Phytochemistry 71 (2010) 1988–1998

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Three nicotine demethylase genes mediate nornicotine biosynthesis in *Nicotiana tabacum* L.: Functional characterization of the *CYP82E10* gene

Ramsey S. Lewis, Steven W. Bowen, Matthew R. Keogh [1], Ralph E. Dewey *

Department of Crop Science, North Carolina State University, Raleigh, NC 27695, USA

ARTICLE INFO

Article history:
Received 27 July 2010
Received in revised form 17 September 2010
Accepted 21 September 2010
Available online 25 October 2010

Keywords:
Nicotiana tabacum
Solanaceae
Gene function
Alkaloids
Nornicotine
Tobacco specific nitrosamine (TSNA)
Cytochrome P450
CYP82E10

ABSTRACT

In most tobacco (*Nicotiana tabacum* L.) plants, nornicotine is a relatively minor alkaloid, comprising about 2–5% of the total pyridine alkaloid pool in the mature leaf. Changes in gene expression at an unstable locus, however, can give rise to plants that produce high levels of nornicotine, specifically during leaf senescence and curing. Minimizing the nornicotine content in tobacco is highly desirable, because this compound serves as the direct precursor in the synthesis of *N*′-nitrosonornicotine, a potent carcinogen in laboratory animals. Nornicotine is likely produced almost entirely via the *N*-demethylation of nicotine, in a process called nicotine conversion that is catalyzed by the enzyme nicotine *N*-demethylase (NND). Previous studies have identified *CYP82E4* as the specific NND gene responsible for the unstable conversion phenomenon, and *CYP82E5v2* as a putative minor NND gene. Here, by discovery and characterization of *CYP82E10*, a tobacco NND gene, is reported. PCR amplification studies showed that *CYP82E10* originated from the *N. sylvestris* ancestral parent of modern tobacco. Using a chemical mutagenesis strategy, knockout mutations were induced and identified in all three tobacco NND genes. By generating a series of mutant NND genotypes, the relative contribution of each NND gene toward the nornicotine content of the plant was assessed. Plants possessing knockout mutations in all three genes displayed nornicotine phenotypes that were much lower (~0.5% of total alkaloid content) than that found in conventional tobacco cultivars. The introduction of these mutations into commercial breeding lines promises to be a viable strategy for reducing the levels of one of the best characterized animal carcinogens found in tobacco products.

1. Introduction

Nicotine represents the predominant alkaloid found in cultivated tobacco varieties, 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 believed to be almost entirely a demethylated derivative of nicotine, and usually represents less than 5% of the total pyridine alkaloid pool. Through changes in gene expression at an unstable locus, tobacco plants that initially produce very low amounts of nornicotine can give rise to progeny that metabolically convert a large percentage (as much as 98%) of leaf nicotine to nornicotine (Mann et al., 1964; Jack et al., 2007). In these tobacco plants, characterized as 'high nicotine converters', the vast majority of nornicotine production occurs during senescence and curing of mature leaf (Wernsman and Matzinger, 1968). For reasons unknown, the spontaneous increase in gene expression at the unstable locus contributing to nicotine conversion is more likely to occur in burley tobaccos than in flue-cured tobaccos. As a consequence of this increased propensity for conversion, burley tobacco varieties are known for having considerably higher levels of nornicotine than flue-cured cultivars (Jack et al., 2007).

There has recently been great interest in elucidating the biochemical and molecular genetic basis of nornicotine biosynthesis in tobacco. This interest is primarily driven by the fact that nornicotine serves as a precursor to one of the most potent animal carcinogens characterized in tobacco products. During the curing and processing of the tobacco leaf, nornicotine can chemically react with nitrosating agents to form the compound *N*′-nitrosonornicotine (NNN), a tobacco-specific nitrosamine (TSNA) that has been strongly linked to several types of cancer in a large array of animal studies (Hecht and Hoffmann, 1990; Hoffmann et al., 1994; Hecht, 2003). In flue-cured tobaccos, NNN and other TSNAs were found to be predominantly formed through the reaction of alkaloids with

Abbreviations: EMS, ethyl methane sulfonate; EST, expressed sequence tag; NND, nicotine *N*-demethylase; NNN, *N*′-nitrosonornicotine; TSNA, tobacco specific nitrosamine.

* Corresponding author. Tel.: +1 919 515 2705; fax: +1 919 515 7959.
E-mail address: ralph_dewey@ncsu.edu (R.E. Dewey).
[1] Present address: Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC 27599, USA.

0031-9422/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2010.09.011

nitrogen oxide species present in combustion gases formed by direct-fired heating systems found in traditional curing barns (Peele and Gentry, 1999). Retrofitting these curing barns with heat-exchangers greatly minimized 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). In contrast, in air-cured burley and dark tobaccos, TSNA formation proceeds primarily through reaction of tobacco alkaloids with nitrite, a process likely facilitated by leaf-borne microbes (Bush et al., 2001). As a means of mitigating NNN accumulation in burley tobaccos, alkaloid assays are conducted on all plants in the production of foundation seed. In this method, individuals that exhibit greater than 3% nicotine to nornicotine conversion are systematically removed prior to seed set (Jack et al., 2007). Although this screening procedure leads to significant reductions in nornicotine (and NNN) in comparison to tobacco populations that have not been screened, this process can never be perfect, since high nornicotine producing converter plants can spontaneously arise with each generation.

In addition to serving as a precursor for NNN, recent studies suggest that the nornicotine found in tobacco products may have other undesirable health consequences. Dickerson and Janda (2002) 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, potentially altering both efficacy and toxicity of these drugs. Furthermore, studies have been reported linking nornicotine found in tobacco products with age-related macular degeneration, birth defects, and periodontal disease (Brogan et al., 2005; Katz et al., 2005).

Nornicotine is synthesized directly from nicotine through the activity of the enzyme nicotine *N*-demethylase (NND) (Fig. 1). An important breakthrough toward understanding the molecular genetic basis of nornicotine biosynthesis came with the characterization of *CYP82E4*, a gene encoding an enzyme of the cytochrome P450 superfamily of monooxygenases that catalyzes the *N*-demethylation of nicotine (Siminszky et al., 2005; Xu et al., 2007). *CYP82E4* is clearly the NND gene associated with the unstable nicotine conversion locus described above. In conventional, or 'nonconverter' tobacco plants, transcript accumulation of *CYP82E4* is negligible, whereas in 'converter' plants, a high level of transcript

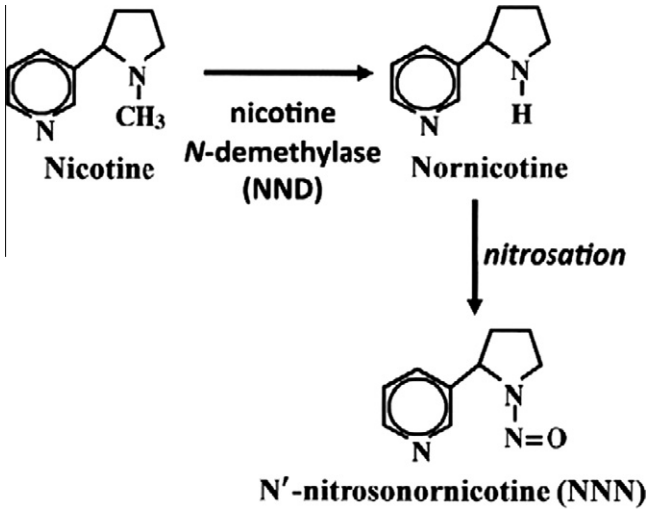

Fig. 1. Chemical structures of nicotine, nornicotine and NNN. Nicotine is enzymatically converted to nornicotine, a compound that can be nitrosated to produce NNN.

accumulation is observed during senescence and air-curing. Furthermore, inactivation of *CYP82E4* through chemical mutagenesis in a strong converter background leads to a reduction in the nicotine to nornicotine conversion rate from an average of 77.5%, to ~3.1%, a level typical of that observed in nonconverter plants (Julio et al., 2008). Analysis of the *CYP82E4* promoter showed that the gene is regulated in a highly senescence specific manner, induced only by abiotic (e.g., air-curing) and biotic (e.g., ethylene and tobacco mosaic virus infection) treatments that result in leaf senescence (Chakrabarti et al., 2008). Interestingly, even though *CYP82E4* maps to the unstable genetic locus responsible for conversion (unpublished results), and the aforementioned *CYP82E4* promoter analysis also supports the premise that *CYP82E4* represents this unstable locus, the molecular mechanism by which *CYP82E4* becomes transcriptionally activated in a converter plant has yet to be elucidated.

CYP82E4 is a member of a small family of genes within the tobacco genome that share over 90% nucleotide and amino acid sequence identity. Despite the high level of shared sequence identity, two other members of this gene family that were initially characterized, *CYP82E2* and *CYP82E3*, were shown to have no NND function (Siminszky et al., 2005). Subsequent investigations of this gene family provided intriguing insights with regard to the molecular evolution of tobacco NND genes. *Nicotiana tabacum* is an allotetraploid species that originated from the interspecific hybridization of species that very closely resemble modern day *N. tomentosiformis* and *N. sylvestris* (Murad et al., 2002; Yukawa et al., 2006). Unlike cultivated tobacco, nornicotine is the predominant alkaloid in both progenitor species. In *N. tomentosiformis*, nicotine to nornicotine conversion occurs constitutively, whereas in *N. sylvestris* nornicotine synthesis is largely confined to senescing leaves (Sisson and Severson, 1990). *CYP82E2* serves as a functional NND gene within the ancestral *N. sylvestris* species, but two amino acid substitution mutations have been introduced that make this gene nonfunctional in *N. tabacum* (Chakrabarti et al., 2007). *CYP82E3* are *CYP82E4* are closely linked, functional NND genes in *N. tomentosiformis*, but a single amino acid substitution in the *CYP82E3* gene and transcriptional inactivation of *CYP82E4* (albeit unstable), have rendered these genes nonfunctional in modern tobacco (Gavilano et al., 2007). In sum, it appears that since speciation of the initial *N. tabacum* allotetraploid(s), mutations have been selected in three different NND genes to transform what would originally have been a nornicotine accumulating species, into one that predominantly accumulates nicotine. Whether the selection and maintenance of this trait was mediated because of use by man, or alternatively, because nicotine may have presented an adaptive advantage over nornicotine in its role as an insecticidal agent, is unknown.

Additional information regarding the role of *CYP82E4* in the production of nornicotine and NNN in tobacco plants was obtained through transgenic plant studies. The expression of an RNAi transgene construct directed against a segment of *CYP82E4* not only eliminated the high nornicotine phenotype of a strong converter line, but also lowered nornicotine levels below that which is observed in nonconverter plants when expressed in either converter or nonconverter plants alike (Gavilano et al., 2006). When evaluated in field studies, the RNAi construct lead to an ~80% decrease in the nornicotine content below that typically observed in nonconverter controls (Lewis et al., 2008). Importantly, a similar decrease in NNN was also observed in the RNAi transgenics versus nontransgenic, nonconverter control lines. These experiments demonstrated the utility of targeting the NND step within the plant as an effective means of lowering the levels of the NNN carcinogen within the cured leaf. The observation that the *CYP82E4* RNAi construct could reduce nornicotine levels below that found in nonconverter plants also opened the possibility that the tobacco genome

may possess another *CYP82E4*-like NND gene(s) that contributes toward the 2–5% nornicotine observed in a typical nonconverter tobacco plant. This supposition was further supported by the characterization of *CYP82E5v2*, a functional NND gene that shares 92.7% DNA sequence identity with *CYP82E4* and is expressed at a low level in the green leaves of both converter and nonconverter plants (Gavilano and Siminszky, 2007).

In this report, we continue our investigation of tobacco NND genes by testing the effects of ethyl methane sulfonate (EMS)-induced knockout mutations in *CYP82E4* and *CYP82E5v2* in altering the nornicotine profile of the plant. These experiments led to the realization that tobacco encodes yet another *CYP82E4*-like NND gene, designated *CYP82E10*. This new NND gene originated from the *N. sylvestris* ancestral parent of tobacco. Combining knockout mutations in all three known NND genes reduced the nornicotine content of the plant to a level virtually equivalent to that observed in our best transgenic RNAi suppression line. The results of this study not only provide a more complete characterization of the *CYP82E* gene family in tobacco, but also demonstrate an effective nontransgenic strategy for lowering the levels of both nornicotine and NNN in tobacco products.

2. Results

2.1. Identification of mutant alleles of CYP82E4 and CYP82E5v2

It was previously shown that transgenic tobacco plants expressing an RNAi construct directed against a 298 bp region of *CYP82E4* displayed greatly reduced levels of nornicotine and NNN in the cured leaf (Lewis et al., 2008). In addition to targeting the specific gene against which the construct was directly made, closely related members of a gene family are also typically silenced using RNAi technologies (Xu et al., 2006). This characteristic of RNAi-mediated gene suppression provided a reasonable explanation of why lower levels of nornicotine were observed in the transgenic tobacco plants expressing the *CYP82E4* RNAi construct (~0.5% nicotine conversion; Lewis et al., 2008) than that which was observed in plants possessing a *CYP82E4* knockout mutation (~3.1% nicotine conversion; Julio et al., 2008). To determine the degree with which the *CYP82E5v2* gene may be responsible for the 2–5% nornicotine content that is commonly found in nonconverter plants, a mutagenesis approach was employed.

To facilitate introduction of random mutations into the tobacco genome, thousands of seeds of the strong converter burley line DH98-325-6 were treated with the chemical mutagen ethyl methane sulfonate (EMS). Approximately 4000 mutagenized plants (M_0 generation) were grown in the field and allowed to self-pollinate. Several seed capsules were collected from each individual plant to create discrete M_1 families, each corresponding to single original M_0 plants. Genomic DNAs were isolated from young leaf tissue of single greenhouse-grown M_1 plants representing each original M_0 individual. To identify plants carrying mutations in the *CYP82E4* gene, specific primers were designed to independently amplify each of the gene's two exons. DNA sequence information was obtained using 96-well high-throughput sequence analysis of the amplification products. Mutations were identified by conducting multi-sequence alignments of all sequences in a given experiment and looking for deviations from the wild type sequence. Any germline mutation that occurred in a parental M_0 plant would be expected to segregate in an 1:2:1 ratio in the M_1 generation for the mutant versus wild type alleles. M_1 plants that were homozygous for an EMS-induced mutation were readily recognized by observed polymorphisms deviating from the wild type sequence (verified by sequence analysis in both directions). M_1 plants that were heterozygous for a mutation in *CYP82E4* would be expected to have the mutant and wild type alleles each representing 50% of the amplification product, a condition that was easily identified through the visual inspection of the corresponding chromatograms.

High-throughput sequence analysis of 672 independent M_1 plants resulted in identification of 11 individuals possessing point mutations within the *CYP82E4* gene (Table 1). All 11 mutations resulted in changes in the predicted amino acid sequence of the encoded protein and are depicted in Fig. 2. Most notable was plant #775 that possessed a mutation at codon 329 which changed a tryptophan codon (TGG) into a premature stop codon (TAG). This mutation would certainly render the encoded protein completely nonfunctional given that the essential heme-binding domain and other highly conserved regions of the enzyme are located downstream of codon 329. Additional plants contained mutations affecting residues within highly conserved motifs shared by the majority of P450 enzymes. Plants #101 and #569 possessed nonconservative substitutions in residues immediately adjacent to the heme-binding Cys amino acid, and plant #164 contained a Pro to Leu substitution within the highly conserved Pro-rich "hinge" region of the protein (Table 1 and Fig. 2).

By germinating seed from each of the M_1 families corresponding to the various mutant lines identified, homozygous mutant

Table 1
EMS treated DH98-325-6 plants possessing mutations in NND genes.

Plant number	Mutation[a]	Amino acid[b]	
CYP82E4 mutants			Enzyme functionality[c]
325-6 #101	C1372T	P458S	Nonfunctional
325-6 #121	G1092T	K364N	Nonfunctional
325-6 #164	C113T	P38L	Wild type
325-6 #321	G601A	E201K	Wild type
325-6 #377	G506A	R169Q	nd[e]
325-6 #569	G1375A	G459R	Nonfunctional
325-6 #506	G886A	E296K	Wild type
325-6 #453	C1280T	T427I	nd
325-6 #775	G986A	W329Stop	Nonfunctional
325-6 #610	G1126A	V376M	~50% of wild type
325-6 #761	G511A	D171N	Wild type
CYP82E5v2 mutants			Enzyme functionality
325-6 #198	C703T	P235S	nd
325-6 #680	C1346T	P449L	nd
325-6 #744	G558A	Silent	nd
325-6 #163	C521T	S174L	nd
325-6 #561	G1458A	Silent	nd
325-6 #154	C1229T	A410V	nd
325-6 #340	G1248A	Silent	nd
325-6 #780	G672A	M224I	nd
325-6 #162	C to T (intron)	Intron	nd
325-6 #799	C215T	P72L	nd
325-6 #540	C427T	L143F	nd
325-6 #1013	G1266A	W422Stop	Nonfunctional
CYP82E10 mutants			Enzyme functionality[d]
325-6 #2476	G235A	G79S	Nonfunctional
325-6 #1512	C319T	P107S	Nonfunctional
325-6 #319	C442T	L148F	nd
325-6 #634	G514A	G172R	nd
325-6 #1460	G524A	S175N	nd
325-6 #1035	G1030A	A344T	Wild type
325-6 #1041	C1141T	P381S	Nonfunctional
325-6 #1251	G1177A	E393K	nd
325-6 #817	G1228A	A410T	Wild type
325-6 #726	G1248A	Silent	nd
325-6 #693	G1250A	R417H	Wild type
325-6 #1442	C1255T	P419S	~25% of wild type
325-6 #1963	G1281A	Silent	nd
325-6 #229	C1317T	Silent	nd
325-6 #1332	C1344T	Silent	nd

[a] Position relative to the start codon of the cDNA sequence.
[b] Position relative to the start Met.
[c] Based on levels of nornicotine observed in plants homozygous for the mutation.
[d] Based on enzyme activity assays conducted in yeast.
[e] Not determined.

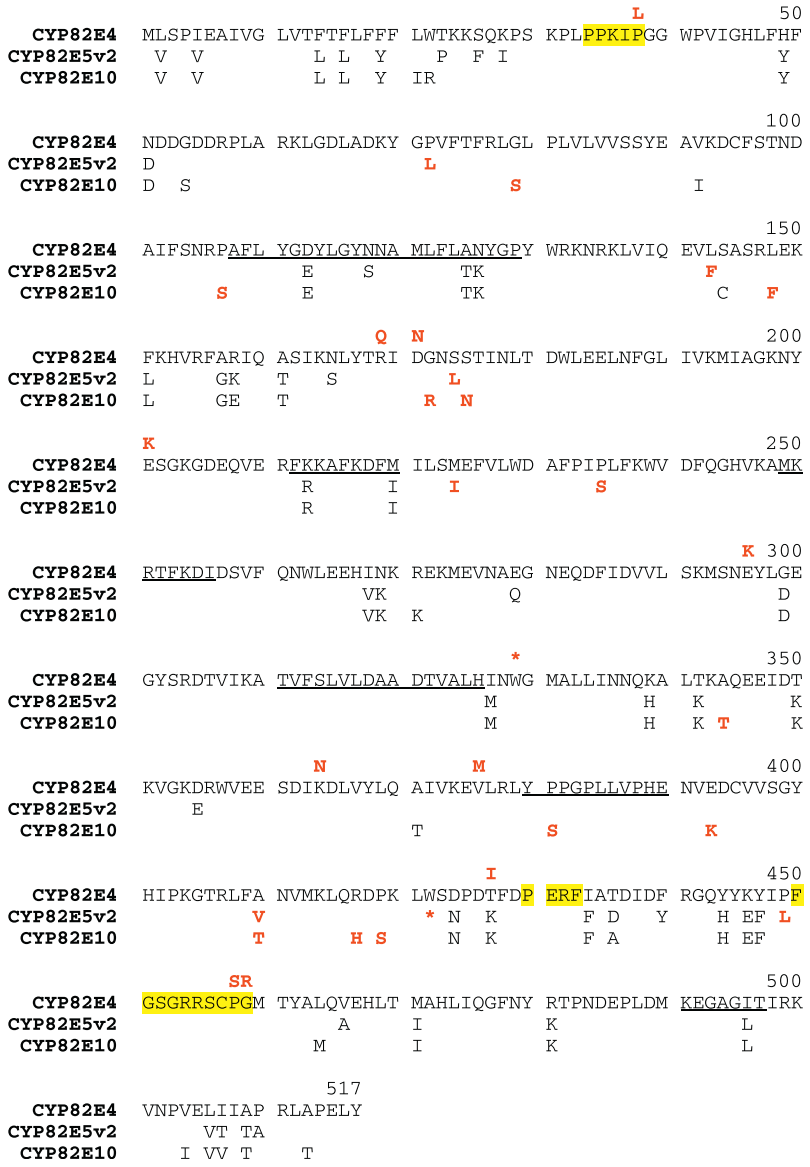

Fig. 2. Alignment of the CYP82E4, CYP82E5v2 and CYP82E10 predicted protein sequences. The complete amino acid sequence for CYP82E4 is shown; for CYP82E5v2 and CYP82E10, the amino acid positions that differ from CYP82E4 are shown in black type. Numbering for the protein sequences are shown on the right. Sequences that are highly conserved in a diversity of plant P450 enzymes are highlighted in yellow. Domains predicted to be involved in substrate recognition by Xu et al. (2007) are underlined. The position and nature of the amino acid substitutions that were discovered in the mutation screen (see Table 1) are shown in red. These substitutions are positioned directly above the amino acid sequence for CYP82E4, and within the peptide sequence regions for CYP82E5v2 and CYP82E10. Red asterisks denote stop codons. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

individuals were recovered for 9 of the 11 observed mutations. An alkaloid analysis was conducted on cured leaves of these individuals to assess the effects of the mutations on the metabolic conversion of nicotine to nornicotine. As expected, plants homozygous for the truncation mutation of line #775 displayed nornicotine levels typical for a nonconverter plant (~3%, data not shown), a result similar to that reported by Julio et al. (2008) for a *CYP82E4* knockout mutation discovered in their mutagenized tobacco population. Likewise, plants homozygous for substitution mutations within the heme-binding region of the enzyme (#101 and #569) gave nonconverter phenotypes, suggesting that these enzymes were nonfunctional as well (Table 1). Interestingly, the mutation within the conserved Pro-rich domain (#164) appeared to have no effect on gene function, and a Lys to Asn substitution (line #121) at a fairly nonconserved region in an area of no known or predicted function appeared to completely disable gene function.

Using the same high-throughput sequencing strategy, 12 individuals possessing mutations in the *CYP82E5v2* gene were identified among a group of 768 M_1 plants screened (Table 1). Four plants contained either silent nucleotide substitutions or a mutation within the intron that did not alter the predicted protein sequence. Of the eight mutations that conferred amino acid changes, the most notable mutation was found in plant #1013. A G to A mutation at position 1266 in plant #1013 changed a Trp codon (TGG) into a premature stop codon (TGA). Although located further downstream than the truncation mutation discovered in *CYP82E4*, the truncated CYP82E5v2 protein that would be produced in plant #1013 would similarly lack the heme-binding domain that is essential for activity for any P450 enzyme (Fig. 2).

2.2. Combining mutant alleles of CYP82E4 and CYP82E5v2

Since EMS mutagenesis was conducted using seed from a strong converter genotype, the only way to accurately determine the contribution of *CYP82E5v2* on plant nornicotine content would be to compare the nornicotine levels of plants possessing a knockout

Table 2
Alkaloid profiles for experimental materials evaluated in 2008 field experiment.

Genotype	Gene targeted	Mutation[b]	Amino acid change	Average				
				% Nicotine[c]	% Nornicotine	% Anabasine	% Anatabine	% Conversion[d]
DH98-325-6 control (15)[a]	Control	-		1.228	2.014	0.016	0.125	62.4
TN 90 LC (14)	Control	-		4.680	0.157	0.022	0.155	3.2
DH98-325-6 RNAi 300-02 #1 (15)	*CYP82E4* and related	-		3.741	0.026	0.017	0.106	0.7
DH98-325-6 #775 (15)	*CYP82E4*	G986A	W329Stop	2.941	0.077	0.013	0.093	2.6
DH98-325-6 #1013 (14)	*CYP82E5v2*	G1266A	W422Stop	1.005	1.876	0.012	0.097	65.2
e4e4/*e5e5* double mutant (9)	*CYP82E4* + *CYP82E5v2*	Double	Double	3.160	0.076	0.015	0.117	2.3

[a] Number in parenthesis 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)] × 100.

mutation in the *CYP82E4* gene with plants containing disabling mutations in both *CYP82E4* and *CYP82E5v2*. To accomplish this, crosses were made that combined the above described truncation mutations in *CYP82E4* (*e4*) and *CYP82E5v2* (*e5*) that originated from lines #775 and #1013, respectively. Molecular genotyping of numerous F_2 individuals derived from the F_1 progeny of an initial #775 × #1013 cross resulted in the identification of nine individuals that were homozygous for both mutations (*e4e4*/*e5e5*).

The effects of the *e4* and *e5* mutations on nornicotine accumulation were evaluated in a field experiment conducted in the summer of 2008. As shown in Table 2, plants homozygous for the #775-derived *e4* mutation averaged 2.6% nicotine to nornicotine conversion. In contrast, >60% conversion was observed in the parental line DH98-325-6, a result typical for a strong converter genotype. These results were very consistent with those reported by Julio et al. (2008), who recorded conversion percentages ranging from 2.82 to 3.37 for plants homozygous for a *CYP82E4* knockout mutant within the strong converter burley genotype BB16NN (parental conversion rates ranged between 68% and 98%). The nicotine conversion percentage in *e4e4* plants was also found not be significantly different ($P = 0.305$) from the 3.2% nicotine conversion observed for plants grown from a seedlot of the commercial variety TN 90 LC (where the LC, or "low converter", designation means that individual plants displaying over 3% nicotine conversion were eliminated during the course of foundation seed production). Thus, debilitating mutation in *CYP82E4* alone appears to be effective in eliminating the problems arising from the unstable genetic phenomenon associated with the generation of converter plants. As expected, the effect, if any, of the #1013-derived *e5* mutation could not be detected when carried within a DH98-325-6 background with an activated *CYP82E4* gene (65.2% conversion).

Despite the fact that *CYP82E5v2* has been shown to encode a functional nicotine demethylase enzyme (Gavilano and Siminszky, 2007), combining the nonfunctional *e5* mutation into plants possessing the knockout *e4* mutation had remarkably little impact on leaf nornicotine levels. Plants homozygous for both mutations (*e4e4*/*e5e5*) averaged 2.3% nicotine conversion, compared with the 2.6% conversion observed for #775 plants homozygous for only the *e4* mutation (Table 2). This modest difference in mean conversion between the two genotypes was not statistically significant ($P = 0.118$). In contrast, a *CYP82E4* RNAi-silenced transgenic line 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. The fact that nornicotine levels can be reduced to a lower level in the RNAi suppression line than in *e4e4*/*e5e5* individuals suggested that yet another gene with high DNA sequence homology to *CYP82E4* must exist within the tobacco genome that contributes toward nornicotine production in the plant.

2.3. Identification of CYP82E10

To identify other genes within the tobacco genome that have the potential of encoding NND enzymes, homology searches using the BLASTN and BLASTX algorithms (Altschul et al., 1990, 1997) were conducted on the *N. tabacum* expressed sequenced tagged (EST) database in GenBank, using the DNA and predicted protein sequences of *CYP82E4* as the respective query sequences. In addition to identifying cDNA sequences corresponding to previously characterized members of the *CYP82E* subfamily (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 (data not shown). This observation suggested that the new *CYP82E* gene may be expressed preferentially in root tissue, a property that could explain why this particular member of the *CYP82E* P450 subfamily had eluded detection previously, since prior efforts focused on the characterization of *CYP82E* genes expressed in leaf tissue (Siminszky et al., 2005; Gavilano and Siminszky, 2007; Xu et al., 2007). Although no individual EST sequence was long enough to cover the entire coding region of this novel gene, a full-length cDNA could be constructed by assembling independent EST sequences that spanned different regions of the gene. The veracity of the predicted sequence was also confirmed by independently amplifying full-length cDNAs from first-strand tobacco root cDNA using PCR primers directed against the ends of the predicted gene (data not shown). In addition, PCR was used to amplify, then sequence the large centrally located intron from a tobacco genomic DNA preparation. This new *CYP82E* gene shares 92.4% nucleotide identity and 91.1% predicted amino acid identity with the tobacco *CYP82E4* cDNA and enzyme, respectively. In keeping with the guidelines for P450 gene nomenclature (Nelson et al., 1996), the new gene was designated *CYP82E10*. Of all the characterized members of the *CYP82E* subfamily, *CYP82E10* displays the highest sequence homology with *CYP82E5v2*, sharing 96.5% identity at the nucleotide level and 95.7% predicted amino acid sequence identity. The DNA sequence of *CYP82E10* has been deposited in GenBank (accession #HM802352) and its predicted protein sequence is shown in Fig. 2.

2.4. CYP82E10 encodes a functional NND enzyme

As detailed in previous publications, sequence homology alone is not a very accurate indicator of gene function for the *CYP82E*

subfamily. *CYP82E2* and *CYP82E3* function as NND enzymes within the respective *N. sylvestris* and *N. tomentosiformis* progenitor species of *N. tabacum*, but not within the tobacco genome (Chakrabarti et al., 2007; Gavilano et al., 2007). Furthermore, of four *CYP82E* member cDNAs analyzed by Xu et al. (2007), only the cDNA that corresponded to *CYP82E4* encoded an enzyme that displayed NND activity. Therefore, in order to establish whether an individual member of this gene family encodes an enzyme capable of demethylating nicotine, functional analysis in either transgenic plants (Siminszky et al., 2005) or yeast (Gavilano and Siminszky, 2007; Xu et al., 2007) is required.

To establish whether *CYP82E10* represents a viable NND 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 when expressed in yeast (Pompon et al., 1996). NND assays were conducted by incubating yeast microsomal membrane preparations with [^{14}C]-nicotine, and resolving the products by thin layer chromatography.

As shown in Fig. 3, no NND activity could be detected using yeast microsomes from the W(R) strain expressing only the pYeDP60 vector. In contrast, very robust NND activity could be measured using 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 calculated an apparent K_m of 3.9 μM nicotine using the microsomal assay (data not shown). This kinetic parameter for CYP82E10 is very similar to the K_ms reported for the CYP82E4 and CYP82E5v2 enzymes when expressed in yeast (Gavilano et al., 2007; Gavilano and Siminszky, 2007; Xu et al., 2007).

2.5. CYP82E10 originated from the N. sylvestris ancestral parent of tobacco

Of the *CYP82E* family members characterized to date, *CYP82E3*, *CYP82E4* and *CYP82E5v2* were shown to be derived from the *N. tomentosiformis* ancestral parent (Gavilano et al., 2007; Gavilano

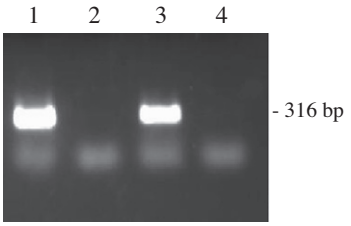

Fig. 4. Electrophoretic separation of PCR amplification products using *CYP82E10*-specific primers with genomic DNAs isolated from *N. sylvestris* (lane 1), *N. tomentosiformis* (lane 2) and *N. tabacum* (lane 3). Lane 4 is a negative control with no added template DNA.

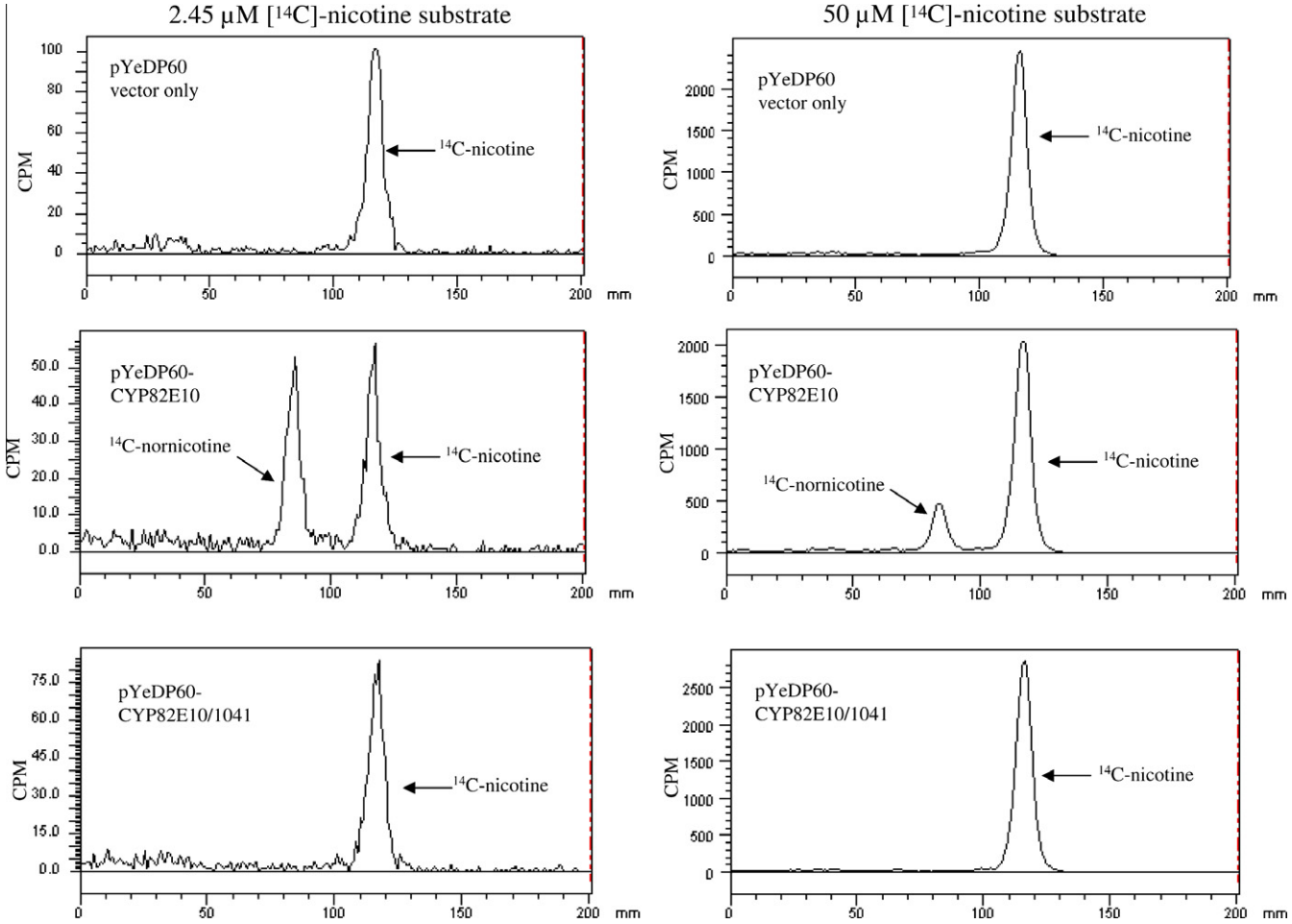

Fig. 3. Thin layer chromatographic data of NND activities of microsomal membranes from yeast cells expressing CYP82E10, the CYP82E10 construct possessing the Ser381Pro mutation from plant #1041, and vector controls. *X*-axis depicts the migration in mm of the radiolabeled compounds on the plate. CPM, counts per minute.

and Siminszky, 2007), and *CYP82E2* was shown to originate from *N. sylvestris* (Chakrabarti et al., 2007). To establish the ancestral origin of *CYP82E10*, gene-specific primers were designed that amplified a 316 bp fragment from exon 2. As shown in Fig. 4, the *CYP82E10*-specific primers amplified a fragment of the expected size when genomic DNAs from *N. tabacum* and *N. sylvestris* were used as templates, but no product could be detected using genomic DNA from *N. tomentosiformis*. DNA sequence analysis established that the fragment amplified from *N. sylvestris* was 100% identical with the same region of the *N. tabacum* *CYP82E10* sequence (data not shown). To further characterize the *CYP82E10* gene from *N. sylvestris*, specific PCR primers where used to amplify a full-length genomic clone. DNA sequence analysis of the *N. sylvestris* *CYP82E10* gene demonstrated 100% sequence conservation across all of exon 2. Three single nucleotide polymorphisms were observed in exon 1, only one of which altered the amino acid sequence (a Thr to Ala substitution at amino acid position 15; data not shown), and 10 polymorphisms (including small insertion/deletion events) were observed within the intron region (data not shown). The DNA and predicted protein sequences of the *N. sylvestris* *CYP82E10* gene have been deposited in GenBank (accession number HQ259993).

2.6. Identification of plants possessing mutant alleles of CYP82E10

In order to assess the specific contribution of *CYP82E10* toward the total nornicotine content of the tobacco plant, a tobacco plant with a knockout mutation within this gene needed to be identified. To identify potentially debilitating mutations in *CYP82E10*, the EMS-mutagenized DH98-325-6 population was again screened by high-throughput DNA sequence analysis using primers that specifically amplified the two *CYP82E10* exons. Sequence analysis of over 1200 M_1 individuals from the mutagenized population resulted in the identification of 15 individuals with mutations in *CYP82E10*, 11 of which caused amino acids substitutions (Table 1). The amino acid residues mutated in these plants are highlighted in Fig. 2. 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 modification corresponding to seven of the 11 mutations shown in Table 1 into the *CYP82E10* cDNA that had been cloned into the pYeDP60 yeast expression vector. Microsomal preparations from yeast strains expressing each of the seven CYP82E10 variants were assayed *in vitro* for NND activity using both non-saturating (2.45 μM) and saturating (50 μM) concentrations of [^{14}C]-nicotine. These experiments showed that mutations found in plants #693, #817 and #1035 did not alter enzyme activity, whereas the mutations found in plants #1041, #1512 and #2476 led 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 Fig. 3. This particular mutation lies within a region predicted by Xu et al. (2007) to be involved in substrate binding (Fig. 2) and was selected for more extensive investigation. To provide additional confirmation that the Pro to Ser substitution at amino acid position 381 that defines the plant #1041 mutation is incompatible with NND function, this same mutation into a *CYP82E4* cDNA that had been similarly cloned into the pYeDP60 vector (Siminszky et al., 2005) was introduced. The results of these assays are displayed in Table 3. Whether introduced into the CYP82E10 or CYP82E4 enzyme, a Ser substitution for Pro at position 381 leads to the complete ablation of NND activity in this assay. Interestingly, although the activities of the wild type CYP82E10 and CYP82E4 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 CYP82E4.

Table 3
NND activity of CYP82E4 and CYP82E10 enzymes possessing the #1041 mutation.

Vector	Bq nornicotine at 2.45 μM [^{14}C]-nicotine substrate[a]	Bq nornicotine at 25 μM [^{14}C]-nicotine substrate
pYeDP60-CYP82E4	408 ± 140[b]	1212 ± 114
pYeDP60-CYP82E4/1041	Not detected	Not detected
pYeDP60-CYP82E10	517 ± 22	3435 ± 105
pYeDP60-CYP82E10/1041	Not detected	Not detected

[a] Bq of [^{14}C]-nornicotine/mg microsomal protein.
[b] Standard deviation of two technical replications.

NND activities in yeast cells expressing wild type CYP82E10 and the #1041 mutant enzyme were 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 *in vitro* and *in vivo* yeast expression assays strongly suggest that CYP82E10 enzyme function is completely abolished by the introduction of the Ser381-Pro mutation discovered in plant #1041.

2.7. Combining mutant alleles of CYP82E10, CYP82E4 and CYP82E5v2

Given that the original #1041 mutation is in a genetic background that contains both a strong converter *CYP82E4* allele as well as a wild type *CYP82E5v2* gene, the only way to accurately assess the independent contribution of *CYP82E10* toward total plant nornicotine content would be to introduce the #1041 mutation into tobacco plants possessing knockout *CYP82E4* 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 (*e4E4*/*e5E5*). F_1 plants heterozygous for all three NND 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 above plants were grown in the field during the summer of 2009. Also included in this study were two of the genotypes tested in the 2008 field trial. Specifically, 10 DH98-325-6 plants homozygous for only the #775 mutation (*e4e4*/*E5E5*/*E10E10*) and 11 DH98-325-6 plants homozygous for both the #775 and #1013 alleles (*e4e4*/*e5e5*/*E10E10*) were included for comparison. As controls, individual plants randomly selected from a commercial seedlot of TN 90 LC, wild type DH98-325-6 individuals, and plants from our best *CYP82E4* RNAi-suppressed transgenic line were also included. After the plants were about an average of 30 cm tall (35 days after transplanting), leaves from similar stalk positions were collected, treated with ethephon, air-cured and analyzed for alkaloid content.

Table 4 and Fig. 5 show the results of the alkaloid analyses for the 2009 field trial. Consistent with previous observations, the *e4* knockout mutation alone negates the strong converter phenotype of line DH98-325-6. In addition, the homozygous mutant *e4* plants also displayed a numerically lower nornicotine accumulation

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

Genotype	Gene(s) targeted	Mutation[b]	Amino acid change	Average				
				% Nicotine[c]	% Nornicotine	% Anabasine	% Anatabine	% Conversion[d]
DH98-325-6 control (8)[a]	Control	–	–	0.133	1.553	0.009	0.085	92.21
TN 90 LC (11)	Control	–	–	1.519	0.104	0.002	0.065	7.15
DH98-325-6 RNAi 300-02 #1 (10)	*CYP82E4* and related	–	–	1.747	0.009	0.003	0.063	0.54
DH98-325-6 #775 (10)	*CYP82E4*	G986A	W329Stop	1.375	0.030	0.002	0.057	2.20
e4e4/e5e5/E10E10 (11)	*CYP82E4 + CYP82E5v2*	Double	Double	1.524	0.036	0.003	0.084	2.34
DH98-325-6 #1041 Homo. (3)	*CYP82E10*	C1141T	P381S	0.082	1.302	0.007	0.073	93.87
E4E4/e5e5/e10e10 (5)	*CYP82E5v2 + CYP82E10*	Double	Double	0.081	1.345	0.010	0.068	94.31
e4e4/E5E5/e10/e10 (4)	*CYP82E4 + CYP82E10*	Double	Double	2.168	0.045	0.004	0.087	2.02
e4e4/e5e5/e10/e10 (5)	*CYP82E4 + CYP82E5v2 + CYP82E10*	Triple	Triple	1.793	0.012	0.003	0.056	0.55

[a] Number in parenthesis 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)] × 100.

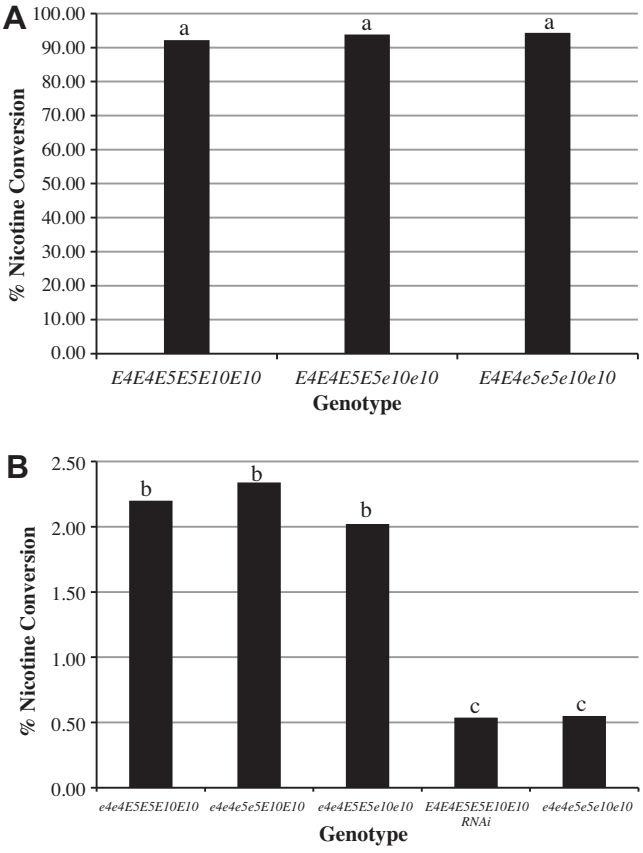

Fig. 5. Mean percent nicotine to nornicotine conversion for burley tobacco plants with different combinations of mutant alleles at the *CYP82E4*, *CYP82E5v2* and *CYP82E10* loci from 2009 field experiment. Genotypes are grouped according to whether they display a high (A) or low (B) level of nicotine conversion. Means with different letters are significantly different at the $P < 0.05$ level.

phenotype compared to plants from the commercial TN 90 LC seedlot (2.2% versus 7.15% conversion, respectively). Though not statistically significant ($P = 0.1005$) the mean differences in nicotine conversion observed in the TN 90 LC plants grown in 2008 versus 2009 underscores the difficulties in attempting to control the conversion problem in burley tobaccos solely through the converter screening procedure. The higher mean conversion percentage recorded for TN 90 LC in the 2009 study was largely driven by two plants that displayed >20% nicotine conversion (data not shown). As observed in the 2008 field trial (Table 2), combining the *e5* mutation with *e4* did not lead to further reductions in nornicotine content. As expected, the *e10* mutation had no impact on the high nornicotine levels conferred by an active *CYP82E4* gene, either alone (*E4E4/E5E5/e10e10* genotypes), or when combined with a mutant *e5* allele (*E4E4/e5e5/e10e10* genotypes) (Fig. 5A). Similar to the *e4e4/e5e5* double mutant results, introducing the *e10* mutation into an *e4e4* background was not effective in reducing nornicotine levels below than that which could be achieved by the *e4* mutation alone (Fig. 5B). The *e4e4/E5E5/e10e10* genotypes averaged 2.02% conversion, a percentage that was not significantly different than the 2.2% mean conversion level observed for *e4e4/E5E5/E10E10* individuals ($P = 0.557$).

Although the *e5* and *e10* mutations did not serve to significantly decrease the nornicotine content of *e4* plants when combined individually, pyramiding all three NND 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. 5B). This represents over a threefold reduction in nicotine conversion beyond that which was mediated by the *e4* 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 were highly significant ($P < 0.0001$). Similar to our investigation of RNAi-mediated suppression of nicotine conversion (Lewis et al., 2008), the 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 (Table 4).

3. Discussion

The tobacco genome contains several genes that display greater than 90% sequence identity to *CYP82E4*, yet the majority of these do not function as NND genes. Although *CYP82E2* and *CYP82E3* encode functional NND enzymes within the *N. sylvestris* and *N. tomentosiformis* progenitor species of tobacco, respectively, mutations have arisen to make these genes nonfunctional within *N. tabacum* (Gavilano et al., 2007; Chakrabarti et al., 2007). Furthermore, database searches reveal that numerous additional *CYP82E4*-like sequences are present within the tobacco genome, but most of these appear to represent pseudogenes, as short deletions are observed that would result in frame-shift mutations (unpublished results). Prior to this study, *CYP82E5v2* was the only

gene other than *CYP82E4* that had been shown to encode a functional NND enzyme within *N. tabacum* genome (Gavilano and Siminszky, 2007). The results presented here define *CYP82E10* as another viable NND gene.

Compared with the very high levels of nornicotine mediated by an activated *CYP82E4* gene in converter tobacco plants, the combined effects of the *CYP82E10* and *CYP82E5v2* genes result in only a small percentage of nicotine conversion (Table 4 and Fig. 5). Thus, *CYP8210* and *CYP82E5v2* could be considered as "minor" NND genes and *CYP82E4* as the "major" NND gene of tobacco. Given the seemingly simple correlation between NND gene expression and nornicotine accumulation in tobacco plants (Siminszky et al., 2005; Gavilano et al., 2007; Xu et al., 2007), the observation that the effects of *CYP82E5v2* and *CYP82E10* gene activity appeared to be dominant, rather than additive, was unexpected. Only the triple mutant combination was effective in lowering nornicotine levels below that which could be achieved by the knockout *e4* mutation alone (Table 4, Fig. 5), suggesting that NND activity derived from *CYP82E5v2* or *CYP82E10* is not rate limiting under the conditions controlling the synthesis of the 2–5% nornicotine routinely found in nonconverter plants, or converter plants carrying a knockout *e4* allele. Except for the period of translocation from its site of synthesis in the root to aerial parts of the plant via xylem tissue (Baldwin, 1989; Hashimoto and Yamada, 2003), nicotine is primarily sequestered within the central vacuole of the cell, a process believed to be facilitated by the recently characterized nicotine transporters of the MATE family (Morita et al., 2009; Shoji et al., 2009). It therefore could be proposed that in healthy, nonsenescent cells, very little nicotine is actually present in the cytosol where it could be exposed to demethylation by the endoplasmic reticulum-localized NND P450 enzymes, and that under the relatively rare set of circumstances where nicotine molecules are found unprotected in the cytosol, CYP82E5v2 or CYP82E10 enzyme levels alone are sufficient in converting these molecules to nornicotine. Nevertheless, such a model is inconsistent with our previous transgenic plant results, whereby simply overexpressing *CYP82E4* using a constitutive promoter was sufficient to drive nicotine conversion to over 90% in young, nonsenescent tobacco leaves (Siminszky et al., 2005). Further investigation will be needed to resolve the apparent dichotomy between the nonadditive nature of *CYP82E5v2* and *CYP82E10*, and the presumed simple correlation between NND activity and plant nornicotine content.

Another interesting observation that arose from this study involves the fact that a small, but very measurable amount of nornicotine can still be found in both *e4e4*/*e5e5*/*e10e10* triple mutant plants and our most effective *CYP82E4* RNAi transgenic event. In both cases, nornicotine in the NND compromised plants represents about 0.5% of the total pyridine alkaloid pool (Table 4). One possibility is that the RNAi-induced silencing mechanism is not 100% effective in the transgenic line, and given the absence of complete tobacco genome information, the possibility of the existence of yet another minor *CYP82E* member NND gene in the triple mutant plants cannot be ruled out. Alternatively, a minor amount of nicotine *N*-demethylation could be catalyzed by a different oxidative enzyme(s) that shares no structural similarity with the CYP82E type NNDs. A third option that warrants consideration involves the possibility that a small amount of nornicotine may be synthesized directly through a pathway that does not include nicotine as an intermediate. The plausibility of such a pathway is supported by the recent characterization of the tobacco methylputrescine oxidase 1 (*MPO1*) gene. MPO enzymes catalyze the oxidative deamination of *N*-methylputrescine to produce *N*-methylaminobutanal, a compound that spontaneously cyclizes to provide the *N*-methylpyrrolinium molecule that forms the pyrrolidine ring of nicotine (Bush et al., 1999). In their characterizations of the tobacco *MPO1* gene product, both Heim et al. (2007) and Katoh et al. (2007) demonstrated that in addition to its preferred *N*-methylputrescine substrate, recombinant MPO1 enzyme preparations could to a lesser degree utilize putrescine to produce an unmethylated pyrrolinium salt. Should the terminal nicotine synthase reaction (an enzyme that has yet to be definitively purified or characterized) show some activity toward a cyclic unmethylated pyrrolinium substrate, then nornicotine could be directly produced, bypassing nicotine altogether. Should this latter scenario prove to be correct, then it is likely that we have reached the limit by which tobacco nornicotine levels can be reduced through the targeting of NND genes.

Through our discovery and characterization of *CYP82E10*, a strategy was developed for reducing the nornicotine content of tobacco plants to levels that have previously only been possible using transgenic approaches. Using an RNAi construct directed against *CYP82E4*, nicotine conversion levels could be lowered from ~2.8%, which is the best that could be achieved after effort had been expended to assure that these plants were all nonconverters, to ~0.5% in the cured leaf. Importantly, the reduction in nornicotine led to a commensurate decrease in NNN formation (Lewis et al., 2008). Although the utilization of varieties possessing this RNAi element would be an effective means of lowering tobacco nornicotine levels, the commercial deployment of any transgenic crop is a daunting task. Among the issues involved are: (1) procurement of (and payment for) licenses for all of the enabling technologies involved with the generation of transgenic plants; (2) long and costly process of getting regulatory approval to allow release of the specific transgene event into grower's fields; and (3) potential rejection of the product by a subset of end-users who may be philosophically opposed to genetically modified organisms. The *e4e4*/*e5e5*/*e10e10* triple mutant materials generated in the current study offer the potential of achieving the same degree of nornicotine minimization as the transgenics without having to confront the obstacles inherent to transgenic crops. Although several generations of backcrossing will be required to eliminate undesirable secondary mutations that are expected to be scattered throughout the genome in any given mutant line, the rapid introgression of the mutant *CYP82E4* genes into elite backgrounds can be greatly facilitated with the assistance of molecular markers specific for each mutation. Tobacco products developed from varieties possessing these NND mutations should contain substantially reduced levels of both nornicotine and NNN, one of the most thoroughly documented animal carcinogens found in these products.

4. Conclusion

NND enzymes catalyze the metabolic conversion of nicotine to nornicotine. Reducing nornicotine levels in tobacco products is a high priority in efforts aimed at reducing the carcinogenic potential of these products as it serves as the precursor in the formation of NNN, a strong animal carcinogen. The discovery of *CYP82E10* as an additional functional NND encoding gene within the tobacco genome has enabled us to develop an effective strategy for minimizing the synthesis of nornicotine, and thereby NNN, within tobacco plants.

5. Experimental

5.1. Development of an EMS mutant population

Seeds of the strong converter burley tobacco line DH98-325-6 were surface-sterilized in 50% bleach for 12 min and rinsed six times in sterile dH_2O. Seeds (80 mg) were placed in a screw cap vial, and imbibed in dH_2O (1 ml) for 12 h. The dH_2O was decanted, and 0.5% EMS (1 ml) (Sigma–Aldrich, St. Louis, MO) was added to

the vial, followed by gentle rotation in the dark for 12 h. The treated seeds were washed with dH_2O (8×) for 5 min each with shaking to remove residual EMS. Washed seeds were placed on filter paper and collected onto a Buchner funnel after a final rinse. After air drying, M_0 seeds were germinated in a greenhouse and subsequently transplanted to the field. Seed capsules were collected from ~4000 field grown M_0 plants.

5.2. Screening for NND mutations

Using screw cap vials (1.5 ml), single leaf discs were collected from ~3000 independent, greenhouse grown M_1 plants. After forcing the leaf material to the bottom of the vials through brief centrifugation, the vials were placed in a shallow amount of liq. N_2 to facilitate tissue grinding using plastic pestles. Genomic DNA was isolated from ground material by adding extraction buffer (325 μl) (200 mM Tris–Cl, pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% SDS), vortexing for 20 s, and incubating tubes at 65 °C for 10 min. Protein precipitation solution (60 μl) (Qiagen, Valencia, CA) was added to the lysed extracts, the samples were vortexed for 20 s and set on ice for 5 min. The extracts were centrifuged at 16,000*g* for 4.5 min and supernatants were transferred to fresh tubes. Nucleic acids were precipitated by adding isopropanol (250 μl), incubating at room temperature for 10 min, followed by centrifugation at 16,000*g* for 10 min. After an ethanol–H_2O (70:30, v/v) wash, pellets were air dried and resuspended in 50 μl TE.

The three tobacco NND genes share a common structure: a 939 bp exon 1 and a 612 bp exon 2 separated by a large intron, whose length varies among the three genes. Achieving PCR amplification specificity for these genes can be problematic when using primers corresponding to the coding regions because of the high sequence identity (>90%) shared among the cDNA sequences of the several members of the *CYP82E* subfamily within the tobacco genome. Unlike the coding region, however, it is not difficult to find sequences within the respective introns that are very specific for each NND gene. To screen for mutations in the individual genes, separate amplification reactions were conducted for exon 1 and exon 2, with specificity assured in each case by having one of the primer pairs targeting a sequence within the intron. The respective forward and reverse primers used for screening the mutant DH98-325-6 population were as follows: *CYP82E4*, 5′-TGGAATTATGC CCATCCTACA-3′ and 5′-CATTAGTGGTTGCACCTGAGG-3′ (exon 1), and 5′-GATGAGATGTGTGCATACTTG-3′ and 5′-CCAAATTAGAAAAA CTCGTACTG-3′ (exon 2); *CYP82E5v2*, 5′-ATTGTAGGACTAGTAACCC TTACAC-3′ and 5′-GAGGCACAAAGAATTCTCATC-3′ (exon 1), and 5′-GAGTAGAGGGATTGTTTCCG-3′ and 5′GTACAATCAAGATAAAAC ATCTAAGG-3′ (exon 2); *CYP82E10*, 5′-GTGATAGTTTGATTCCCAA GTGC-3′ and 5′-CTCCCAAAGTTAGATTAGTCCG-3′ (exon 1), and 5′-A GGTCGCGCTGATTCTTG-3′ and 5′-AGATGAATACCCATCTATCTAGG AGT-3′ (exon 2). PCR was conducted in a 96-well format with each reaction containing template (1 μl), each primer (10 pmol), 200 mM dNTP, 1.5 mM MgCl, and 1.4 units high fidelity Taq DNA polymerase (Roche) in a final volume of 25 μl. DNA amplification for *CYP82E4* was performed using an initial 3 min denaturation at 94 °C, followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1.5 min, followed by a final 7 min 72 °C extension. PCR conditions for *CYP82E5v2* and *CYP82E10* were the same except the annealing temperatures used were 53 °C and 54 °C, respectively. The PCR reactions were cleaned up using Montage PCRm96 filter units (Millipore, Billerica, MA), and subsequently analyzed by cycle sequencing using Big Dye Version 3.1 (at 0.175× concentration) according to the manufacturer's protocol (Applied Biosystems, Foster City, CA). The primers used in the sequencing reactions were the same as the forward and reverse amplification primers specified above. Sequence reactions were cleaned up using Performa DTR cartridges (EdgeBioSystems, Gaithersburg, MD), and sequencing was conducted using a high-throughput Applied Biosystems 3730 capillary sequencer. Sequences were aligned using the Clustal W algorithm from the Vector NTI software package (Invitrogen, Carlsbad, CA).

5.3. Yeast assays

To facilitate the expression of *CYP82E10* in yeast, the full-length cDNA sequence was amplified from tobacco root cDNA using primers that introduced 5′ restriction sites to assist cloning into the pYeDP60 yeast expression vector (Pompon et al., 1996). The specific PCR primers used were 5′-AAGCTTGGATCCATGGTTTCTCC CGTAGAAGCCATC-3′ (forward) and 5′-AAGCTTGAATTCTTAATAAA GCTCAGGTGTCAGGCGAGGCGTAATTAC-3′ (reverse). The QuikChange Lightning Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) was used to introduce the mutations discovered in *CYP82E10* from the EMS mutagenesized population (Table 1) into the pYeDP60-CYP82E10 vector (and the pYeDP60-CYP82E4 vector for the #1041 mutation) as outlined by the manufacturer. Yeast strain W(R) was transformed with the yeast expression vectors according to Geitz et al. (1992) and microsomal extracts prepared as described by Pompon et al. (1996). NND activity was assayed in a reaction mix (150 μl) containing 500 μg microsomal protein, 0.75 mM NADPH, 50 mM phosphate buffer (pH 7.2) and differing amounts of [pyrrolidine-2-^{14}C] nicotine (Moravek Biochemicals, Brea, CA) as indicated in Section 2.5. Samples were incubated at room temperature for 45 min and the reactions were stopped by adding acetone (50 μl). Reactions products were resolved by thin layer chromatography and quantified using a Bioscan System 400 imaging scanner (Washington, DC) as previously described (Siminszky et al., 2005; Gavilano and Siminszky, 2007).

5.4. Amplification of a genomic DNA of CYP82E10 from ancestral progenitors

Total genomic DNA was isolated from young leaf material of *N. tomentosiformis, N. sylvestris* and *N. tabacum* (DH98-325-6) by following the protocol provided with the FastDNA Kit (MP Biomedical, Solon, OH). The following primers were used that we had previously established as mediating specific amplification of a 316 bp fragment of *CYP82E10* from *N. tabacum*: 5′-GTTAAA GAAGTGTTACGATTATATC-3′ (forward) and 5′-AACCCTGGATCAA GTGTGCG-3′ (reverse). PCR reactions contained genomic DNA (100 ng), each primer (20 pmol), 200 mM dNTP, 1.5 mM MgCl, and 2.5 U of Taq DNA polymerase (New England Biolabs) in a final volume of 50 μl. DNA amplification was performed using an initial denaturing step of 94 °C for 5 min, followed by 35 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min, then a final 7 min 72 °C extension. Amplification products were resolved on a 1% agarose gel and visualized by ethidium bromide staining. Full-length genomic sequence information from *CYP82E10* from *N. sylvestris* was obtained by amplifying and sequencing two overlapping PCR products. The 5′ half of the gene was amplified using 5′-GTGA-TAGTTTGATTCCCAGTGC-3′ (forward) and 5′-CAAGAATCAGCGC-GACCT-3′ (reverse); the 3′ half of the gene was amplified using 5′-CGGACTAATCTAACTTTGGGAG-3′ (forward) and 5′-AGATGAA-TACCCATCTATCTAGGAGT-3′ (reverse).

5.5. Alkaloid analysis

For the 2008 field study, plants were grown according to production practices used for burley tobacco in North Carolina. At maturity, plants were stalk cut and air cured in a protected shelter. After curing, alkaloid analyses were performed on ground leaf samples generated from the fourth leaf from the top of each stalk. Plants from the 2009 field trial were assayed by taking leaves from

mid-stalk positions of plants 35 days post-transplanting, dipping them in a solution of 2% ethephon, and allowing them to senesce (turn yellow) in the laboratory for approximately 14 days. Leaf samples were oven-dried and ground to pass through a 40-mesh screen. Alkaloid contents were determined using an Autosystem XL Gas Chromatograph with Prevent™ (Perkin–Elmer, Shelton, CT) as detailed by Jack et al. (2007). Individual alkaloid contents (nicotine, nornicotine, anatabine, and anabasine) were reported as percentages on a dry-tobacco-weight basis. Percent nicotine conversion was calculated as [% nornicotine/(% nornicotine + % nicotine)] × 100. Statistical comparisons between genotypic group means were conducted using *t*-tests executed by PROC TTEST of SAS 9.1 (SAS Institute, Cary, NC).

5.6. Molecular genotyping

Molecular assays were required to identify members of the mutant NND genotypic groups tested in the 2008 and 2009 field studies (Tables 2 and 4). Genomic DNAs (isolated using the method described in Section 5.2) from plants segregating for the various NND mutations were subjected to high-throughput DNA sequence analysis to reveal the NND genotype for each plant. To identity plants possessing the *e4* mutation derived from line #775, the following PCR primers were used to specifically amplify exon 2 of the *CYP82E4* gene: 5′-GGTCTAAAACGTGTGTTTGCTTT-3′ (forward) and 5′-CTG GAGATTGTGGTACAACCAG-3′ (reverse). Sequencing of the amplification products was determined using 5′-GATGAGATGTGTGCA TACTTG-3′ as the sequencing primer. Molecular genotyping for the *e5* and *e10* mutations originating from plants #1013 and #1041, respectively, was conducting using the exon 2-specific primer pairs for *CYP82E5v2* and *CYP82E10* that were described in Section 5.2. The sequencing primer for the *CYP82E5v2*-specific product was the same as the reverse primer used for amplification. For the *CYP82E10*-specific PCR product, an interior primer (5′-GTTTAGTCTTGGA TGCTGCG-3′) closer to the site of the diagnostic mutation was used as the sequencing primer. All PCR and sequencing reactions were conducted in 96-well format using the same reaction conditions and protocols described in Section 5.2.

Acknowledgements

We are very grateful to Lowell Bush (University of Kentucky) for conducting the alkaloid analyses. We also thank Carol Griffin for excellent technical assistance. We gratefully acknowledge Denis Pompon and Philippe Urban for the pYeDP60 yeast expression vector and the W(R) yeast strain. This work was supported in part by Altria Client Services and Philip Morris International.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., 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., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acids Res. 25, 3389–3402.

Baldwin, I.T., 1989. Mechanism of damage-induced alkaloid production in wild tobacco. J. Chem. Ecol. 15, 1661–1680.

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

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

Bush, L.P., Hempfling, W.P., Burton, H., 1999. Biosynthesis of nicotine and related compounds. In: Gorrod, J.W., Jacob, P. (Eds.), Analytical Determination of Nicotine and Related Compounds and Their Metabolites. Elsevier, New York, pp. 13–44.

Bush, L.P., Cui, M., Shi, H., Burton, H.R., Fannin, F.F., Lei, L., 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., 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.

Chakrabarti, M., Bowen, S.W., Coleman, N.P., Meekins, K.M., Dewey, R.E., Siminszky, B., 2008. *CYP82E4*-mediated nicotine to nornicotine conversion in tobacco is regulated by a senescence-specific signaling pathway. Plant Mol. Biol. 66, 415–427.

Dickerson, T.J., 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., 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.

Gavilano, L.B., Coleman, N.P., Burnley, L.-E., Bowman, M.L., Kalengamaliro, N.E., Hayes, A., Bush, L., Siminszky, B., 2006. Genetic engineering of *Nicotiana tabacum* for reduced nornicotine content. J. Agric. Food. Chem. 54, 9071–9078.

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

Geitz, D., Stjean, A., Woods, R.A., Schiestl, R.H., 1992. Improved method for high-efficiency transformation of intact yeast cells. Nucl. Acids Res. 20, 1425.

Hashimoto, T., Yamada, Y., 2003. New genes in alkaloid metabolism and transport. Curr. Opin. Biotechnol. 14, 163–168.

Hecht, S.S., 2003. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nat. Rev. Cancer 3, 733–744.

Hecht, S.S., Hoffmann, D., 1990. The relevance of tobacco specific nitrosamines to human cancer. Cancer Surv. 8, 273–294.

Heim, W.G., Sykes, K.A., Hildreth, S.B., Sun, J., Lu, R.-H., Jelesko, J.G., 2007. Cloning and characterization of a *Nicotiana tabacum* methylputrescine oxidase transcript. Phytochemistry 68, 454–463.

Hoffmann, D., Brunnemann, K.D., Prokopczyk, B., 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., 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., Dorlhac de Borne, F., 2008. Reducing the content of nornicotine in tobacco via targeted mutation breeding. Mol. Breed. 21, 369–381.

Katoh, A., Shoji, T., Hashimoto, T., 2007. Molecular cloning of *N*-methylputrescine oxidase from tobacco. Plant Cell Physiol. 48, 550–554.

Katz, J., Caudle, R.M., Bhattacharyya, I., Stewart, C.M., 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., 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.

Mann, T.J., Weybrew, J.A., Matzinger, D.F., Hall, J.F., 1964. Inheritance of the conversion of nicotine to nornicotine in varieties of *Nicotiana tabacum* L. and related amphidiploids. Crop Sci. 4, 349–353.

Morita, M., Shitan, N., Sawada, K., Van Montagu, M.C.E., Inze, D., Rischer, H., Goossens, A., Oksman-Caldentey, K.-M., Moriyama, Y., Yazaki, K., 2009. Vacuolar transport of nicotine is mediated by a multidrug and toxic compound extrusion (MATE) transporter in *Nicotiana tabacum*. Proc. Natl. Acad. Sci. USA 106, 2447–2452.

Murad, L., Lim, K.Y., Christopodulou, V., Matyasek, R., Lichtenstein, C.P., Kovarik, A., Leitch, A.R., 2002. The origin of tobacco's T genome is traced to a particular lineage within *Nicotiana tomentosiformis* (*Solanaceae*). Am. J. Bot. 89, 921–928.

Nelson, D.R., Koymans, L., Kamataki, T., Stegeman, J.J., Feyereisen, R., Waxman, D.J., Waterman, M.R., Gotoh, O., Coon, M.J., Estabrook, R.W., Gunsalus, I.C., Nebert, D.W., 1996. P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. Pharmacogenetics 6, 1–42.

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

Pompon, D., Louerat, B., Bronine, A., Urban, P., 1996. Yeast expression of animal and plant P450s in optimized redox environments. Methods Enzymol. 272, 51–64.

Shoji, T., Inai, K., Yazaki, Y., Sato, Y., Takase, H., Shitan, N., Yazaki, K., Goto, Y., Toyooka, K., Matsuoka, K., Hashimoto, T., 2009. Multidrug and toxic compound extrusion-type transporters implicated in vacuolar sequestration of nicotine in tobacco roots. Plant Physiol. 149, 708–718.

Siminszky, B., Gavilano, L., Bowen, S.W., 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.

Sisson, V.A., Severson, R.F., 1990. Alkaloid composition of the *Nicotiana* species. Beitr. Tabakforsch. 14, 327–399.

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

Xu, P., Zhang, Y., Kang, L., Roossinck, M.J., Mysore, K.S., 2006. Computational estimation and experimental verification of off-target silencing during posttranscriptional gene silencing in plants. Plant Physiol. 142, 429–440.

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

Yukawa, M., Tsudzuki, T., Sugiura, M., 2006. The chloroplast genome of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*: complete sequencing confirms that the *Nicotiana sylvestris* progenitor is the maternal genome donor of *Nicotiana tabacum*. Mol. Genet. Genom. 275, 367–373.