Functional Analysis of Nicotine Demethylase Genes Reveals Insights into the Evolution of Modern Tobacco*

Received for publication, October 10, 2006, and in revised form, November 6, 2006 Published, JBC Papers in Press, November 13, 2006, DOI 10.1074/jbc.M609512.220

Lily B. Gavilano‡, Nicholas P. Coleman†, Steven W. Bowen§, and Balazs Siminszky†

From the ‡Department of Plant and Soil Sciences, University of Kentucky, Lexington, Kentucky 40546-0312 and the §Crop Science Department, North Carolina State University, Raleigh, North Carolina 27695

Tobacco (Nicotiana tabacum L.) is a natural allotetraploid derived from the interspecific hybridization between ancestral Nicotiana sylvestris and Nicotiana tomentosiformis. The majority of cultivated tobacco differs from both of its progenitor species in that tobacco typically contains nicotine as the primary alkaloid, in contrast to its two progenitors that accumulate nornicotine in the senescing leaves. However, most, if not all, tobacco cultivars possess an unstable mutation, commonly referred to as the conversion locus, that when activated mediates the conversion of a large percentage of nicotine to nornicotine in the senescing leaf. We have recently identified CYP82E4, a tobacco nicotine N-demethylase gene whose expression was highly induced during senescence in plants that have converted, and CYP82E3, a closely related homolog that exhibited no nicotine N-demethylase activity. In this study, domain swapping and site-directed mutagenesis studies identified a single amino acid change that fully restored nicotine N-demethylase activity to CYP82E3. An examination of the N. tomentosiformis orthologs of CYP82E3 and CYP82E4 revealed that both are functional nicotine N-demethylase genes in N. tomentosiformis. Collectively, our results suggest that a single base pair mutation in CYP82E3 and transcriptional suppression of CYP82E4 played important roles in the evolution of the alkaloid profile characteristic of modern tobacco.

Pyridine alkaloids, such as nicotine, nornicotine, anabasine, and anatabine, have been detected in numerous species of the Nicotiana genus (1) where they are thought to function as herbivore deterrents (2). The production of pyridine alkaloids is elicited by herbivore attacks (3, 4) through a jasmonic acid-induced signaling pathway (5, 6). Several studies have demonstrated that the alkaloid composition within the Nicotiana genus is species-specific and highly variable, and in the majority of the species a single alkaloid predominates (1, 7). For example, Nicotiana tomentosiformis primarily accumulates nornicotine, whereas in Nicotiana glauca anabasine is the major alkaloid (1). Interestingly, the alkaloid profiles of some Nicotiana are different in green versus senescing (yellow) leaves. For instance, in the green leaves of Nicotiana sylvestris nicotine is the primary alkaloid, but a substantial portion of nicotine is converted into nornicotine once the leaves senesce through air-curing (1).

In cultivated tobacco alkaloid composition can vary because of the presence of an unstable locus, termed the “converter locus” (8). In commercial tobacco leaves that harbor an inactive converter locus, the alkaloid pool typically contains about 95% nicotine and 3% nornicotine (7). These plants are frequently referred to as “nonconverters.” In some of the individuals, however, the converter locus becomes activated, and a large portion of nicotine is metabolized to nornicotine (9) giving rise to plants termed “converters.” Similar to N. sylvestris, nicotine to nornicotine metabolism in converter tobacco is induced during leaf senescence (10). The accumulation of nornicotine in converter plants is undesirable, because nornicotine is a biochemical precursor of N’-nitrosonornicotine, a tobacco-specific nitrosamine that has been shown to exhibit carcinogenic properties in laboratory animals (11–13). In addition, nornicotine has been recently recognized as an elicitor of protein glycation (14) and other biochemical processes that are detrimental to human health (15, 16).

Nicotiana tabacum is a natural allotetraploid derived from the interspecific hybridization of ancestral N. sylvestris and N. tomentosiformis (17–22), species that convert nicotine to nornicotine in senescing and green leaves, respectively (1). As a result, the emergence of the nicotine-accumulating (nonconverter) phenotype of N. tabacum from its two progenitors would have required the inactivation of nornicotine-accumulating traits originating from both ancestral parents. Mann et al. (8) suggested that the nicotine-accumulating phenotype of nonconverter tobacco was developed by the successive inactivation of the two dominant conversion loci donated by the nornicotine-containing parental species, but to date such loci have not been identified.

Another interesting question relates to the origin of the unstable conversion locus present in converter tobacco. In theory, either of the two progenitor species could have donated the conversion locus to tobacco, because nicotine to nornicotine conversion occurs in both N. sylvestris and N. tomentosiformis at one developmental stage. Classical genetic studies conducted by Mann et al. (8) demonstrated that the unstable conversion locus found in converter tobacco is allelic to the green leaf conversion locus of N. tomentosiformis and inherited as a dominant gene (8, 9, 23). The observation that the unstable conversion locus was originated from green-leaf converter N. tomentosiformis was surprising, because the reversion of the conversion

* This work was supported by Philip Morris U. S. A. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Dept. of Plant and Soil Sciences, University of Kentucky, 1405 Veterans Dr., Lexington, KY 40546-0312. Tel.: 859-257-5020 (ext. 80761); Fax: 859-257-7874; E-mail: balazs@uky.edu.
Evolution of Nicotine-containing Tobacco

locus to the ancestral type in tobacco confers senescing leaf conversion phenotype characteristic to *N. sylvestris*. The apparent inconsistency between the time of nornicotine production and the origin of the unstable conversion locus in tobacco suggested that evolution of the alkaloid profile of modern tobacco involved complex genetic changes.

We have previously isolated a closely related gene family, designated the CYP82E gene subfamily, whose collective transcript accumulation was up-regulated in converter versus nonconverter tobacco (24). Of these cDNAs, CYP82E4 was shown to mediate the metabolic conversion of nicotine to nornicotine when overexpressed in tobacco and yeast (24). Expression analysis of CYP82E4 revealed a >80-fold increase of transcript accumulation in the cured leaves of converter versus nonconverter plants demonstrating that CYP82E4 plays a central role in mediating nicotine N-demethylation in senescing leaves of tobacco (25). To distinguish between orthologous P450 genes from tobacco and *N. tomentosiformis*, an NTab or NTom prefix is attached, respectively, to the name of the genes in the remaining part of the paper (e.g. NTabCYP82E4 or NTomCYP82E4).

Because not all members of the NTabCYP82E gene subfamily exhibited nicotine N-demethylase (NND) activity (24), the original goal of this study was to define the structural elements responsible for NND activity in the closely related P450 isoforms of the NTabCYP82E gene subfamily. Our results showed that a single base amino acid substitution in the coding region of another closely related isoform, NtabCYP82E3, is sufficient to confer NND activity to the encoded protein. DNA sequence and gene expression data also demonstrated that *N. tomentosiformis* possesses two active NND genes, NtomCYP82E3 expressed strongly in the green leaves and NtomCYP82E4 specifically up-regulated during senescence. We propose that during the evolution of modern tobacco, NtabCYP82E3 was impaired by a point mutation, and NtabCYP82E4 was inactivated by an unstable mutation resulting in the transcriptional suppression of the gene.

**EXPERIMENTAL PROCEDURES**

Plant Materials—Green-leaf converter SC58(C<sub>1</sub>C<sub>1</sub>) tobacco line was generated by crossing an interspecific hybrid between *N. tomentosiformis* and Flue-cured tobacco variety 402 with Flue-cured tobacco cultivar SC58 followed by eight cycles of backcrossing of the progeny to the SC58 parent. In each generation, green-leaf converter individuals were selected and used in the subsequent cross with cultivar SC58. Finally, the conversion locus (C<sub>4</sub>) was fixed to homozygosity by self-pollinating the heterozygous introgressed lines and identifying individuals whose progeny no longer segregated for the green-leaf conversion trait SC58 (8). Other tobacco cultivars included in the study were full-sib double haploid burley lines DH98-325-5 (nonconverter) and DH98-325-6 (senescing-leaf converter). All genotypes were originally grown for 30 days in a growth chamber and then transferred to a controlled environment greenhouse equipped with supplemental lighting providing a 14/10-h light/dark cycle. Leaf samples were collected 30 days after transplanting. For curing, tobacco leaves were dipped in 0.2% ethephon, dried for 30 min at room temperature, and stored in dark until they turned yellow.

Construction and Analysis of the NtabCYP82E3/NtabCYP82E4 Chimeras—Chimeric clones were generated by digesting the NtabCYP82E3 and NtabCYP82E4 cDNAs with restriction enzymes that recognize restriction sites conserved between the two cDNAs and replacing the resulting NtabCYP82E4 restriction fragment with the corresponding NtabCYP82E3 region. NtabCYP82E3/NtabCYP82E4 chimeras were subcloned into the pYeDP60 yeast expression vector, and the resulting constructs were introduced to yeast strain WAT11 as described (26). Yeast culturing, transgene induction, and preparation of yeast microsomes were performed as described previously (27). Microsomal fractions were incubated for 45 min at 23 °C in 50 mM phosphate buffer, pH 7.1, containing 0.5 mg of microsomal protein, 2.45 μM [pyrrolidine-2-<sup>14</sup>C]nicotine (Moravek Biochemicals, Brea, CA), and 0.75 mM NADPH in a total reaction volume of 150 μl. The reactions were arrested with 50 μl of acetone, and 50 μl of the reaction mixture was spotted on a 250-μm Whatman K<sub>6</sub>F silica plate. Conversion of nicotine to nornicotine was quantified by localizing and integrating the radioactive traces of nicotine with a Bioscan System 400 imaging scanner. Site-directed mutagenesis was conducted with the QuikChange™ site-directed mutagenesis kit as described by the manufacturer (Stratagene, La Jolla, CA).

Isolation and DNA Sequence Analysis of the CYP82E3 and CYP82E4 Orthologs—Plant genomic DNA was extracted by grinding a 100-mg fresh leaf sample in 500 μl of homogenization buffer (100 mM Tris-HCl, pH 8.8, 1 M KCl, 10 mM EDTA) using the FastPrep FP120 Cell Disrupter (Qbiogene, Inc. Carlsbad, CA). P450s were amplified in a PCR containing 1 μl of leaf extract, 2 μM each primer, 200 μM each of dNTP, and 1.5 mM MgCl<sub>2</sub> in a 50-μl final reaction volume. The primers used for the isolation of the various genomic P450 fragments are listed in Table 1. DNA amplification was performed in 30 cycles of denaturing at 95 °C for 15 s, primer annealing at 55 °C for 30 s, and primer extension at 72 °C for 30 s. For the allele-specific amplification of the CYP82E3 orthologs, a 61 °C annealing temperature was used. PCR products were visualized on an ethidium bromide-stained 1% agarose gel and directly cloned into the pGEM<sup>®</sup> Easy cloning vector (Promega Corp., Madison, WI). DNA sequence of the clones were determined by the chain termination method (28). Sequences were aligned using the ClustalW algorithm (29).

Isolation of NtomCYP82E3 and NtomCYP82E4 cDNAs—Total RNA was extracted from plant leaves using the TRIzol reagent following the manufacturer’s instructions (Invitrogen). Purified RNA was treated with RNase-free DNase (TURBO DNA-free; Ambion, Austin, TX). First strand cDNA was synthesized using 5 μg of total RNA using the StrataScript first-strand synthesis system (Stratagene, Cedar Creek, TX). PCR amplification, cloning of the PCR products, and DNA sequencing of the P450 cDNAs were performed as described for the isolation of the genomic P450 fragments. The nucleotide sequences of the primers used in the PCR are listed in Table 1.

The abbreviations used are: NND, nicotine N-demethylase; RT, real time.
Kinetic Analysis of NtabCYP82E3- and NtabCYP82E4-mediated Nicotine Metabolism—Yeast microsomes were prepared as described previously (27). Microsomal protein concentrations were determined by the Bradford protein assay (Bio-Rad). Nicotine concentration ranges used in the study were 3–467 and 0.6–13 μM. Microsomal protein concentrations was confirmed by ligating the RT-PCR products into the pGEM-T Easy vector and sequencing 20 randomly selected clones of the glyceraldehyde-3-phosphate dehydrogenase fragment of the glyceraldehyde-3-phosphate dehydrogenase gene subfamily, play a major role in the conversion of nicotine to nornicotine (24). We isolated five unique members of the CYP82E2 gene subfamily, of which only one isoform, NtabCYP82E4, was shown to encode an NND. NtabCYP82E4, the most closely related homolog to NtabCYP82E4, shares 95% predicted amino acid sequence identity with NtabCYP82E4 but conferred no NND activity when expressed in yeast or tobacco (24).

The high amino acid sequence homology yet distinctly different NND activity between NtabCYP82E3 and NtabCYP82E4 provided us with an opportunity to identify unique structural elements that bestow NND activity to NtabCYP82E4. Our hypothesis was that one or few of the 30 polymorphic amino acids between NtabCYP82E3 and NtabCYP82E4 represent catalytically important residues for NND activity in NtabCYP82E4. To identify segments of the coding region that selectively impart NND activity to NtabCYP82E4, we employed domain swapping, a strategy that creates a fusion polypeptide from two (or more) protein sequences and thus facilitates the identification of domains that control the expression of traits (e.g. function) differentially encoded in the parent molecules.

The high degree of sequence homology between NtabCYP82E3 and NtabCYP82E4 enabled us to identify three restriction sites (PmlI, SwaI, and EcoRV) located in the same position in both cDNAs (Fig. 1). Using these restriction sites as junction points, we replaced the 459-, 711-, and 1201-bp restriction sites (PmlI, SwaI, and EcoRV) located in the same position. P450 cDNA concentration was calculated based on the transcript-specific calibration curve normalized to the internal standard. The purity of PCR products was determined by melting-curve analysis as described previously (31). DNA sequence of the amplified cDNA fragments was confirmed by ligating the RT-PCR products into the pGEM-T Easy vector and sequencing 20 randomly selected clones of each cDNA fragment.

RESULTS

The C330W Amino Acid Substitution Confers NND Activity to the NtabCYP82E3 cDNA—We have demonstrated previously that a group of P450 isoymes, designated the CYP82E2

| Name | Nucleotide Sequence | Polarity | Note |
|------|---------------------|----------|------|
| E3-UT | GATCAATTGGATAGTTGATGTTTGTG | + | Used to isolate full-length genomic NtomCYP82E3 |
| E3-UT | TTTGAATTTACGTTGATGTTTGTG | + | Used to amplify exon 2 of CYP82E3 orthologs |
| E3-IntF | TGGGTTGGTTGTTGATTG | + | Used for allele-specific PCR to determine the nucleotide sequence at N° of the CYP82E3 orthologs |
| E3-IntR | GATCTAGGTTGTTGATGTTTGTG | + | Used to isolate full-length NtomCYP82E3 cDNA |
| E3-cDNAF | TTATCAATAATGCCATTCCC | + | Used to isolate full-length NtomCYP82E3 cDNA |
| E3-cDNAR | TGGGTTGGTTGTTGATTG | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
| E3-R1 | GAGAATCTCGTTGTTGTTTCA | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
| E3-R2 | CCGTCGCCAGGAAAGT | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
| E3-RTF | CTATATGCTCCATCAGGAGG | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
| E3-RTR | GGTTTTTCCGGTAGAA | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
| E3-CW | GCCCTTATATTTTTACGCCATAT | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
| E3-Cys | TTATCAATAATGCCATTCCC | + | Used to isolate full-length NtomCYP82E3 cDNA |
| E3-Trp | TTTGAATTTACGTTGATGTTTGTG | + | Used to amplify exon 1 of the CYP82E3 orthologs |
| E3-DBAF | ATGCGTTCCTGATCCATGACGG | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
| E3-DBAN | TGGAATTATGCCCATCCTACA | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
| E3-R1 | GAGAATCTCGTTGTTGTTTCA | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
| E3-R2 | CCGTCGCCAGGAAAGT | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
| E3-RTF | CTATATGCTCCATCAGGAGG | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
| E3-RTR | GGTTTTTCCGGTAGAA | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
| E3-CW | GCCCTTATATTTTTACGCCATAT | + | Used for expression analysis of the CYP82E3 orthologs by RT-PCR |
Evolution of Nicotine-containing Tobacco

To further define amino acid residues essential for NND activity, site-directed mutagenesis experiments were conducted. The SwaI-EcoRV region of NtabCYP82E4 encodes 164 amino acids, 7 of which are polymorphic to NtabCYP82E3 (Fig. 2). The relatively few polymorphic amino acids encoded within the SwaI-EcoRV region of the NtabCYP82E4 and NtabCYP82E3 genes made it feasible to employ a mutation stacking approach for identifying catalytically important amino acids. Codons encoding all polymorphic amino acid residues in the SwaI-EcoRV fragment of the NtabCYP82E3 coding region were progressively substituted for the corresponding codons of the NtabCYP82E4 cDNA and expressed in yeast. Microsomal extracts isolated from yeast cells enriched in the S-E chimeric protein catalyzed the nicotine N-demethylation confirming the notion that the SwaI-EcoRV fragment of NtabCYP82E4 encodes amino acids necessary for NND activity.

To test this hypothesis, we attempted to amplify the CYP82E3 orthologs from N. tomentosiformis and N. sylvestris genomic DNA using NtabCYP82E3-specific primers. PCR products were detected only when genomic DNA isolated from N. tomentosiformis was used as template suggesting that in tobacco CYP82E3 originated from the N. tomentosiformis ancestral parent (data not shown). DNA sequence analysis of NtomCYP82E3 predicted a Trp amino acid residue at position 330 lending support to our hypothesis that N. tomentosiformis CYP82E3 encodes a functional NND—The observation that a single C330W substitution confers NND activity to NtabCYP82E3 led us to the hypothesis that the inactivation of NtabCYP82E3 by the W330C mutation may represent one of the genetic changes responsible for the evolution of nicotine-retaining N. tabacum from its nicotine N-demethylating ancestors.

To test this hypothesis, we attempted to amplify the CYP82E3 orthologs from N. tomentosiformis and N. sylvestris genomic DNA using NtabCYP82E3-specific primers. PCR products were detected only when genomic DNA isolated from N. tomentosiformis was used as template suggesting that in tobacco CYP82E3 originated from the N. tomentosiformis ancestral parent (data not shown). DNA sequence analysis of NtomCYP82E3 predicted a Trp amino acid residue at position 330 lending support to our hypothesis that NtomCYP82E3 encodes a functional NND gene (Table 2). In addition to the predicted C330W mutation, a second substitution (K118N) polymorphism is neutral given that NtomCYP82E3 also exhibits an NND activity similar to that displayed by the C330W NtomCYP82E3 mutant (Table 3).

N. tomentosiformis CYP82E3 Is a Second NND Gene in N. tomentosiformis—To further characterize the genes involved in nicotine N-demethylation in N. tomentosiformis, we sought to determine whether NtomCYP82E4 also encodes a viable NND protein. NtomCYP82E4 was considered a likely NND candidate, because NtabCYP82E4 has been shown previously to mediate the nicotine to nornicotine conversion in senescing tobacco leaves (24).

To investigate the role of NtomCYP82E4 in nicotine N-demethylation, NtomCYP82E4 was isolated from N. tomentosiformis genomic DNA, and the nucleotide sequence of the gene was compared with that of NtabCYP82E4 derived from tobacco cul-

![Image](https://example.com/image.png)

**TABLE 2**

| Sample       | Phenotype | N<sup>330</sup> (AA<sup>118</sup>) | N<sup>330</sup> (AA<sup>115</sup>) |
|--------------|-----------|----------------------------------|----------------------------------|
| DH98-325-5   | NC        | C (Cys)                          | ND                               |
| DH98-325-6   | SC        | C (Cys)                          | A (Lys)                          |
| SC58         | NC        | C (Cys)                          | A (Lys)                          |
| SC58(C,C<sub>1</sub>) | GC     | G (Trp)                          | T (Asn)                          |
| N. tomentosiformis | GC  | G (Trp)                          | T (Asn)                          |

* Nucleotide numbering is based on cDNA sequences.
whether a correlation exists between NtomCYP82E4 and the green-leaf converter phenotype of N. tomentosiformis from yeast cells transformed with the cDNAs catalyzing the conversion of nicotine to nornicotine, whereas microsomes expressing the empty plasmid produced no DNA amplification was detected when the alternative primer. Accumulation of PCR products was evident when the C 990 allele-specific primer was used to amplify the CYP82E3 orthologs isolated from various tobacco cultivars and N. tomentosiformis. The abbreviations used are as follows: N, nucleotide; AA, amino acid.

**TABLE 3**

| Protein          | NtCYP82E3 | NtCYP82E4 | NtomCYP82E3 | NtomCYP82E4 |
|------------------|-----------|-----------|-------------|-------------|
| K_m,app (μM)    | 48.1 ± 8.1| 9.1 ± 2.4 | 48.1 ± 8.1  | 9.1 ± 2.4   |
| V_max (nmol/min/mg protein) | 7.1 ± 0.32 | 1.0 ± 0.12 | 0.15 ± 0.017 | 0.11 ± 0.015 |
| N. tomentosiformis | | | | |

Assays were repeated three times. Values following ± represent S.E.

**TABLE 4**

| Sample | NtabCYP82E3 | NtomCYP82E3 | NtomCYP82E4 |
|--------|-------------|-------------|-------------|
| DH98-325-5 | T (Phe) | C | C |
| DH98-325-6 | T (Phe) | C | C |
| SC58    | T (Phe) | C | C |
| SC58(C,C) | A (Tyr) | T | T |
| N. tomentosiformis | A (Tyr) | T | T |

The abbreviations used are as follows: N, nucleotide; AA, amino acid.

**FIGURE 3** Allele-specific PCR amplification of CYP82E3 orthologs isolated from various tobacco cultivars and N. tomentosiformis (N.tom). CG letters indicate the nucleotide residue at N990 complementary to the allele-specific reverse primers E3-Cys (left) and E3-Trp (right). Lanes 1–4 show the PCR products generated by the amplification of purified NtomCYP82E3 and NtabCYP82E3 templates, respectively. The abbreviation used is as follows: 325-6, DH98-325-6.

Evolution of Nicotine-containing Tobacco

The 56A→T nucleotide polymorphism encodes for a Phe→Tyr amino acid residue in tobacco and a Tyr in N. tomentosiformis. The GenBank accession number of the genomic sequence of NtomCYP82E3 is EF042307.

Similar to NtomCYP82E3, NtomCYP82E4 also exhibited NND activity when expressed in yeast (Table 3). To compare catalytic competence of NtomCYP82E3 and NtomCYP82E4 for nicotine N-demethylation, we determined the kinetic parameters of the two proteins. Microsomal fractions isolated from yeast cells transformed with the NtomCYP82E3 or NtomCYP82E4 cDNAs catalyzed the conversion of nicotine to nornicotine, whereas microsomes expressing the empty plasmid did not produce nornicotine. Although both proteins exhibited high affinity toward nicotine, the binding affinity and V_max values of NtomCYP82E3 were about 5 and 4 times higher than that of NtomCYP82E4, respectively (Table 3).

Both NtomCYP82E3 and NtomCYP82E4 Correlate with the Green-leaf Converter Phenotype—A distinguishing feature of the converter phenotype between N. tomentosiformis and tobacco is the accumulation of nornicotine in the green, nonsenescent leaf of N. tomentosiformis. To investigate whether a correlation exists between NtomCYP82E3 and/or NtomCYP82E4 and the green-leaf converter phenotype of N. tomentosiformis, we extended our sequence analysis of the CYP82E3 and CYP82E4 orthologs to nonconverter individuals of the Flu-cured variety SC58 and the green-leaf converter SC58(C,C) tobacco line. SC58(C,C) was developed by crossing an interspecific hybrid between N. tomentosiformis and Flue-cured variety 402 with cultivar SC58, followed by backcrosses to the SC58 recurrent parent (8). During each generation, selection was made based on the dominant green-leaf converter trait originating from N. tomentosiformis. Theoretically, SC58(C,C) plants should contain less than 0.5% of the N. tomentosiformis haploid genome, including the selected DNA fragment containing the green-leaf converter locus (C_1).

If the NtomCYP82E3 and/or NtomCYP82E4 genes are associated with the green-leaf converter phenotype, one would expect to find these genes within the SC58(C,C) line. DNA sequence analysis of the CYP82E3 and CYP82E4 orthologs showed that both NtomCYP82E3 and NtomCYP82E4 were present in SC58(C,C) (Tables 2 and 4). As expected, the NtabCYP82E3 and NtabCYP82E4 genes were detected in all pure tobacco genotypes tested, including SC58, DH98-325-5, and DH98-325-6 (Tables 2 and 4). These results suggest that the presence of the NtomCYP82E3 and NtomCYP82E4 genes correlates with NND activity in the green leaves of SC58(C,C) tobacco.

Allele-specific primers based on the mutant-defining polymorphism at position 990 were also effective in distinguishing the origin of the specific CYP82E3 alleles in the materials tested. PCRs were performed using reverse primers specific for C_990 or G_990 in combination with an intron-specific forward primer. Accumulation of PCR products was evident when the C_990 allele-specific primer was used to amplify the NtabCYP82E3 fragment from the SC58, DH98-325-5, and DH98-325-6 plants or when the G_990 primer was used with SC58(C,C) and N. tomentosiformis DNA (Fig. 3). In contrast, no DNA amplification was detected when the alternative allele-specific primers were added. These results provide additional confirmation that the CYP82E3 variant found in SC58(C,C) originated from N. tomentosiformis.

Expression Analysis of the CYP82E3 and CYP82E4 Orthologs in DH98-325-6 and SC58(C,C) Tobacco—Discovery of the N. tomentosiformis orthologs of CYP82E3 and CYP82E4 in the introgression line SC58(C,C) suggest that the NND genes per se may be responsible for mediating the green-leaf converter phenotype common to N. tomentosiformis and SC58(C,C). To better understand the regulatory roles of the CYP82E3 and CYP82E4 orthologs in nonnicotine biosynthesis, we measured the transcript accumulation of NtabCYP82E3/NtabCYP82E4 and NtomCYP82E3/NtomCYP82E4 in green and senescing leaves of tobacco lines DH98-325-6 and SC58(C,C), respectively, using an allele-specific RT-PCR strategy. NtabCYP82E3 and NtabCYP82E4 were weakly expressed in the green leaves of cultivar DH98-325-6 (Table 5). Consistent with our previous observations, however, NtabCYP82E4 transcription was sharply induced in the cured leaves. No enhancement of NtabCYP82E3 expression was observed in cured leaves of DH98-325-6 (Table 5).

Transcript accumulation of NtomCYP82E3 and NtomCYP82E4 was markedly different in line SC58(C,C).
Expression of NtomCYP82E3 was high in the green leaves but rapidly declined during senescence (Table 5). Although transcript accumulation of NtomCYP82E4 in the green leaves of SC58(C1,C4) plants was about 30-fold higher than that of NtabCYP82E4 in the pure tobacco line, it was still about 20-fold lower than that of NtomCYP82E3. Amplification specificity of the cDNAs was confirmed by DNA sequence analysis of 20 randomly selected clones derived from each RT-PCR product. In all cases, the clones yielded the identical DNA sequence to the predicted cDNA fragments confirming that the PCRs were allele-specific (data not shown). These results demonstrate that although both NtomCYP82E3 and NtomCYP82E4 are expressed in the green leaves of SC58(C1,C4), the expression of NtomCYP82E3 clearly predominates. Furthermore, the transcription of NtomCYP82E4 originating from N. tomentosiformis is strongly inducible by senescence in the SC58(C1,C4) line, similar to the manner in which NtabCYP82E4 is induced in converter tobacco lines. Finally, the expression of both NtabCYP82E3 and NtabCYP82E4 in green leaves is lower than that of their N. tomentosiformis orthologs.

**DISCUSSION**

Domain swapping and site-directed mutagenesis experiments were performed to determine the structural basis for the catalytic activity of NtabCYP82E4. These studies revealed that a single amino acid substitution (C330W) was necessary and sufficient to confer NND activity to NtabCYP82E3 and the fact that NND genes are theoretically possible between N. tomentosiformis and N. tabacum. An observation that a single amino acid mutation could confer NND activity to NtabCYP82E3 and the fact that NND genes in N. tabacum confirm the notion that NND genes represent the converter loci described by genetic studies (Tables 2 and 4 and Fig. 3). The finding that the SC58(C1,C4) plants, which were selected based on their green-leaf converter phenotype, contain the senescence-inducible NtomCYP82E4 gene in addition to the constitutively expressed NtomCYP82E3 indicate that NtomCYP82E3 and NtomCYP82E4 are physically linked, an observation consistent with the genomic organization of other closely related P450 isoforms that are often arranged in tandem repeats (32). Because of the close genetic linkage among the NND genes in N. tomentosiformis, nicotine conversion appears to be inherited as a single Mendelian trait as observed by several investigators (8, 9, 23).

Expression analysis of the CYP82E3 and CYP82E4 orthologs demonstrated that NtomCYP82E3 is highly expressed in the green leaves and NtomCYP82E4 is highly induced during leaf senescence. Although variations in the expression profiles of the NND genes are theoretically possible between N. tomentosiformis and SC58(C1,C4), the similar expression patterns of NtomCYP82E4 in SC58(C1,C4) and NtabCYP82E4 in converter tobacco confirms the notion that NtabCYP82E4 represents the unstable locus, derived from N. tomentosiformis, that is responsible for the genetic conversion of tobacco (Table 5).

Collectively, these results support the model proposed in Fig. 4. In N. tomentosiformis at least two active NNDs, NtomCYP82E3 and NtomCYP82E4, are located on the chromosomal region defined by the C4 locus. These genes likely
arose by the duplication of an ancestral NND gene followed by sequence diversification. At the onset of amphiploidy, both progenitor species donated a set of conversion factors to tobacco. Because the acquired conversion loci included NtomCYP82E3 and NtomCYP82E4, ancestral tobacco had likely emerged as a green-leaf converter plant. During the evolution of tobacco, mutations at several conversion loci affected the alkaloid composition of the plant. These lesions include a stable small or large scale mutation, at the N. sylvestris-originating C_S locus, whose regulation appears to mirror that of the senescence-inducible CYP82E4 gene, and the W330C mutation at the C_T locus originating from N. tomentosiformis. In addition to the stable mutations, an unstable mutation in the N. tomentosiformis-derived CYP82E4 locus led to the inactivation of this NND gene by transcriptional suppression. Similar to NtabCYP82E4, the transcriptions of NtabCYP82E3 was also greatly reduced. These and possibly other mutations affecting additional conversion loci gave rise to nicotine-producing modern nonconverter tobacco, and the transcriptional reactivation of the NtabCYP82E4 locus played a major role in the emergence of converter tobacco. This model also suggests the mechanism by which the conversion locus derived from the green-leaf converter N. tomentosiformis conferred the senescing-leaf converter phenotype to converter tobacco. Because the C_T locus donated by N. tomentosiformis includes both the constitutively expressed NtomCYP82E3 and senescence-induced NtomCYP82E4 genes, the inactivation of NtabCYP82E3 in tobacco restricted nornicotine production to the senescing leaves.

Our model can be extended to explain the differences in the nicotine alkaloid profiles of other Nicotiana species. Sisson et al. (1) classified numerous species of the Nicotiana genus as nicotine, green leaf nornicotine, or curing-specific nornicotine accumulators. One possibility is that nicotine accumulators possess a dysfunctional CYP82E3 and CYP82E4, green leaf nornicotine accumulators carry a functional CYP82E3 with or without CYP82E4, and curing-specific nornicotine accumulators contain a functional CYP82E4 and a dysfunctional CYP82E3.

The driving force behind the evolution of nicotine-accumulating tobacco remains an open question. High nicotine and low nornicotine tobacco plants could have been selected by man at the beginning of tobacco use. However, Wernsman and Matzinger (10) challenged this hypothesis arguing that man could not have selected nicotine-accumulating tobacco because of the presence of the two conversion loci donated by the N. sylvestris and N. tomentosiformis. Given our results that a minimum of three mutations was required for the evolution of nicotine-accumulating tobacco, the selection of these plants by planned breeding alone appears to be even more unlikely. Nicotine-containing tobacco could have also arisen through natural selection, because nicotine exhibits superior insecticidal activity to nornicotine in several insects (33–35). Alternatively, it is possible that the adaptive values of nicotine and nornicotine are equal, and nicotine N-demethylation is a dispensable function in tobacco. Interestingly, in the Repandae section of the Nicotiana genus nornicotine serves as a precursor for N-acyl-nornicotine, a compound that exhibits 1000-fold higher insecticidal activity against nicotine-resistant Manduca sexta than nicotine (36–38). Other sections of Nicotiana, however, lack the acyltransferase enzyme necessary for N-acyl-nornicotine biosynthesis; therefore, nornicotine production provides no adaptive advantage in these species (36). Similar to members of the Repandae section, an early progenitor of Nicotiana could have benefited from nornicotine production, but the diminished adaptive value of the compound could have led to the loss of NND function in tobacco.

Acknowledgments—We thank Earl Wernsman for the DH98-325-5 and DH98-325-6 tobacco seed. We also thank Denis Pompon and Philippe Urban for their kind gifts of the pYedP60 yeast expression vector and the WAT11 yeast strain. We thank Ramsey Lewis for helpful discussions.

REFERENCES

1. Sisson, V. A., and Severson, R. F. (1990) Beitr. Tabakforsch. 14, 327–339
2. Metcalf, R. L., and Metcalf, R. A. (1992) Destructive and Useful Insects, 5th Ed., McGraw-Hill Inc., New York
3. Baldwin, I. T., and Ohnmeiss, T. E. (1993) J. Chem. Ecol. 19, 1143–1153
4. Baldwin, I. T. (1988) Oecologia 75, 367–370
5. Halitschke, R., and Baldwin, I. T. (2003) Plant J. 36, 794–807
6. Baldwin, I. T., Schmelz, E. A., and Ohnmeiss, T. E. (1994) J. Chem. Ecol. 20, 2139–2157
7. Saitoh, F., Noma, M., and Kawashima, N. (1985) Phytochemistry 24, 477–480
8. Mann, T. I., Weybrew, J. A., Matzinger, D. F., and Hall, J. L. (1964) Crop Sci. 4, 349–353
9. Burk, L. G., and Jeffrey, R. N. (1958) Tob. Sci. 2, 139–141
10. Wernsman, E. A., and Matzinger, D. F. (1968) Tob. Sci. 12, 226–228
11. Hecht, S. S. (1998) Chem. Res. Toxicol. 11, 559–603
12. Hecht, S. S. (2003) Nat. Rev. Cancer 3, 733–744
13. Hecht, S. S., and Hoffmann, D. (1989) Cancer Surv. 8, 273–294
14. Dickerson, T. J., and Janda, K. D. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 15084–15088
15. Katz, J., Caudle, R. M., Bhattacharya, I., Stewart, C. M., and Cohen, D. M. (2005) J. Periodontol. 76, 1171–1174
16. Brogan, A. P., Dickerson, T. J., Boldt, G. E., and Janda, K. D. (2005) Proc. Natl. Acad. Sci. U. S. A. 102, 10433–10438
17. Murad, L., Lim, K. Y., Christopodulou, V., Matyasek, R., Lichtenstein, C. P., Kovarik, A., and Leitch, A. R. (2002) Ann. J. Bot. 89, 921–928
18. Yukawa, M., Tsuzuki, T., and Sugira, M. (2006) Mol. Genet. Genomics 275, 367–373
19. Ren, N., and Timko, M. P. (2001) Genome 44, 559–571
20. Riechers, D. E., and Timko, M. P. (1999) Plant Mol. Biol. 41, 387–401
21. Goodspeed, T. H. (1954) in The Genus Nicotiana (Verdoorn, F., ed), 5th Ed., Chronica Botanica, Vol. 16, pp. 1–536, Waltham, MA
22. Bogani, P., Lio, P., Intrieri, M. C., and Buatti, M. (1997) Mol. Phylogenet. Evol. 7, 62–70
23. Griffith, R. B., Valleau, W. D., and Stokes, G. W. (1955) Science 121, 343–344
24. Siminszky, B., Gavilano, L., Bowen, S. W., and Dewey, R. E. (2005) Proc. Natl. Acad. Sci. U. S. A. 102, 14919–14924
25. Gavilano, L. B., Coleman, L. P., Burnley, L., Bowman, M. L., Kalengama-liro, N. E., Hayes, A., Bush, L. P., and Siminszky, B. (2006) J. Agric. Food Chem., 54, 9071–9078
26. Gietz, D., Stige, A., Woods, R. A., and Schiestl, R. H. (1992) Nucleic Acids Res. 20, 1425–1425
27. Pompon, D., Louerat, B., Bronine, A., and Urban, P. (1996) Methods Enzymol. 272, 51–64
28. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
Evolution of Nicotine-containing Tobacco

29. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673–4680
30. Morrison, T. B., Weiss, J. J., and Wittwer, C. T. (1998) BioTechniques 24, 954–962
31. Ririe, K. M., Rasmussen, R. P., and Wittwer, C. T. (1997) Anal. Biochem. 245, 154–160
32. Paquette, S. M., Bak, S., and Feyereisen, R. (2000) DNA Cell Biol. 19, 307–317
33. Yamamoto, I., Soeda, Y., Kamimura, H., and Yamamoto, R. (1968) Agric. Biol. Chem. 32, 1341–1348
34. Siegler, E. H., and Bowen, C. V. (1946) J. Econ. Entomol. 39, 673–674
35. Riah, O., Dousset, J. C., Courriere, P., Baziard-Mouysset, G., and Ecalle, R. (1997) Nat. Prod. Lett. 11, 37–45
36. Huesing, J., Jones, D., Deverna, J., Myers, J., Collins, G., Severson, R., and Sisson, V. (1989) J. Chem. Ecol. 15, 1203–1217
37. Laue, G., Preston, C. A., and Baldwin, I. T. (2000) Planta 210, 510–514
38. Self, L. S., Guthrie, F. E., and Hodgson, E. (1964) J. Insect Physiol. 10, 907–914