid mutation that negatively affects conversion in senescent leaves.

In accordance with the present invention, the expression of a CYP82E10 nicotine demethylase of the present invention is inhibited if the protein level of the CYP82E10 polypeptide is statistically lower than the protein level of the same CYP82E10 polypeptide in a plant that has not been genetically modified or mutagenized to inhibit the expression of that CYP82E10 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 CYP82E10 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%, or less than 5% of the protein level of the same CYP82E10 polypeptide in a plant that is not a mutant or that has not been genetically modified to inhibit the expression of that CYP82E10 polypeptide and which has been cultured and harvested using the same protocols. The expression level of the CYP82E10 polypeptide may be measured directly, for example, by assaying for the level of the CYP82E10 transcript or CYP82E10 polypeptide expressed in the tobacco plant or

plant part, or indirectly, for example, by measuring the conversion of nicotine to nornicotine in the tobacco 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 and RNA differentiation assays. Methods for determining the activity of a targeted CYP82E10 polypeptide in converting nicotine to nornicotine are known in the art and described elsewhere herein below, and include, but are not limited to, alkaloid analysis using gas chromatography.

The present invention provides methods for reducing the level of nornicotine, or reducing the rate of conversion of nicotine to nornicotine, in a tobacco plant, or plant part thereof. The methods comprise introducing into the genome of a tobacco plant a mutation within at least one allele of each of at least three nicotine demethylase genes, wherein the mutation reduces expression of the nicotine demethylase gene, and wherein a first of these nicotine demethylase genes encodes a root-specific nicotine demethylase involved in the metabolic conversion of nicotine to nornicotine in a tobacco plant or a plant part thereof. In some embodiments, the root-specific nicotine demethylase is CYP82E10 or variant thereof. In other embodiments, these methods comprise introducing into the genome of a tobacco plant a mutation within at least one allele of a nicotine demethylase gene encoding CYP82E10 or variant thereof, and a mutation within at least one allele of a nicotine demethylase encoding CYP82E4 or variant thereof, and/or a nicotine demetyylase encoding CYP82E5 or variant thereof.

A number of approaches have been used to combine mutations in one plant including sexual crossing. A plant having a favorable mutation in a *CYP82E10* gene that inhibits the nicotine demethylases activity in roots can be crossed with a plant having a favorable mutation in a *CYP82E4v2* gene that inhibits the nicotine demethylase activity in senescent leaves, or be crossed with a plant having a favorable mutation in a *CYP83E5v2* gene that inhibits nicotine demethylase activity in green leaves to produce a plant having reduced nicotine to nornicotine conversion. In preferred embodiments, crosses are made in order to introduce a favorable mutation within a *CYP82E10, CYP82E4v2, and CYP82E5v2* gene within the same plant. In this manner, a plant having a favorable mutation in a *CYP82E10* gene that inhibits the nicotine demethylases activity in roots is crossed with a plant having a favorable mutation in a *CYP82E4v2* gene that inhibits the nicotine demethylase activity in senescent leaves and a favorable mutation in a *CYP83E5v2* gene that inhibits nicotine demethylase activity in green leaves. Alternatively, a plant having a favorable mutation in a *CYP82E4v2* gene that inhibits the

nicotine demethylase activity in senescent leaves is crossed with a plant having a favorable mutation in a *CYP82E10* gene that inhibits the nicotine demethylase activity in roots and a favorable mutation in a *CYP83E5v2* gene that inhibits nicotine demethylase activity in green leaves. In yet another embodiment, a plant having a favorable mutation in a *CYP82E5v2* gene that inhibits the nicotine demethylase activity in green leaves is crossed with a plant having a favorable mutation in a *CYP82E10* gene that inhibits the nicotine demethylase activity in roots and a favorable mutation in a *CYP83E4v2* gene that inhibits nicotine demethylase activity in senescent leaves. By introducing a favorable mutation into each of these nicotine demethylases genes it is possible to produce a plant having reduced nicotine to nornicotine conversion rates with conversion levels lower than about 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, or 0.7%.

In a more preferred embodiment, a plant having one or more favorable mutations that results in a modification of the CYP82E10 polypeptide at position 79, 107, 382, or 419 (where the numbering is according to SEQ ID NO:2) can be crossed with a plant having one or more favorable mutations that results in a modification of the CYP82E4 polypeptide at position 329, 364, 376, 382, or 458 and/or having one or more favorable mutations that results in a modification of the CYP82E5 polypeptide at position 422 or 449 to produce a plant with conversion levels lower than 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, or 0.7%. A particularly preferred conversion level of nicotine to nornicotine can be between 0.05% - 0.4%, between 0.1 - 0.6%, between 0.1% - 0.3%, between 0.1% - 0.5%, between 0.1% - 0.4%, between 0.1% - 0.7%, between 0.1% - 1.0%, between 0.1% - 1.1%, between 0.1% - 1.2%, between 0.1% - 1.3%, between 0.1% - 1.4%, or between 0.1% - 1.5%. Any mutation of a polynucleotide of the present invention that results in a truncation of the CYP82E10, CYP83E4, or CYP83E5 polypeptide before a conserved heme-binding motif will inhibit the enzyme and can be used in a cross described above. The domains of cytochrome P450 proteins are known in the art. See, for example, Xu *et al.* (2007) *Physiologia Plantarum* 129:307-319, hereby incorporated by reference. By crossing plants having a nonfunctional or inhibited *CYP82E10* gene with plants having a nonfunctional or inhibited *CYP82E4v2* gene, a nonfunctional or inhibited *CYP82E5v2* gene, or nonfunctional or inhibited *CYP82E4v2* and *CYP82E5v2* genes, nornicotine levels can be reduced in a tobacco plant.

The activity of a CYP82E10, CYP82E4, or CYP82E5 nicotine demethylase polypeptide in converting nicotine to nornicotine in a tobacco plant or plant part is inhibited according to the present invention if this conversion activity is statistically lower

than conversion activity of the same nicotine demethylase polypeptide in a tobacco plant or plant part that has not been genetically modified to inhibit the conversion activity of that nicotine demethylase polypeptide and which has been cultured and harvested using the same protocols. In particular embodiments, activity of a nicotine demethylase polypeptide in converting nicotine to nornicotine in a modified tobacco 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 2%, or less than 1% of the conversion activity of the same nicotine demethylase polypeptide in a tobacco plant that has not been genetically modified to inhibit the expression of that nicotine demethylase polypeptide and has been cultured and harvested using the same protocols. The activity of a nicotine demethylase polypeptide in converting nicotine to nornicotine in a tobacco plant or plant part is eliminated according to the invention when it is not detectable by the assay methods described elsewhere herein. Methods of determining the activity of a nicotine demethylase polypetide in converting nicotine to nornicotine in a tobacco plant using gas chromatography are disclosed in the examples here in below.

In some embodiments, the favorable mutation is introduced into a tobacco plant or plant part using a mutagenesis approach, and the introduced mutation 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.

A modified tobacco plant containing a favorable mutation in a nicotine demethylase described herein has a reduced level of conversion of nicotine to nornicotine. In particular embodiments, conversion of nicotine to nornicotine in a modified 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 2%, or less than 1% of the conversion in a tobacco plant that that has not been genetically modified to inhibit the expression of that nicotine demethylase polypeptide and which has been cultured and harvested using the same protocols. In some embodiments, the modified tobacco plant is a converter tobacco

plant. In other embodiments, the modified tobacco plant is a nonconverter tobacco plant. In some embodiments, the modified tobacco plant has a conversion rate lower than the rate observed in commercial non-converter tobacco plants.

According to the present invention, changes in levels, ratios, activity, or distribution of CYP82E10 polypeptides of the present invention, or changes in tobacco 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, for example, by mutagenesis, has been affected as to the nicotine demethylase polypeptide of interest, or is a tobacco plant or plant part that is descended from a tobacco 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 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 tobacco 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 tobacco 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). 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.

In some embodiments, the activity of a nicotine demethylase polypeptide of the present invention may be reduced or eliminated by disrupting the gene encoding the nicotine demethylase polypeptide. The invention encompasses mutagenized plants that

carry mutations in nicotine demethylase genes, where the mutations reduce expression of the nicotine demethylase gene or inhibit the activity of an encoded nicotine demethylase polypeptide of the present invention.

In other embodiments, the activity of a nicotine demethylase polypeptide of the present invention is reduced or eliminated by disrupting the gene encoding the nicotine demethylase polypeptide. The gene encoding the nicotine demethylase 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 nicotine demethylase activity or mutations in CYP82E10, alone or in combination with mutations in CYP82E4 or CYP82E5.

Transposon tagging may be used to reduce or eliminate the activity of one or more CYP82E10 nicotine demethylase polypeptides of the present invention. Transposon tagging comprises inserting a transposon within an endogenous nicotine demethylase gene to reduce or eliminate expression of the nicotine demethylase polypeptide.

Methods for the transposon tagging of specific genes in plants are well known in the art. See, for example, Maes *et al.* (1999) *Trends Plant Sci.* 4:90-96; Dharmapuri and Sonti (1999) *FEMS Micerobiol. Lett.* 179:53-59; Meissner *et al.* (2000) *Plant J.* 22:265-274; Phogat *et al.* (2000) *J. Biosci.* 25:57-63; Walbot (2000) *Curr. Opin. Plant Biol.* 2:103-107; Gai *et al.* (2000) *Nucleic Acids Res.* 28:94-9b; Fitzmaurice *et al.* (1999) *Genetics* 153:1919-1928).

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, using mutagenic or carcinogenic compounds including 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* 213:472-481; Okubara *et al.* (1994) *Genetics* 137:867-874; and Quesada *et al.* (2000) *Genetics* 154:421-4315; 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 nicotine demethylase 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 inhibiting the metabolic function of the encoded protein. Conserved residues of plant nicotine demethylase polypeptides suitable for mutagenesis with the goal to eliminate activity of a nicotine demethylase polypeptide in converting nicotine to nornicotine in a tobacco plant or plant part have been described. See Figure 1A-C of U.S. Patent Application Publication No. 2009/0205072 A1, herein incorporated by reference in its entirety, where the residues that differ from the other P450 polypeptides are shaded in grey. The conserved residue is that which is not shaded in grey at each position. Such mutants can be isolated according to well-known procedures.

In another embodiment of this invention, dominant mutants can be used to trigger RNA silencing due to gene inversion and recombination of a duplicated gene locus. See, for example, Kusaba *et al.* (2003) *Plant Cell* 15:1455-1467.

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 the *CYP82E10* gene 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 the CYP82E10 polypeptides described herein, compositions of the invention include an intron sequence in the *CYP82E10* gene sequence that can be used in screening methods. While not bound by any mechanism of action, the CYP82E10 gene(s) may represent the only member(s) of the cytochrome P450 family involved in the metabolic conversion of nicotine to nornicotine in roots of tobacco. For certain applications it would be useful to have a means of diagnostically differentiating this specific member of the cytochrome P450 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 may be valuable as a permanently nonconverter resource. Such dysfunctional sequences may include those encoding the polypeptides set forth in SEQ ID NO: 11, 12, or 13. A method to specifically assay for such genotypes (e.g. deletion mutants,

rearrangements, and the like) could serve as a powerful tool. The present invention includes primers designed to specifically amplify exon 1 and exon 2 of *CYP82E10* where one of the two primer pairs corresponds to the intron between the exons. Examples of primers useful to amplify the exons of *CYP82E10* include SEQ ID NO: 35 with SEQ ID NO: 36 and SEQ ID NO: 37 with SEQ ID NO: 38. These same primers can be used for sequence analysis of the products.

Because the intron regions of genes are typically less conserved than exons, it is predicted that the use of an intron-specific probe would better enable one to distinguish the gene(s) corresponding to the *CYP82E10* gene from the other members of the *CYP82E* family. The use of a *CYP82E10* intron-specific probe, and/or the PCR primers used to generate products 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.

Tobacco Plants, Plant Parts, and Products Having Reduced Nornicotine and NNN Content

The CYP82E10 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 CYP82E10 nicotine demethylases that are involved in the metabolic conversion of nicotine to nornicotine in a plant. In this manner, favorable mutations can be introduced into the *CYP82E10* gene of interest. The methods of the invention do not depend on a particular method for introducing the favorable mutation into the *CYP82E10* nicotine demethylase gene.

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, attenuate, bigelovii, clevelandii, excelsior, forgetiana, glauca, glutinosa, langsdorffii, longiflora, obtusifolia, palmeri, paniculata, plumbaginifolia, qudrivalvis, repanda, rustica, suaveolens, sylvestris, tabacum, tomentosa, trigonophylla, and x sanderae*. The present invention can also be practiced using 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 plameri, 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. In some embodiments, the tobacco plant of interest is a Burley, Virginia, flue-cured, air-cured, fire-cured, Oriental, or a dark tobacco plant.

The tobacco plants and varieties 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 tobacco plant, or plant part thereof, comprising a mutation in a gene encoding a CYP82E10 nicotine demethylase, wherein said mutation results in reduced expression or function of said CYP82E10 nicotine demethylases, and a reduced amount of nornicotine and N'-nitrosonornicotine. As used herein, the term "a reduced amount" or "a reduced level" is intended to refer to an amount of nornicotine and/or N'-nitrosonornicotine in a plant of the present invention 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 genetically modified for reduced nornicotine and/or N'-nitrosonornicotine. 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 conversion of nicotine to nornicotine can be less than 0.3%, less than 0.5%, less than 0.7%, between 0.1% - 0.5%, between 0.1% - 0.4%, between 0.1% - 0.7%, or between 0.1% - 1.0% in plants, plant parts, and products of the present invention, and more specifically in plants, plant parts having mutations in *CYP82E10, CYP82E4v2,* and *CYP825v2*.

The term "tobacco products" as used herein include, but are not limited to, smoking materials (e.g., any cigarette, including a cigarillo, a non-ventilated or vented recess filter cigarette, a cigar, pipe tobacco), smokeless products (e.g., snuff, chewing tobacco, biodegradable inserts (e.g., gum, lozenges, dissolving strips)). See, for example, U.S. Patent 2005/0019448, herein incorporated by reference. The present invention also encompasses a range of tobacco product blends that can be made by combining conventional tobacco with differing amounts of the low nornicotine and/or

N'-nitrosonornicotine 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'-nitrosonornicotine 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 derived from the 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 citations mentioned in the following discussion are provided at the close of the Experimental section.

Background

The knowledge that *CYP82E4v2* represents the nicotine demethylase locus responsible for the high nornicotine accumulation observed in Converter plants (Siminszky et al., 2005), opened the door for nontransgenic, as well as transgenic, approaches toward overcoming the conversion problem and lowering the nornicotine content of the senescent, cured leaf. Specifically, it became possible for researchers to generate tobacco populations that had been exposed to a chemical mutagen, and select for individuals possessing nonfunctional alleles at the *CYP82E4v2* locus. In fact, three independent groups have already generated nonconverting tobacco lines based on this strategy (Dewey et al., 2007; Xu et al., 2007b; Julio et al., 2008).

As previously reported, a tobacco plant designated 775 was identified from an EMS-mutagenized population of Burley line DH98-325-6 and shown to possess a knockout mutation within the *CYP82E4v2* gene (Dewey et al., 2007). In the summer of 2008, plants homozygous for the 775 mutation were grown at the Upper Coastal Plains research station in Rocky Mount, NC, and air-cured according to standard industry practice. Alkaloid analysis of these materials was conducted using the "LC Protocol" described by Jack et al. (2007). As shown in Table 1, plants possessing the 775 mutation averaged 2.6% nicotine to nornicotine conversion, In contrast, >60% conversion was observed in the parental line DH98-325-6, a strong converter genotype. Nearly identical results were reported by Julio et al. (2008), who recorded conversion percentages ranging from 2.82 to 3.37 for plants homozygous for a *cyp82e4v2* knockout mutant within the strong converter burley genotype BB16NN (parental conversion rates ranged between 68 - 98%). Thus, debilitating mutations in *CYP82E4v2* alone appear to be effective in eliminating the problems arising from the unstable genetic phenomenon associated with the generation of Converter plants.

Table 1. Alkaloid profiles for experimental materials evaluated in 2008 field experiment. Percentage values represent an average.

Genotype	Gene Targeted	Mutation[b]	Amino Acid Change	% Nicotine[c]	% Nornicotine	% Anabasine	% Anatabine	% Conversion[d]
DH98-325-6 control (15)[a]	Control	-	-	1.228	2.014	0.016	0.125	62.4
TN90LC (14)	Control	-	-	4.680	0.157	0.022	0.155	3.2
DH98-325-6 RNAi 300-08 #1 (15)	*CYP82E4v2* and related	-	-	3.351	0.040	0.016	0.101	1.2
DH98-325-6 RNAi 300-02 #1 (15)	*CYP82E4v2* and related	-	-	3.741	0.026	0.017	0.106	0.7
DH98-325-6 #775 Homo. (15)	*CYP82E4v2*	G986A	W329Stop	2.941	0.077	0.013	0.093	2.6
DH98-325-6 #1013 Homo. (14)	*CYP82E5v2*	G1266A	W422Stop	1.005	1.876	0.012	0.097	65.2
DH98-325-6 Double Homozygous Mutant (9)		Double	Double	3.160	0.076	0.015	0.117	2.3

[a]Number in parentheses indicates total number of plants analyzed.

[b]Numbering relative to start codon of cDNA sequence.

[c]Percentages were calculated on a dry tobacco weight basis.

[d]Percentage nicotine conversion equals [% nornicotine/ (% nornicotine + % nicotine)] X 100.

Although the utilization of tobacco plants possessing the 775, or comparable, mutations in *CYP82E4v2* can be an effective means of eliminating the introduction of Converter plants within tobacco populations, a low, but significant amount of nornicotine remains in these plants. Given that nicotine to nornicotine conversion rates as low as 0.45% were observed in transgenic plants expressing an RNAi-based construct directed against *CYP82E4v2* (Lewis et al., 2008), it was apparent that at least one other gene with high DNA sequence homology to *CYP82E4v2* must be responsible for the majority of the nornicotine synthesis that is observed within both Nonconverter plants and Converter plants possessing an inactivated *CYP82E4v2* gene. This possibility was further supported by the discovery of *CYP82E5v2*, a gene that shares 92.7% DNA sequence identity with *CYP82E4v2* that was also shown to encode a functional nicotine demethylase enzyme (Dewey et al., 2007; Gavilano and Siminszky, 2007). The *CYP82E5v2* nicotine demethylase gene is expressed at low levels in green tobacco leaves of Converter and Nonconverter plants alike, in contrast to *CYP82E4v2* which is expressed at very high levels, but only in the leaves of Converter plants during senescence and air-curing.

As outlined in Dewey et al. (2007), screening of an EMS-mutagenized DH98-325-6 tobacco population lead to the identification of an individual (plant 1013) possessing a knockout mutation in *CYP82E5v2*. To determine the impact of the non-functional *cyp82e5v2* allele on nornicotine accumulation, crosses were made that combined the mutations from plants 775 and 1013. Molecular genotyping of numerous F_2 individuals derived from the F_1 progeny of the initial cross resulted in the identification of nine individuals that were homozygous for both mutations (*e4e4/e5e5*). These nine plants were also included in the 2008 field trial. Despite the fact the *CYP82E5v2* has been shown to encode a functional nicotine demethylase enzyme (Dewey et al., 2007; Gavilano and Siminszky, 2007), combining the dysfunctional *cyp82e5v2* mutation with the knockout *cyp82e4v2* mutation had remarkably little impact on leaf nornicotine levels. As shown in Table 1, plants homozygous for the double mutation (*e4e4/e5e5*) averaged 2.3% nicotine conversion, compared with an average of 2.6% conversion for plants possessing only the *cyp82e4v2* mutation (*e4e4*). The modest difference in mean conversion between the two genotypes was not statistically significant ($P = 0.118$). In contrast, one of the *CYP82E4v2* RNAi-silenced transgenic lines that was included in this study averaged 0.7% conversion, an amount significantly lower ($P < 0.001$) than that obtained from either the *e4e4* or *e4e4/e5e5* genotypes. Thus, another gene with high homology to *CYP82E4v2* must exist within the tobacco genome that contributes toward nornicotine production in the plant.

Example 1: Isolation and characterization of the cyp82e10 nicotine demethylase gene.

To identify other genes in the tobacco genome that have the potential of encoding nicotine demethylase enzymes, homology searches using the BLASTN and BLASTX algorithms (Altschul et al., 1990, 1997) were directed against the *N. tabacum* expressed sequenced tagged (EST) databases in GenBank, using the DNA and protein sequences of *CYP82E4v2* as the respective query sequences. In addition to identifying cDNA sequences corresponding to previously characterized members of the *CYP82E* superfamily (such as *CYP82E2, CYP82E3* and *CYP82E5v2*), seven ESTs were discovered that did not align perfectly with any previously characterized member of this gene family. Interestingly, all seven of the ESTs originated from either root-specific cDNA libraries, or cDNA libraries made up of mixed tissues that included roots. This observation suggested that the new *CYP82E* gene is expressed specifically in root tissue, a property that could explain why this particular member of the *CYP82E* P450 superfamily has eluded detection previously, as prior efforts have focused on the characterization of *CYP82E* genes expressed in leaf tissue. Because no individual EST sequence was long enough to cover the entire coding region of this novel gene, PCR primers were designed that enabled amplification of the entire cDNA sequence from first-strand cDNA that had been generated from RNA isolated from tobacco root tissue. In addition, primers were used to amplify the corresponding genomic region of the gene that includes a central, large intron. This novel *CYP82E* cDNA shares 92.4% nucleotide identity with the tobacco *CYP82E4v2* cDNA, and a 91.1% predicted identity at the amino acid level. In keeping with the guidelines for P450 gene nomenclature, this new gene was designated *CYP82E10.* Of all the characterized members of the *CYP82E* superfamily, *CYP82E10* displays that highest sequence similarity with *CYP82E5v2*, sharing 96.5% nucleotide identity at the cDNA level and 95.7% predicted amino acid sequence identity. The DNA sequence of *CYP82E10* and its predicted protein sequence are shown in Figure 1.

Although the cDNAs of the various *CYP82E* family members tend to be highly conserved, the genomic versions of these genes show much greater sequence diversity. This is due primarily to the substantial sequence divergence observed within the large, central intron. An alignment of *CYP82E4v2*, *CYP82E5v2*, and *CYP82E10* genomic sequences is shown in Figure 2. As calculated using the EMBOSS Pairwise Alignment algorithm (www.ebi.ac.uk/Tools/emboss/align/index.html), the *CYP82E4v2* and

CYP82E10 genes share 78.3% nucleotide identity, and *CYP82E10* is 84.9% identical to the *CYP82E5v2* gene as they exist within the tobacco genome (*CYP82E4v2* and *CYP82E5v2* genomic sequences share 75% identity).

As detailed in several publications, most of the genes of the *CYP82E* superfamily that are found in the tobacco genome do not encode functional nicotine demethylase enzymes (Siminszky et al., 2005; Chakrabarti et al., 2007; Dewey et al., 2007; Gavilano et al., 2007; Xu et al., 2007a). Therefore, sequence homology alone is not a very accurate indicator of gene function for the *CYP82E* family. Instead, expression analysis in either transgenic plants (Siminszky et al., 2005) or in yeast (Gavilano and Siminszky, 2007; Xu et al., 2007a) has become the established means for determining whether individual members of this gene family encode nicotine demethylase activity.

To determine whether *CYP82E10* functions as a nicotine demethylase gene, its cDNA was cloned into the yeast expression vector pYeDP60 and transformed into yeast strain W(R). Strain W(R) is a yeast cell line that was engineered to overexpress the yeast NADPH-dependent P450 reductase, an enzyme that serves as the direct electron donor to P450s; this system greatly enhances the detection of foreign P450 enzyme activities that are expressed in yeast (Pompon et al., 1995). Nicotine demethylase assays were conducted by incubating yeast microsomal membrane preparations with [^{14}C]-nicotine, and resolving the products by thin layer chromatography as described in Siminszky et al. (2005).

As shown in Figure 3, no nicotine demethylase activity could be detected using yeast microsomes from the W(R) strain expressing only the pYeDP60 vector. In contrast, a very robust nicotine demethylase activity could be measures from microsomes derived from yeast cells expressing the *CYP82E10* cDNA. By measuring CYP82E10 enzyme activity across a wide range of [^{14}C]-nicotine concentrations, a substrate saturation curve was established and an apparent K_m of 3.9 μM nicotine was calculated using the microsomal assay. This kinetic parameter for CYP82E10 is very similar to the K_ms reported for the CYP82E4v2 and CYP82E5v2 enzymes when similarly expressed in yeast (Gavilano et al., 2007; Gavilano and Siminszky, 2007; Xu et al., 2007a).

Example 2: Identification of plants possessing mutant alleles of *CYP82E10.*

In order to accurately assess the specific contribution of *CYP82E10* toward the total nornicotine content of the tobacco plant, it was necessary to: (1) identify a tobacco plant with a knockout mutation within this gene; and (2) combine this mutation with the

cyp82e4v2 and *cyp82e5v2* mutations originating from plants 775 and 1013, respectively. To identify potentially debilitating mutations in *CYP82E10*, the EMS-mutagenized DH98-325-6 population was screened by high-throughput DNA sequence analysis using primers that specifically amplify portions of *CYP82E10* (without simultaneously amplifying other members of the *CYP82E* superfamily). To specifically amplify exon 1 of *CYP82E10*, the following PCR primers were used: 5'-GTGATAGTTTGATTCCCAAGTGC-3' (forward) and 5'-CTCCCAAAGTTAGATTAGTCCG-3' (reverse); specific amplification of exon 2 was achieved using the primers 5'-AGGTCGCGCTGATTCTTG-3' (forward) and 5'-AGATGAATACCCATCTATCTAGGAGT-3' (reverse). To ensure maximal specificity, the reverse primer for exon 1 and the forward primer for exon 2 correspond to sequences within the *CYP82E10* intron (Fig. 1). PCR amplification and sequence analysis of the mutagenized plants was conducted using a 96-well format as described in Dewey et al. (2007).

High-throughput sequence analysis of over 1,200 individuals from the mutagenized tobacco population resulted in the identification of 15 individuals with mutations in *CYP82E10*. The most notable of these are shown in Table 2. The nucleotide and amino acid residues mutated in these plants are also highlighted in Figure 1. Although no truncation mutations were observed among these individuals, in several cases, mutations were identified that altered an amino acid residue within a highly conserved region of the enzyme. To determine the effects of a particular mutation on CYP82E10 enzyme activity, site-directed mutagenesis was used to introduce the specific mutations corresponding to seven of the nine mutations shown in Table 2 into the *CYP82E10* cDNA within the pYeDP60 yeast expression vector. Microsomal preparations from yeast strains expressing each of the seven CYP82E10 variants were assayed *in vitro* for nicotine demethylase activity using both non-saturating (2.45 μM) and saturating (50 μM) concentrations of [^{14}C]-nicotine. The results from the yeast expression assays showed that mutations found plants 693, 817 and 1035 did not alter enzyme activity, whereas the mutations found in plants 1041, 1512 and 2476 lead to complete enzyme inactivation. The mutation observed in plant 1442 resulted in a 75% reduction in activity compared to the wild type CYP82E10 enzyme.

The thin layer chromatographic data for the *in vitro* yeast expression assays corresponding to the plant 1041 mutation are shown in Figure 3. This particular mutation was selected for more extensive investigation. To provide additional confirmation that the Pro to Ser substitution at amino acid position 382 that defines the plant 1041 mutation is

incompatible with nicotine demethylation function, this same mutation was introduced into a *CYP82E4v2* cDNA that had been similarly cloned into the pYeDP60 vector. The results of these yeast assays are displayed in Table 3. Whether introduced into the CYP82E10 or CYP82E4v2 enzymes, a Ser substitution for Pro at position 382 leads to the complete ablation of nicotine demethylase activity in this assay. Interestingly, although the activities of the wild type CYP82E10 and CYP82E4v2 enzymes were comparable at the non-saturating [^{14}C]-nicotine concentration (2.45 μM), at the 25 μM substrate level, the rate of [^{14}C]-nornicotine synthesis was nearly three times greater in microsomal preparations possessing the CYP82E10 enzyme than preparations containing CYP82E4v2.

Table 2. EMS treated lines of DH98-325-6 with mutations in the *CYP82E10* gene.

Plant Number	Mutation[a]	Amino Acid Change	Activity of Mutant Enzyme[b]
2476	G235A	G79S	Not detected
1512	C319T	P107S	Not detected
319	C442T	L148F	Not tested
634	G514A	G172R	Not tested
1035	G1030A	A344T	100%
1041	C1141T	P382S	Not detected
817	G1228A	A410T	100%
693	G1250A	R417H	100%
1442	C1255T	P419S	25%

[a]In reference to the start codon of the CYP82E10 cDNA sequence.

[b]Relative to the wild type enzyme when expressed in yeast.

Table 3. Nicotine demthylase activity of CYP82E4v2 and CYP82E10 enzymes possessing the 1041 mutation (Pro382Ser).

Vector	CPM nornicotine at 2.45 μM [^{14}C]-nicotine substrate[a]	CPM nornicotine at 50.0 μM [^{14}C]-nicotine substrate
pYeDP60-CYPE4v2	1,813 ± 623[b]	5,383 ± 505
pYeDP60-CYPE4v2/1041	Not detected	Not detected
pYeDP60-CYPE10	2,296 ± 99	15,253 ± 465
pYeDP60-CYPE10/1041	Not detected	Not detected

[a]Counts per minute of [^{14}C]-nornicotine/mg microsomal protein.

[b]Standard deviation of two technical replications.

Nicotine demethylase activities of wild type and 1041 mutant *CYP82E10*-expressing yeast cells was also assayed *in vivo*. Yeast cultures were shaken overnight in the presence of 55 μM [^{14}C]-nicotine, extracted with methanol and analyzed by thin layer chromatography. [^{14}C]-nornicotine could be detected in the extracts of yeast expressing wild type *CYP82E10*, but not the 1041 mutant version of the gene (data not shown). Cumulatively, the yeast expression assays strongly suggest that CYP82E10 enzyme function is completely abolished by the introduction of the 1041 mutation.

Example 3: Combining mutant alleles of *cyp82e10, cyp82e4v2* and *cyp82e5v2.*

Given that the original 1041 mutation is in a genetic background (DH98-325-6) that contains both a strong converter *CYP82E4v2* allele as well as a wild type *CYP82E5v2* gene, the only way to accurately assess the specific contribution of *CYP82E10* toward total plant nornicotine content is to introduce the 1041 mutation into tobacco plants possessing knockout *CYP82E4v2* and *CYP82E5v2* mutations as well. To accomplish this, plants heterozygous for the 1041 mutation (*e10E10*) were crossed with plants heterozygous for both the 775 and 1013 mutations described above (*e4E4/e5E5*). The latter plants represent progeny of from the cross 775/1013//TN90/3/TN90/4/TN90. F_1 plants heterozygous for all three nicotine demethylase mutations (*e4E4/e5E5/e10E10*) were identified by molecular genotyping, and allowed to self-pollinate. Molecular genotyping was also used to screen over 400 F_2 progeny and subsequently group them into the following genotypic classes: *E4E4/E5E5/e10e10* (3 plants total); *e4e4/E5E5/e10e10* (4 plants total); *E4E4/e5e5/e10e10* (5 plants total); and *e4e4/e5e5/e10e10* (5 plants total).

All of the plants described above were transplanted and grown in the field at the Upper Coastal Plains research station in Rocky Mount, NC in the summer of 2009. Also included in this study were two of the genotypes tested in the 2008 field trial shown in Table 1. Specifically, ten DH98-325-6 plants homozygous for only the *cyp82e4v2* mutation (*e4e4/E5E5/E10E10*) and eleven DH98-325-6 plants possessing the double homozygous *e4e4/e5e5/E10E10* genotype were included for comparison. As controls, individual plants randomly selected from a commercial "low converter" seedlot (TN90LC), wild type DH98-325-6 individuals, and plants from one of the best *CYP82E4v2* RNAi-suppressed transgenic lines was also included in the study. After the plants were about an average of 30 cm tall (35 days after transplanting) leaves from similar stalk positions where collected, treated with ethephon and air-cured according to the protocol established by Jack et al. (2007). Alkaloid content of the cured leaf materials was determined by gas chromatography as described in the same protocol.

Table 4 and Figure 4 shows the results of the alkaloid analyses for the 2009 field trial. Consistent with previous observations, the *cyp82e4v2* knockout mutation alone negates the strong converter phenotype of line DH98-325-6, and also confers a substantially lower nornicotine accumulation phenotype than plants from the commercial TN90LC seed (2.2% conversion versus 7.1%, respectively). As observed in the 2008 field trial (Table 1), combining the *cyp82e5v2* mutation with *cyp82e4v2* did not lead to further reductions in nornicotine content. In fact, the mean nicotine conversion for the *e4e4/E5E5/E10E10* plants was actually lower than that observed for *e4e4/e5e5/E10E10* individuals (2.2% versus 2.3%), though this slight difference was not statistically significant. As expected, the *cyp82e10* mutation had no impact on the high nornicotine levels conferred by an active *CYP82E4v2* gene, either alone (*E4E4/E5E5/e10e10* genotypes), or when combined with a mutant *cyp82e5v2* allele (*E4E4/e5e5/e10e10* genotypes) (Fig. 4A). Similar to the *cyp82e4v2* and *cyp82e5v2* double mutant results (Tables 1 and 4), introducing *cyp82e10* into a *cyp82e4v2* background was not effective in reducing nornicotine levels below than that which could be achieved by the *cyp82e4v2* mutation alone (Fig. 4B). The *e4e4/E5E5/e10e10* genotypes averaged 1.85 % conversion which was not significantly different than the 2.2% mean conversion levels observed for *e4e4/E5E5/E10E10* individuals ($P = 0.235$).

Table 4. Alkaloid profiles for experimental materials evaluated in 2009 field experiment. Measurements taken from leaves harvested 35 days after transplanting. Percentage values represent an average.

Genotype	Gene Targeted	Mutation[b]	Amino Acid Change	% Nicotine[c]	% Nornicotine	% Anabasine	% Anatabine	% Conversion[d]
DH98-325-6 control (8)[a]	Control	-	-	0.133	1.553	0.009	0.085	92.21
TN90LC (11)	Control	-	-	1.519	0.104	0.002	0.065	7.15
DH98-325-6 RNAi 300-02 #1 (10)	*CYP82E4v2* and related	-	-	1.747	0.009	0.003	0.063	0.54
DH98-325-6 #775 Homo. (10)	*CYP82E4v2*	G986A	W329Stop	1.375	0.030	0.002	0.057	2.20
DH98-325-6 Double Homo. Mutant (11)	*CYP82E4v2* *CYP82E5v2*	Double	Double	1.524	0.036	0.003	0.084	2.34
DH980325-6 #1041 Homo. (3)	*CYP82E10*	C1141T	P382S	0.082	1.302	0.007	0.073	93.87
DH98-325-6 Double Homo. Mutant (5)	*CYP82E5v2* *CYP82E10*	Double	Double	0.081	1.345	0.010	0.068	94.31
DH98-325-6 Double Homo. Mutant (4)	*CYP82E4v2* *CYP82E10*	Double	Double	2.168	0.045	0.004	0.087	1.85
DH98-325-6 Triple Homo. Mutant (5)	*CYP82E4v2* *CYP82E5v2* *CYP82E10*	Triple	Triple	1.793	0.012	0.003	0.056	0.55

[a]Number in parentheses indicates total number of plants analyzed. [b]Numbering relative to start codon of cDNA sequence.
[c]Percentages were calculated on a dry tobacco weight basis. [d]Percentage nicotine conversion equals [% nornicotine/ (% nornicotine + % nicotine)] X 100.

Although the *cyp82e5v2* and *cyp82e10* mutations did not serve to significantly decrease the nornicotine content of *cyp82e4v2* plants when combined individually, pyramiding all three nicotine demethylase mutations had a very notable effect. Nicotine to nornicotine conversion in triple mutant plants (*e4e4/e5e5/e10e10*) averaged only 0.55 %, a percentage virtually identical to the 0.54% observed in the RNAi-suppressed transgenic line ($P = 0.893$; Fig. 4B). This represents over a 3-fold reduction in nicotine conversion beyond that which was mediated by the *cyp82e4v2* mutation alone. Statistically, the differences in percent nicotine conversion (and nornicotine accumulation as a percentage of total dry weight) between *e4e4/E5E5/E10E10* and *e4e4/e5e5/e10e10* genotypes was highly significant ($P < 0.0001$). Similar to the investigation of RNAi-mediated suppression of nicotine conversion (Lewis et al., 2008), the present nontransgenic alteration of nicotine demethylase activities in the tobacco plant did not appear to significantly alter the content of the minor alkaloid species anatabine and anabasine.

The effects of pyramiding the three independent nicotine demethylase gene mutations were also tested in a field trial conducted during the 2010 growing season. For this study, the crosses were conducted entirely within the DH98-325-6 genetic background (in contrast to the 2009 study where a TN90 parent was also used). Molecular genotyping was again used to create every possible combination needed to determine the respective contributions of each *CYP82E* locus on the nornicotine phenotype. Alkaloid data were collected on tobacco plants that were grown to maturity and cured according to standard industry practice. As shown in Table 5, a high level of nicotine conversion (ranging from 52.4 – 65.59%) was observed in all genotypes homozygous for a wild type *CYP82E4v2* gene (genotypes *E4E4/E5E5/E10E10, E4E4/e5e5/E10E10, E4E4/E5E5/e10e10,* and *E4E4/e5e5/e10e10*). Plants homozygous for just the *cyp82e4v2* mutation (*e4e4/E5E5/E10E10*) averaged 2.91% nicotine to nornicotine conversion. Similar to the 2009 results, the effects of the *cyp82E5v2* and *cyp82E10* mutations were not additive, and were only manifest when all three mutant loci were pyramided together. DH98-325-6 (*e4e4/E5E5/e10e10*) plants averaged 2.89% conversion and DH98-325-6 (*e4e4/e5e5/E10E10*) individuals averaged 2.52%, values that were not statistically different than that observed with the *cyp82e4v2* mutation alone. In contrast, the reduction in nornicotine

observed in the triple mutant DH98-325-6 (*e4e4/e5e5/e10e10*) genotype (1.11% nicotine conversion) was 2.6-fold lower than that attained via the *cyp82e4v2* mutation alone. The reduction in nicotine conversion attributable to the triple mutant combination was highly significant ($P<0.001$) compared with either *cyp82e4v2* alone or any double mutant combination.

Table 5. Alkaloid profiles for DH98-325-6 genotypes possessing different mutation combinations at the *CYP82E4v2* (*E4*), *CYP82Ev25* (*E5*), and *CYP82E10* (*E10*) loci. Data are averaged over five replications and generated from analysis of composite ground samples of the fourth and fifth leaves from the top of the plant.

Genotype	Nicotine (%)	Nornicotine (%)	Anabasine (%)	Anatabine (%)	Nicotine Conversion (%)
DH98-325-6 *E4E4 E5E5 E10E10*	1.76	2.46	0.02	0.17	58.66
DH98-325-6 *e4e4 E5E5 E10E10*	2.61	0.08	0.01	0.09	2.91
DH98-325-6 *E4E4 e5e5 E10E10*	1.08	2.06	0.02	0.14	65.59
DH98-325-6 *E4E4 E5E5 e10e10*	1.40	1.96	0.01	0.13	59.30
DH98-325-6 *e4e4 e5e5 E10E10*	3.25	0.09	0.02	0.16	2.89
DH98-325-6 *e4e4 E5E5 e10e10*	3.59	0.09	0.01	0.12	2.52
DH98-325-6 *E4E4 e5e5 e10e10*	1.59	1.72	0.01	0.09	52.40
DH98-325-6 e*4e4 e5e5 e10e10*	4.18	0.05	0.02	0.13	1.11

Alkaloid percentages were calculated on a dry weight basis

Percentage nicotine conversion equals [% nornicotine/(%nornicotine + % nicotine)] X 100

Conclusions

Through the present discovery and characterization of a new nicotine demethylase gene, *CYP82E10*, it has been possible to develop a strategy for reducing the nicotine conversion rates (and thus nornicotine levels) in commercial grade air-cured tobacco plants to levels that have previously only been possible using transgenic approaches. This non-GMO based technology can reduce the levels of nornicotine to a degree similar to that which has been achieved using transgenic strategies, yet offers the tremendous advantage of serving as a means for developing ultra-low nornicotine tobacco varieties while bypassing the substantial hurdles associated with the commercialization of transgenic crops, such as: (1) negotiating and paying licensing fees for the several enabling technologies required for generating transgenic plants; (2) avoiding the lengthy time and onerous costs associated with the deregulation of a transgenic event; and (3) encountering the possibility of product rejection by end users philosophically opposed to GMOs. The discovery reported here represents a major advancement in our ability to lower the levels of one of the most well documented strong carcinogens found in tobacco products, in comparison with the previously described non-GMO strategies that only targeted mutations in the *CYP82E4v2* nicotine demethylase gene (Julio et al., 2008; Xu et al., 2007b) or combined *CYP82E4v2* and *CYP82E5v2* mutations (Dewey et al., 2007). Using transgenic technologies, it was previously demonstrated that lowering nicotine conversion levels from ~2.6% to ~0.5% in the cured leaf lead to a commensurate reduction in the NNN content of the leaf as well (Lewis et al., 2008). One would expect to see similar reductions in the NNN content from tobacco leaves containing the triple mutant combination (*e4e4/e5e5/e10e10*) described in this report. Although originally targeted for air-cured tobaccos, this technology will be of benefit to flue-cured varieties as well. As heat exchangers age, their ability to remove NO_x gases during flue-curing can decrease. Furthermore, recent studies have shown that a considerable amount of TSNA formation can occur during the storage of the cured leaf. Minimizing nornicotine levels through the introduction of the triple mutant combination in flue-cured varieties can act as a safeguard against NNN formation either during storage or as a consequence of inefficient heat exchange during the curing process.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402.

Brogan, A.P., Dickerson, T.J., Boldt, G.E. and Janda, K.D. 2005. Altered retinoid homeostasis catalyzed by nicotine metabolite: implications in macular degeneration and normal development. *Proc. Natl. Acad. Sci. USA* 102: 10433-10438.

Boyette, M.D. and Hamm, L.A. 2001. Results of year 2000 TSNA sampling program in flue-cured tobacco. *Rec. Adv. Tob. Sci.* 27: 17-22.

Bush, L.P., Cui, M., Shi, H., Burton, H.R., Fannin, F.F., Lei, L. and Dye, N. 2001. Formation of tobacco-specific nitrosamines in air-cured tobacco. *Rec. Adv. Tob. Sci.* 27: 23-46.

Chakrabarti, M., Meekins, K.M., Gavilano, L.B. and Siminszky, B. 2007. Inactivation of the cytochrome P450 gene *CYP82E2* by degenerative mutations was a key event in the evolution of the alkaloid profile of modern tobacco. *New Phytol.* 175: 565-574.

Dewey, R.E., Siminszky, B., Bowen, S.W. and Gavilano, L. 2007. Alteration of tobacco alkaloid content through modification of specific cytochrome P450 genes. U.S. Patent Application 60/987,243.

Dickerson, T.J. and Janda, K.D. 2002. A previously undescribed chemical link between smoking and metabolic disease. *Proc. Natl. Acad. Sci. USA* 99: 15084-15088.

Gavilano, L.B., Coleman, N.P., Bowen, S.W. and Siminszky, B. 2007. Functional analysis of nicotine demethylase genes reveals insights into the evolution of modern tobacco. *J. Biol. Chem.* 282: 249-256.

Gavilano, L.B. and Siminszky, B. 2007. Isolation and characterization of the cytochrome P450 gene *CYP82E5v2* that mediates nicotine to nornicotine conversion in the green leaves of tobacco. *Plant Cell Physiol.* 48: 1567-1574.

Hecht, S.S. 1998. Biochemistry, biology and carcinogenicity of tobacco-specific N-nitrosamines. *Chem. Res. Toxicol.* 11: 559-603.

Hecht, S.S. 2003. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nature Rev.* 3: 733-744.

Hecht, S.S. and Hoffmann, D. 1990. The relevance of tobacco specific nitrosamines to human cancer. *Cancer Surveys* 8: 273-294.

Hoffmann, D., Brunnemann, K.D., Prokopczyk, B. and Djordjevic, M.V. 1994. Tobacco-specific N-nitrosamines and Areca-derived N-nitrosamine chemistry, biochemistry, carcinogenicity and relevance to humans. *J. Toxicol. Environ. Health* 41: 1-52.

Jack, A., Fannin, N. and Bush, L.P. 2007. Implications of reducing nornicotine accumulation in burley tobacco: Appendix A - The LC Protocol. *Rec. Adv. Tob. Sci.* 33: 58-79.

Julio, E., Laporte, F., Reis, S., Rothan, C. and Dorlhac de Borne, F. 2008. Reducing the content of nornicotine in tobacco via targeted mutation breeding. *Mol. Breed.* 21: 369-381.

Katz, J., Caudle, R.M., Bhattacharyya, I., Stewart, C.M. and Cohen, D.M. 2005. Receptor for advanced glycation end product (RAGE) upregulation in human gingival fibroblasts incubated with nornicotine. *J. Periodontol.* 76: 1171-1174.

Lewis, R.S., Jack, A.M., Morris, J.W., Robert, V.J.M., Gavilano, L., Siminszky, B., Bush, L.P., Hayes, A.J. and Dewey, R.E. 2008. RNAi-induced suppression of nicotine demethylase activity reduces levels of a key carcinogen in cured tobacco leaves. *Plant Biotech. J.* 6: 346-354.

Peele, D.M. and Gentry, J.S. 1999. Formation of tobacco-specific nitrosamines in flue-cured tobacco. CORESTA Meeting, Agro-Phyto Groups, Suzhou, China.

Pompon, D., Louerat, B., Bronne, A. and Urban, P. 1995. Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol.* 272: 51-64.

Siminszky, B., Gavilano, L., Bowen, S.W. and Dewey, R.E. 2005. Conversion of nicotine to nornicotine in *Nicotiana tabacum* is mediated by CYP82E4, a cytochrome P450 monooxygenase. *Proc. Natl. Acad. Sci. USA* 102: 14919-14924.

Wernsman, E.A. and Matzinger, D.F. 1968. Time and site of nicotine conversion in tobacco. *Tob. Sci.* 12: 226-228.

Xu, D., Shen, Y., Chappell, J., Cui, M. and Nielsen, M. 2007a. Biochemical and molecular characterization of nicotine demethylase in tobacco. *Physiol. Plantarum* 129: 307-319.

Xu, D., Nielsen, M.T. and Shen, Y. 2007b. Tobacco plants having a mutation in a nicotine demethylase gene. U.S. Patent Application 20070199097.

Many modifications and other embodiments of the inventions set forth herein will come to mind to one skilled in the art to which these inventions pertain having the benefit of the teachings presented in the foregoing descriptions and the associated drawings. Therefore, it is to be understood that the inventions are 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 list of embodiments and appended claims. Although specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

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.

THAT WHICH IS CLAIMED

1. A tobacco plant, or plant part thereof, comprising a mutation in a gene encoding a CYP82E10 nicotine demethylase, wherein said mutation results in reduced expression or function of said CYP82E10 nicotine demethylase.

2. The tobacco plant, or plant part thereof, according to claim 1, wherein said CYP82E10 nicotine demethylase is selected from the group consisting of the sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, and 9.

3. The tobacco plant, or plant part thereof, according to claim 1 or 2, wherein said mutation results in a modification of said CYP82E10 nicotine demethylase occurring at a position selected from the group consisting of amino acid residues 79, 107, 382, 419, and any combination thereof, wherein said numbering is according to SEQ ID NO:2.

4. The tobacco plant, or plant part thereof, according to claim 3, wherein said mutation is selected from the group consisting of:

a) a serine substitution for the glycine residue at position 79;

b) a serine substitution for the proline residue at position 107;

c) a serine substitution for the proline residue at position 382;

d) a serine substitution for the proline residue at position 419; and

e) any combination thereof.

5. The tobacco plant, or plant part thereof, according to any of claims 1-4, further comprising a mutation in a gene encoding a CYP82E4 nicotine demethylase, wherein said mutation results in reduced expression or function of said CYP82E4 nicotine demethylase.

6. The tobacco plant, or plant part thereof, according to claim 5, wherein said CYP82E4 nicotine demethylase is selected from the sequence set forth in SEQ ID NO:14, 15, 16, 17, 18, 19, and 20.

7. The tobacco plant, or plant part thereof, according to claim 5 or 6, wherein said mutation results in a modification of said CYP82E4 nicotine demethylase occurring at a position selected from the group consisting of amino acid residues 329, 364, 376, 382, and 458, wherein said numbering is according to SEQ ID NO:14.

8. The tobacco plant, or plant part thereof, according to claim 7, wherein said mutation is selected from the group consisting of:

a) a stop codon substitution for the tryptophan residue at position 329;

b) an asparagine substitution for the lysine residue at position 364;

c) a methionine substitution for the valine residue at position 376;

d) a serine substitution for the proline residue at position 382;

d) a serine substitution for the proline residue at position 458; and

e) any combination thereof.

9. The tobacco plant, or plant part thereof, according to any of claims 1-8, further comprising a mutation in a gene encoding a CYP82E5 nicotine demethylase, wherein said mutation results in reduced expression or function of said CYP82E5 nicotine demethylase.

10. The tobacco plant, or plant part thereof, according to claim 9, wherein said CYP82E5 nicotine demethylase is selected from the sequence set forth in SEQ ID NO:26, 27, 28, 29, 30, 31, and 32.

11. The tobacco plant, or plant part thereof, according to claim 9 or 10, wherein said mutation results in a modification of said CYP82E5 nicotine demethylase occuring at a position selected from the group consisting of amino acid residues 422 and 449, wherein said numbering is according to SEQ ID NO:26.

12. The tobacco plant, or plant part thereof, according to claim 11, wherein said mutation is selected from the group consisting of:

a) a stop codon substituted for the tryptophan residue at position 422;

b) a leucine substituted for the proline residue at position 449; and

c) any combination thereof.

13. The tobacco plant, or plant part thereof, according to any of claims 9-12, comprising a mutation in said CYP82E10 nicotine demethylase gene and said CYP82E4 nicotine demethylase gene.

14. The tobacco plant, or plant part thereof, according to any of claims 1-13, wherein said tobacco plant, or plant part thereof, is homozygous for said mutation.

15. The tobacco plant, or plant part thereof, according to claim 14, wherein said CYP82E10 nicotine demethylase comprises a mutation at position 382, said CYP82E4 nicotine demethylase comprises a mutation at position 329, and said CYP82E5 nicotine demethylase comprises a mutation at position 422, wherein said numbering is according to SEQ ID NO:2, 14, and 26, respectively.

16. The tobacco plant, or plant part thereof, according to claim 15, wherein said mutation is selected from the group consisting of:

a) a serine substitution for the proline residue at position 382;

b) a stop codon substitution for the tryptophan residue at position 329;

c) a stop codon substitution for the tryptophan residue at position 422; and

d) any combination thereof.

17. The tobacco plant, or plant part thereof, according to any of claims 13-16, wherein said plant or plant part thereof has less than 1.5% conversion of nicotine to nornicotine.

18. The tobacco plant, or plant part thereof, according to claim 17, wherein said plant or plant part thereof has no more than 0.5% conversion of nicotine to nornicotine.

19. Seed of the tobacco plant according to any of claims 1-18, or progeny thereof.

20. A tobacco product prepared from a tobacco plant, or plant part or progeny thereof, according to any of claims 1-19.

21. A method for reducing a carcinogenic potential of a tobacco product, said method comprising preparing said tobacco product from a tobacco plant, or plant part or progeny thereof, according to any of claims 1-18.

22. A method for reducing the level of nornicotine, or reducing the rate of conversion of nicotine to nornicotine, in a tobacco plant, or a plant part thereof, said method comprising introducing into the genome of said plant a mutation within at least one allele of each of at least three nicotine demethylase genes, wherein said mutation reduces expression of said nicotine demethylase gene, and wherein a first of said nicotine demethylase genes encodes a root-specific nicotine demethylase involved in the metabolic conversion of nicotine to nornicotine in a tobacco plant or a plant part thereof.

23. The method of claim 22, wherein said root-specific nicotine demethylase is a CYP82E10 nicotine demethylase comprising an amino acid sequence selected from the group consisting of:

a) the amino acid sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, 9, or 10; and

b) an amino acid sequence having at least 98% sequence identity to the amino acid sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, 9, or 10.

24. The method of claim 23, wherein said amino acid sequence for said CYP82E10 nicotine demethylase has a substitution at an amino acid residue in a position selected from the group consisting of residues 79, 107, 382, 419, and any combination thereof, where the numbering is according to SEQ ID NO:2.

25. The method of claim 24, wherein said substitution at position 79, 107, 382, or 419 is a serine residue.

26. The method of any one of claims 22-25, wherein a second of said nicotine demethylase genes encodes a CYP82E4 nicotine demethylase.

27. The method of claim 26, wherein said CYP82E4 nicotine demethylase comprises an amino acid sequence selected from the group consisting of:

a) the amino acid sequence set forth in SEQ ID NO:14, 15, 16, 17, 18, 19, 20, or 21; and

b) an amino acid sequence having at least 98% sequence identity to the sequence set forth in SEQ ID NO:14, 15, 16, 17, 18, 19, 20, or 21.

28. The method of claim 27, wherein said amino acid sequence for said CYP82E4 nicotine demethylase has a substitution at an amino acid residue in a position selected from the group consisting of residues 329, 364, 382, 458, and any combination thereof, where the numbering is according to SEQ ID NO:14.

29. The method of claim 28, wherein said substitution at position 329 is a stop codon, said substitution at position 364 is an asparagine residue, said substitution at position 382 is a serine residue, said substitution at position 458 is a serine residue, or any combination thereof.

30. The method of any one of claims 22-29, wherein a third of said nicotine demethylase genes encodes a CYP82E5 nicotine demethylase.

31. The method of claim 30, wherein said CYP82E5 nicotine demethylase comprises an amino acid sequence selected from the group consisting of:

a) the amino acid sequence set forth in SEQ ID NO:26, 27, 28, 29, 30, 31, or 32; and

b) an amino acid sequence having at least 98% sequence identity to the sequence set forth in SEQ ID NO: 26, 27, 28, 29, 30, 31, or 32.

32. The method of claim 31, wherein said amino acid sequence for said CYP82E5 nicotine demethylase has a substitution at an amino acid residue in a position selected from the group consisting of residues 422 and 449, and any combination thereof, where the numbering is according to SEQ ID NO:26.

33. The method of claim 32, wherein said substitution at position 422 is a stop codon, said substitution at position 449 is a leucine residue, or any combination thereof.

34. The method of any one of claims 22-33, wherein said plant or plant part thereof is homozygous for said mutation.

35. The method of any one of claims 22-34, wherein said introducing comprises a breeding protocol.

36. The method of any one of claims 22-35, wherein said plant is a Burley, Virginia, flue-cured, air-cured, fire-cured, Oriental, or a dark tobacco plant.

37. The tobacco plant, or plant part thereof, according to any of claims 1-19, wherein said tobacco plant is a Burley, Virginia, flue-cured, air-cured, fire-cured, Oriental, or a dark tobacco plant.

38. A method for identifying a tobacco plant with low levels of nornicotine, said method comprising screening a DNA sample from a tobacco plant of interest for the presence of a mutation in SEQ ID NO:1 or 3.

39. The method according to claim 38, wherein said tobacco plant is a nonconverter.

40. The method according to claim 38 or 39, wherein said screening is carried out using a sequence selected from the group consisting of SEQ ID NOS:1, 3, 35, 36, 37, and 38.

41. The method according to any one of claims 38-40, further comprising screening said DNA sample, or another DNA sample from said tobacco plant of interest, for the presence of a mutation in SEQ ID NO:14, the presence of a mutation in SEQ ID NO:26, or the presence of a mutation in SEQ ID NO:14 and SEQ ID NO:26.

42. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

a) a nucleotide sequence comprising SEQ ID NO:1, 3, or 4;

b) a nucleotide sequence comprising a fragment of at least 20 consecutive nucleotides of SEQ ID NO:1, 3, or 4;

c) a nucleotide sequence having at least 97% sequence identity to the entirety of the sequence set forth in SEQ ID NO:1, wherein said polynucleotide encodes a polypeptide involved in the metabolic conversion of nicotine to nornicotine in a plant;

d) a nucleotide sequence encoding a polypeptide selected from the group consisting of SEQ ID NOS:2 and 5-13, or a fragment thereof comprising at least 115 contiguous residues;

e) a nucleotide sequence encoding a polypeptide having at least 98% sequence identity to the sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, 9, 10, 11, 12, or 13; and

f) a nucleotide sequence that is complementary to the sequence according to any of preceding items (a) through (e).

43. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

a) an amino acid sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, 9, 10, 11, 12, or 13;

b) an amino acid sequence that is at least 98% identical to an amino acid sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, 9, 10, 11, 12, or 13; and

c) an amino acid sequence that is a fragment of the amino acid sequence set forth in SEQ ID NO:2, 5, 6, 7, 8, 9, 10, 11, 12, or 13, wherein said fragment comprises at least 115 contiguous residues of the amino acid sequence of SEQ ID NO:2, 5, 6, 7, 8, 9, 10, 11, 12, or 13.

44. A tobacco plant, or plant part thereof that is homozygous for a mutation in a gene encoding a CYP82E10 nicotine demethylase, a gene encoding a CYP82E4 nicotine demethylase, and a gene encoding a CYP82E5 nicotine demethylase, wherein said mutation results in reduced expression or function of said CYP82E10, CYP82E4, and CYP82E5 nicotine demethylase, wherein said CYP82E10 nicotine demethylase comprises a mutation at position 382, said CYP82E4 nicotine demethylase comprises a mutation at position 329, and said CYP82E5 nicotine demethylase comprises a mutation at position 422, wherein said numbering is according to SEQ ID NO:2, 14, and 26, respectively.

45. A mutation in a gene encoding a CYP82E10 nicotine demethylase, wherein said mutation results in reduced expression or function of said CYPe2E10 nicotine demethylase.

46. A plant having a mutation in a *CYP82E10* gene that inhibits nicotine demethylase activity in roots, a mutation in a *CYP82E4v2* gene that inhibits nicotine demethylase activity in senescent leaves, and a mutation in a *CYP83E5v2* gene that inhibits nicotine demethylase activity in green leaves.

```
  1 ttttcaatttttgttacttttgtatttatcatattattatgcatagccctaaattatcta

 61 taaaagggaagttggtgatagtttgattcccaagtgcttttctaaaaatccataATGGTT
                                                          M  V   2

121 TCTCCCGTAGAAGCCATCGTAGGACTAGTAACTCTTACACTTCTCTTCTACTTCATACGG
     S  P  V  E  A  I  V  G  L  V  T  L  T  L  L  F  Y  F  I  R  22

181 ACCAAAAAATCTCAAAAACCTTCAAAACCATTACCACCGAAAATCCCCGGAGGGTGGCCG
     T  K  K  S  Q  K  P  S  K  P  L  P  P  K  I  P  G  G  W  P  42

241 GTAATCGGCCATCTTTTCTATTTCGATGACGACAGCGACGACCGTCCATTAGCACGAAAA
     V  I  G  H  L  F  Y  F  D  D  D  S  D  D  R  P  L  A  R  K  62

301 CTCGGAGACTTAGCTGACAAATACGGCCCGGTTTTCACTTTTCGGCTAGGCCTTCCGCTT
     L  G  D  L  A  D  K  Y  G  P  V  F  T  F  R  L  G  L  P  L  82

361 GTGTTAGTTGTAAGCAGTTACGAAGCTATAAAAGACTGCTTCTCTACAAATGATGCCATT
     V  L  V  V  S  S  Y  E  A  I  K  D  C  F  S  T  N  D  A  I 102

421 TTCTCCAATCGTCCAGCTTTTCTTTATGGCGAATACCTTGGCTACAATAATGCCATGCTA
     F  S  N  R  P  A  F  L  Y  G  E  Y  L  G  Y  N  N  A  M  L 122

481 TTTTTGACAAAATACGGACCTTACTGGCGAAAAAATAGAAAATTAGTCATTCAGGAAGTT
     F  L  T  K  Y  G  P  Y  W  R  K  N  R  K  L  V  I  Q  E  V 142

541 CTCTGTGCTAGTCGTCTCGAAAAATTGAAGCACGTGAGATTTGGTGAAATTCAGACGAGC
     L  C  A  S  R  L  E  K  L  K  H  V  R  F  G  E  I  Q  T  S 162

601 ATTAAGAATTTATACACTCGAATTGATGGAAATTCGAGTACGATAAATCTAACCGATTGG
     I  K  N  L  Y  T  R  I  D  G  N  S  S  T  I  N  L  T  D  W 182

661 TTAGAAGAATTGAATTTTGGTCTGATCGTGAAAATGATCGCTGGGAAAAATTATGAATCC
     L  E  E  L  N  F  G  L  I  V  K  M  I  A  G  K  N  Y  E  S 202

721 GGTAAAGGAGATGAACAAGTGGAGAGATTTAGGAAAGCGTTTAAGGATTTTATAATTTTA
     G  K  G  D  E  Q  V  E  R  F  R  K  A  F  K  D  F  I  I  L 222

781 TCAATGGAGTTTGTGTTATGGGATGCTTTTCCAATTCCATTGTTCAAATGGGTGGATTTT
     S  M  E  F  V  L  W  D  A  F  P  I  P  L  F  K  W  V  D  F 242

841 CAAGGCCATGTTAAGGCCATGAAAAGGACATTTAAGGATATAGATTCTGTTTTTCAGAAT
     Q  G  H  V  K  A  M  K  R  T  F  K  D  I  D  S  V  F  Q  N 262

901 TGGTTAGAGGAACATGTCAAGAAAAAAGAAAAAATGGAGGTTAATGCAGAAGGAAATGAA
     W  L  E  E  H  V  K  K  K  E  K  M  E  V  N  A  E  G  N  E 282

961 CAAGATTTCATTGATGTGGTGCTTTCAAAAATGAGTAATGAATATCTTGATGAAGGCTAC
     Q  D  F  I  D  V  V  L  S  K  M  S  N  E  Y  L  D  E  G  Y 302
```

FIGURE 1B

1021 TCTCGTGATACTGTCATAAAAGCAACAGTGTTTgtaagttcatctcatttttcatttatt
S R D T V I K A T V F 313

1081 ctttgaggaatagacaggttaatagtaatttaagtaattagattatctaaatactaagga

1141 tgagtaaatatggcaaaaatatagaatgataaatggaaaaggatgataatttttatgcc

1201 cggactaatctaactttgggagttaaagcacttcctaccaatagggacttttcttcaagc

1261 tcgatcttgatgaaactctgtggttaaaaaaatgagatatanccaattataattgataga

1321 ataaaactttattactcccattgagcataacaaaacaaaaaaaagtaaagggacttcttc

1381 tctttttagggagaaattctttgattgtttgttaatatagattcatgtttttttttatt

1441 tctaataataattgtgcttgaatc**aggtcgcgctgattcttg**gcttttagcagcaatag

1501 agtcaaagctaatatacatattatttggttttcgaataagttatactgaaattatataat

1561 acgggtattaaataataacatgattatttataggatatgctttttttattgggtaaatat

1621 atttttttaattaaaaatgaaatatacaagtaaggtataaaacactatttgattttaca

1681 ctagataaatttgccctcgtacatctctaagagaagagctgaaataaatgaattttaaat

1741 ttcagaaaaaaataaattcattagtataatgagatgtcgatacttgacaattactatact

1801 aactagaacaaggttcagcagatagtgacgctaacctatttttgtattgaattattctaa

1861 tttgtccacagAGTTTAGTCTTGGATGCTGCGGACACAGTTGCTCTTCACATGAATTGGG
S L V L D A A D T V A L H M N W 329

1921 GAATGGCATTATTGATAAACAATCAACATGCCTTGAAGAAAGCGCAAGAAGAGATAGATA
G M A L L I N N Q H A L K K A Q E E I D 349

1981 AAAAAGTTGGTAAGGATAGATGGGTAGAAGAGAGTGATATTAAGGATTTGGTATACCTCC
K K V G K D R W V E E S D I K D L V Y L 369

2041 AAACTATTGTTAAAGAAGTGTTACGATTATATCCACCGGGACCTTTATTAGTACCCCATG
Q T I V K E V L R L Y P P G P L L V P H 389

2101 AAAATGTAGAGGATTGTGTTGTTAGTGGATATCACATTCCTAAAGGGACTAGACTATTCG
E N V E D C V V S G Y H I P K G T R L F 409

2161 CGAACGTTATGAAATTACAGCGCGATCCTAAACTCTGGTCAAATCCTGATAAGTTCGATC
A N V M K L Q R D P K L W S N P D K F D 429

FIGURE 1C

```
2221 CAGAGAGATTTTTCGCTGCTGATATTGACTTTCGTGGTCAACACTATGAGTTTATCCCAT
     P  E  R  F  F  A  A  D  I  D  F  R  G  Q  H  Y  E  F  I  P   449

2281 TTGGTTCTGGAAGACGATCTTGTCCGGGGATGACTTATGCAATGCAAGTGGAACACCTAA
     F  G  S  G  R  R  S  C  P  G  M  T  Y  A  M  Q  V  E  H  L   469

2341 CAATCGCACACTTGATCCAGGGTTTCAATTACAAAACTCCAAATGACGAGCCCTTGGATA
     T  I  A  H  L  I  Q  G  F  N  Y  K  T  P  N  D  E  P  L  D   489

2401 TGAAGGAAGGTGCAGGATTAACTATACGTAAGGTAAATCCTATAGAAGTGGTAATTACGC
     M  K  E  G  A  G  L  T  I  R  K  V  N  P  I  E  V  V  I  T   509

2461 CTCGCCTGACACCTGAGCTTTATtaaaatctaagatgttttatcttggttgatcattgtt
     P  R  L  T  P  E  L  Y                                       517

2521 taatactcctagatagatgggtattcatctatcttttaaaattaattgtcagtacgagt

2581 gtttctaatttggtaagtttgtaacaacaagtaaagaaggattgtgctagtatgta
```

FIGURE 2A

```
                                                                                                                         100
CYP82E10   aaattatcta taaaagggaa gttggtgata gtttgattcc caagtgcttt tctaaaaatc cataATGGTT TCTCCCGTAG AAGCCATCGT AGGACTAGTA
CYP82E5v2  .......... .......... .......... ....gattcc caagttcttt tctaaaactc cataATGGTT TCTCCCGTAG AAGCCATTGT AGGACTAGTA
CYP82E4v2  .......... ....aaggaa gttgccgata gttatattct caacttctta tctaaaaatc cataATGCTT TCTCCCATAG AAGCCATTGT AGGACTAGTA
                                                                                                                         200
CYP82E10   ACTCTTACAC TTCTCTTCTA CTTCATACGG ACCAAAAAAT CTCAAAAACC TTCAAAACCA TTACCACCGA AAATCCCCGG AGGGTGGCCG GTAATCGGCC
CYP82E5v2  ACCCTTACAC TTCTCTTCTA CTTCCTATGG CCCAAAAAAT TTCAAATACC TTCAAAACCA TTACCACCGA AAATTCCCGG AGGGTGGCCG GTAATCGGCC
CYP82E4v2  ACCTTCACAT TTCTCTTCTT CTTCCTATGG ACAAAAAAAT CTCAAAAACC TTCAAAACCC TTACCACCGA AAATCCCCGG AGGATGGCCG GTAATCGGCC
                                                                                                                         300
CYP82E10   ATCTTTTCTA TTTCGATGAC GACAGCGACG ACCGTCCATT AGCACGAAAA CTCGGAGACT TAGCTGACAA ATACGGCCCG GTTTTCACTT TTCGGCTAGG
CYP82E5v2  ATCTTTTCTA CTTCGATGAT GACGGCGACG ACCGTCCATT AGCTCGAAAA CTCGGAGACT TAGCTGACAA ATACGGCCCG GTTTTCACTT TCCGGCTAGG
CYP82E4v2  ATCTTTTCCA CTTCAATGAC GACGGCGACG ACCGTCCATT AGCTCGAAAA CTCGGAGACT TAGCTGACAA ATACGGCCCC GTTTTCACTT TTCGGCTAGG
                                                                                                                         400
CYP82E10   CCTTCCGCTT GTGTTAGTTG TAAGCAGTTA CGAAGCTATA AAAGACTGCT TCTCTACAAA TGATGCCATT TTCTCCAATC GTCCAGCTTT TCTTTATGGC
CYP82E5v2  CCTTCCGCTT GTGTTAGTTG TAAGCAGTTA CGAAGCTGTA AAAGACTGCT TCTCTACAAA TGACGCCATT TTCTCCAATC GTCCAGCTTT TCTTTACGGT
CYP82E4v2  CCTTCCCCTT GTCTTAGTTG TAAGCAGTTA CGAAGCTGTA AAAGACTGTT TCTCTACAAA TGACGCCATT TTTTCCAATC GTCCAGCTTT TCTTTACGGC
                                                                                                                         500
CYP82E10   GAATACCTTG GCTACAATAA TGCCATGCTA TTTTTGACAA AATACGGACC TTACTGGCGA AAAAATAGAA AATTAGTCAT TCAGGAAGTT CTCTGTGCTA
CYP82E5v2  GAATACCTTG GCTACAGTAA TGCCATGCTA TTTTTGACAA AATACGGACC TTATTGGCGA AAAAATAGAA AATTAGTCAT TCAGGAAGTT CTCTCTGCTA
CYP82E4v2  GATTACCTTG GCTACAATAA TGCCATGCTA TTTTTGGCCA ATTACGGACC TTACTGGCGA AAAAATCGAA AATTAGTTAT TCAGGAAGTT CTCTCCGCTA
                                                                                                                         600
CYP82E10   GTCGTCTCGA AAAATTGAAG CACGTGAGAT TTGGTGAAAT TCAGACGAGC ATTAAGAATT TATACACTCG AATTGATGGA AATTCGAGTA CGATAAATCT
CYP82E5v2  GTCGTCTCGA AAAATTGAAG CACGTGAGAT TTGGTAAAAT TCAAACGAGC ATTAAGAATT TATACACTCG AATTGATGGA AATTCGAGTA CGATAAATCT
CYP82E4v2  GTCGTCTCGA AAAATTCAAA CACGTGAGAT TTGCAAGAAT TCAAGCGAGC ATTAAGAATT TATATACTCG AATTGATGGA AATTCGAGTA CGATAAATTT
                                                                                                                         700
CYP82E10   AACCGATTGG TTAGAAGAAT TGAATTTTGG TCTGATCGTG AAAATGATCG CTGGGAAAAA TTATGAATCC GGTAAAGGAG ATGAACAAGT GGAGAGATTT
CYP82E5v2  AACTGATTGG TTAGAAGAAT TGAATTTTGG TCTGATCGTG AAAATGATCG CTGGGAAAAA TTATGAATCC GGTAAAGGAG ATGAACAAGT GGAGAGATTT
CYP82E4v2  AACTGATTGG TTAGAAGAAT TGAATTTTGG TCTGATCGTG AAGATGATCG CTGGAAAAAA TTATGAATCC GGTAAAGGAG ATGAACAAGT GGAGAGATTT
                                                                                                                         800
CYP82E10   AGGAAAGCGT TTAAGGATTT TATAATTTTA TCAATGGAGT TTGTGTTATG GGATGCTTTT CCAATTCCAT TGTTCAAATG GGTGGATTTT CAAGGCCATG
CYP82E5v2  AGGAAAGCGT TTAAGGATTT TATAATTTTA TCAATGGAGT TTGTGTTATG GGATGCTTTT CCAATTCCAT TGTTCAAATG GGTGGATTTT CAAGGCCATG
CYP82E4v2  AAGAAAGCGT TTAAGGATTT TATGATTTTA TCAATGGAGT TTGTGTTATG GGATGCATTT CCAATTCCAT TATTTAAATG GGTGGATTTT CAAGGGCATG
                                                                                                                         900
CYP82E10   TTAAGGCCAT GAAAAGGACA TTTAAGGATA TAGATTCTGT TTTTCAGAAT TGGTTAGAGG AACATGTCAA GAAAAAAGAA AAAATGGAGG TTAATGCAGA
CYP82E5v2  TTAAGGCCAT GAAAAGGACA TTTAAGGATA TAGATTCTGT TTTTCAGAAT TGGTTAGAGG AACATGTCAA GAAAAGAGAA AAAATGGAGG TTAATGCACA
CYP82E4v2  TTAAGGCTAT GAAAAGGACT TTTAAAGATA TAGATTCTGT TTTTCAGAAT TGGTTAGAGG AACATATTAA TAAAAGAGAA AAAATGGAGG TTAATGCAGA
                                                                                                                        1000
CYP82E10   AGGAAATGAA CAAGATTTCA TTGATGTGGT GCTTTCAAAA ATGAGTAATG AATATCTTGA TGAAGGCTAC TCTCGTGATA CTGTCATAAA AGCAACAGTG
CYP82E5v2  AGGGAATGAA CAAGATTTCA TTGATGTGGT GCTTTCAAAA ATGAGTAATG AATATCTTGA TGAAGGTTAC TCTCGTGATA CTGTCATAAA AGCAACAGTG
CYP82E4v2  AGGGAATGAA CAAGATTTCA TTGATGTGGT GCTTTCAAAA ATGAGTAATG AATATCTTGG TGAAGGTTAC TCTCGTGATA CTGTCATTAA AGCAACGGTG
```

FIGURE 2B

```
                                                                                                                   1100
CYP82E10   TTTgtaagtt catct..cat ttttcattt. at...tct.. ..ttgaggaa tagacaggtt aatagtaatt ta.agtaa.. .ttagattat ctaaatacta
CYP82E5v2  TTTgtaagtt cattt.tcat ttttcatta. ttcagtctga ttttgaggaa tagacaggtt aataataatt ta.agtaa.. .ttagattat ctaaatacta
CYP82E4v2  TTTgtaagtt catctgtcat ttttcattta ttcactttta ttttgaggag cagacatgtt aataataatt tggagcaact gtaaagttat ctatgtgtac
                                                                                                                   1200
CYP82E10   aggatgagta aatatggcaa aaatatagaa tgataaatgg aaaag.gatg ataatttttt atgcccggac taatctaa.. ..ctttggga gttaaa.gca
CYP82E5v2  aggatgatta tatatagtaa aaatgtagaa tgataaatgg aaaaaagatg agaatttttt gtgcctcgac taatctatat atctttggga gttaaaagtg
CYP82E4v2  aggtt.cgag cctcaggtgc aaccactaat gcttgtatta gattatgttg tctgcatcat ac.ccctaat tggagtgtgg ctcttcccga accct..gca
                                                                                                                   1300
CYP82E10   cttcctacca atagggactt ttcttca.ag ctcga..... .........t cttgatgaaa .......... .......... .ctctgtggt taaaa..aaa
CYP82E5v2  cttc..acca aaggggactt ttcctcatag ctcaagttag aagtttgatt atagatgaaa gagtatttat cacttcacga actctgatga taaaagtaaa
CYP82E4v2  atgctggatg ctggatgctt tatgtatcag actgac.... .......ctt tttgttaaac .......... .......... tatctaaata ctaaggatga
                                                                                                                   1400
CYP82E10   tgagata.ta accaattata attgatagaa taaaacttta ttactcccat tgagcataac aaaacaaaaa aaagtaaagg gacttcttct ctttttt..a
CYP82E5v2  tgagata.ta accagttata attgatagaa taaaacttca ttactcccat tgagcataaa aaaaaaagta aaag.....g gacttcttct cttttttttta
CYP82E4v2  tgatttaata aaaatataga atggtaaaca gaaaaagatg agattatttt tggggctata tggattcgcc cgggctttgg gaggtaaaac ggtatctacc
                                                                                                                   1500
CYP82E10   gggagaaatt ctttgattgt ttgttaa.ta tagattcatg tttttttttt. .....atttc taataataat tgtgcttgaa tcaggtcgcg ctgattcttg
CYP82E5v2  gggagaaatt ctttaattgt ttgttaaata tagattcatg tttttttttt cttctatttc taataataat ggttcttgaa tcaggtcg.. .......ttg
CYP82E4v2  agttgagact ttactccaga actttatctc gagagctctg aataaaaatg aaatagtatt taccactcca aaatctttga tggtaaaaag atgagatata
                                                                                                                   1600
CYP82E10   gctttttagc agcaatagag tcaaa..gct aatatacata ttatttggtt ttcgaataag ttatactgaa attatataat acgggtatta aataataaca
CYP82E5v2  actttgtagc agcaatatag tcaaa..gct aatatccatg ttatttggtt ttcgaacaag ttatactgaa attatata.t acgggtatta aataataaca
CYP82E4v2  acctcttata attgattgaa ccacgttgat agaataaaac ttcttta..c tcccattcag cataagaaaa atgaaaccaa acgg.....a attcttctct
                                                                                                                   1700
CYP82E10   tgattattta taggatatgc tttttttatt gggtaaatat .......... .......... .......... .......... .......... ..........
CYP82E5v2  ttattattta taggatatac tttttttatt gggtaaatat tacaacaaca acaactgact cagtaaaatt ttactagtgg ggtatgggga gggtagtgtg
CYP82E4v2  tttttа...g ggggaaattc cttaattgct tgttgaatat agattcatgt cg.......t tattctattt ttaataatga tgaaaatcaa ........ta
                                                                                                                   1800
CYP82E10   .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CYP82E5v2  tatgcagacc ttaccсctac cccgaaggag tagagggatt gtttccgaaa gaccctcggc tcaagaaaac aaaaagagac aatatcagta ccaccacaga
CYP82E4v2  tagtcaaagt taatacttat gtcatttggt ttgcggacaa gttat..att ggaactatat aatacgtcta ttatagaata gtgattattt agaggatata
                                                                                                                   1900
CYP82E10   .......... .......... .a.tttttt taattaaaaa tgaaatatac aagtaaggta taaaacac.t atttgatttt acactagata aatttgccct
CYP82E5v2  tcatattatt aggtaaatgt ta.ttttatt gaattaaaga tgaaatatac aggtaaggta taaaacgtgt atttgatttt acactagata aatttgacct
CYP82E4v2  cattttttttt ggataaatat ttgatttatt ggattaaaaa tagaatatac aggtaaggtc taaaacgtgt gtttgctttt acactaaata aacttgacct
                                                                                                                   2000
CYP82E10   cgtacatctc taagagaaga gctgaaataa atgaatt... ttaaatttca gaaaaaaata aattcattag tataatgaga tgtc..gata cttgacaatt
CYP82E5v2  cgtacatctc taagagaa.a gctgaaataa atgaatt... ttaaattt.. ..aaaaaaaa aattcattag tataatgaga tgtg..cata cttgacaatt
CYP82E4v2  cgtacaattc taagaaaata tttgaaataa atgaattatt ttattgttaa tcaattaaaa aaatcatagt atagatgaga tgtgtgcata cttgacaata
```

FIGURE 2C

```
                                                                                                               2100
CYP82E10   actatactaa ctagaacaag gttcagcaga tagtgacgct aacctatttt tgtattgaat tat....... ........tc taatttgtcc acagAGTTTA
CYP82E5v2  actatactaa atagaacaag gttcggcaga tagtgacact aacctacttt tgtattgaat tatccttttt aattttattc taatttgtct acagAGTTTG
CYP82E4v2  actatactaa ctaaaacaag gtatgtgaat aattgatatt ..cctttttt .......aat tat....... ........tc tttttt.... ccagAGTTTG
                                                                                                               2200
CYP82E10   GTCTTGGATG CTGCGGACAC AGTTGCTCTT CACATGAATT GGGGAATGGC ATTATTGATA AACAATCAAC ATGCCTTGAA GAAAGCGCAA GAAGAGATAG
CYP82E5v2  GTCTTGGATG CTGCGGACAC AGTTGCTCTT CACATGAATT GGGGAATGGC ATTACTGATA AACAATCAAC ATGCCTTGAA GAAAGCACAA GAAGAGATCG
CYP82E4v2  GTCTTGGATG CAGCAGACAC AGTTGCTCTT CACATAAATT GGGGAATGGC ATTATTGATA AACAATCAAA AGGCCTTGAC GAAAGCACAA GAAGAGATAG
                                                                                                               2300
CYP82E10   ATAAAAAAGT TGGTAAGGAT AGATGGGTAG AAGAGAGTGA TATTAAGGAT TTGGTATACC TCCAAACTAT TGTTAAAGAA GTGTTACGAT TATATCCACC
CYP82E5v2  ATAAAAAAGT TGGTAAGGAA AGATGGGTAG AAGAGAGTGA TATTAAGGAT TTGGTCTACC TCCAAGCTAT TGTTAAAGAA GTGTTACGAT TATATCCACC
CYP82E4v2  ACACAAAAGT TGGTAAGGAC AGATGGGTAG AAGAGAGTGA TATTAAGGAT TTGGTATACC TCCAAGCTAT TGTTAAAGAA GTGTTACGAT TATATCCACC
                                                                                                               2400
CYP82E10   GGGACCTTTA TTAGTACCCC ATGAAAATGT AGAGGATTGT GTTGTTAGTG GATATCACAT TCCTAAAGGG ACTAGACTAT TCGCGAACGT TATGAAATTA
CYP82E5v2  AGGACCTTTA TTAGTACCTC ATGAAAATGT AGAGGATTGT GTTGTTAGTG GATATCACAT TCCTAAAGGG ACTAGACTAT TCGCGAACGT TATGAAATTG
CYP82E4v2  AGGACCTTTG TTAGTACCAC ACGAAAATGT AGAAGATTGT GTTGTTAGTG GATATCACAT TCCTAAAGGG ACAAGATTAT TCGCAAACGT CATGAAACTG
                                                                                                               2500
CYP82E10   CAGCGCGATC CTAAACTCTG GTCAAATCCT GATAAGTTCG ATCCAGAGAG ATTTTTCGCT GCTGATATTG ACTTTCGTGG TCAACACTAT GAGTTTATCC
CYP82E5v2  CAGCGCGATC CTAAACTCTG GTCAAATCCT GATAAGTTTG ATCCAGAGAG ATTCTTCGCT GATGATATTG ACTACCGTGG TCAGCACTAT GAGTTTATCC
CYP82E4v2  CAACGTGATC CTAAACTCTG GTCTGATCCT GATACTTTCG ATCCAGAGAG ATTCATTGCT ACTGATATTG ACTTTCGTGG TCAGTACTAT AAGTATATCC
                                                                                                               2600
CYP82E10   CATTTGGTTC TGGAAGACGA TCTTGTCCGG GGATGACTTA TGCAATGCAA GTGGAACACC TAACAATCGC ACACTTGATC CAGGGTTTCA ATTACAAAAC
CYP82E5v2  CATTTGGTTC TGGAAGACGA TCTTGTCCGG GGATGACTTA TGCATTACAA GTGGAACACC TAACAATAGC ACATTTGATC CAGGGTTTCA ATTACAAAAC
CYP82E4v2  CGTTTGGTTC TGGAAGACGA TCTTGTCCAG GGATGACTTA TGCATTGCAA GTGGAACACT TAACAATGGC ACATTTGATC CAAGGTTTCA ATTACAGAAC
                                                                                                               2700
CYP82E10   TCCAAATGAC GAGCCCTTGG ATATGAAGGA AGGTGCAGGA TTAACTATAC GTAAGGTAAA TCCTATAGAA GTGGTAATTA CGCCTCGCCT GACACCTGAG
CYP82E5v2  TCCAAATGAC GAGCCCTTGG ATATGAAGGA AGGTGCAGGA TTAACTATAC GTAAAGTAAA TCCTGTAGAA GTGACAATTA CGGCTCGCCT GGCACCTGAG
CYP82E4v2  TCCAAATGAC GAGCCCTTGG ATATGAAGGA AGGTGCAGGC ATAACTATAC GTAAGGTAAA TCCTGTGGAA CTGATAATAG CGCCTCGCCT GGCACCTGAG
                                                                                                               2800
CYP82E10   CTTTATTAAa atctaagatg ttttatcttg gttgatcatt gtttaatact cctagataga tgggtattca tctatctttt taaaattaat tgtcagtacg
CYP82E5v2  CTTTATTAAa accttagatg ttttatcttg attgtac..t aatatatata gcagaaaa.. .......... .......... .......... ..........
CYP82E4v2  CTTTATTAAa acctaagatc tttcatcttg gttgatcatt gtataatact cctaaa.... tggatattca tttacctttt atcaattaat tgtcagtacg
                                                                     2860
CYP82E10   agtgtttcta atttggtaag tttgtaacaa caagtaaaga aggattgtgc tagtatgta.
CYP82E5v2  .......... .......... .......... .......... .......... ..........
CYP82E4v2  agtttttcta atttggtaca tttgtaataa taagtaaaga ataattgtgc taatatataa
```

FIGURE 3

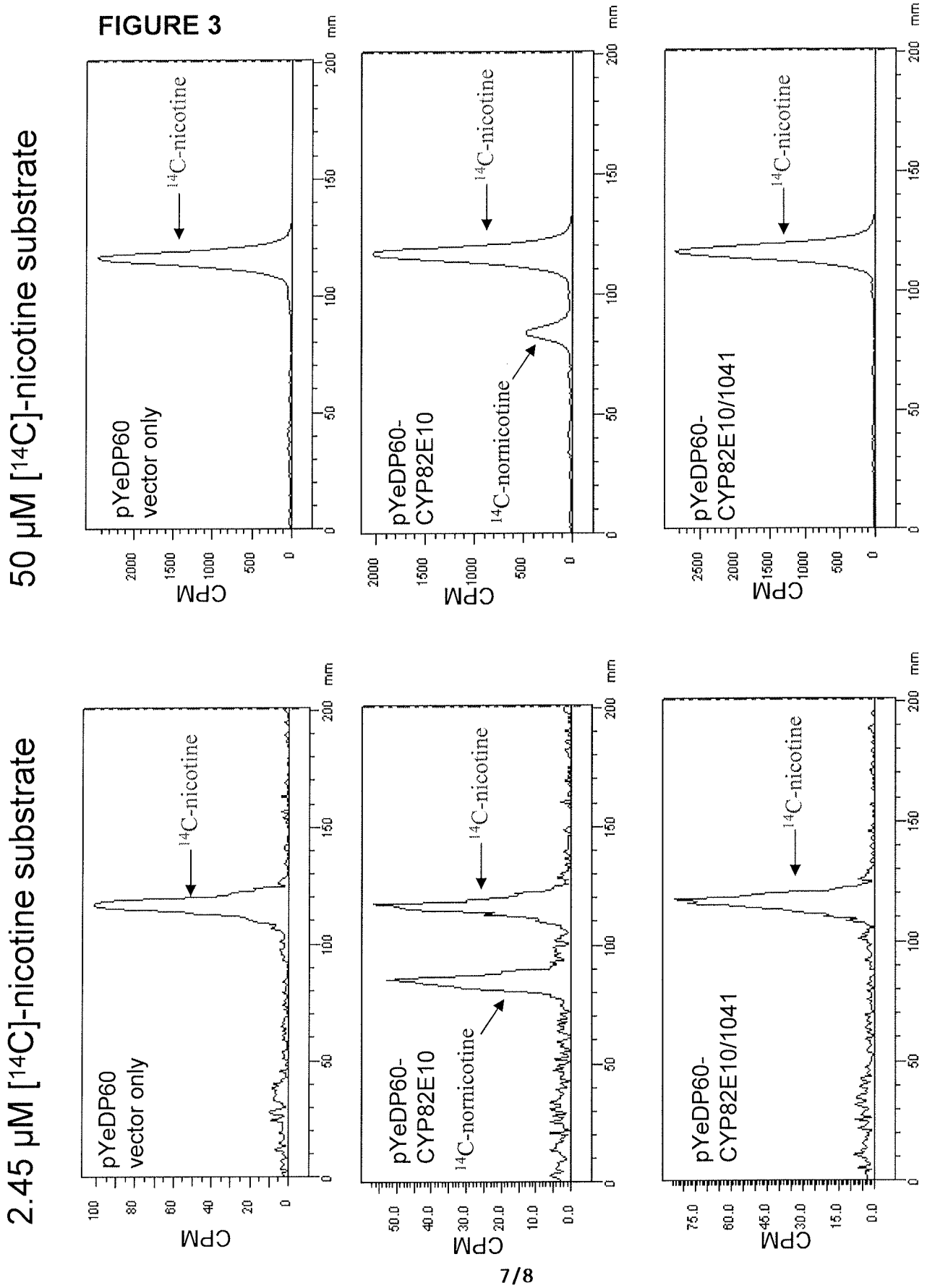
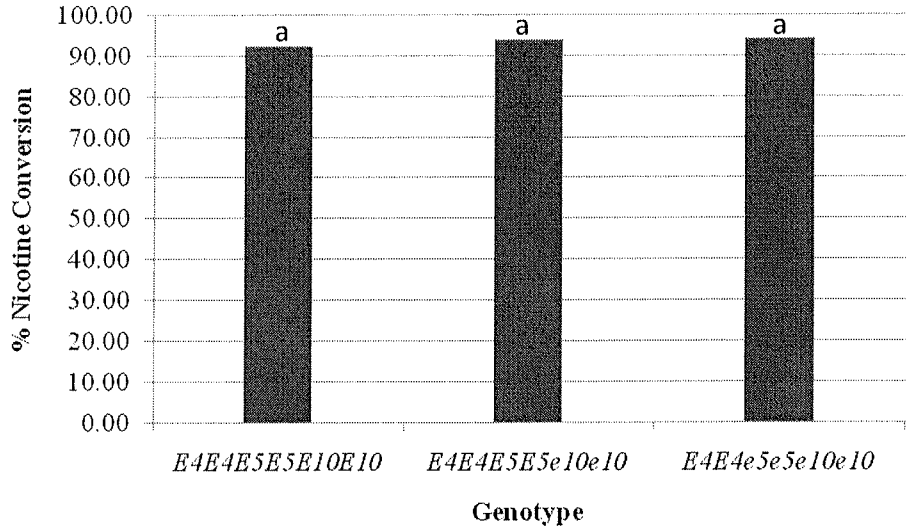

50 µM [14C]-nicotine substrate
2.45 µM [14C]-nicotine substrate
pYeDP60
vector only
pYeDP60-
CYP82E10
pYeDP60-
CYP82E10/1041
14C-nicotine
14C-normicotine
CPM
mm

FIGURE 4A



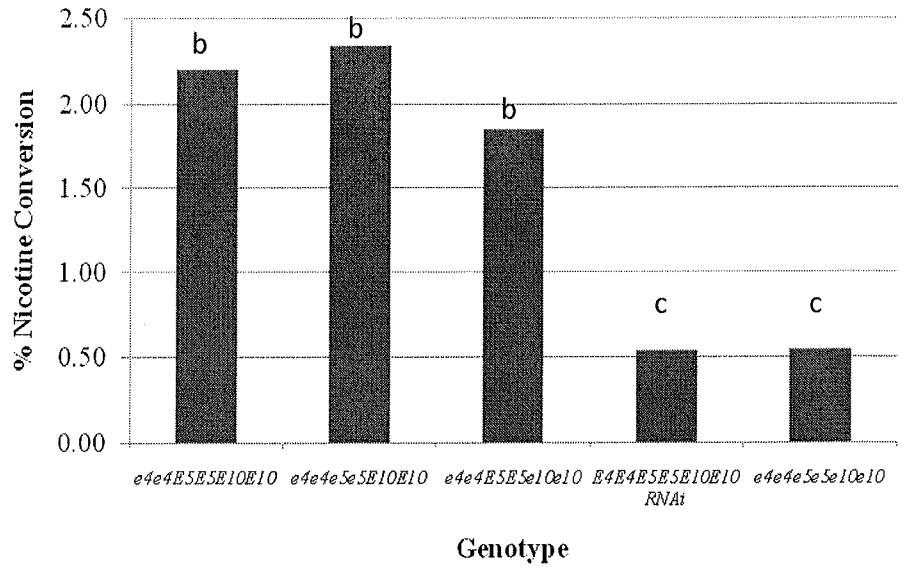


FIGURE 4B

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2011/021088

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/82 C12N15/53 A01H5/12
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, BIOSIS, Sequence Search, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2009/064771 A2 (UNIV NORTH CAROLINA STATE [US]; UNIV KENTUCKY RES FOUND [US]; DEWEY RA) 22 May 2009 (2009-05-22) sequences 1,2,4	42,43
X	DATABASE EMBL [Online] 20 June 2008 (2008-06-20), "CHO_OF040xc14f1.ab1 CHO_OF Nicotiana tabacum genomic 5', genomic survey sequence.", XP002629029, retrieved from EBI accession no. EMBL:ET701936 Database accession no. ET701936 the whole document	42

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21 March 2011	12/04/2011

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Maddox, Andrew

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2011/021088

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] 24 July 2009 (2009-07-24), "Nicotiana tabacum cDNA, clone: TBK02GR0013_1_H12, 5'-end sequence.", XP002629030, retrieved from EBI accession no. EMBL:FS405337 Database accession no. FS405337 the whole document -----	42
X	DATABASE EMBL [Online] 22 June 2008 (2008-06-22), "CHO_OF3490xi05r1.ab1 CHO_OF3 Nicotiana tabacum genomic 3', genomic survey sequence.", XP002629031, retrieved from EBI accession no. EMBL:FH058766 Database accession no. FH058766 the whole document -----	42
X	DATABASE EMBL [Online] 24 June 2008 (2008-06-24), "CHO_OF4790xg08f1.ab1 CHO_OF4 Nicotiana tabacum genomic 5', genomic survey sequence.", XP002629032, retrieved from EBI accession no. EMBL:FH555036 Database accession no. FH555036 the whole document -----	42
X,P	LEWIS RAMSEY S ET AL: "Three nicotine demethylase genes mediate nornicotine biosynthesis in Nicotiana tabacum L Functional characterization of the CYP82E10 gene", PHYTOCHEMISTRY (AMSTERDAM), vol. 71, no. 17-18, December 2010 (2010-12), pages 1988-1998, XP002629033, ISSN: 0031-9422 the whole document -/--	1-37, 42-46

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	-& DATABASE EMBL [Online] 8 November 2010 (2010-11-08), "Nicotiana tabacum cultivar DH98-325-6 nicotine N-demethylase (CYP82E10) gene, complete cds.", XP002629034, retrieved from EBI accession no. EMBL:HM802352 Database accession no. HM802352 the whole document	
A	CHAKRABARTI MANOHAR ET AL: "Inactivation of the cytochrome P450 gene CYP82E2 by degenerative mutations was a key event in the evolution of the alkaloid profile of modern tobacco", NEW PHYTOLOGIST, CAMBRIDGE UNIVERSITY PRESS, CAMBRIDGE, GB, vol. 175, no. 3, 1 August 2007 (2007-08-01), pages 565-574, XP002524136, ISSN: 0028-646X, DOI: DOI:10.1111/J.1469-8137.2007.02116.X [retrieved on 2007-05-30] the whole document	1-46
A	GAVILANO LILY B ET AL: "Functional analysis of nicotine demethylase genes reveals insights into the evolution of modern tobacco", JOURNAL OF BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY FOR BIOCHEMISTRY AND MOLECULAR BIOLOGY, INC, US, vol. 282, no. 1, 5 January 2007 (2007-01-05), pages 249-256, XP002524135, ISSN: 0021-9258, DOI: DOI:10.1074/JBC.M609512200 [retrieved on 2006-11-13] the whole document	1-46
A	LEWIS RAMSEY S ET AL: "RNA interference (RNAi)-induced suppression of nicotine demethylase activity reduces levels of a key carcinogen in cured tobacco leaves.", PLANT BIOTECHNOLOGY JOURNAL, vol. 6, no. 4, May 2008 (2008-05), pages 346-354, XP002629035, ISSN: 1467-7652 the whole document	1-46

-/--

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2011/021088

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Dr. Nelson: "P450:CYP82", Metabolomics.JP , 20 June 2009 (2009-06-20), XP002629107, Retrieved from the Internet: URL:http://metabolomics.jp/mediawiki/index.php?title=P450:CYP82&printable=yes the whole document -----	1-46

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2011/021088

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2009064771 A2	22-05-2009	CN 101939431 A	05-01-2011
		EP 2220231 A2	25-08-2010
		US 2009205072 A1	13-08-2009