ABSTRACT

ALLEN, ZACHARY K. Insight into the Genetic Control of Nicotine Biosynthesis in *Nicotiana tabacum*. (Under the direction of Dr. Ramsey S. Lewis).

Tobacco (*Nicotiana tabacum*) produces secondary metabolites such as nicotine, nornicotine, anatabine, and anabasine. Of these natural products, nicotine holds significant social relevance as it is the primary addictive component in various tobacco products such as combustible cigarettes. In tobacco plants, nicotine can exist in either the (*R*)- or (*S*)-enantiomeric form, with the (*S*)-enantiomer predominating. The (*S*)-enantiomer is reported to have increased biological activity and physiological potency in humans relative to the (*R*) enantiomer. Due to nicotine being the primary addictive chemical in combustible cigarettes, there is potential for regulatory authority-mandated lowering of nicotine levels in cigarette filler to be below 0.4 mg g^{-1}. This could be a problem for the tobacco industry as there is currently no commercial tobacco variety that consistently produces lower than 0.4 mg g^{-1} of nicotine when grown under conventional field production conditions. Gaining knowledge about the genes coding for enzymes and transcription factors that regulate or control nicotine biosynthesis may lead to genetic strategies to modify the accumulation of this alkaloid in tobacco plants.

In Chapter 1, research is reported on the relative impact of different isoforms of the BBL enzyme which are involved in one of the final stages of nicotine biosynthesis. Evaluation of lines of the standard burley tobacco cultivar TN 90 carrying induced mutations in genes coding for these BBL enzymes and genes coding for nicotine demethylase enzymes suggests that *BBLa* and *BBLb* are preferentially involved in the biosynthesis of (*S*)-nicotine, while *BBLc* is preferentially involved in the biosynthesis of (*R*)-nicotine. Through a combination of mutation breeding and transgene overexpression, we have developed tobacco genotypes that preferentially accumulate

(*R*)-nicotine. The data presented in this chapter sheds light on the genetic control of alkaloid enantiomer accumulation in *N. tabacum*.

In Chapter 2, ten nearly isogenic versions of tobacco cultivar K326 differing for the presence/absence of genetic variability influencing alkaloid accumulation were evaluated in field experiments for total alkaloid accumulation, cured leaf yields, and insect and disease resistance. In addition, nearly isogenic lines differing for the presence/absence of recessive alleles at the *Nic1* and *Nic2* loci, as well as for the presence/absence of a deleterious 5 bp mutation in *Myc2a*, were grown and evaluated in a greenhouse experiment to generate root transcriptome information to gain insight on the categories of genes transcriptionally regulated by the Nic1, Nic2, and Myc2a transcription factors. K326 *nic1/nic2/myc2a/Ct* was found to accumulate below the suggested 0.4 mg g^{-1} level of nicotine in cured leaves. This came at the expense of increased susceptibility to insects and soil-borne diseases, however. Interestingly, K326 222 was observed to exhibit significantly improved resistance to both black shank and bacterial wilt. Differentially expressed genes were identified for all comparisons between K326 *per se*, K326 *nic1/nic2,* K326 *myc2a,* and K326 *nic1/nic2/myc2a*. Results support previous studies that suggest Myc TFs such as Myc2a to be significantly involved in the regulation of nicotine biosynthesis, as well as provide insight into the advantages that certain genetic modifications can have on lowering nicotine content in tobacco plants.

Insight into the Genetic Control of Nicotine Biosynthesis in *Nicotiana tabacum*

By
Zachary Allen

A thesis submitted to the Graduate Faculty of

North Carolina State University

in partial fulfillment of the

requirements for the degree of

Master of Science

Crop Science

Raleigh, NC
2024

APPROVED BY:

Ramsey S. Lewis
Committee Chair

Ralph E. Dewey

Clyde E. Sorenson

BIOGRAPHY

Zachary Allen was born in 1998 to parents Suzanne and Gregory in Chapel Hill, NC. He attended the University of North Carolina at Wilmington, Appalachian State University, and Asheville-Buncombe Technical Community College before transferring to North Carolina State University. While being an undergraduate student studying Agroecology and Sustainable Research and Production, he worked in Dr. Jim Holland's maize breeding lab and Dr. Collen Doherty's biochemistry lab. It did not take long for him to decide that a career in plant breeding was his goal and that it would be advantageous to earn a master's degree in this field. Immediately after earning his bachelor's degree, he began plant breeding research as a MS student in the tobacco breeding and genetics lab at North Carolina State University under the direction of Dr. Ramsey Lewis. After completion of his degree, Zachary would like to find work in a related field in the Research Triangle Park.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Literature Review

Nicotine

Nicotine is a plant-derived natural compound that has been extensively researched and that is found in substantial amounts in tobacco (*Nicotiana tabacum*), as well as in various other species within the *Nicotiana* genus. In addition to nicotine, tobacco accumulates smaller amounts of other alkaloids such as anatabine, anabasine, and nornicotine. Nicotine is the predominant alkaloid found in the majority of tobacco varieties, constituting over 90% of the total alkaloid pool. Nicotine is mostly present in the tobacco plant in the (*S*) enantiomeric form, although (*R*)-nicotine can also exist in small amounts. In tobacco plants, nicotine is synthesized in the roots and subsequently translocated via the xylem to the leaves and other aerial plant parts where it is stored and used as a plant defense mechanism against herbivores (Dawson, 1942). Nicotine also holds significant social relevance as it is the primary addictive component in various tobacco products such as combustible cigarettes. The United States Food and Drug Administration (FDA) has identified and listed 93 'harmful or potentially harmful' chemical constituents found in tobacco and tobacco smoke due to their known links to carcinogenesis and addiction, as well as respiratory, cardiovascular, reproductive, and developmental toxicity (WHO, 2015). For these reasons, there has been considerable interest in tobacco harm reduction research and strategies for reducing the health burden of tobacco use. Most recently, the World Health Organization (WHO) has recommended the lowering of nicotine levels in cigarette filler to below 0.4 mg g^{-1} (WHO, 2015). Conventional tobacco cultivars typically contain between 10 – 50 mg g^{-1} nicotine on a dry weight basis, with nicotine concentrations fluctuating based on market type, growing environment, plant genetics, and stalk position (Lewis et al., 2020).

Alkaloid Biosynthesis

Alkaloids constitute a category of nitrogen-containing compounds, primarily alkaline in nature, that are typically derived biochemically from amino acids. This process involves a sequence of reactions guided by metabolic enzymes sourced from diverse protein families (Shoji, 2016). In tobacco, biosynthesis of nicotine and nornicotine requires two pathways: the pyridine ring pathway which starts from L-aspartic acid, and the pyrrolidine ring pathway which starts from arginine or ornithine (Dewey & Xie, 2013). Nicotinic acid (NA) is formed through the pyridine ring pathway, while *N*-methyl-Δ^1-pyrrolinium is formed through the pyrrolidine ring pathway, and nicotine is the product of the condensation of these two ring structure compounds (Dewey & Xie, 2013).

Pyridine Ring Pathway

Nicotinic acid is produced via the pyridine ring pathway, which begins with the oxidation of aspartic acid to form α-iminosuccinic acid in a reaction catalyzed by L-aspartate oxidase (AO) (Guo et al., 2021). Quinolinate synthase (QS) facilitates the condensation of glyceraldehyde 3-phosphate with α-iminosuccinic acid in the following step, resulting in the formation of quinolinic acid (Katoh et al., 2006). Quinolinate phosphoribosyl transferase (QPT) plays a crucial role in the subsequent step by catalyzing the reaction between quinolinic acid and phosphoribosyl pyrophosphate. This enzymatic process leads to the formation of nicotinic acid mononucleotide (NAMN) (Wagner, 1984). Although the precise pathway of converting NAMN to nicotinic acid is still unknown, three potential pathways have been identified. Possible pathways include those where (1) NAMN adenyltransferase and nicotinamide adenine dinucleotide (NAD) synthetase catalyze the cycle, (2) nicotinic acid phosphoribosyltransferase facilitates direct phosphoribose transfer, and (3) nicotinic acid riboside kinase and nucleoside phosphorylase are involved in successive ribosylation and phosphorylation via an intermediary nicotinic acid riboside (Ashihara et al., 2005; Noctor et al., 2006; Zheng et al., 2004). In the pyridine ring pathway, QPT has been identified as the rate-limiting enzyme (Saunders & Bush, 1978). *N. tabacum* has two genes in its genome that encode for QPT, and they are usually referred to as *QPT1* and *QPT2*. *QPT1* is believed to be expressed in all plant tissues while *QPT2* is predominantly transcribed in root tissues and exhibits an increase in expression in response to wounding or Methyl Jasmonate (MeJA) treatments (Shoji & Hashimoto, 2011; Ryan et al., 2012). Firko (2001) described the use of antisense technology targeting *QPT2* to develop a low-nicotine line designated as "Vector 21-41." This line was evaluated in field experiments in comparison with wild-type Burley 21. Leaves harvested from the antisense line contained just 0.14 mg g^{-1} of nicotine which was estimated to be a 92.5% reduction in nicotine accumulation (Firko, 2001). Since *QPT2* is root-specific, there may have been minimal impact on the growth and yield of "Vector 21-41" (Firko, 2001). Khan et al. (2017) targeted sequences conserved in both *QPT1* and *QPT2* using an RNA interference (RNAi) method, and reductions in nicotine and anabasine accumulation were observed in addition to significant alterations in growth patterns and phenotypes of the transgenic tobacco plants. Most recently, Smith et al. (2022) used CRISPR/Cas9 technology to silence *QPT2* in K326 and TN90 tobacco varieties. The authors reported 91% - 96% reductions in nicotine content and similar reductions in nornicotine content. Anatabine and anabasine concentrations were not reported because their levels were too low to quantify in the majority of *qpt2* mutant individuals. In the greenhouse environment, only modest differences in growth phenotypes were observed in mutant *qpt2* individuals compared to the controls. However, when the same genotypes were grown in the field environment, the authors reported extremely suppressed growth among all *qpt2* mutant individuals. Reportedly, none of the *qpt2* mutant individuals grew taller than 30 cm, nor did they flower, making it not possible to collect meaningful alkaloid data (Smith et al., 2022). These results call into question the true nature of previously reported "Vector 21-41", and indicate that knocking out *QPT2* function is not a viable strategy for obtaining low-nicotine tobacco cultivars.

Pyrrolidine Ring Pathway

There are two paths in the initial steps of the pyrrolidine ring pathway that lead to the formation of putrescine, the first being the direct decarboxylation of ornithine via ornithine decarboxylase (ODC) (Guo et al., 2021). The second route involves a three-step process in which arginine is first decarboxylated by arginine decarboxylase (ADC) to form agmatine which, in turn, is hydrolyzed by agmatine iminohydrolase to form the intermediate *N*-carbamoyl putrescine. Lastly, *N*-carbamoylputrescine amidohydrolase catalyzes the formation of putrescine (Guo et al., 2021). Once putrescine has formed, it is catalyzed by putrescine methyltransferase (PMT) to form *N*-methylputrescine (MP), which is then converted to 4-methylaminobutanal through oxidative deamination catalyzed by *N*-methylputrescine oxidases (MPO) (Guo et al., 2021). To complete the pyrrolidine ring structure, 4-methylaminobutanal spontaneously cyclizes to produce the *N*-methyl-Δ^1-pyrrolinium cation (Dewey & Xie, 2013).

All of the aforementioned enzymes in the pyrrolidine pathway have been well studied at the molecular level and manipulated to alter the nicotine accumulation in tobacco (Guo et al., 2021). *NtODC1* and *NtODC2* are two genes that encode for ODC. DeBoer et al. (2011) observed a reduction in nicotine levels and increases in anatabine levels when *ODC* transcription was downregulated in transgenic tobacco plants. The same lab did not observe meaningful changes in alkaloid levels when ADC transcription was downregulated (DeBoer et. al, 2011).

It remains unclear how the pyrrolidine and pyridine rings are condensed in the late stages of nicotine biosynthesis. Two oxidoreductases, coded by *A622* and *BBL* gene families are hypothesized to be involved in the final stages of nicotine biosynthesis. Their precise biochemical roles in the pathway have yet to be determined, however (Kajikawa et al., 2011). Further analysis of these two gene families is discussed in future sections of this review.

Nic1/Nic2

The *Nic1* (originally called the *A* locus) and *Nic2* (originally called the *B* locus) loci have been well studied for their effect on alkaloid accumulation in tobacco. In the 1930s, several low-nicotine accumulating tobacco genotypes were identified amongst Cuban cigar tobacco varieties. Since these findings, it has been of great interest to scientists to utilize these materials in the development of low-nicotine tobacco cultivars. The first attempt was made by Legg et al. (1969) in which Burley 21, a commercially acceptable variety that contains leaf nicotine content between 3.5 to 4.0% on a dry weight basis, was used as the recurrent parent in a backcross breeding program to transfer the low alkaloid trait from a Cuban cigar variety. The product of this backcrossing effort was a line designated as LA Burley 21 which has similar morphological characteristics to Burley 21, but exhibits low total alkaloid accumulation (Legg et al. 1969). Through genetic analysis of Burley 21 and LA Burley 21, Legg et al. (1969) identified variability at two unlinked loci (*Nic1*and *Nic2*) as being primarily responsible for the low-nicotine phenotype. The two loci were shown to have dose-dependent effects on nicotine content in tobacco plants, with the *Nic1* locus having a much larger phenotypic effect than the *Nic2* locus

(Hibi et al., 1994; Chintapakorn & Hamill, 2003). In tobacco root cultures, the wild type (*Nic1/Nic1 Nic2/Nic2*) genotype exhibited a nicotine content of 20 mg g^{-1} dry weight, while the concentration in *nic1/nic1 nic2/nic2* mutant materials was found to be less than 4 mg g^{-1} (Hibi et al., 1994). This low alkaloid accumulation comes at the expense of other desirable agronomic traits, however. Through the evaluation of 10 tobacco cultivars and their corresponding low alkaloid isolines, Chaplin and Weeks (1976) observed significant decreases in yield, cured leaf quality, and reducing sugars associated with the low alkaloid trait.

A follow-up study to better characterize the dosage effects was conducted by Legg and Collins (1971), who generated nearly-isogenic lines with different alkaloid levels through either selfing or backcrossing to LA Burley 21 F_2 plants that contained alkaloid levels expected of the desirable genotypes. The lines they generated were: high alkaloid (*Nic1/Nic1 Nic2/Nic2*), high intermediate (*Nic1/Nic1 nic2/nic2*), low intermediate (*nic1/nic1 Nic2/Nic2*), and low alkaloid (*nic1/nic1 nic2/nic2*) (Legg and Collins, 1971). Results from the alkaloid analysis revealed that the *Nic1* locus had a dosage effect 2.4 times that of the *Nic2* locus. It is important to note that this study was conducted across three very different testing locations concerning typical burley tobacco growing environmental conditions. Despite these very different growing conditions, no significant genotype x environment interactions were observed (Legg & Collins, 1971).

The effects that the genetic variability at the *Nic1* and *Nic2* loci have on expression of genes associated with nicotine biosynthesis have been characterized via enzymatic assays. PMT, MPO, and QPT are three enzymes whose levels of activity in root tissues have been shown to be strongly correlated with nicotine accumulation (Mizusaki et al., 1973). Saunders and Bush (1979) conducted enzymatic assays on these three enzymes to better characterize their levels of activity related to genetic variability at the *Nic1* and *Nic2* loci. The assay results from the

Saunders and Bush (1979) study concluded that recessive alleles at either locus significantly reduced the activity of all three enzymes, suggesting that the *Nic* loci code for transcriptional regulators of these nicotine biosynthetic enzymes, although the exact nature of this regulation was not well characterized.

Hibi et al. (1994) later conducted a study involving burley tobacco genotypes with all the different possible allelic combinations at the *Nic1* and *Nic2* loci. They identified two transcript classes that were significantly downregulated in double homozygous recessive individuals and designated them as *A411* and *A622* (Hibi et. al, 1994). *A411* was determined to be the gene that codes for PMT. Although the precise function of *A622* was not characterized, it was determined to be a PIP-family isoflavone reductase-like protein (Kajikawa et al., 2009). Both *A411* and *A622* showed transcript levels that correlated well with both nicotine accumulation as well as with activity of enzymes previously described to have major roles in the nicotine biosynthetic pathway (Saunders and Bush, 1979). The fact that enzymatic activity of PMT and QPT, as well as transcript levels of *A622* were simultaneously decreased by recessive alleles at either the *nic1* or *nic2* loci, provides further evidence that multiple enzymes involved in nicotine biosynthesis are affected by the variability at the *nic1*/*nic2* loci (Saunders and Bush, 1979). Hibi et al. (1994) also reported observing strong synergistic effects between the two recessive alleles and suggested that the *Nic1* and *Nic2* loci are likely duplicates that originated from tobacco's two progenitors species *N. sylvestris* and *N. tomentosiformis*.

Kidd et al. (2006) attempted to identify additional genes that may be regulated by genes at the *Nic1* and *Nic2* loci through a gene transcript analysis of root tissues from Burley 21 and LA Burley 21. In total, 33 genes were identified based on observing changes in expression levels resulting from MeJA and/or indolebutyric acid (IBA) treatment. Of the 33 identified genes, only four (*A622, PMT, QPT*, and *QS*) genes were found to be involved in the biosynthesis of nicotine, while the remaining identified genes were commonly known to be upregulated in response to stress (Kidd et al., 2006).

Through transcriptome-based identification of genes at the *Nic2* locus, Shoji et al. (2010) revealed this locus to be comprised of a cluster of transcription factors (TF) genes from the ethylene response factor (ERF) subfamily. Most notably, seven *ERF* genes were identified and observed to be deleted in the *nic2* recessive mutant genotype of LA Burley 21 (Shoji et al., 2010). Of these seven Nic2-ERFs, ERF 189 was concluded to have the largest functional role in directly activating at least 12 genes associated with nicotine biosynthesis due to specific binding to the promoter regions of these genes (Shoji et al., 2010). Although suppression of the Nic2-ERFs reduced nicotine content in tobacco, significant nicotine accumulation still occurred due to the major effect of genes residing at the *Nic1* locus (Kajikawa et al., 2017).

The *Nic1* locus was found to contain several *ERF* genes related to those found at *Nic2,* including *ERF199* which is a homologue of *ERF189* (Qin et al., 2021). *ERF199* was identified as playing a major role in alkaloid biosynthesis in tobacco due to its high expression levels in nicotine producing roots. The expression of both of these *ERFs* is most strongly, but not exclusively, found in the roots and neither are induced by salt stress as compared to all the other identified *ERFs* in which expression levels are low in the roots and induced by abiotic stress (Shoji and Hashimoto, 2015; Kajikawa et al., 2017). Hayashi et al. (2020) used CRISPR/Cas9-mediated editing to introduce knockout mutations of *ERF189* and *ERF199* in tobacco, which resulted in drastic reductions in the levels of tobacco alkaloids accumulated and with no reported major effect on plant growth. Consistent with the findings of Kajikawa et al. (2017), loss of function of only *ERF189* has a much milder effect on the alkaloid phenotype of the plant as compared to loss of function of *ERF199*. Most recently, Shoji et al. (2022) evaluated 11 previously identified low-nicotine tobacco genotypes and characterized the sequence differences of the *Nic1*/*Nic2* alleles. Of the 11 evaluated lines, 6 contained deletions of either *ERF*, and 4 of the 6 lines contained the *ERF199* deletion. No line contained deletions of both *ERFs*, however. Sequencing analysis suggested that deletions of distal genomic regions can suppress the expression of *ERF199*, and the naturally occurring *nic1* allele in some germplasm contains deletions only at positions distal to *ERF199* (Shoji et at., 2022). The same study concluded the function of *ERF189* and *ERF199* is nearly equivalent to each other and that there is no reciprocal regulation between these two genes. Due to the deletions of genes being distal to *ERF199*, it is reasonable to believe that there are additional variations of the *Nic1* allele that may contribute to the low-nicotine trait.

PMT

The first dedicated step of the nicotine biosynthesis pathway involves PMT, which serves as the key catalyst for the transition from primary to secondary metabolite production in alkaloid-producing plants (Mizusaki et al., 1972, 1973). The preceding stages in the pathway are common in plants, where putrescine can be produced from either arginine or ornithine, with arginine being the primary source (Hashimoto & Yamada, 1994). In species that do not require *N*-methylputrescine (MP) for the synthesis of alkaloids, putrescine undergoes conversion into different polyamines like spermidine and spermine, which play a role as growth regulators. However, in species that produce alkaloids such as nicotine and tropanes, a portion of putrescine is directed towards these specific pathways (Hashimoto & Yamada, 1994).

Five *PMT* gene cDNA fragments were isolated by screening a tobacco genomic library and designated as *NtPMT1a*, *NtPMT1b*, *NtPMT2, NtPMT3*, and *NtPMT4* (Riechers & Timko, 1999). All five of the full-length cDNA fragments were very similar except for exon 1, where a highly conserved 11 amino acid repeat motif is found (Hashimoto & Yamada, 1994). Differences in the repeat motif have led to better characterizations of the origin of these *PMT* isoforms found in the tobacco genome. *PMT* sequence information was obtained for three *PMT* isoforms from *N. sylvestris (NsPMT1, NsPMT2, and NsPMT3)*, and one *PMT* isoform from both *N. tomentosiformis (NtomPMT)* and *N. otophora (NotoPMT)* (Hashimoto & Yamada, 1994; Riechers & Timko, 1999). Analysis of differences between the repeat motifs found on exon 1 led to the conclusion that *NtPMT2*, *NtPMT3,* and *NtPMT4* originated from *N. sylvestris, NtPMT1a* originated from *N. tomentosiformis,* and *NtPMT1b* likely originated from *N. otophora* (Hashimoto & Yamada, 1994; Riechers & Timko, 1999). Exon 1 of the *N. otophora PMT* gene (*NtPMT1b*) was shown to be identical in length to exon 1 of the *N. tomentosiformis PMT* gene (*NtPMT1a*), but there are seven base changes and five amino acid sequence differences between the two. Due to these significant sequence differences, these two genes likely did not originate from the same genome and therefore provided support for a model in which *N. tabacum* is the product of a cross between *N. sylvestris* and a *N. tomentosiformis* × *N. otophora* hybrid (Riechers & Timko, 1999).

All five *PMT* genes were shown to be preferentially expressed in the root tissues through Northern blot and RT-PCR analysis (Hibi et al., 1994; Hashimoto et al., 1998; Riechers & Timko, 1999). The expression of the *PMT* genes is differentially regulated in high and low-alkaloid tobacco varieties, and induced by the removal of inflorescence, wounding, and MeJA treatment (Hibi et al., 1994; Riechers & Timko, 1999). Plants in which the *PMT* gene family was silenced via antisense, inverse repeat silencing, or RNAi exhibited up to 97% reductions in nicotine accumulation, and a further reduction in nicotine was observed when the roots were treated with MeJA (Chintapakorn & Hamill, 2003; Steppuhn et al., 2004; Wang et al., 2008). Interestingly, when *PMT* expression was downregulated, anatabine accumulation doubled in the roots which may be a mechanism for detoxification due to nicotinic acid having limited opportunity to condense with methylputrescine in these suppressed plants (Chintapakorn & Hamill, 2003). Following topping, *PMT* expression was only stimulated in the standard Burley 21 variety and not in the LA Burley 21 variety, further suggesting that the expression of *PMT* is under the control of the *Nic1* and *Nic2* loci (Riechers & Timko, 1999).

MPO

The second dedicated step of the nicotine biosynthesis pathway involves the conversion of *N*-methylputrescine to 4-methylaminobutanal through oxidative deamination catalyzed by *N*-methylputrescine oxidases (MPO) (Guo et al., 2021). In the tobacco genome, MPO enzymes are encoded by two homologous genes *MPO1* and *NtDAO1* (Heim et al., 2007; Katoh et al., 2007). *MPO1* was isolated by designing degenerate primers using conserved amino acid sequences from copper amine oxidases identified in various plant species (Heim et al., 2007). In the same year, Katoh et al. (2007) identified the same *MPO1* gene by analyzing nicotine biosynthesis-associated genes that were down-regulated in the *nic1*/*nic1 nic2/nic2* mutant genotypes. In addition to *MPO1*, Katoh et al. (2007) discovered a second *MPO* gene (*NtDAO1*), that has a predicted protein product 88% identical to that of MPO1. The function of MPO was confirmed by in *vitro* enzyme assays of the recombinant MPO1 protein produced in *E. coli*, and MPO activity was shown to be most specific to *N*-methylputrescine (Katoh et al., 2007). *NtDAO1* expression in *E.*

coli was minimal in comparison to *MPO1*, which is consistent with the Shoji & Hashimoto (2008) study that showed *MPO1* expression levels in tobacco root hairs to be higher than that of *NtDAO1*. The expression of *MPO1* is controlled by MeJA induction and relies on the presence of ERF transcription factors. Conversely, *NtDAO1* is constitutively expressed at relatively low levels in tobacco (Naconsie et al., 2014). Studies that analyzed phenotypic effects of silencing the expression of *MPO1* in hairy roots via targeted RNAi reported a 93%-99% reduction in nicotine accumulation accompanied by an increase in anatabine production (Shoji & Hashimoto, 2008).

A622

One of the two oxidoreductases required for the later steps of nicotine biosynthesis is A622, although its exact biochemical function remains unknown. In studies analyzing the expression levels of the *A622* gene family, it was revealed to have joint expression levels with that of *PMT* (Hibi et al., 1994). In addition, *A622* transcript levels were coordinately upregulated with genes that play a crucial role in nicotine synthesis when subjected to foliar wounding and MeJA treatment in both *Nic1*/*Nic1 Nic2/Nic2* (*AABB*) and *nic1*/*nic1 nic2/nic2* (*aabb*) genotypes (Cane et al., 2005).

Several studies have been conducted using RNAi constructs to suppress the expression of *A622*. DeBoer et al. (2009) found that in hairy roots of *N. glauca* containing an RNAi construct, anabasine and nicotine accumulation were nearly eliminated. Unlike the controls in this experiment, MeJA treatment did not restore the ability to accumulate either alkaloid in the transgenic lines. These results confirmed that *A622* plays a role in the pyridine alkaloid synthesis. To support these findings, Kajikawa et al. (2009) found that downregulating *A622* through RNAi significantly reduced the accumulation of nicotine and other alkaloids as well as severely inhibiting root growth. Furthermore, Burner et al. (2022) reported that in double homozygous *A622* mutant versions of K326, nicotine levels were severely reduced and levels of nornicotine, anabasine, and anatabine were mostly undetectable. This further supports previously mentioned studies that suggest the role of the A622 enzyme in the biosynthesis of the tobacco pyridine alkaloids.

MYC2

In addition to genes coding for ERF transcription factors identified at the *Nic1* and *Nic2* loci, *MYC2* genes also code for a class of TFs known to be associated with inducing nicotine biosynthesis in tobacco. This class of TFs has a basic helix-loop-helix domain and was originally identified as positive regulators for the expression of several JA-inducible genes involved in alkaloid biosynthesis in *N. benthamiana* (Todd et al., 2010). Four genes (*MYC1a, MYC1b, MYC2a*, and *MYC2b*) make up the *MYC2* gene family. Shoji and Hashimoto (2011) showed that the *N. tabacum MYC2* orthologs function very similarly to the genes originally identified in *N. benthamiana*, and further confirmed their involvement in nicotine biosynthesis, specifically. The same study reported that MYC2 TFs recognize the G-box sequences located in promoter regions of major nicotine biosynthesis genes (Shoji and Hashimoto, 2011). When MYC2 expression was suppressed through RNAi-mediated induction, transcription of JA-responsive genes was suppressed in the tobacco roots, as well as *ERF* genes known to be associated with the *Nic2* locus (Shoji and Hashimoto, 2008). Sui et al. (2020) observed an 80% reduction in nicotine accumulation in tobacco leaves possessing an induced knockout mutation in *MYC2a*. This study did not find the *MYC2a* gene to be regulated by the *Nic1*/*Nic2* loci. The authors reported a downregulation of *ERF189* when *MYC2a* was silenced, however, suggesting that *MYC2a* functions upstream of ERF transcription factors coded by the *Nic2* locus in influencing nicotine biosynthesis (Sui et al., 2021). Yang et al. (2023) confirmed the previous results that nicotine accumulation decreases when *MYC2a* is silenced by observing similar results when *MYC2a/b* was co-silenced in *N. attenuata*. Interestingly, this study reported increased growth and fitness of *MYC2a/b* co-silenced plants, compared to the control. Although the increased growth was accompanied by decreased metabolic potential, Yang et al. (2023) reported that co-silenced *MYC2a/b* plants can adjust their metabolic productivity based on the growth/defense status of their neighbors as previous studies have suggested for other species (Broz et al., 2010; Yang et al., 2023).

Through QTL mapping, Burner et al. (2022) identified a naturally-occurring 5 base pair deletion in the *MYC2a* gene in tobacco accession TI 313 (Burner et al., 2022). The result of this deletion is a reduction in nicotine levels. These plants were not tested in large-scale field evaluations, however, nor were plants evaluated in which the *MYC2a* mutation was combined with the recessive *nic1*/*nic2* alleles. A nearly isogenic line (NIL) of K326 that is homozygous for the *myc2a* mutation as well as a NIL that is homozygous for the *myc2a* mutation in combination with the recessive *nic1*/*nic2* alleles have since been developed and are currently being field tested. The results and details of that study are presented in Chapter 3 of this thesis.

BBL

The second oxidoreductase that is believed to participate in a final key step of the nicotine biosynthesis pathway is coded by a gene family known as the *Berberine Bridge Like* (*BBL)* gene family (Kajikawa et al., 2011). The BBL enzyme is a flavin-containing oxidase that is believed to be involved in the final oxidation step for nicotine production, although its exact function is presently unknown (Kajikawa et al., 2011). There are six genes in this particular gene family, with *BBL-a*, *BBL-b*, and *BBL-c* being expressed to significant degrees, while *BBL-d1*, *BBL-d2,* and *BBL-e* are expressed at greatly reduced levels (Kajikawa et al., 2017). Kajikawa et al. (2011) identified these genes based on their lack of expression in the LA Burley 21 variety containing the *nic1*/*nic2* mutations as compared to wild-type Burley 21. Transgenic approaches involving the use of RNAi-based constructs to suppress *BBL* expression in hairy roots resulted in a 90% reduction in nicotine content (Kajikawa et al., 2011). When *BBL* expression was suppressed, there was a noticeable accumulation of the novel metabolite dihydrometanicotine (DMN) that was confined in the root tissues (Kajikawa et al., 2011). Lewis et al. (2015) later revealed that the DMN alkaloid is also present in the leaves of transgenic RNAi tobacco plants, although the levels were very low. Previously, De Clercq and Truhaut, (1962) identified DMN as being a metabolite of nicotine, and Neurath et al. (1966) reported the presence of DMN in cigarette smoke, but it had never been reported in tobacco plants. The structural arrangement of DMN, where the *N*-(methylamino)butyl moiety is linked to the C-3 position of the pyridine ring, indicates that the fusion of the pyridine ring and the *N*-methylpyrroline ring takes place before the BBL-catalyzed oxidation of an unidentified reaction intermediate (Kajikawa et al., 2011).

Most recently, Lewis et al. (2020) developed burley tobacco genotypes with different combinations of EMS-induced mutations in the three major *BBL* genes. In addition, nearly isogenic versions of a historically important burley tobacco cultivar (TN 90) homozygous for deleterious mutations in three known *Nicotine Demethylase* (*NND*) genes (designated as TN90 SRC) were analyzed in combination with different combinations of the *BBL* knockout mutations. The results showed that leaf nicotine was significantly reduced in genotypes homozygous for combined EMS-induced mutations in *BBL-a* and *BBL-b*, with an additional slight reduction in lines homozygous for combined mutations in *BBL-a*, *BBL-b*, and *BBL-c* (Lewis et al., 2020). This reduction of nicotine was accompanied by a significant decrease in both cured leaf yield and quality, however. In the SRC genetic background, only the double and triple homozygous mutant genotypes (TN 90 SRC 220 and TN 90 SRC 222) showed slightly significant reductions in nicotine content compared to TN 90 SRC (Lewis et al., 2020). No genetic combination reduced nicotine levels to zero, suggesting that there may be an additional minor pathway contributing to nicotine biosynthesis in tobacco. The role of BBL enzymes on the enantiomeric selectivity of nicotine is discussed in detail in a dedicated chapter of this thesis.

NND

Nornicotine is the demethylated derivative of nicotine and typically represents less than 5% of the total pyridine alkaloid pool in tobacco (Bush et al., 2001). Nicotine Demethylase (NND) is the enzyme that facilitates the conversion of nicotine to nornicotine. In *N. tabacum*, this process predominantly occurs during senescence and curing of mature leaves (Wernsman and Matzinger, 1968). As the tobacco leaf undergoes curing and processing, a chemical reaction can take place between nornicotine and nitrosating agents, resulting in the creation of a compound known as *N′*-nitrosonornicotine (NNN). NNN is the most abundant tobacco-specific nitrosamine (TSNA) found in tobacco smoke and has been strongly linked to several types of cancer in numerous animal studies (Hecht & Hoffmann, 1989; Hecht, 2003). Recent studies have suggested that in addition to nornicotine being the precursor for NNN, nornicotine that is found in tobacco smoke may have additional undesirable health effects. In a study conducted by Dickerson and Janda (2002), it was shown that nornicotine induces abnormal protein glycation within cellular structures. When the plasma of smokers and nonsmokers was analyzed for concentrations of nornicotine-modified proteins, the plasma of smokers was found to contain a much higher concentration. Additionally, the study demonstrated that nornicotine can chemically modify commonly prescribed steroid medications like prednisone through covalent bonding. This process has the potential to impact both the effectiveness and potential side effects of these drugs. Two studies have also suggested that nornicotine found in tobacco products may be linked to age-related macular degeneration, birth defects, and periodontal disease (Borgan et al., 2005; Katz et al., 2005).

For these reasons, it has been of great interest to tobacco breeders to better understand the nicotine conversion process and to reduce it in commercial tobacco cultivars. The conversion of nicotine to nornicotine can be an unstable trait in *N. tabacum*, particularly in the burley tobacco market class (Mann et al., 1964). This means that in a field of a commercial tobacco cultivar, individual plants can contain commercially undesirable rates of nicotine conversion.

An important breakthrough towards understanding the genetic basis for nornicotine biosynthesis was made with the discovery of the *CYP82E* gene family. This gene family was identified by conducting a microarray analysis on genes that were showing varying expression patterns among pairs of isogenic *N. tabacum* lines with distinct conversion rates (referred to as 'converter' and 'nonconverter' plants) (Siminszky et al., 2005). In total, six genes were identified as being members of this gene family, all sharing > 90% sequence and amino acid identity. *CYP82E2* and *CYP82E3* were the first genes in this family to be identified, but were shown to have no NND function within *N. tabacum* (Siminszky et al., 2005).

Subsequent research has provided interesting insights into the molecular evolution of the NND genes in tobacco. In the species that are believed to be the diploid ancestral progenitors of allotetraploid *N. tabacum* (*N. tomentosiformis* and *N. sylvestris),* nornicotine is the predominant alkaloid as opposed to nicotine. In *N. sylvestris, CYP82E2* serves as a functional NND gene, but a two-amino acid substitution has caused this gene to become nonfunctional in modern-day *N. tabacum* (Charkrabarti et al., 2007). In *N. tomentosiformis*, *CYP82E3* and *CYP82E4* were identified as being closely linked and functional *NND* genes. A single amino acid mutation in *CYP82E3* gene and the transcriptional repression of *CYP82E4* has led to the loss of functionality of these genes in modern tobacco, however (Gavilano et al., 2007). This suggests that following the speciation of the initial *N. tabacum* allotetraploid(s), there may have been selection for mutations in three NND genes to alter a species that originally predominantly accumulated nornicotine into one that predominantly accumulates nicotine.

Interestingly, RNAi-mediated suppression of *CYP82E4* resulted in an approximately 80% reduction in nicotine conversion as compared to controls (Lewis et al., 2008). This observation led researchers to believe that there may be additional *CYP82E4*-like *NND* genes in the tobacco genome. This supposition was supported when *CYP82E5v2* was identified and characterized as being a functional *NND* gene that shares 92.7% DNA sequence identity with *CYP82E4* and has very low transcription levels in green leaves in both converter and nonconverter genetic backgrounds (Gavilano & Siminszky, 2007).

Lewis et al. (2010) conducted a follow-up study to determine the effects on the nornicotine profile of a tobacco plant carrying EMS-induced knockout mutations in both *CYP82E4* and *CYP82E5v2*. When individual plants contained a combination of knockout mutation of both *CYP* genes, they were not significantly different from individuals containing the knockout mutation of only *CYP82E4*. This suggests that the *CYP82E4* gene is the dominant gene in the converter phenotype. Interestingly, plants containing the combination knockout mutation displayed a conversion rate that was about three times greater than what was seen in the *CYP82E4* RNAi-suppressed lines. Because a broader set of genes was suppressed by the RNAi vector compared to the EMS mutation, it was reasonable to believe there was yet another *CYP82E4*-like *NND* gene. *CYP82E10* was characterized as being the previously unidentified *NND* gene (Lewis et al., 2010). Currently, there are three genes in *N. tabacum* known to encode enzymes that demethylate nicotine to form nornicotine: *CYP82E4, CYP82E5v2*, and *CYP82E10* (Lewis et al., 2010). All three of these genes are part of the cytochrome P450 superfamily of monooxygenases in which they get the "CYP" nomenclature (Siminszky et al., 2005; Xu et al., 2007). In individual plants containing EMS-induced combination knockout mutations of all three *CYP* genes, nornicotine content was reduced to levels seen in the best RNAi suppression lines (Lewis et al. 2010). This study concludes that altering these three NND genes is a potential non-transgenic solution to reducing NNN accumulation in tobacco.

References

Ashihara, H. S. (2005). *De novo* and salvage biosynthetic pathways of pyridine nucleotides and nicotinic acid conjugates in cultured plant cells. *Plant Science*, *169*(1), 107-114.

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. *Proceedings of the National Academy of Sciences, 102*(30), 10433-10438.

Broz, A. K., Broeckling, C. D., De-la-Peña, C., Lewis, M. R., Greene, E., Callaway, R. M., Vivanco, J. M. (2010). Plant neighbor identity influences plant biochemistry and physiology related to defense. *BMC Plant Biology*, *10*(1), 1-14.

Burner, N., Kernodle, S. P., Steede, T., & Lewis, R. S. . (2022). Editing of *A622* genes results in ultra-low nicotine whole tobacco plants at the expense of dramatically reduced growth and development. *Molecular Breeding, 42*(4), 20.

Burner, N., McCauley, A., Pramod, S., Frederick, J., Steede, T., Kernodle, S. P., & Lewis, R. S. (2022). Analyses of diverse low alkaloid tobacco germplasm identify naturally occurring nucleotide variability contributing to reduced leaf nicotine accumulation. *Molecular Breeding, 42*(1), 4.

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. *Recent Advances in Tobacco Science, 27*, 23-46.

Cane, K. A., Mayer, M., Lidgett, A. J., Michael, A. J., & Hamill, J. D. (2005). Molecular analysis of alkaloid metabolism in *AABB* v. *aabb* genotype *Nicotiana tabacum* in response to wounding of aerial tissues and methyl jasmonate treatment of cultured roots. *Functional Plant Biology*, *32*(4), 305-320.

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 Phytologist*, *175*(3), 565-574.

Chaplin, J. F. (1975). Registration of LAFC 53 tobacco germplasm (Reg. No. GP13). *15*(2), 282-282.

Chaplin, J. F., & Weeks, W. W. (1976). Association between percent total alkaloids and other traits in flue-cured tobacco. *Crop Science, 16*(3), 416-418.

Chintapakorn, Y., & Hamill, J. D. (2003). Antisense-mediated down-regulation of putrescine *N*-methyltransferase activity in transgenic *Nicotiana tabacum* L. can lead to elevated levels of anatabine at the expense of nicotine. *Plant Molecular Biology, 53*, 87-105.

Dawson, R. F. (1942). Accumulation of nicotine in reciprocal grafts of tomato and tobacco. *American Journal of Botany*, 66-71.

DeBoer, K. D., Dalton, H. L., Edward, F. J., & Hamill, J. D. (2011). RNAi-mediated down-regulation of ornithine decarboxylase (ODC) leads to reduced nicotine and increased anatabine levels in transgenic *Nicotiana tabacum* L. *Phytochemistry*, *72*(4-5), 344-355.

Declercq, M., & Truhaut, R. (1962). Sur les mecanismes daction de la nicotine-effet de la nicotine et de ses metabolites sur la biosynthese du diphosphopyridine-nucleotide (DPN). *Bulletin De La Societe De Chime Biologique, 44*(2-3), 227.

Dewey, R. E., & Xie, J. (2013). Molecular genetics of alkaloid biosynthesis in *Nicotiana tabacum*. *Phytochemistry, 94*, 10-27.

Dickerson, T. J., & Janda, K. D. (2002). A previously undescribed chemical link between smoking and metabolic disease. *Proceedings of the National Academy of Sciences, 99*(23), 15084-15088.

Firko, M. J., USDA, A., & Firko, D. D. (2001). Vector Tobacco (USA) Ltd. *Revue/cl*, 1-15.

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. *Journal of Biological Chemistry, 282*(1), 249-256.

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 and Cell Physiology, 48*(11), 1567-1574.

Guo, Y., Hiatt, E., Bonnet, C., Kudithipudi, C., Lewis, R. S., Shi, H., Patra, B., Zhao, X., de Borne, F. D., & Gilles, T. (2021). Molecular regulation and genetic manipulation of alkaloid accumulation in tobacco plants. *Studies in Natural Products Chemistry, 70*, 119-149.

Hashimoto, T., Shoji, T., Mihara, T., Oguri, H., Tamaki, K., Suzuki, K. I., & Yamada, Y. (1998). Intraspecific variability of the tandem repeats in *Nicotiana* putrescine *N*-methyltransferases. *Plant Molecular Biology, 37*, 25-37.

Hashimoto, T., & Yamada, Y. (1994). Alkaloid biogenesis: molecular aspects. *Annual Review of Plant Biology, 45*(1), 257-285.

Hecht, S. S. (2003). Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nature Reviews Cancer, 3*(10), 733-744.

Hecht, S. S., & Hoffmann, D. (1989). The relevance of tobacco-specific nitrosamines to human cancer. *Cancer Surveys, 8*(2), 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*(4), 454-463.

Hibi, N., Higashiguchi, S., Hashimoto, T., & Yamada, Y. (1994). Gene expression in tobacco low-nicotine mutants. *The Plant Cell*, *6*(5), 723-735.

Kajikawa, M., Hirai, N., & Hashimoto, T. (2009). A PIP-family protein is required for biosynthesis of tobacco alkaloids. *Plant Molecular Biology, 69*, 287-298.

Kajikawa, M., Shoji, T., Kato, A., & Hashimoto, T. (2011). Vacuole-localized berberine bridge enzyme-like proteins are required for a late step of nicotine biosynthesis in tobacco. *Plant Physiology*, *155*(4), 2010-2022.

Katoh, A., Shoji, T., & Hashimoto, T. (2007). Molecular cloning of N-methylputrescine oxidase from tobacco. *Plant and Cell Physiology, 48*(3), 550-554.

Katoh, A., Uenohara, K., Akita, M., & Hashimoto, T. (2006). Early steps in the biosynthesis of NAD in Arabidopsis start with aspartate and occur in the plastid. *Plant Physiology, 141*(3), 851-857.

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. *Journal of Periodontology, 76*(7), 1171-1174.

Khan, S., Pandey, S. S., Jyotshna, Shanker, K., Khan, F., & Rahman, L. U. (2017). Cloning and functional characterization of *quinolinic acid phosphoribosyl transferase* (*QPT*) gene of *Nicotiana tabacum*. *Physiologia Plantarum, 160*(3), 253-265.

Kidd, S. K., Melillo, A. A., Lu, R., Reed, D. G., Kuno, N., Uchida, K., Furuya, M., & Jelesko, J. G. (2006). The *A* and *B* loci in tobacco regulate a network of stress response genes, few of which are associated with nicotine biosynthesis. *Plant Molecular Biology, 60*, 699-716.

Legg, P. D., Chaplin, J. F., & Collins, G. B. (1969). Inheritance of percent total alkaloids in *Nicotiana tabacum* L. Populations derived from crosses of low alkaloid lines with burley and flue-cured varieties. *Journal of Heredity, 60*(4), 213-217.

Legg, P. D., & Collins, G. B. (1971). Inheritance of percent total alkaloids in *Nicotiana tabacum* L. II. Genetic effects of two loci in Burley 21× LA Burley 21 populations. *Canadian Journal of Genetics and Cytology, 13*(2), 287-291.

Legg, P. D., Collins, G. B., & Litton, C. C. (1970). Registration of LA Burley 21 tobacco germplasm 1 Registration No. (GP 8). *Crop Science, 10*(2), 212-212.

Lewis, R. S., Bowen, S. W., Keogh, M. R., & Dewey, R. E. (2010). Three nicotine demethylase genes mediate nornicotine biosynthesis in *Nicotiana tabacum* L.: functional characterization of the *CYP82E10* gene. *Phytochemistry, 71*(17-18), 1988-1998.

Lewis, R. S., Drake-Stowe, K. E., Heim, C., Steede, T., Smith, W., & Dewey, R. E. (2020). Genetic and agronomic analysis of tobacco genotypes exhibiting reduced nicotine accumulation due to induced mutations in *berberine bridge like* (*BBL*) genes. *Frontiers in Plant Science, 11*, 368.

Lewis, R. S., Jack, A. M., Morris, J. W., Robert, V. J., Gavilano, L. B., Siminszky, B., Bush, L. P., Hayes, A. J., & Dewey, R. E. (2008). RNA interference (RNAi)-induced suppression of nicotine demethylase activity reduces levels of a key carcinogen in cured tobacco leaves. *Plant Biotechnology Journal*, *6*(4), 346-354.

Lewis, R. S., Lopez, H. O., Bowen, S. W., Andres, K. R., Steede, W. T., & Dewey, R. E. (2015). Transgenic and mutation-based suppression of a *berberine bridge enzyme-like* (*BBL*) gene family reduces alkaloid content in field-grown tobacco. *PloS One, 10*(2), e0117273

Mann, T. J., Weybrew, J. A., Matzinger, D. F., & Hall, J. L. (1964). Inheritance of the conversion of nicotine to nornicotine in varieties of *Nicotiana tabacum* L. and related amphiploids. *Crop Science*, *4*(4), 349-353.

Mizusaki, S., Tanabe, Y., Noguchi, M., & Tamaki, E. (1972). *N*-methylputrescine oxidase from tobacco roots. *Phytochemistry, 11*(9), 2757-2762.

Mizusaki, S., Tanabe, Y., Noguchi, M., & Tamaki, E. (1973). Changes in the activities of ornithine decarboxylase, putrescine *N*-methyltransferase and *N*-methylputrescine oxidase in tobacco roots in relation to nicotine biosynthesis. *Plant and Cell Physiology, 14*(1), 103-110.

Naconsie, M., Kato, K., Shoji, T., & Hashimoto, T. (2014). Molecular evolution of *N*-methylputrescine oxidase in tobacco. *Plant and Cell Physiology*, *55*(2), 436-444.

Neurath, G., Dünger, M., Gewe, J., Lüttich, W., & Wichern, H. (1966). Unterschung der Flüchtigen Basen des Tabakrauches. *Beitr Tabakforsch, 3*(563-571), 220.

Noctor, G., Guillaume, Q., and Bertrand, G. (2006). NAD (P) synthesis and pyridine nucleotide cycling in plants and their potential importance in stress conditions. *Journal of Experimental Botany, 57*(8), 1603-1620.

Qin, Q., Humphry, M., Gilles, T., Fisher, A., Patra, B., Singh, S. K., Li, D., & Yang, S. (2021). NIC1 cloning and gene editing generates low-nicotine tobacco plants. *Plant Biotechnology Journal, 19*(11), 2150.

Riechers, D. E., & Timko, M. P. (1999). Structure and expression of the gene family encoding putrescine *N*-methyltransferase in *Nicotiana tabacum*: new clues to the evolutionary origin of cultivated tobacco. *Plant Molecular Biology, 41*, 387-401.

Ryan, S. M., Cane, K. A., DeBoer, K. D., Sinclair, S. J., Brimblecombe, R., & Hamill, J. D. (2012). Structure and expression of the *quinolinate phosphoribosyltransferase* (*QPT*) gene family in *Nicotiana*. *Plant Science, 188*, 102-110.

Saunders, J. W., & Bush, L. P. (1979). Nicotine biosynthetic enzyme activities in *Nicotiana tabacum* L. genotypes with different alkaloid levels. *Plant Physiology, 64*(2), 236-240.

Shoji, T. (2016). Alkaloid Biosynthesis and Regulation in Plants. *Plant Specialized Metabolism* (pp. 97-130). CRC Press.

Shoji, T., & Hashimoto, T. (2011). Recruitment of a duplicated primary metabolism gene into the nicotine biosynthesis regulon in tobacco. *The Plant Journal, 67*(6), 949-959.

Shoji, T., & Hashimoto, T. (2015). Stress-induced expression of *NICOTINE2*-locus genes and their homologs encoding Ethylene Response Factor transcription factors in tobacco. *Phytochemistry, 113*, 41-49.

Shoji, T., & Hashimoto, T. (2011). Tobacco MYC2 regulates jasmonate-inducible nicotine biosynthesis genes directly and by way of the *NIC2*-locus ERF genes. *Plant and Cell Physiology, 52*(6), 1117-1130.

Shoji, T., & Hashimoto, T. (2008). Why does anatabine, but not nicotine, accumulate in jasmonate-elicited cultured tobacco BY-2 cells? *Plant and Cell Physiology, 49*(8), 1209-1216.

Shoji, T., Kajikawa, M., & Hashimoto, T. (2010). Clustered transcription factor genes regulate nicotine biosynthesis in tobacco. *The Plant Cell, 22*(10), 3390-3409.

Shoji, T., Moriyama, K., Sierro, N., Ouadi, S., Ivanov, N. V., Hashimoto, T., & Saito, K. (2022). Natural and induced variations in transcriptional regulator genes result in low-nicotine phenotypes in tobacco. *The Plant Journal, 111*(6), 1768-1779.

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. *Proceedings of the National Academy of Sciences, 102*(41), 14919-14924.

Sisson, V. A., & Severson, R. F. (1990). Alkaloid composition of the *Nicotiana* species. *Beitr. Tabakforsch. Int, 14*(6), 327-339.

Smith, W. A., Matsuba, Y., & Dewey, R. E. (2022). Knockout of a key gene of the nicotine biosynthetic pathway severely affects tobacco growth under field, but not greenhouse conditions. *BMC Research Notes, 15*(1), 1-6.

Steppuhn, A., Gase, K., Krock, B., Halitschke, R., Baldwin, I. T. (2004). Nicotine's defensive function in nature. *PLoS Biology, 2*(8), e217.

Sui, X., He, X., Song, Z., Gao, Y., Zhao, L., Jiao, F., Kong, G., Li, Y., Han, S., & Wang, B. (2021). The gene *NtMYC2a* acts as a 'master switch'in the regulation of JA-induced nicotine accumulation in tobacco. *Plant Biology, 23*(2), 317-326.

Todd, A. T., Liu, E., Polvi, S. L., Pammett, R. T., & Page, J. E. (2010). A functional genomics screen identifies diverse transcription factors that regulate alkaloid biosynthesis in *Nicotiana benthamiana*. *The Plant Journal, 62*(4), 589-600.

University of Tennessee Agricultural Experiment Station; Heggestad, Howard E.; Clayton, E. E.; Neas, M. O. (1960). *Development of Burley 21, the first wildfire-resistant tobacco variety, including results of variety trials.* Knoxville: Bulletins.

Wagner, R., & Wagner, K. G. (1984). Determination of *quinolinic acid phosphoribosyl-transferase* in tobacco. *Phytochemistry, 23*(9), 1881-1883.

Wang, P., Liang, Z., Zeng, J., Li, W., Sun, X., Miao, Z., & Tang, K. (2008). Generation of tobacco lines with widely different reduction in nicotine levels via RNA silencing approaches. *Journal of Biosciences, 33*, 177-184.

Wernsman, E. A., & Matzinger, D. F. (1968). Time and site of nicotine conversion in tobacco. *Tobacco Sci, 12*, 226-228.

World Health Organization (WHO). (2015). Advisory note: global nicotine reduction strategy: WHO Study Group on Tobacco Product Regulation.

Xu, D., Shen, Y., Chappell, J., Cui, M., & Nielsen, M. (2007). Biochemical and molecular characterizations of nicotine demethylase in tobacco. *Physiologia Plantarum, 129*(2), 307-319.

Yang, C., Bai, Y., Halitschke, R., Gase, K., Baldwin, G., & Baldwin, I. T. (2023). Exploring the metabolic basis of growth/defense trade-offs in complex environments with *Nicotiana attenuata* plants cosilenced in *NaMYC2a/b* expression. *New Phytologist*, *238*(1), 349-366.

Zheng, X. Q., Nagai, C., & Ashihara, H. (2004). Pyridine nucleotide cycle and trigonelline (N-methylnicotinic acid) synthesis in developing leaves and fruits of *Coffea arabica*. *Physiologia Plantarum, 122*(4), 404-411.

Chapter 1: Reversing the Enantiomeric Composition of Nicotine in *Nicotiana tabacum*

Abstract

Numerous plant species produce an abundance of secondary metabolites that typically exist in optically pure form but can also exist as enantiomeric mixtures. Interestingly, enantiomers can confer different biological activities. In *Nicotiana tabacum*, nicotine can exist in either the (*R*)- or (*S*)-enantiomeric form, with the (*S*)-enantiomer predominating. (*R*)-nicotine typically accounts for just 0.02% of the total nicotine pool on a dry leaf weight basis in air-cured burley tobacco. (*S*)-nicotine has been reported to have increased biological activity and physiological potency in humans relative to the *R* enantiomer. The mechanisms by which nicotine enantiomers accumulate remain largely unknown. We have researched the relative impact of different isoforms of the BBL enzyme which are involved in one of the final stages of nicotine biosynthesis. Evaluation of lines of the standard burley tobacco cultivar TN 90 carrying induced mutations in the genes coding for these BBL enzymes and genes coding for nicotine demethylase enzymes suggests that *BBLa* and *BBLb* are preferentially involved in the biosynthesis of (*S*)-nicotine, while *BBLc* is preferentially involved in the biosynthesis of (*R*)-nicotine. Through a combination of mutation breeding and transgene overexpression, we have developed tobacco genotypes that preferentially accumulate (*R*)-nicotine. The data presented in this chapter sheds light on the genetic control of alkaloid enantiomer accumulation in *N. tabacum*.

INTRODUCTION

Tobacco (*Nicotiana tabacum*), along with numerous other plant species, produces an abundance of secondary metabolites which are often referred to as natural products. These natural products are not required for survival but are usually of importance due to their involvement in interactions with other organisms such as pollinators, herbivores, or microorganisms. Some plant natural products can be isolated for their commercial value as fragrances, antimicrobial agents, or pharmacological products. In tobacco plants, nicotine is synthesized in the roots and subsequently translocated via the xylem to the leaves and other aerial plant parts where it is stored and plays a role in plant defense mechanisms against herbivores (Dawson, 1942). Nicotine also has significant social relevance as it is the primary addictive component in various tobacco products such as combustible cigarettes. Many natural products, including nicotine, can exist in the natural world as enantiomers, a pair of molecules that are mirror images of one another but that cannot be superimposed one upon the other. In most cases, natural products accumulate in optically pure form, meaning only one enantiomeric form is produced. However, enantiomeric mixtures can occasionally be found in a given species. In these cases, however, one enantiomer is still usually more prevalent than the other (Finefield et al., 2012). Interestingly, enantiomers can produce different aromas or confer different biological activities. For example, the limonene (-)-(4*S*) - enantiomer produces a turpentine or lemon odor, whereas the (+)-(4*R*)-limonene enantiomer produces the fragrance of oranges and can confer insecticidal properties (Finefield et al., 2012). Another example is with the enantiomers of gossypol (produced in cotton seeds) in which both molecules are effective insect herbivore repellants. However, the (*S*)-enantiomer has antispermatogenic effects on mammals whereas the (*R*)-enantiomer does not (Lin et. al, 2023). In cases of extraction and commercialization of plant natural products, there may be occasions where it is advantageous to alter the natural biosynthesis of a molecule of interest from one enantiomer to another.

The highly studied pyridine alkaloids of tobacco include nicotine, nornicotine, anatabine, and anabasine. Nicotine typically represents greater than 90% of the total alkaloid pool in tobacco and the remaining 10% consists of nornicotine, anatabine, and anabasine. All four of these alkaloids can exist naturally in the tobacco plant as either (*R*)- or (*S*)-enantiomeric forms that differ at the 2'C position of the pyrrolidine ring (Cai et al., 2012). (*S*)-enantiomers are naturally more prevalent than the (*R*)-enantiomers, and (*R*)-nicotine accounts for about 0.02% of the total nicotine pool in air-cured burley tobacco on a dry leaf weight basis (Zhang et al., 2018). The (*R*)-enantiomer of nicotine has been reported to be less biologically active than the (*S*)-enantiomer, and the derived TSNA (*S*)-*N*'-nitrosonornicotine (NNN) has been suggested to be more carcinogenic than the (*R*)-enantiomer (Mclntee and Hecht, 2000; Pogocki et al., 2007; Rosencrans and Young, 2017). Researchers have been uncertain as to the mechanism of (*R*)-nicotine biosynthesis, although some have speculated that it is derived from spontaneous chemical reactions (Vollheyde et. al, 2023).

Although the mechanisms of accumulation of the tobacco alkaloid enantiomers are largely unknown, some research has been reported on the control of accumulation of the enantiomers of nornicotine. Nornicotine is present in the leaves of tobacco plants as an enantiomeric mixture, with the (*R*)-enantiomer accounting for between 10% - 30% of the total nornicotine pool. Nornicotine is the demethylated derivative of nicotine and is mostly (if not entirely) produced by the enzymatic *N*-demethylation of nicotine (Lewis et al., 2010). There have been three genes identified (*CYP82E4*, *CYP82E5v2*, and *CYP82E10*) that code for nicotine demethylase enzymes that are active in the leaves or roots of the tobacco plant. These enzymes exhibit enantiomeric selectivity by demethylating (*R*)-nicotine approximately 3, 10, and 10 times faster than (*S*)-nicotine, respectivley (Cai et al., 2012).

Very little research has been reported that provides insight into the accumulation of the specific enantiomers of nicotine in tobacco. Substantial work has been carried out to identify and characterize the genes and enzymes involved in tobacco nicotine biosynthesis, however. This knowledge has been applied in several studies to generate tobacco genotypes with reduced nicotine accumulation (Lewis, 2018). One example is the *Berberine Bridge Like* (*BBL*) gene family that codes for enzymes involved in one of the final stages of nicotine biosynthesis (Kajikawa et al., 2011). This gene family is comprised of six genes, with *BBLa, BBLb,* and *BBLc* being expressed to significant degrees, and *BBL-d1, BBL-d2,* and *BBL-e* being expressed at greatly reduced levels (Kajikawa et al., 2017). Lewis et al. (2020) observed very significant reductions in leaf nicotine content in genotypes homozygous for EMS-induced knockout mutations in *BBLa* and *BBLb* and genotypes homozygous for combined mutations in *BBLa*, *BBLb*, and *BBLc*. The addition of mutations in *BBL-d1*, *BBL-d2*, and *BBL-e* was reported to have no additional detectable effect on lowering nicotine levels (Lewis et al., 2020). No genotype evaluated in this study exhibited nicotine content reduced to zero, suggesting that there may be an alternate minor biosynthetic pathway contributing to nicotine biosynthesis in tobacco which does not involve BBL enzymes. A recent study conducted by Vollheyde et al. (2023) eliminated functional BBL enzyme production in *Nicotiana benthamiana* and also observed residual amounts of nicotine being produced in the homozygous knockout genotypes. The residual amount of nicotine was reported to be racemic, and the authors suggested that the residual nicotine derives from a non-stereoselective, spontaneous chemical reaction (Vollheyde et. al, 2023). Interestingly, in the study of Lewis et al. (2020), genotypes homozygous for combined knockout mutations in *BBLa* and *BBLb* exhibited a notable rise in leaf nornicotine content.

The initial objective of this study was to determine the enantiomeric composition of the tobacco alkaloids for genotypes with different combinations of *BBL* knockout mutations in genotypes with and without additional deleterious mutations in the three *Nicotine Demethylase* (*NND*) genes *CYP82E4*, *CYP82E5v2*, and *CYP82E10*. The results observed provide insights into the respective contributions of the *BBL* and *NDM* genes to the biosynthesis of alkaloid enantiomers in tobacco. Initial findings suggested that the BBLc isoform may be preferentially associated with the biosynthesis of (*R*)-alkaloid enantiomers. A second objective was therefore to use a combination of mutation breeding and genetic engineering to produce further evidence of the role that this particular BBL isoform plays in the biosynthesis of the (*R*)-enantiomers of the tobacco alkaloids. A third objective was to determine if tobacco genotypes could be generated with normal total alkaloid content but where the (*R*)-enantiomer of nicotine is more prevalent than the (*S*)-enantiomer.

MATERIALS AND METHODS

Development & Testing of Nearly Isogenic Lines Carrying Mutations in CYP82E4, CYP82E5, CYP82E10, BBLa, BBLb, and BBLc

Previously, nearly isogenic versions of the standard burley tobacco cultivar ‘TN 90’ differing for the presence/absence of ethyl methanesulfonate-induced (EMS) knockout mutations in the three *NND* genes as well as in the three mostly highly expressed *BBL* genes were developed and described (Lewis et al., 2020). Initially, a version of TN 90 designated as ‘TN 90 SRC’ (**S**table **R**educed **C**onverter) was developed by transferring knockout mutations in the three

NND genes (*CYP82E4*, *CYP82E5*, and *CYP82E10*) (Lewis et al., 2010) to TN 90 using seven generations of backcrossing followed by two generations of self-pollinations to establish a triple homozygous mutant genotype (Lewis et al., 2010). The combination of knockout mutations in the three *NND* genes leads to a dramatic reduction in the demethylation of nicotine and therefore results in significantly reduced nornicotine accumulation in cured tobacco leaves (Lewis et al., 2010).

The backcross breeding method was then used to transfer knockout mutations in the three most highly expressed *BBL* genes (*BBLa*, *BBLb*, and *BBLc*) involved in one of the final stages of nicotine biosynthesis to both the TN 90 and TN 90 SRC genetic backgrounds. Backcross breeding was carried out to the BC_7F_1 generation for both lines, at which time self-pollination of triple heterozygous *BBL* mutant genotypes was used to produce BC_7F_2 progenies. BC_7F_2 individuals were genotyped to identify all eight possible homozygous *BBL* mutant combinations in both TN 90 and TN 90 SRC, which were then self-pollinated to produce BC_7F_3 seed of the nearly isogenic lines (NILs) for each genetic background. Numerical nomenclature is hereafter used to refer to each of these genotypes, where ‘0’ indicates the homozygous ‘wild-type’ condition at a particular locus and a ‘2’ designates the homozygous mutant condition at a particular locus. For example, the designation ‘TN 90 220’ indicates a line homozygous for the mutant genotype at the *BBLa* and *BBLb* loci, and homozygous for the wild-type allele at the *BBLc* locus.

The genetic materials described above were previously evaluated for yield, cured leaf quality, and alkaloid composition at two locations representative of burley tobacco-producing environments in North Carolina during the 2016 and 2017 growing seasons: the Mountain Research Station at Waynesville, and the Upper Mountain Research Station at Laurel Springs.

Materials were evaluated in comparison with TN 90 and TN 90 SRC (Lewis et al., 2010). Cured leaf alkaloid enantiomer concentrations were not determined at this time. Remnant dried and ground-cured leaf samples from the Laurel Springs environment in 2017 were used for tobacco alkaloid enantiomer quantifications described hereafter. The experimental design used in this environment was a randomized complete block design with four replications. Experimental units were single 20-plant rows managed according to recommended production practices for burley tobacco production in North Carolina. Row and plant spacing were 122 cm and 46 cm, respectively. Plots were stalk cut at maturity, speared onto sticks, and air-cured in structures protected from rainfall. After curing, leaves were stripped into four stalk positions and weighed. Plot chemical samples were 50 g cured leaf-weighted composite samples.

Chemical Analyses

Oven-dried composite cured leaf samples were ground to pass through a 1-mm sieve and analyzed for total nicotine, nornicotine, anatabine, and anabasine. Tobacco samples were prepared by weighing approximately 0.2000 g of ground leaf into 50 ml Erlenmeyer flasks. Two ml of a 2 N NaOH solution was added to each sample. Flasks were swirled gently to thoroughly moisten the ground samples and were allowed to sit for 15 min. Ten ml of methyl tert-butyl ether (MTBE) solution containing an internal standard consisting of 0.40 mg ml^{-1} quinolone was added to each flask. Stoppers covered in aluminum foil were placed into each flask before shaking at high rpm for 2.5 h. Flasks were allowed to sit overnight to permit layer separation, and aliquots of the MTBE layer were subjected to gas chromatographic analysis as outlined by Lewis et. al., (2015).

Quantitative determinations of nicotine, nornicotine, anabasine, and anatabine in ground leaf samples were made via gas chromatographic analyses using an Agilent HP 6890 GC-FID and a 30 m DB-5MS column (J & W 0.53 mm i.d and 1.50 µm film thickness) with a 1 m deactivated guard column in front. The carrier gas was Helium at a linear gas velocity of approximately 38 cm/s. The injector was set at 250°C and the detector at 325°C. The analysis consisted of a temperature program from 110°C held for 1 minute to 200°C at 10°C/min followed by an increase to 300°C at 25°C/min. Temperature was held at 300°C for 10 min. The identity of compounds of interest was verified by comparison of retention times to that of authentic standards and/or confirmed by GC-MS analysis using an HP 5890 GC equipped with an HP5972 series MSD detector. The GC-MS system was equipped with a 30 m Restek RTX-5MS column (0.25 mm i.d., 0.25 µm film thickness) and operated using the same instrument parameters used in the GC-FID analyses. The ionization voltage was 70 eV. GC-FID quantification of compounds utilized an internal standard method. Results were reported as percentages on a dry-tobacco-weight basis.

Enantiomer ratios for nicotine, anatabine, and anabasine were determined using an Agilent 6890N gas chromatograph equipped with an Agilent 5973N mass selective detector and an Agilent 7683 autosampler (Santa Clara, CA). Automatic injections of 1 µL were performed in split mode with a split ratio of 10:1. The separation column consisted of two columns connected with a Restek union (Bellefonte, PA). The first column was an Agilent J&W Cyclodex-B capillary column (30 meters long, 0.25 mm I.D., 0.25 µm film thickness). The second column was a Restek Rt-bDEXsm capillary column (30 meters long, 0.25 mm I.D., and 0.25 µm film thickness). The oven temperature was initially set at 70°C and was then ramped to 150°C at 0.5 °C min^{-1}, followed by an increase to 210°C at 10°C min^{-1} which was maintained for 15 minutes.

The column flow rate was held constant at 1.3 mL min^{-1}. The injection port, mass spectrometry source, and transfer line were held at 230°C. The mass spectrometry quad was held at 150°C. The mass selective detector was operated in selective ion monitoring mode with m/z set at 84 for nicotine and 105 for anatabine and anabasine.

Enantiomer ratios for nornicotine were determined by initial derivatization, whereby a 0.6 mL tobacco extract in MTBE was treated with 20 µL pyridine and 10 µL isobutyl chloroformate. After sitting for 30 min., solutions were injected into the GC-MS system described above. The column was a Restek Rt-bDEXsm capillary column (30 meters long, 0.25 mm I.D., and 0.25 µm film thickness). The oven temperature was initially set at 50°C with a hold time of two min and was then ramped to 160°C at 30°C min^{-1} with a hold time of 9 min. This was followed by an increase to 205°C at 1°C min^{-1}, and an increase to 225°C at 20°C min^{-1} which was then maintained for 4.3 min. The inlet was held at 230°C in splitless injection mode (1 µL) with a constant flow of 1 mL min^{-1}. The mass selective detector was operated in selective ion monitoring mode with m/z set at 191.

Generation of Transgenic Tobacco Lines Overexpressing BBLc

To produce further evidence of the role *BBLc* plays in the preferential biosynthesis of the (*R*)-enantiomer of nicotine, as well as to determine if tobacco genotypes could be generated with normal total alkaloid content but where the (*R*)-enantiomer of nicotine prevails, we developed transgenic tobacco lines overexpressing *BBLc*. pUC57 plasmids were ordered from GeneScript (Piscataway, NJ) containing the laboratory synthesized wild-type *BBLc* cDNA. Two pMDC32 overexpression vectors were provided by Dr. Micheal Timko (University of Virginia). One pMDC32 vector contained the ‘constitutive’ cauliflower mosaic virus 35S promoter, and the other vector contained the *PMT1a* promoter. The *PMT1a* promoter was used because it is derived from the tobacco nicotine biosynthetic gene *Putrescine N-methyltransferase 1* (*PMT-1*), and is preferentially expressed in the root tissues of tobacco plants, the known site of tobacco alkaloid biosynthesis (Hibi et al., 1994; Hashimoto et al., 1998; Riechers & Timko, 1999). The 35S promoter was used because it was believed *a priori* to have some activity in the root tips (Battraw and Hall, 1990). Xu and Timko (2004) identified the GAG regulatory motif which is a composite regulatory region in the *NtPMT1a* promoter that is comprised of a G-box, an AT-rich motif, and a GCC-box-like element that is required for JA induction of *NtPMT1a* (Sears et al., 2014). Both promoters were fused with four copies of the GAG regulatory motif to establish p4GAG-35S-*BBLc* and p4GAG-PMT1a-*BBLc*.

Plasmids were ordered from Intact Genomics (St. Louis, MO) and introduced into the electrocompetent *Agrobacterium tumefaciens* strain LBA4404. *Agrobacterium tumefaciens* cells carrying the pMDC32 vector containing one of the two constructs were used to introduce corresponding T-DNA to the leaf tissue of three selected tobacco genotypes: TN 90 SRC, TN 90 SRC 220, and TN 90 SRC 222.

Young leaf tissues from all three *N. tabacum* genetic backgrounds were excised, sterilized, cut into smaller disks, and wounded in the middle. Disks were inoculated in a sterile petri dish with the resuspended *A. tumefaciens* culture (An et al., 1986). A few disks from each explant were left uninoculated to serve as regeneration controls. Regardless of inoculation treatment, leaf disks were transferred to petri dishes containing T-medium and were allowed to incubate at 25^{o}C with an 18h L/6h D light cycle.

Following this incubation period, the leaf disks were transferred to selection media (R-media) with and without antibiotics depending on whether they were regeneration controls or not. Selection media plates were incubated under the same conditions as used in the previous incubation step. Leaf disks were transferred to new media every two weeks and checked for mold growth daily. Successfully transformed leaf cells generated callus tissue and subsequently shoots. Shoots that were established and that exhibited a normal morphology were trimmed of callus tissue and transferred to plates containing R-medium with antibiotics. Each transferred shoot was labeled based on the plate and leaf disk it was derived from. For example, P1 1a and P1 1b are from leaf disk #1 on plate #1 and P1 2a, and P1 2b are from leaf disk #2 on plate #1. When the regenerated shoots were large enough, they were transferred to 125 mL narrow-mouth Erlenmeyer flasks containing T-medium. Regenerated shoots were potted once roots were well-established and designated as T_0 plants.

DNA was isolated from young leaf tissue from each of the T_0 plants following the modified cetyltrimethylammonium bromide (CTAB) procedure described by Johnson et al. (1995). Maceration of leaf tissue was facilitated by the FastPrep®-96 (MP Biomedicals, Irvine, CA) and DNA was quantified using the Hoefer DQ-300 Fluorometer (Hoefer, Inc., San Francisco, CA). Isolated DNA was normalized to approximately 50 ng uL^{-1} by diluting with $T_{10}E_1$ pH 8.0 (Low TE) buffer.

The transgenic status of regenerated plants was verified by PCR-based amplification of the *nptII* selectable marker sequence. Amplification was confirmed through gel electrophoresis on 1% agarose in 1x TAE buffer. Ethidium bromide was added to the gel before solidification for visualization of the fragments under UV light.

Development and Screening of T_1 Families

One hundred twenty-six T_0 plants (20 from each construct/genotype + 1 regeneration control for each group) were identified as containing the *nptII* gene as well as the *BBLc* cDNA. These plants were grown in a greenhouse under natural light conditions until flowering at which time one leaf from the middle stalk of each plant was collected, dried, ground, and analyzed for alkaloid composition using previously described methodology (Lewis et al., 2015).

T_0 plants were self-pollinated to produce T_1 generation seed and also testcrossed with K326 to identify T_0 plants that likely contained single transgene insertions. Approximately 100 testcross progeny for each transgenic plant were surface sterilized and seeded onto plates containing sterile MS inorganic media containing 100 mg L^{-1} kanamycin. Individual plants that likely contained a single segregating transgene insertion were indicated by a 1:1 ratio of kanamycin-resistant to kanamycin-susceptible seedlings. Seven plants from T_1 families corresponding to T_0 individuals predicted to carry single transgene insertions and that exhibited elevated nicotine content relative to controls were potted in a greenhouse (21 families, approximately 4 per genotype/construct), grown to flowering, self-pollinated to produce T_2 seed, and also testcrossed to K326. Testcross progeny were germinated on MS inorganic medium containing kanamycin to identify those T_1 individuals that did not segregate for antibiotic resistance, meaning the T_1 individual was homozygous for the transgene.

Field Evaluation of T_2 Families

A total of 21 non-segregating transgenic T_2 families representing 21 independent transgenic events were selected for evaluation of alkaloid profiles during the 2023 growing season in North Carolina (Table 1.2). Also included in the field evaluation were TN 90 LC and the non-transgenic parental control lines TN90 SRC 220, TN90 SRC 222, and TN90 SRC.

The field experiment was conducted at the Central Crops Research Station in Clayton, North Carolina. The experimental design was a randomized complete block design with twenty replications. Experimental units were single plants and were managed according to recommended production practices for burley tobacco production in North Carolina. Row and plant spacing were 122 cm and 46 cm, respectively. When 50% of the plants in the field were flowering, all plants were topped and treated with Prime Plus to control suckers. Twenty-one days after topping, the top leaf was harvested from each plant, oven-dried, ground, and analyzed for alkaloid and enantiomer composition as previously described above.

Data Analyses

Analyses of variance appropriate for analyzing data from a randomized complete block design were carried out for all measured molecules using PROC MIXED of SAS 9.3 (SAS Institute, Cary, NC) according to Littell et al. (1996). Means separations were performed by using Fisher's protected least significant difference test (alpha $\alpha = 0.05$). Enantiomeric ratios were determined for all alkaloids by dividing the fraction of the (*R*)-enantiomer (given in mg g^{-1} dry weight basis) by the total concentration (mg g^{-1}) of the alkaloid of interest.

RESULTS

Alkaloid Enantiomeric Composition of Genetic Materials Carrying Mutations in Nicotine Demethylase and BBL Genes

In line with the findings from a multi-location data analysis conducted by Lewis et al. (2020), average total nicotine and total anatabine levels of TN 90 220, TN 90 222, TN 90 SRC 220, and TN 90 SRC 222 in the Laurel Springs environment were significantly ($P < 0.05$) lower than their respective control lines (TN 90 and TN 90 SRC) (Figure 1.1 and 1.3). For average total anabasine content, only the triple homozygous *BBL* mutant (222) versions of TN 90 and TN 90 SRC exhibited significantly lower levels compared to the controls (Figure 1.4). Interestingly, a 322% increase in total nornicotine was observed for the double homozygous *BBL* mutant line TN90 220 as compared to TN90 *per se*. A 71% decrease in average nornicotine content was observed in TN 90 222 compared to the same control line (Figure 1.2), however. For no genotype were alkaloid levels reduced to zero.

For the alkaloid enantiomeric fractions for the *BBL* mutant NILs of the TN 90 genotype, (*R*)-nicotine was only measurable in TN 90 220 (Figure 1.1). In the TN 90 SRC genetic background, (*R*)-nicotine was measurable in all NILs except those that were homozygous for mutations in *BBLc* (TN 90 SRC lines 002, 022, and 202). The (*R*)-nicotine fractions were low (0.07 to 0.12) in all genotypes that accumulated (*R*)-nicotine, except for TN 90 SRC 220 and TN 90 SRC 222, which displayed significant increases in (*R*)-nicotine accumulation, reaching fractions of 0.84 and 0.67, respectively.

(*R*)-anatabine fractions were in the range of 0.13 to 0.18 for all genotypes in both the TN 90 and TN 90 SRC genetic backgrounds, except the double or triple homozygous *BBL* mutant lines TN 90 220, TN 90 SRC 220, and TN 90 SRC 222. For these three genotypes, (*R*)-anatabine fractions were approximately 0.50 (Figure 1.3). Similar observations were made for anabasine, in which (*R*)-anabasine fractions for 220 and 222 genotypes of both genetic backgrounds were significantly higher and in the range of 0.53 to 0.67.

(*R*)-nornicotine was detectable in all tested genotypes. However, (*R*)-nornicotine fractions were less than 0.50 for all genotypes except for TN 90 220, TN 90 222, TN 90 SRC 220, and TN 90 SRC 222.

Overexpression of BBLc in Selected Genetic Backgrounds

Twenty primary transformants (T_0 generation) were regenerated from each of three genetic backgrounds (TN 90 SRC, TN 90 SRC 220, and TN 90 SRC 222) and evaluated for alkaloid levels under greenhouse conditions to identify those with possible elevated total nicotine levels and increased (*R*)-enantiomeric fractions relative to non-engineered control plants. Of these primary transformants, twenty-one were selected to be advanced for the second generation (T_1 generation) based on nicotine content, (*R*)/(*S*) enantiomer ratios, and segregation ratios of the transgene in testcross progeny (Table 1.1). Eight lines from three genetic backgrounds carrying the construct containing the *PMT-1a* promoter were advanced, and thirteen lines with the construct containing the *35S* promoter were advanced (Table 1.1). T_1 plants from each of these twenty-one selected transgenic events were grown out in families of seven to identify individuals homozygous for the transgene insertion. At least one plant from each T_1 family was identified as being homozygous for a transgene insertion after testcross seeds were germinated on MS inorganic medium containing kanamycin. These plants were self-pollinated to produce non-segregating T_2 lines.

Field Evaluations of T_2 Lines

Twenty-one selected transgenic burley tobacco genotypes were evaluated in the Clayton field environment for their alkaloid profiles along with four non-transgenic control lines (Table 1.2). A means separation test with (α=0.05) was performed on the alkaloid data from the tested transgenic and control lines for total nicotine, nornicotine, anatabine, and anabasine levels. Due to the intensive nature of enantiomer determinations, these analyses were only carried out on selected transgenic lines and associated controls from the first 10 replicates of the 20 replicate field experiment. No transgenic line of the TN 90 SRC background exhibited significantly (P < 0.05) higher levels of total nicotine compared to the control line. However, we selected the two lines that exhibited numerically higher nicotine levels compared to the TN 90 SRC control (Table 1.2). Two lines from each of the TN 90 SRC 220, and TN 90 SRC 222 genetic backgrounds exhibited significantly (P < 0.05) higher levels of nicotine compared to their respective control lines (Table 1.2) and were selected for determination of enantiomeric fractions.

Average total nicotine levels of the two selected TN 90 SRC 220 transgenic T_2 lines (P2 4a and P13 9a) exhibited significantly (P < 0.05) greater amounts of nicotine compared to TN 90 SRC 220 *per se*, exhibiting 23.1 and 21.9 mg g^{-1} respectively (Figure 1.5). Similarly, TN 90 SRC 222 transgenic T_2 lines (P1 8a and P9 8a) exhibited nicotine levels of 13.5 and 14.4 mg g^{-1} respectively, which were significantly (P < 0.05) greater as compared to the TN 90 SRC 222 non-transgenic control *per se* (Figure 1.5). Neither of the selected TN 90 SRC transgenic lines (P4 6a or P4 1a) exhibited statistically different levels of nicotine compared to the TN 90 SRC control (Figure 1.5). Total nornicotine levels were not statistically different for any of the transgenic lines as compared to any of their corresponding non-engineered control lines, and only TN 90 LC exhibited significantly ($P < 0.05$) greater amounts of nornicotine compared to all other lines (Figure 1.6). For total anabasine content, TN 90 SRC 220 transgenic lines P2 4a and P13 9a exhibited significantly ($P < 0.05$) greater amounts of anabasine compared to their corresponding non-transgenic control (Figure 1.7). None of the other tested transgenic lines had statistically different anabasine levels compared to their respective control lines (Figure 1.7). Compared to the TN 90 SRC control line, transgenic line P4 6a exhibited significantly ($P < 0.05$) less anatabine, and P4 1a exhibited significantly ($P < 0.05$) greater anatabine content (Figure 1.8). Anatabine levels for the transgenic lines of the TN 90 SRC 220 and TN 90 SRC 222 backgrounds were not statistically different compared to their corresponding control lines (Figure 1.8).

For the enantiomeric fractions of the selected *BBLc* overexpressing transgenic lines, every transgenic line exhibited significantly ($P < 0.05$) greater fractions of (*R*)-nicotine compared to their respective control lines (Figure 1.5). Notably, transgenic *BBLc* overexpressing lines of the TN 90 SRC 220 genetic background exhibited the highest (*R*)-nicotine fractions in the field experiment, ranging from 0.84 to 0.85 (Figure 1.5). (*R*)-nicotine fractions of transgenic lines of the TN 90 SRC 222 genetic background were also high and ranged from 0.82 to 0.84 (Figure 1.5).

(*R*)-nornicotine fractions of TN 90 SRC and TN 90 SRC 222 transgenic T_2 lines (P4 6a, P4 1a, P1 8a, and P9 8a) ranged from 0.89 to 0.90 and were significantly ($P < 0.05$) greater than their corresponding non-transgenic control lines (Figure 1.6). There were no statistically significant differences observed for (*R*)-nornicotine fractions between the transgenic and non-transgenic lines for the TN 90 SRC 220 genetic background, however (Figure 1.6).

For anabasine, TN 90 SRC 220 and TN 90 SRC 222 transgenic T_2 lines (P2 4a, P13 9a, P1 8a, and P9 8a) exhibited significantly ($P < 0.05$) greater levels of (*R*)-anabasine compared to their respective non-transgenic control lines (Figure 1.7). Neither TN 90 SRC transgenic T_2 line (P4 6a or P4 1a) exhibited levels of (*R*)-anabasine that were statistically different from that of the TN 90 SRC control (Figure 1.7).

Enantiomeric fractions for anatabine were not analyzed due to a large amount of missing data because many determinations were below the level of quantification.

DISCUSSION

The results reported here support previous findings that tobacco plants with induced homozygous mutations in at least the *BBLa* and *BBLb* genes display significantly decreased nicotine levels. No genotype evaluated in this study had nicotine accumulation reduced to zero, however. Since plants with additional mutations in *BBL-d1*, *BBL-d2*, and *BBL-e* do not exhibit further reductions in nicotine content, there may be an alternative minor pathway involved in the accumulation of a small amount of nicotine in the tobacco plant (Lewis et al., 2020).

Vollheyde et al. (2023) have suggested a role for non-stereoselective, non-enzymatic, spontaneous chemical reaction in the formation of the residual nicotine in mutant low-nicotine lines. The results further validate earlier observations, that tobacco plants with triple homozygous induced deleterious mutations in the three *NND* genes accumulate significantly reduced total nornicotine levels (Lewis et. al, 2010). Interestingly, analysis of enantiomeric ratios for the tobacco alkaloids in the studied materials with different *BBL* mutant combinations brings attention to the possibility that different BBL isoforms may act preferentially to biosynthesize different enantiomers.

It has been known that tobacco plants naturally accumulate just a small fraction of (*R*)-nicotine in its leaves, representing only about 0.02% of the total nicotine pool in air-cured burley tobacco (Zhang et al., 2017). No research has previously been reported that provides insight into where this small fraction of (*R*)-nicotine may be derived from. However, our results indicate that (*R*)-nicotine increases in *BBL* mutant NILs of TN 90 SRC of the 020, 200, 220, and 222 genotypes. Our results further demonstrate that additional increases in (*R*)-nicotine are observed in *BBLc* overexpressing transgenic plants of the TN 90 SRC 220 and 222 genotypes. In the non-transgenic TN 90 SRC genetic background, the fraction of (*R*)-nicotine becomes easily quantifiable in all isolines ***except*** those of the 002, 022, and 202 genetic composition (those in which *BBLc* is inactivated). This suggests that non-genetically altered tobacco cultivars (such as TN 90 LC) actually biosynthesize significant amounts of (*R*)-nicotine. It is only measurable in very small amounts, however, because the vast majority of it is converted to nornicotine by *CYP82E4*, *CYP82E5*, and *CYP82E10* that preferentially catalyze the demethylation of the (*R*)-isoform (Cai et al., 2012). The true nature of the tobacco plant's ability to biosynthesize (*R*)-nicotine is only revealed when there is reduced nicotine demethylase activity.

It is noteworthy that (*R*)-nicotine was only detectable in a TN 90 isoline in which *BBLa* and *BBLb* were inactivated, and in which *BBLc* remained active. In addition, nicotine, nornicotine, and anabasine enantiomeric compositions were reversed in TN 90 SRC mutant isolines in which BBLa and BBLb activity was abolished. These observations can be explained by a model in which *BBLc* is involved in the preferential biosynthesis of the (*R*)-enantiomer, while *BBLa* and *BBLb* are involved in the preferential biosynthesis of the (*S*)-enantiomers of these alkaloids. When the *BBLc* overexpression construct is introduced into the TN 90 SRC 220 and TN 90 SRC 222 genetic backgrounds, (*R*)-nicotine fractions increase to over 80%, providing further evidence that *BBLc* is involved in the preferential biosynthesis of (*R*)-nicotine. In research described here and elsewhere (Lewis et al., 2020), nicotine accumulation was not reduced to zero in tobacco genotypes in which *BBL* activity was abolished. As previously noted, a small amount of nicotine in tobacco could be biosynthesized through pathways that do not involve the BBL enzymes, and could even be derived from spontaneous chemical reactions that don't involve enzymes at all. When *BBL* and known *NDM* activity is abolished in TN 90 SRC 222 genotypes, it seems that an alternative biochemical step(s) may lead to racemic nicotine accumulation as previously suggested for *N. benthamiana* (Vollheyde et. al, 2023).

This aforementioned model also explains the significant increase in nornicotine observed in NILs in which *BBLc* is the only active *BBL* gene because the *NND* genes are known to preferentially demethylate (*R*)-nicotine to nornicotine (Cai et al., 2012). When *BBLc* is the only active *BBL* gene, there is a much greater amount of (*R*)-nicotine that can be converted into nornicotine. At least some of the increase in nornicotine may be a result of NND enzymes preferentially acting on a prevalence of (*R*)-nicotine.

Although the precise function of the BBL enzymes is not fully understood, they are known to be involved in one of the later stages of nicotine biosynthesis (Kajikawa et al., 2011). One metabolite that increases in abundance in *BBL* triple homozygous (222) mutants is dihydrometanicotine (DMN) (Kajikawa et al., 2011; Lewis et al., 2015). The DMN molecule is achiral, and it seems plausible that BBL enzymes utilize DMN as a substrate in the formation of chiral nicotine molecules.

In some living organisms, separate enzymatic pathways can give rise to enantiomeric metabolites. Alternatively, enantiomers may arise from individual enzymes that lack both substrate specificity and stereospecificity (Finefeld et al., 2012). *N. tabacum* is an allotetraploid species believed to have originated from the hybridization of the progenitor species *N. sylvestris* and *N. tomentosiformis,* followed by chromosome doubling (Lewis and Nicholson, 2007). Due to this evolutionary history, tobacco usually possesses multiple copies of a given gene. Gene duplication resulting from polyploidization may allow for different gene copies to undergo evolutionary changes that result in the production of different enantiomers. Perhaps, different progenitor species may have evolved tendencies to accumulate different enantiomers. To thoroughly investigate this question, however, the *NND* genes likely need to be inactive to determine the true potential of the given species to accumulate one or the other enantiomer. It is interesting to contemplate whether evolutionary selection pressures have influenced alkaloid enantiomeric composition in the *Nicotiana* genus over time as a response to some enantiomers conferring differing biological properties against herbivores. It seems reasonable to believe that the (*S*)-enantiomer of nicotine became the prevalent isoform in tobacco due to it being more toxic and displaying greater pharmacological effects as compared to the (*R*)-enantiomer (Yildiz, 2004).

The research described here demonstrates the ability to use a combination of mutation breeding and genetic engineering to further increase the accumulation of a potentially desired enantiomeric isoform for a given plant molecule. Tobacco genotypes described here exhibited a reversal in enantiomeric composition of nicotine, although no line contained 100% (*R*)-nicotine. This is possibly due to the ability of BBLc to produce some level of (*S*)-nicotine or due to non-enzymatic steps that produce racemic nicotine.

Although (*R*)-nicotine is suggested to have reduced stimulative properties, there is currently no known commercial value for tobacco plants that produce an abundance of this enantiomer. There are cases in the natural world, however, where it might be desirable to reverse the enantiomeric composition of a given molecule in order to better meet industrial or pharmacological needs. For example, reversing the enantiomeric composition of gossypol in cotton seeds to the non-toxic form was carried out to address health concerns within that industry (Lin et. al, 2023). The research described here demonstrates the possibility of reversing the enantiomeric composition of a given molecule using a combination of mutation breeding and genetic engineering. In addition, this research adds to the body of knowledge relating to the biosynthesis of tobacco alkaloids and, in particular, the role that the *BBL* genes play in the biosynthesis of the alkaloid enantiomers.

Finally, the results described here also reveal the usefulness of using the 4GAG-enhanced 35S promoter to drive transgene expression in tobacco root tips, the site of nicotine biosynthesis. All of the transgene overexpression lines that were selected for advanced study utilized the 4GAG-enhanced 35S promoter. By this measure of comparison, the 35S promoter was more effective than the *PMT-1a* promoter for driving expression of a nicotine biosynthesis transgene in tobacco roots.

References

Afanador, L. K., Haley, S. D., & Kelly, J. D. (1993). Adoption of a “mini prep” DNA extraction method for RAPD marker analysis in common bean (*phaseolus vulgaris* L.). *Annu. Rep. Bean Improv. Coop, 36*, 10-11.

An, G., Watson, B. D., & Chiang, C. C. (1986). Transformation of tobacco, tomato, potato, and *Arabidopsis thaliana* using a binary ti vector system. *Plant Physiology, 81*(1), 301-305.

Battraw, M. J., & Hall, T. C. (1990). Histochemical analysis of CaMV 35S promoter-β-glucuronidase gene expression in transgenic rice plants. *Plant Molecular Biology, 15*, 527-538.

Cai, B., Siminszky, B., Chappell, J., Dewey, R. E., & Bush, L. P. (2012). Enantioselective demethylation of nicotine as a mechanism for variable nornicotine composition in tobacco leaf. *Journal of Biological Chemistry, 287*(51), 42804-42811.

Dawson, R. F. (1942). Accumulation of nicotine in reciprocal grafts of tomato and tobacco. *American Journal of Botany, ,* 66-71.

Finefield, J. M., Sherman, D. H., Kreitman, M., & Williams, R. M. (2012). Enantiomeric natural products: Occurrence and biogenesis. *Angewandte Chemie International Edition, 51*(20), 4802-4836.

Hashimoto, T., Shoji, T., Mihara, T., Oguri, H., Tamaki, K., Suzuki, K., & Yamada, Y. (1998). Intraspecific variability of the tandem repeats in *Nicotiana* putrescine *N*-methyltransferases. *Plant Molecular Biology, 37*, 25-37.

Hibi, N., Higashiguchi, S., Hashimoto, T., & Yamada, Y. (1994). Gene expression in tobacco low-nicotine mutants. *The Plant Cell, 6*(5), 723-735.

Johnson, E., Miklas, P. N., Stavely, J. R., & Martinez-Cruzado, J. C. (1995). Coupling-and repulsion-phase RAPDs for marker-assisted selection of PI 181996 rust resistance in common bean. *Theoretical and Applied Genetics, 90*, 659-664.

Kajikawa, M., Shoji, T., Kato, A., & Hashimoto, T. (2011). Vacuole-localized berberine bridge enzyme-like proteins are required for a late step of nicotine biosynthesis in tobacco. *Plant Physiology, 155*(4), 2010-2022.

Lewis, R. S., & Nicholson, J. S. (2007). Aspects of the evolution of *Nicotiana tabacum* L. and the status of the United States *Nicotiana* germplasm collection. *Genetic Resources and Crop Evolution, 54*, 727-740.

Lewis, R. S. (2019). Potential mandated lowering of nicotine levels in cigarettes: A plant perspective. *Nicotine and Tobacco Research, 21*(7), 991-995.

Lewis, R. S., Bowen, S. W., Keogh, M. R., & Dewey, R. E. (2010). Three nicotine demethylase genes mediate nornicotine biosynthesis in *Nicotiana tabacum* L.: Functional characterization of the *CYP82E10* gene. *Phytochemistry, 71*(17-18), 1988-1998.

Lewis, R. S., Drake-Stowe, K. E., Heim, C., Steede, T., Smith, W., & Dewey, R. E. (2020). Genetic and agronomic analysis of tobacco genotypes exhibiting reduced nicotine accumulation due to induced mutations in *berberine bridge like* (*BBL*) genes. *Frontiers in Plant Science, 11,* 368.

Lewis, R. S., Lopez, H. O., Bowen, S. W., Andres, K. R., Steede, W. T., & Dewey, R. E. (2015). Transgenic and mutation-based suppression of a *berberine bridge enzyme-like* (*BBL*) gene family reduces alkaloid content in field-grown tobacco. *PloS One, 10*(2), e0117273.

Lin, J., Fang, X., Li, J., Chen, Z., Wu, W., Guo, X., Liu, N., Huang, J., Chen, F., & Wang, L. (2023). Dirigent gene editing of gossypol enantiomers for toxicity-depleted cotton seeds. *Nature Plants, 9*(4), 605-615.Littell, R., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). SAS system for mixed models. SAS inst. *Inc., Cary, NC, 633*

McIntee, E. J., & Hecht, S. S. (2000). Metabolism of *N*'-nitrosonornicotine enantiomers by cultured rat esophagus and *in vivo* in rats. *Chemical Research in Toxicology, 13*(3), 192-199.

Murashige, T., & Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia Plantarum, 15*(3), 473-497.

Pogocki, D., Ruman, T., Danilczuk, M., Danilczuk, M., Celuch, M., & Wałajtys-Rode, E. (2007). Application of nicotine enantiomers, derivatives and analogues in therapy of neurodegenerative disorders. *European Journal of Pharmacology, 563*(1-3), 18-39.

Riechers, D. E., & Timko, M. P. (1999). Structure and expression of the gene family encoding putrescine *N*-methyltransferase in *Nicotiana tabacum*: New clues to the evolutionary origin of cultivated tobacco. *Plant Molecular Biology, 41*, 387-401.

Rosecrans, J. A., & Young, R. (2018). Discriminative stimulus properties of S (−)-nicotine: "a drug for all seasons". *The Behavioral Neuroscience of Drug Discrimination,*, 51-94.

Sears, M. T., Zhang, H., Rushton, P. J., Wu, M., Han, S., Spano, A. J., & Timko, M. P. (2014). NtERF32: A non-*NIC2* locus AP2/ERF transcription factor required in jasmonate-inducible nicotine biosynthesis in tobacco. *Plant Molecular Biology, 84*, 49-66.

Vollheyde, K., Dudley, Q. M., Yang, T., Oz, M. T., Mancinotti, D., Fedi, M. O., Heavens, D., Linsmith, G., Chhetry, M., & Smedley, M. A. (2023). An improved nicotiana benthamiana bioproduction chassis provides novel insights into nicotine biosynthesis. *New Phytologist, 240*(1), 302-317.

Xu, B., & Timko, M. (2004). Methyl jasmonate induced expression of the tobacco putrescine *N*-methyltransferase genes requires both G-box and GCC-motif elements. *Plant Molecular Biology*, *55*, 743-761.

Yildiz, D. (2004). Nicotine, its metabolism and an overview of its biological effects. *Toxicon, 43*(6), 619-632.

Zhang, H., Pang, Y., Luo, Y., Li, X., Chen, H., Han, S., Jiang, X., Zhu, F., Hou, H., & Hu, Q. (2018). Enantiomeric composition of nicotine in tobacco leaf, cigarette, smokeless tobacco, and e-liquid by normal phase high-performance liquid chromatography. *Chirality, 30*(7), 923-931.

Tables and Figures:

Table 1.1 Selected Genotypes Based on Greenhouse Data

Plant I.D	Genotype	Nicotine	Fraction R-Nic	Testcross Transgene Segregation
		mg g^{-1}		(R/NR)[b]
		----------------T_o-------------------		
P8 3a	TN90 SRC BBL-C/p4GAG-PMT1a-GPTV	10.62	0.11	41/46
P4 3a	TN90 222 SRC BBL-C/p4GAG-PMT1a-GPTV	11.17	0.88	63/21
P5 4a	TN90 222 SRC BBL-C/p4GAG-PMT1a-GPTV	7.54	0.81	22/41
P5 1a	TN90 222 SRC BBL-C/p4GAG-PMT1a-GPTV	8.91	0.87	43/42
P6 6a	TN90 222 SRC BBL-C/p4GAG-PMT1a-GPTV	6.57	0.83	39/37
P1 6a	TN90 220 SRC BBL-C/p4GAG-PMT1a-GPTV	12.41	0.9	26/64
P7 6a	TN90 220 SRC BBL-C/p4GAG-PMT1a-GPTV	7.83	0.83	33/42
P8 3b	TN90 220 SRC BBL-C/p4GAG-PMT1a-GPTV	10.37	0.89	30/32
P3 9a	TN90 SRC BBL-C/p4GAG-35S-GPTV	25.45	0.06	32/52
P6 3a	TN90 SRC BBL-C/p4GAG-35S-GPTV	27.65	0.07	41/48
P4 6a**[a]	TN90 SRC BBL-C/p4GAG-35S-GPTV	24.31	0.07	33/45
P1 7a	TN90 SRC BBL-C/p4GAG-35S-GPTV	29.32	0.06	45/44
P4 1a**[a]	TN90 SRC BBL-C/p4GAG-35S-GPTV	20.25	0.07	44/53
P1 8a**[a]	TN90 222 SRC BBL-C/p4GAG-35S-GPTV	11.59	0.88	55/42
P9 8a**[a]	TN90 222 SRC BBL-C/p4GAG-35S-GPTV	9.4	0.87	42/33
P12 9a**[a]	TN90 222 SRC BBL-C/p4GAG-35S-GPTV	12.21	0.9	27/52
P2 6a	TN90 222 SRC BBL-C/p4GAG-35S-GPTV	11.34	0.84	37/34
P4 8a	TN90 220 SRC BBL-C/p4GAG-35S-GPTV	26.57	0.92	68/17
P2 4a**[a]	TN90 220 SRC BBL-C/p4GAG-35S-GPTV	15.54	0.9	26/42
P13 9a**[a]	TN90 220 SRC BBL-C/p4GAG-35S-GPTV	12.38	0.89	34/55
P9 5a	TN90 220 SRC BBL-C/p4GAG-35S-GPTV	24.68	0.92	20/43
Control	TN90 LC	–	–	–
Control	TN90 SRC 220	–	–	–
Control	TN90 SRC 222	–	–	–
Control	TN90 SRC	–	–	–

[a]** Represent lines that have been selected for enantiomeric composition analysis
[b]Count of seedlings resistant (R) and not resistant (NR) to kanamyacin

Table 1.2 Lines Selected for Enantiomeric Analysis.

Entry	Genotype	Plant I.D.	Nic	Nor	Anat	Anab
			mg g^{-1}	mg g^{-1}	mg g^{-1}	mg g^{-1}
25	TN90 SRC	-	31.481	0.179	0.455	0.130
1	TN90 SRC BBL-C/p4GAG-PMT1a-GPTV	P8 3a	24.696	0.158	0.306	0.092
9	TN90 SRC BBL-C/p4GAG-35S-GPTV	P3 9a	26.166	0.172	0.295	0.100
10	TN90 SRC BBL-C/p4GAG-35S-GPTV	P6 3a	29.603	0.182	0.536	0.127
11	TN90 SRC BBL-C/p4GAG-35S-GPTV	P4 6a	30.102	0.186	0.376	0.119
12	TN90 SRC BBL-C/p4GAG-35S-GPTV	P1 7a	28.544	0.185	0.375	0.111
13	TN90 SRC BBL-C/p4GAG-35S-GPTV	P4 1a	33.097	0.207	0.564	0.143
23	TN90 SRC 220	-	11.703	0.259	0.039	0.125
6	TN90 220 SRC BBL-C/p4GAG-PMT1a-GPTV	P1 6a	14.351	0.299	0.049	0.117
7	TN90 220 SRC BBL-C/p4GAG-PMT1a-GPTV	P7 6a	3.316	0.107	0.122	0.114
8	TN90 220 SRC BBL-C/p4GAG-PMT1a-GPTV	P8 3b	13.152	0.257	0.048	0.140
18	TN90 220 SRC BBL-C/p4GAG-35S-GPTV	P4 8a	10.180	0.262	0.095	0.119
19	TN90 220 SRC BBL-C/p4GAG-35S-GPTV	P2 4a	20.039	0.381	0.061	0.164
20	TN90 220 SRC BBL-C/p4GAG-35S-GPTV	P13 9a	18.697	0.384	0.059	0.154
21	TN90 220 SRC BBL-C/p4GAG-35S-GPTV	P9 5a	15.187	0.335	0.071	0.132
24	TN90 SRC 222	-	5.765	0.147	0.023	0.079
2	TN90 222 SRC BBL-C/p4GAG-PMT1a-GPTV	P4 3a	4.287	0.111	0.060	0.092
3	TN90 222 SRC BBL-C/p4GAG-PMT1a-GPTV	P5 4a	12.621	0.243	0.036	0.122
4	TN90 222 SRC BBL-C/p4GAG-PMT1a-GPTV	P5 1a	3.553	0.100	0.074	0.082
5	TN90 222 SRC BBL-C/p4GAG-PMT1a-GPTV	P6 6a	9.907	0.265	0.032	0.108
14	TN90 222 SRC BBL-C/p4GAG-35S-GPTV	P1 8a	13.331	0.261	0.055	0.116
15	TN90 222 SRC BBL-C/p4GAG-35S-GPTV	P9 8a	14.065	0.311	0.047	0.118
16	TN90 222 SRC BBL-C/p4GAG-35S-GPTV	P12 9a	11.205	0.271	0.072	0.115
17	TN90 222 SRC BBL-C/p4GAG-35S-GPTV	P2 6a	6.341	0.154	0.030	0.092
22	TN90 LC	-	26.638	1.042	0.416	0.126
LSD (α=0.05)			4.54	0.17	0.07	0.03

*Lines highlighted in grey represent those selected for enantiomeric analysis

Figure 1.1 Nicotine enantiomer determinations for *BBL* mutant combinations in the TN 90 (A) and TN 90 SRC (B) genetic backgrounds from the 2017 field experiment. Red bar color and numbers above each bar indicates the fraction of (*R*)-nicotine. Enantiomeric fraction means with different letters are significantly different from each other at the $P < 0.05$ level of significance. Total nicotine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance.

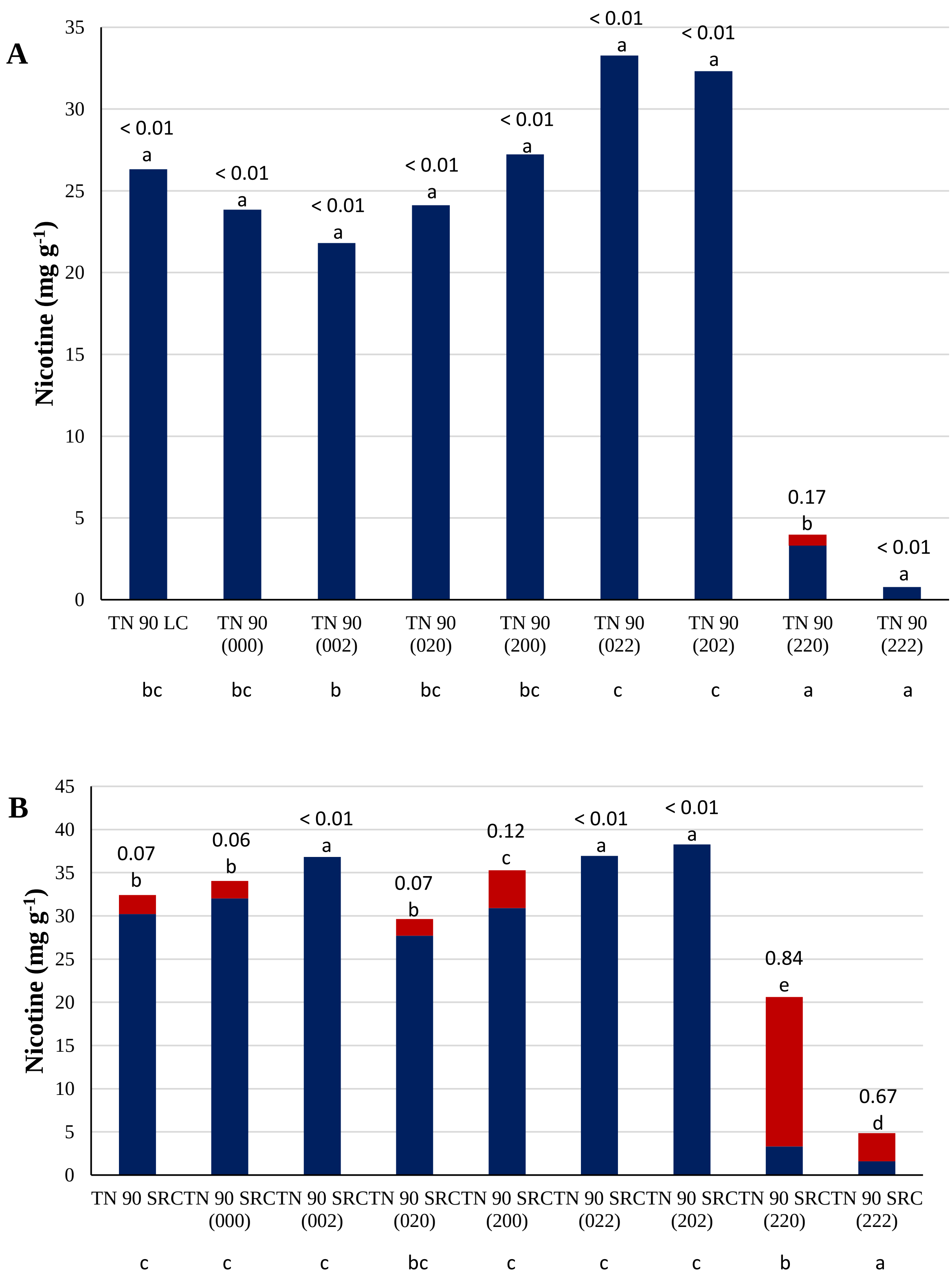
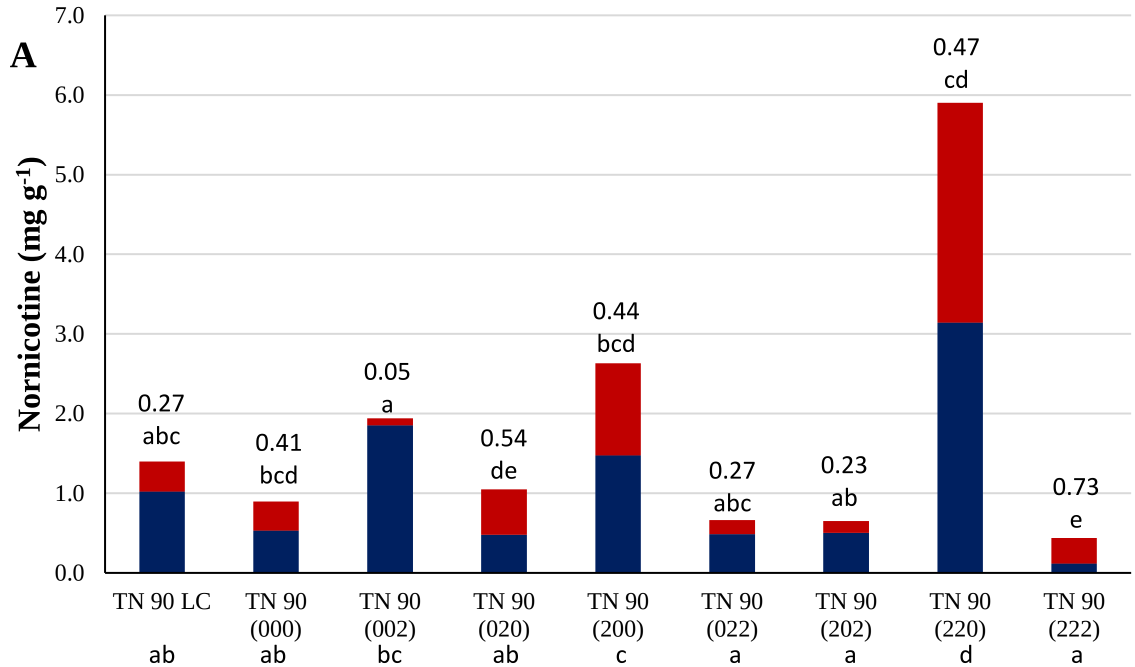

A
Nicotine (mg g-1)
35
30
25
20
15
10
5
0
< 0.01
a
< 0.01
a
< 0.01
a
< 0.01
a
< 0.01
a
< 0.01
a
< 0.01
a
0.17
b
< 0.01
a
TN 90 LC
TN 90 (000)
TN 90 (002)
TN 90 (020)
TN 90 (200)
TN 90 (022)
TN 90 (202)
TN 90 (220)
TN 90 (222)
bc
bc
b
bc
bc
c
c
a
a
B
Nicotine (mg g-1)
45
40
35
30
25
20
15
10
5
0
0.07
b
0.06
b
< 0.01
a
0.07
b
0.12
c
< 0.01
a
< 0.01
a
0.84
e
0.67
d
TN 90 SRC
TN 90 SRC (000)
TN 90 SRC (002)
TN 90 SRC (020)
TN 90 SRC (200)
TN 90 SRC (022)
TN 90 SRC (202)
TN 90 SRC (220)
TN 90 SRC (222)
c
c
c
bc
c
c
c
b
a

Figure 1.2 Nornicotine enantiomer determinations for *BBL* mutant combinations in the TN 90 (A) and TN 90 SRC (B) genetic backgrounds from the 2017 field experiment. Red bar color and numbers above each bar indicates the fraction of (*R*)-nornicotine. Enantiomeric fraction means with different letters are significantly different from each other at the $P < 0.05$ level of significance. Total nornicotine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance.


A
Nornicotine (mg g-1)
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
0.27
abc
0.41
bcd
0.05
a
0.54
de
0.44
bcd
0.27
abc
0.23
ab
0.47
cd
0.73
e
TN 90 LC
TN 90 (000)
TN 90 (002)
TN 90 (020)
TN 90 (200)
TN 90 (022)
TN 90 (202)
TN 90 (220)
TN 90 (222)
ab
ab
bc
ab
c
a
a
d
a

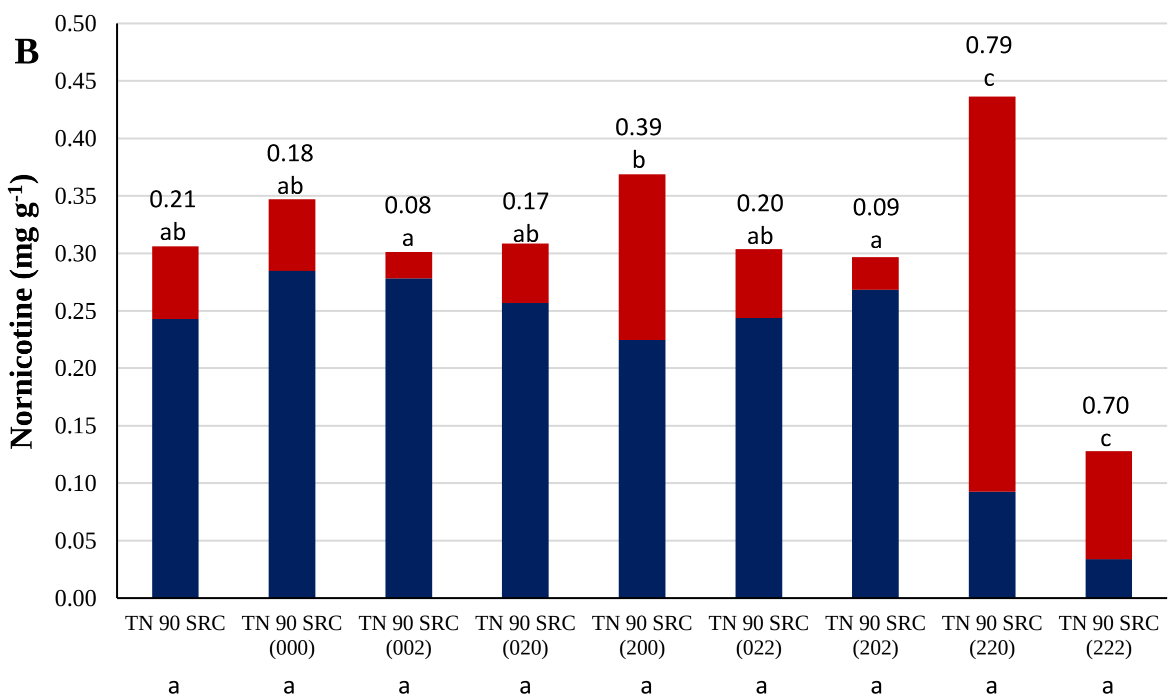
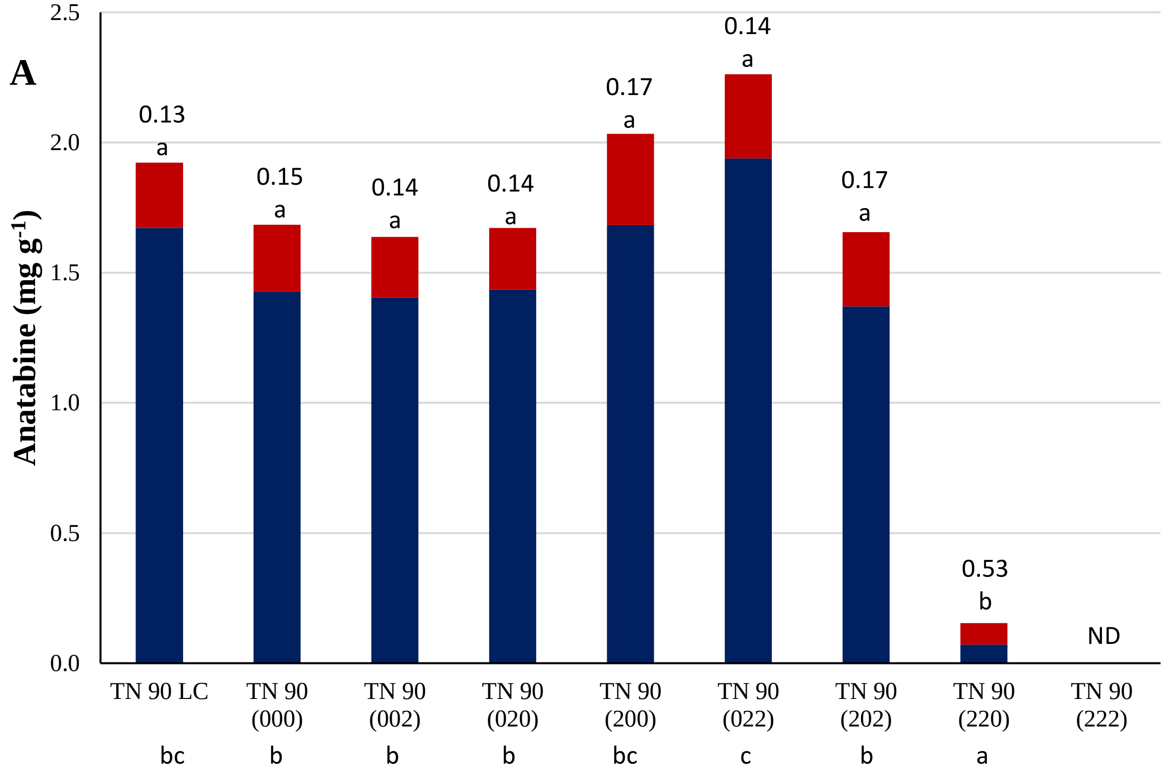

B
Nornicotine (mg g-1)
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.21
ab
0.18
ab
0.08
a
0.17
ab
0.39
b
0.20
ab
0.09
a
0.79
c
0.70
c
TN 90 SRC
TN 90 SRC (000)
TN 90 SRC (002)
TN 90 SRC (020)
TN 90 SRC (200)
TN 90 SRC (022)
TN 90 SRC (202)
TN 90 SRC (220)
TN 90 SRC (222)
a
a
a
a
a
a
a
a
a

Figure 1.3 Anatabine enantiomer determinations for *BBL* mutant combinations in the TN 90 (A) and TN 90 SRC (B) genetic backgrounds from the 2017 field experiment. Red bar color and numbers above each bar indicates the fraction of (*R*)-anatabine. Enantiomeric fraction means with different letters are significantly different from each other at the $P < 0.05$ level of significance. Total anatabine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance. ND = non-detectable.


A
Anatabine (mg g-1)
2.5
2.0
1.5
1.0
0.5
0.0
0.13
a
0.15
a
0.14
a
0.14
a
0.17
a
0.14
a
0.17
a
0.53
b
ND
TN 90 LC
TN 90 (000)
TN 90 (002)
TN 90 (020)
TN 90 (200)
TN 90 (022)
TN 90 (202)
TN 90 (220)
TN 90 (222)
bc
b
b
b
bc
c
b
a

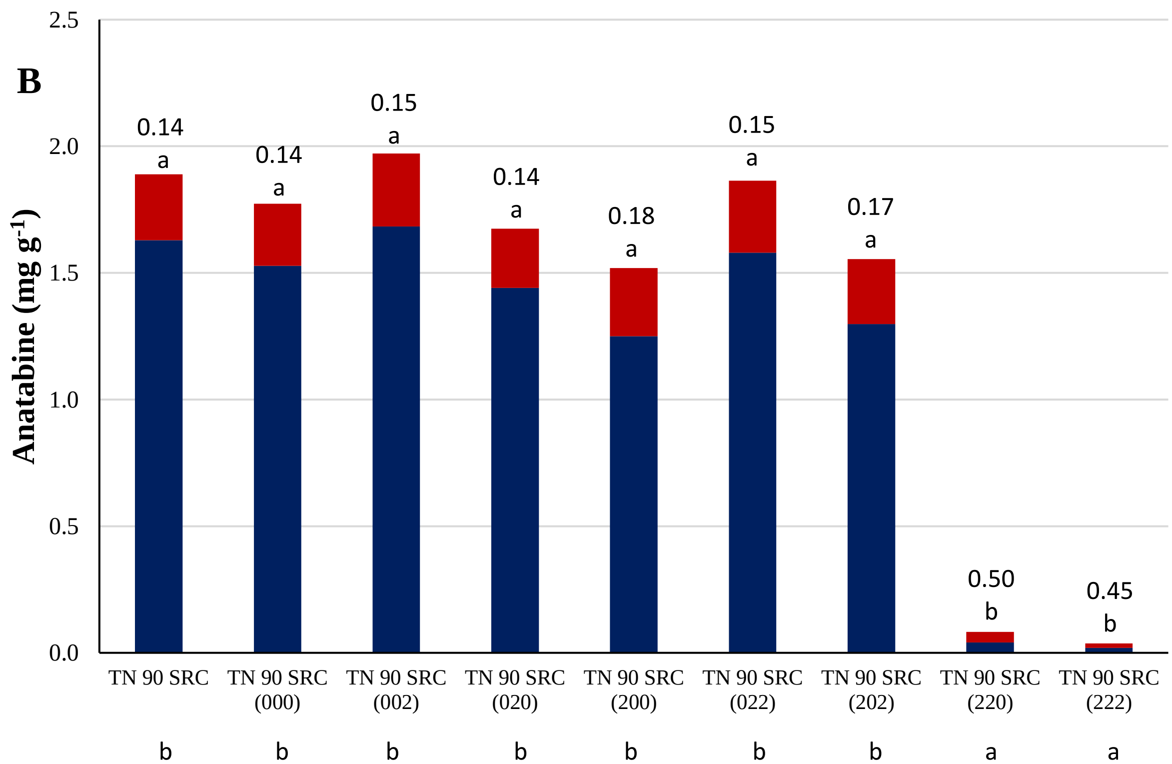
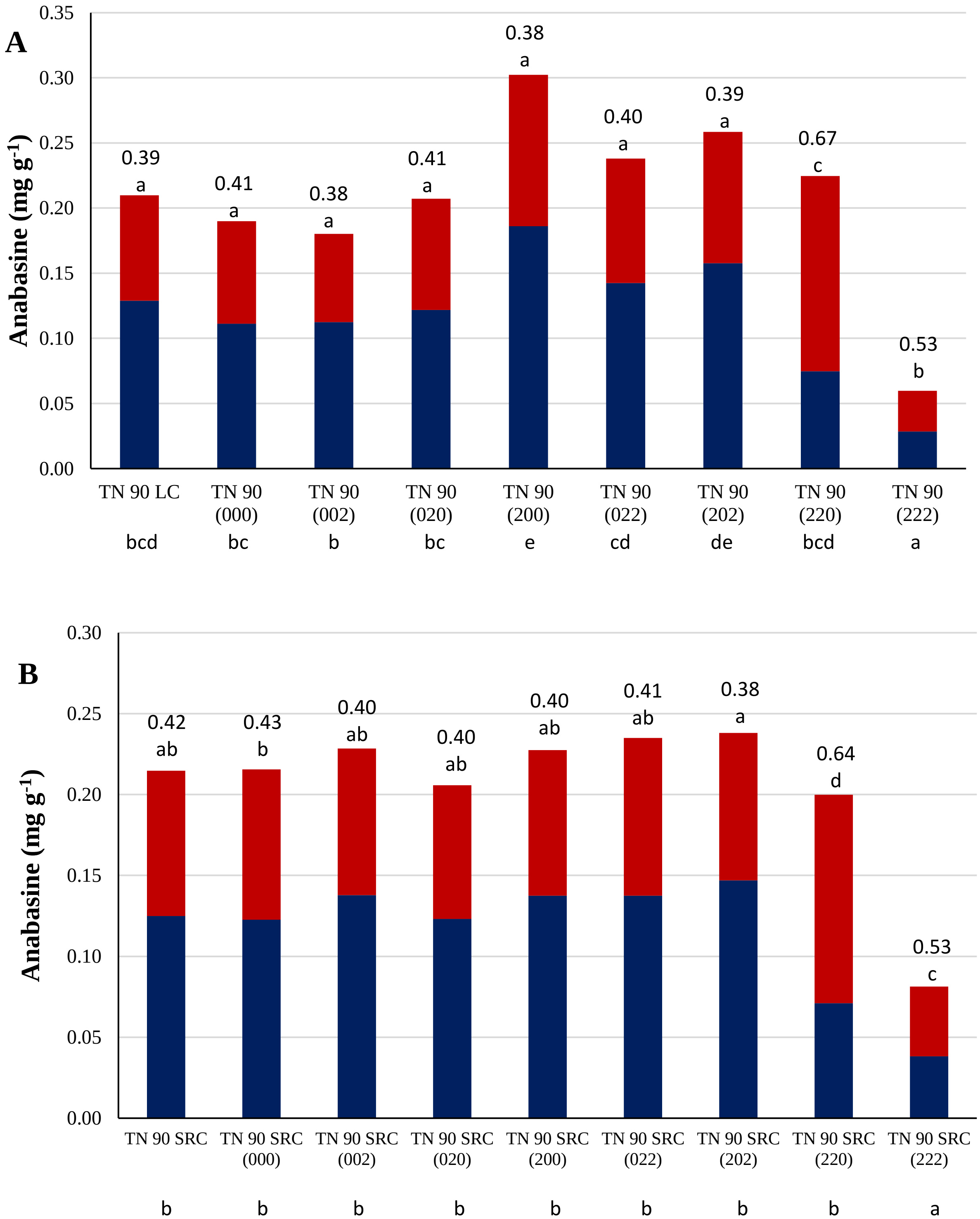

B
Anatabine (mg g-1)
2.5
2.0
1.5
1.0
0.5
0.0
0.14
a
0.14
a
0.15
a
0.14
a
0.18
a
0.15
a
0.17
a
0.50
b
0.45
b
TN 90 SRC
TN 90 SRC (000)
TN 90 SRC (002)
TN 90 SRC (020)
TN 90 SRC (200)
TN 90 SRC (022)
TN 90 SRC (202)
TN 90 SRC (220)
TN 90 SRC (222)
b
b
b
b
b
b
b
a
a

Figure 1.4 Anabasine enantiomer determinations for *BBL* mutant combinations in the TN 90 (A) and TN 90 SRC (B) genetic backgrounds from the 2017 field experiment. Red bar color and numbers above each bar indicates the fraction of (*R*)-anabasine. Enantiomeric fraction means with different letters are significantly different from each other at the $P < 0.05$ level of significance. Total anabasine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance.



A
Anabasine (mg g-1)
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.39
a
0.41
a
0.38
a
0.41
a
0.38
a
0.40
a
0.39
a
0.67
c
0.53
b
TN 90 LC
TN 90 (000)
TN 90 (002)
TN 90 (020)
TN 90 (200)
TN 90 (022)
TN 90 (202)
TN 90 (220)
TN 90 (222)
bcd
bc
b
bc
e
cd
de
bcd
a
B
Anabasine (mg g-1)
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.42
ab
0.43
b
0.40
ab
0.40
ab
0.40
ab
0.41
ab
0.38
a
0.64
d
0.53
c
TN 90 SRC
TN 90 SRC (000)
TN 90 SRC (002)
TN 90 SRC (020)
TN 90 SRC (200)
TN 90 SRC (022)
TN 90 SRC (202)
TN 90 SRC (220)
TN 90 SRC (222)
b
b
b
b
b
b
b
b
a

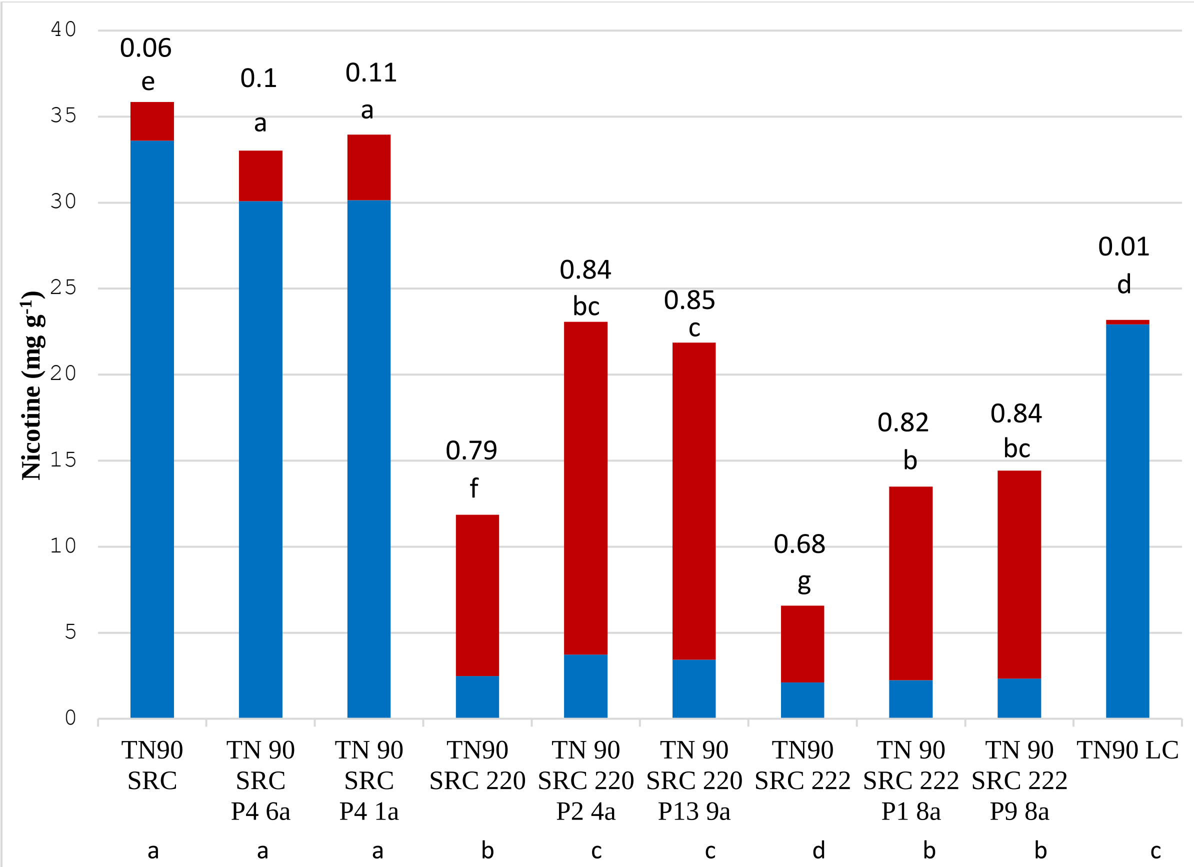


Figure 1.5 Nicotine enantiomer determinations for transgenic *BBLc* overexpressing lines in the TN 90 SRC genetic background and four control lines. Red bar color and numbers above each bar indicate the fraction of (*R*)-nicotine. Enantiomeric fraction means with different letters are significantly different from each other at the $P < 0.05$ level of significance. Total nicotine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance.

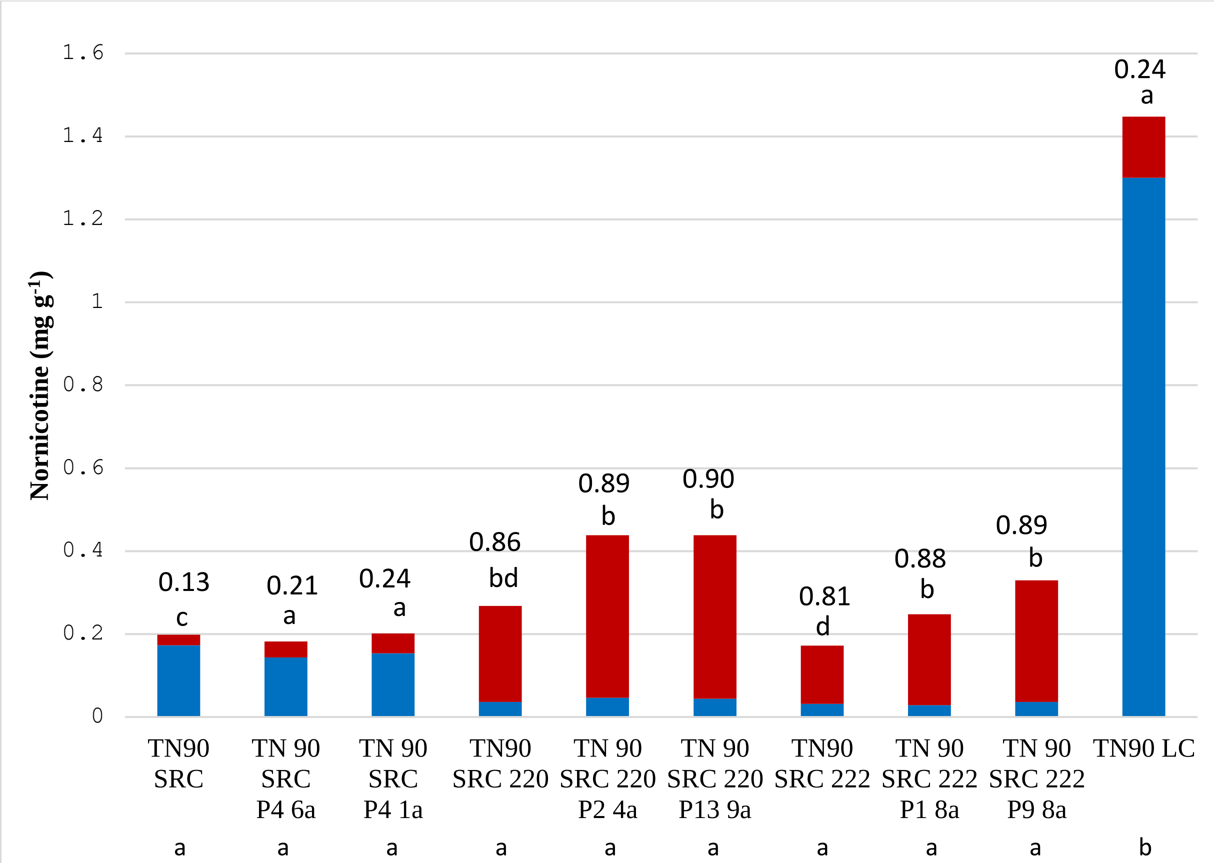


Figure 1.6 Nornicotine enantiomer determinations for transgenic *BBLc* overexpressing lines in the TN 90 SRC genetic background and four control lines. Red bar color and numbers above each bar indicate the fraction of (*R*)-nornicotine. Enantiomeric fraction means with different letters are significantly different from each other at the $P < 0.05$ level of significance. Total nornicotine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance.

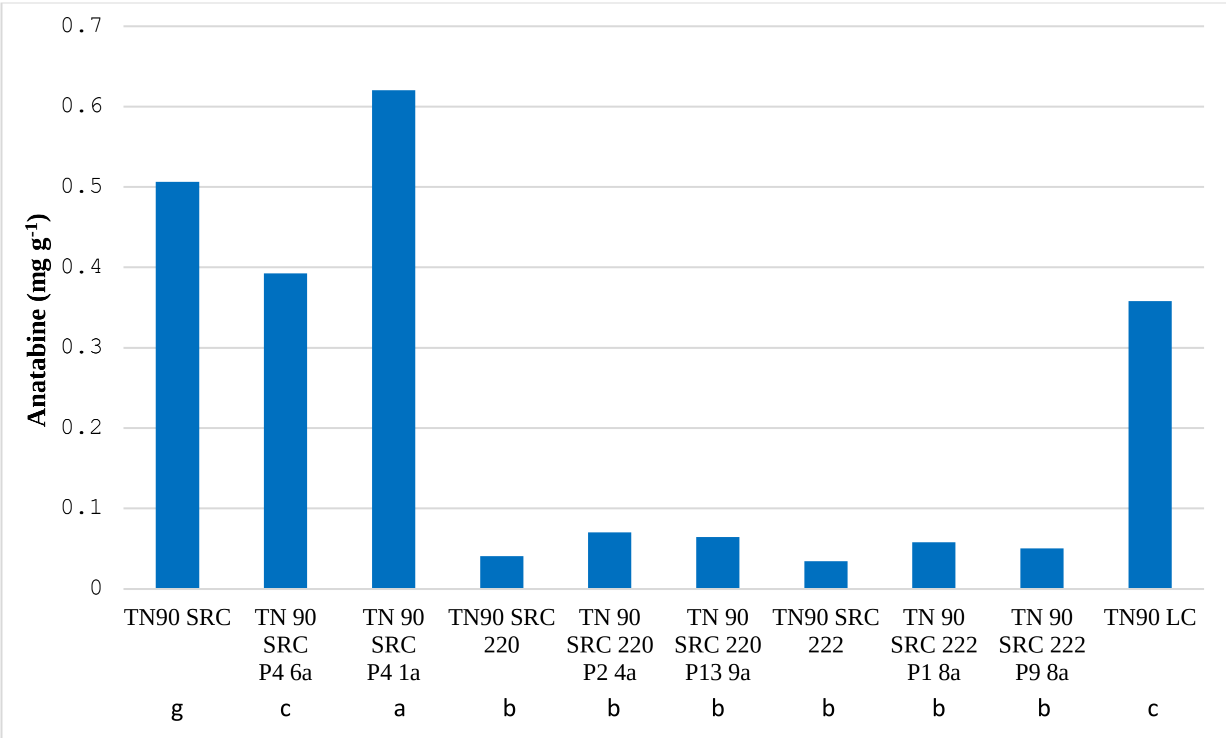


Figure 1.7 Total anatabine determinations for transgenic *BBLc* overexpressing lines in the TN 90 SRC genetic background and four control lines. Total anatabine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance.

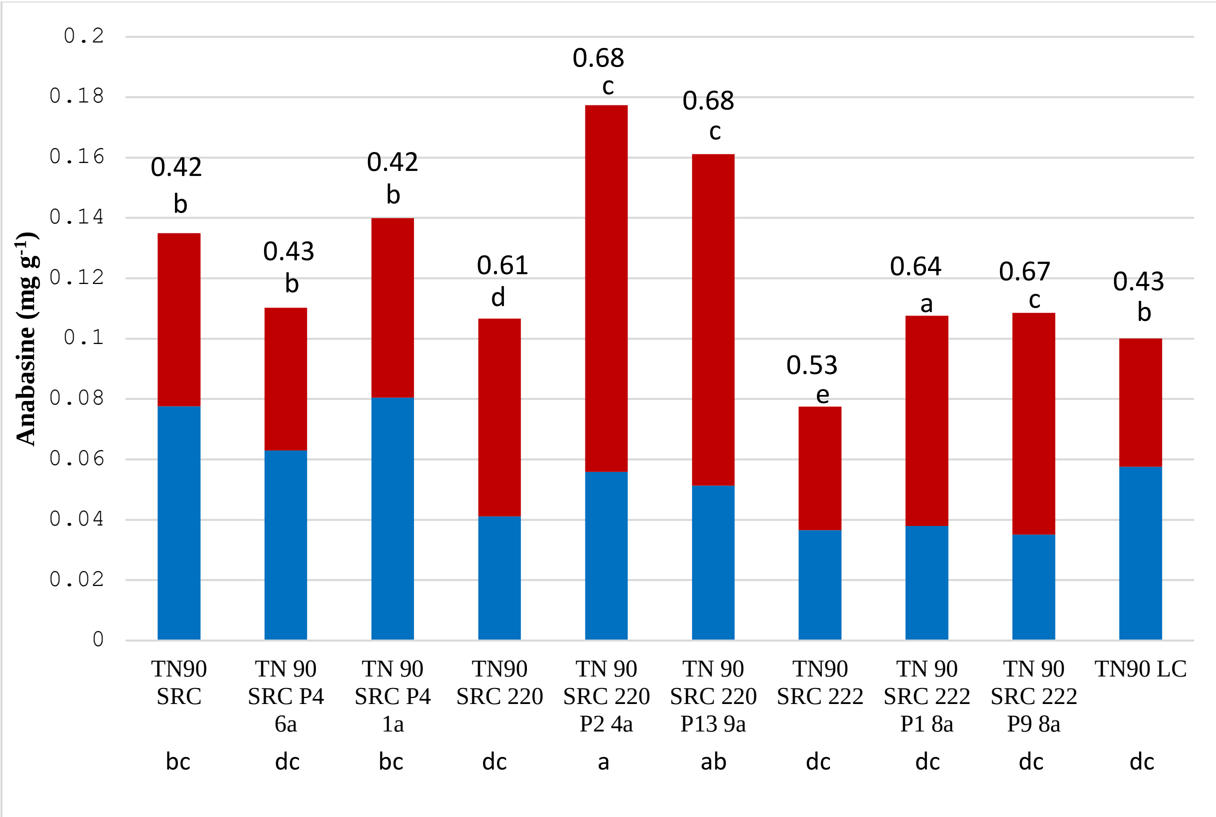


Figure 1.8 Anabasine enantiomer determinations for transgenic *BBLc* overexpressing lines in the TN 90 SRC genetic background and four control lines. Red bar color and numbers above each bar indicate the fraction of (*R*)-anabasine. Enantiomeric fraction means with different letters are significantly different from each other at the $P < 0.05$ level of significance. Total anabasine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance.

Chapter 2: Evaluation of Low-Alkaloid Nearly Isogenic Lines of Tobacco for Pest and Pathogen Resistance and Differential Gene Expression

Abstract

Nicotine is the primary addictive component in various tobacco products, and therefore is a very important natural product both commercially and socially. Potential upcoming regulations that could mandate the lowering of nicotine levels in cigarette filler to below 0.4 mg g^{-1} has caused significant interest in gaining an understanding of the genetic and biochemical control of nicotine biosynthesis in tobacco. Currently, there is no commercial tobacco variety that consistently produces below the recommended level when grown under conventional field production practices. Ten nearly isogenic versions of K326 differing for the presence/absence of genetic variability influencing accumulation were evaluated in field environments for total alkaloid accumulation, cured leaf yields, and insect and disease resistance. In addition, nearly isogenic lines differing at the presence/absence of recessive alleles at the *Nic1* and *Nic2* loci, as well as for the presence/absence of a deleterious 5 bp mutation in *Myc2a*, were evaluated in a greenhouse experiment to generate root transcriptome information to gain insight on the categories of genes transcriptionally regulated by the Nic1, Nic2, and Myc2a transcription factors. K326 *nic1/nic2/myc2a/Ct* was found to produce below the suggested 0.4 mg g^{-1} level, but this reduction came at the expense of increased susceptibility to insects and soil-borne diseases. Interestingly, K326 222 was observed to have significantly improved resistance to both black shank and bacterial wilt. Differentially expressed genes were identified for all comparisons between K326 *per se*, K326 *nic1/nic2*, K326 *myc2a*, and K326 *nic1/nic2/myc2a*. Results support previous studies that suggest Myc TFs such as Myc2a to be significantly involved in the

regulation of nicotine biosynthesis, as well as provides insight into the advantages that certain genetic modifications can have on lowering nicotine content in tobacco plants.

INTRODUCTION

Of the numerous natural products produced by tobacco (*Nicotiana tabacum*) plants, nicotine is perhaps the most important both commercially and socially due to this molecule being the primary addictive component in various tobacco products such as combustible cigarettes. For this reason, there is significant interest in gaining an understanding of the genetic and biochemical control of nicotine biosynthesis in this species. Additionally, the World Health Organization (WHO) has recommended the lowering of nicotine levels in cigarette filler to below 0.4 mg g^{-1} (WHO, 2015). Conventional tobacco typically produces between 10 mg g^{-1} to 50 mg g^{-1} of nicotine on a dry leaf basis (Lewis et al., 2020). This could be a problem for the tobacco industry as there is currently no commercial tobacco variety that consistently produces, on average, less than 0.4 mg g^{-1} nicotine under conventional field production conditions. One of the major grower/industry concerns with the cultivation of low-nicotine tobacco is the possibility of increased susceptibility to biotic and abiotic stresses. It is advantageous for low nicotine tobacco cultivars to possess disease and insect resistance to protect the crop with minimal inputs. This may be very difficult to achieve, however, due to the removal of one of the primary plant defense mechanisms. Another concern is the yield and quality characteristics of low-nicotine tobacco genotypes, which has been a historical problem for such varieties. Leaf nicotine content is a somewhat difficult trait to breed for, due to its accumulation being highly influenced by the environment, production practices, stalk position, as well as genetics (Wu et al., 2013). Genetic approaches used to develop low-nicotine tobacco varieties include the use of: (1) inherent genetic variability within *N. tabacum* or closely related species, (2) genetic diversity induced via gene editing or mutation treatment, and (3) novel variations produced through genetic engineering (Lewis et al., 2019). Using naturally occurring variability to achieve below 0.4 mg g^{-1} nicotine on a dry weight basis is ideal because it avoids regulatory issues that can be associated with the use of technologies such as gene editing and genetic engineering. Gaining insight into the genetic control of the low-nicotine trait could benefit the development of commercially acceptable low-nicotine tobacco varieties.

A substantial amount of research has been carried out to understand the nicotine biosynthetic pathway (Dewey and Xie, 2013; Guo et al., 2021), although the complete process is still not fully understood. Gaining knowledge about the genes coding for enzymes and transcription factors that regulate or control nicotine biosynthesis may lead to genetic strategies to modify the accumulation of this alkaloid in tobacco plants. Genes that have already been identified to be involved in the later stages of nicotine biosynthesis include the *Berberine Bridge Like* (*BBL*) gene family and the *A622* gene family. Previous studies have suggested that A622 enzymes likely play a role prior to the BBL enzymes in the steps between nicotinic acid biosynthesis and the synthesis of the major tobacco alkaloids (Kajikawa et al., 2009). It seems reasonable to believe there may be a gene(s) associated with one of the final stages of nicotine biosynthesis that has yet to be identified, however.

The majority of genes currently known to be involved in tobacco alkaloid biosynthesis were previously identified through differential expression experiments involving *Nic1*/*Nic1 Nic2*/*Nic2* and *nic1*/*nic1 nic2*/*nic2* isolines. The *Nic1* and *Nic2* genomic regions contain genes coding for multiple Ethylene Response Factor (ERF) transcription factors which positively regulate the expression of structural genes associated with the nicotine biosynthetic pathway. A recent study conducted by Sui et al. (2023) revealed that in *nic1*/*nic2* mutant tobacco plants, the expression of genes involved in response to abiotic and biotic stresses are also significantly down-regulated (Sui et. al, 2023).

Although substantial work has been carried out to identify the structural genes associated with nicotine biosynthesis, these studies may not have revealed a gene(s) coding for a possible currently unknown enzyme involved in one of the final stages of nicotine biosynthesis. It is feasible that the possible unknown gene(s) is differentially expressed between the *Nic1/Nic2* isolines and simply was not identified in previous studies (Katoh et. al, 2007; Shoji et. al, 2009; Shoji et. al, 2010). Alternatively, expression of the possible unknown gene(s) may not be regulated by the *Nic1/Nic2* regulatory loci. The expression of this unknown gene could potentially be regulated by another regulatory factor such as a Myc transcription factor, which has been reported to regulate an expanded suite of genes (Sui et. al, 2021). Previously, through the use of QTL mapping, Burner et al. (2022) identified a naturally occurring 5 base pair (bp) deletion in the *Myc2a* gene in tobacco accession TI 313. The result of this mutation is significantly reduced nicotine levels in tobacco plants.

We previously generated a BC_6F_3 nearly isogenic version (NIL) of K326 homozygous for the recessive alleles at the *Nic1* and *Nic2* loci derived from LAFC53 designated as K326 *nic1*/*nic2*. In addition, we have developed a BC_6F_3 nearly isogenic version of K326 that is homozygous for the deleterious mutation in *Myc2a* derived from TI 313 designated as K326 *myc2a*. We have also developed a BC_6F_3 NIL that, in addition to the *myc2a* mutation, is also homozygous for the recessive *nic1* and *nic2* alleles. This NIL was designated as K326 *nic1/nic2/myc2a*. We anticipated that the K326 *nic1/nic2/myc2a* NIL might exhibit a greater number of differentially expressed genes relative to K326 *nic1*/*nic2*. The additional mutation in this NIL could potentially assist in the identification of an additional gene(s) involved in nicotine biosynthesis. In addition to the three aforementioned NIL versions of K326, we have developed seven additional low-nicotine accumulating NIL versions of K326 possessing different combinations of genes affecting nicotine biosynthesis using backcross breeding methodologies.

The first objective of this research was to analyze the ten nearly isogenic versions of K326 in field environments for total alkaloid accumulation, cured leaf yields, and insect and disease resistance. The second objective of this research was to carry out an RNA sequencing experiment involving the NILs differing for the presence/absence of recessive alleles at the *Nic1* and *Nic2* loci, as well as for the presence/absence of the deleterious 5 bp mutation in *Myc2a,* to generate root transcriptome information to gain insight on the categories of genes transcriptionally regulated by the Nic1, Nic2, and Myc2a transcription factors. It was thought that such data might also contribute to the identification of an additional gene(s) involved in nicotine biosynthesis.

MATERIALS AND METHODS

Plant Materials

Popular flue-cured tobacco cultivar K326 and ten low nicotine-accumulating K326 backcross-derived lines (Table 2.0) were evaluated for this experiment. K326 *nic1/nic2* is a BC_6F_3 NIL version of K326 homozygous for the recessive alleles at the *Nic1* and *Nic2* loci derived from LAFC53. K326 *myc2a* is a BC_6F_3 NIL version of K326 that contains the naturally occurring 5 bp deletion in *Myc2a* derived from TI 313 (Burner et al., 2022). A BC_6F_3 NIL homozygous for the recessive *nic1* and *nic2* alleles derived from LAFC53 and the 5 bp *myc2a* mutation was designated as K326 *nic1/nic2/myc2a.* K326 222 was previously described as being homozygous for EMS-induced mutations in the three most highly expressed *BBL* genes (*BBL-a, BBL-b, BBL-c*) that encode for enzymes involved in one of the final stages of nicotine biosynthesis (Lewis et al., 2020). K326 HIC is a version of K326 selected by the NCSU tobacco breeding program that stably converts a high proportion (greater than 85%) of nicotine into nornicotine. K326 222 HIC is a stable high high-converting version of K326 222. K326 222 *nic1/nic2* is a previously described (Lewis et al., 2020) version of K326 homozygous for the EMS-induced mutations in the *BBL* genes as well the recessive alleles at the *Nic1* and *Nic2* loci derived from LAFC53. K326 *nic1/nic2/ct* contains *CYP83E3*, a highly active *Nicotine N-demethylase* (*NND)* gene introgressed from *N. tomentosiformis* (Gavilano et al., 2007), and the recessive *nic1* and *nic2* alleles derived from LAFC53. In addition to the recessive *nic1* and *nic2* alleles and the *Ct* gene, K326 *nic1/nic2/myc2a/Ct* also contains the 5 bp deletion in *myc2a*. Lastly, K326 *nic1/nic2/cs* possesses the recessive *nic1* and *nic2* alleles as well as the functional *CYP82E2 NND* gene, derived from *N. sylvestris*.

Screening Allelic Variability at Nic1/Nic2 and Myc2a Loci

K326 BC_6F_3 individuals segregating for the presence/absence of the 5 bp *myc2a* mutation were screened via a cleaved amplified polymorphic sequence (CAPS) marker assay. Primers were designed to amplify *myc2a* using polymerase chain reaction (PCR) (Table 2.1), using K326 as the negative control and TI 313 as the positive control. PCR products were digested using the *Hpy188I* (New England Biolabs, USA) restriction enzyme. Presence or absence was confirmed through gel electrophoresis on 1% agarose in 1x TAE buffer. Ethidium bromide was added to the gel prior to solidification for visualization of the fragments under UV light. Kompetitive Allele Specific PCR (KASP) marker assays were performed using six primer pairs for allele discrimination at the *Nic1* locus (Nic1-1_s9, Nic1-5, Nic1-9, Nic1-13_s21, Nic1-16, Nic1-20). Two primer pairs were used to assay for the presence/absence of genes (*ERF115* and *ERF189*) at the *Nic2* locus (Tables 2.2, 2.3).

Field Evaluations

The ten K326-derived NILs were evaluated in field experiments for yield, disease and insect resistance, and cured leaf alkaloid composition in comparison with K326. Experiments were carried out at four North Carolina field locations (Upper Coastal Plain Research Station, Rocky Mount; Cunningham Research Station, Kinston; Oxford Tobacco Research Station, Oxford; and the Central Crops Research Station, Clayton) during 2023. Experiments to evaluate the materials for yield and leaf alkaloid levels were conducted in Rocky Mount. Disease resistance evaluations were conducted in Kinston and Rocky Mount, and the insect resistance trials were conducted in Clayton and Oxford. Experimental units for yield and alkaloid evaluations consisted of single 20-plant rows and experimental units for the disease and insect resistance trials were single 12-plant rows managed according to standard flue-cured production practices for North Carolina. Intra-row spacing was 56 cm at all four locations, while inter-row spacing was 122 cm at the Oxford and Rocky Mount locations and 112 cm at the Kinston locations. The experimental design was a randomized complete block design at each location. The cured leaf alkaloid composition and yield experiment consisted of three replications, and the disease and insect resistance studies consisted of eight replications in each environment.

Yield and Alkaloid Analysis

For the yield experiment, leaves were harvested in four separate harvests (primings) and flue-cured. Each priming was weighed to generate yield data. Yield was measured in grams per plot and then converted to kilograms per hectare (kg ha^{-1}). Fifty-gram cured leaf composite samples were prepared for each plot by compositing cured leaf from each priming on a weighted-mean basis. Oven-dried samples were ground to pass through a 1-mm sieve and analyzed for alkaloid profiles (expressed as a percentage of dry weight) as previously outlined by Lewis et al. (2015).

Disease Resistance Analysis

The ten selected low alkaloid lines and K326 were evaluated for black shank resistance at two North Carolina locations naturally infested with the black shank pathogen *Phytophthora nicotianae* (Lower Coastal Plain Research Station, Kinston, NC and Upper Coastal Plain Research Station, Rocky Mount, NC) in 2023. The set of eleven lines was also evaluated for resistance to bacterial wilt, caused by the soil-borne pathogen *Ralstonia solanacearum*, in a disease nursery at the Upper Coastal Research Plain Research Station during 2023. Starting at approximately 40 days after transplanting, the number of plants killed by black shank or bacterial wilt was recorded. This continued throughout the growing season at intervals of approximately 14 days. At the end of the growing season, the area under the disease progress curve (AUDPC) was calculated for each plot using the trapezoidal method of Madden et al. (2007).

Insect Resistance Analysis

K326 and the ten NILs were evaluated for resistance to budworms (*Heliothis virescens*) at the Central Crops Research Station. Approximately 40 days after transplanting, the number of budworms per plant was calculated for the first 10 plants (after the first plant in each row) in each plot in 7-10 day intervals until the plants started to flower and counts were averaged over all the weeks. These eleven lines were also evaluated for resistance to tobacco hornworms (*Manduca sexta*) at the Central Crops Research Station. Beginning in early August, hornworms were counted on the first 10 plants (after the first plant in each row) in each plot in the night. Hornworms were illuminated for counting using an Escolite UV Flashlight Black Light, 51 LED 395 nM Ultraviolet Blacklight. Counts were taken in approximately 7-day intervals for 5 consecutive weeks and counts were averaged over the five weeks.

The eleven lines were also evaluated for resistance to aphids (*Myzus persicae*) at the Oxford Tobacco Research Station. The number of plants infested with aphids was counted on the first 10 plants (after the first plant in each row) in each plot in 7-10 day intervals until aphids were no longer present. Plants were counted as being infested with aphids if there were 50 or more aphids on one leaf of the plant. For the analysis, the number of infested plants each week was averaged over all weeks for each plot.

Data Analysis

Analyses of variance appropriate for analyzing a randomized complete block design were carried out for the aforementioned experiments using PROC MIXED of SAS 9.3 (SAS Institute, Cary, NC, United States) according to Littell et al. (1996). Environment and replication were considered as random effects, while genotype was considered as a fixed effect. Mean separations were performed via appropriate F-tests or by using Fisher's protected least significant difference test ($\alpha = 0.05$).

RNA-seq Data Generation

Four to six plants from K326 and each of its corresponding NILs, K326 *nic1/nic2*, K326 *myc2a*, and K326 *nic1/nic2/myc2a*, were grown in the greenhouse in pots positioned within 5-gallon buckets with enough water to submerge the bottom ~1 inch of each pot (Table 2.4). This was done to allow for efficient collection of root tissues for RNA extraction after they grew through the holes at the bottom of the pot. The water levels were checked twice daily to ensure the bottom of the pots was continuously submerged. Plants were grown until around the time of flowering and then topped. Eighteen hours after topping, roots were excised from each plant, transferred to liquid nitrogen for rapid freezing, and eventually stored at -80°C until RNA purification using the RNeasy mini kit (Qiagen) with DNase treatment step. Total RNA of three to five biological replicates for each line was submitted to Novogene (Sacramento, CA) for RNA sequencing (RNA-seq) library construction and sequencing using NovaSeq PE150 (Novagene) and the 150 bp paired-end read mode (PE150). The top two leaves were collected 21 days after topping, dried, and ground for alkaloid evaluations.

RNA-seq Read Mapping, Differentially Expressed Gene (DEG) Identification, and Gene Ontology Enrichment Analyses

Raw sequencing reads were stored as fastq files and quality checking was performed using FastQC (Andrews, 2010). The raw reads were trimmed to remove poor quality bases (>Q30) using fastq-mcf (Aronesty 2011). After quality trimming was conducted, clean reads were mapped to the reference genome assembly of *N. tabacum* K326 (Nitab-v4.5_genome_Scf_Edwards2017.fasta; Edwards et al. 2017) using hisat2-2.1.0 (Kim et al. 2015). Mapped read data was converted into BAM file format and subsequently filtered using Samtools (version 1.6) to maintain proper paired and high-quality mapped reads (Li et al., 2009). RNA-seq read counts mapping to transcript sequences of annotated gene models were extracted from sorted and indexed BAM files using the multicov function of the Bedtools package (version v2.23.0, Quinlan et al., 2010) according to gene model information converted from available gene annotations from Edwards et al. (2017).

RNA-seq raw read counts for each gene model were analyzed using DESeq2, an R package for data normalization, data exploration, and DEG identification (Love et al., 2014). In the output CSV file, DEGs were filtered by screening all genes with false discovery rate (FDR) < 0.05, and log(2) fold change < 1 for relative expression levels between all pairwise comparisons amongst K326, K326 *nic1/nic2*, K326 *myc2a*, and K326 *nic1/nic2/myc2a*.
DEGs were subjected to Gene Ontology (GO) enrichment analysis using AgriGO.v2 (Du et al., 2010) against Nitab4.5 ID (Solgenomics) as the background with default settings.

RESULTS

Low Alkaloid NIL Field Data

Alkaloid data was collected for the 10 low nicotine accumulating NILs of K326, in addition to K326 *per se*, at the Upper Coastal Plain Research Station in Rocky Mount, North Carolina. Consistent with levels typically observed for conventional flue-cured tobacco cultivars (Lewis et al., 2020), K326 contained 22.42 mg g^{-1} of nicotine in its leaves on a dry weight basis (Figure 2.1). All 10 of the tested genotypes accumulated significantly lower ($P < 0.05$) leaf nicotine levels compared to the K326 control line. For K326 HIC, nicotine levels were reduced to 5.43 mg g^{-1}, however, nornicotine levels for this genotype increased dramatically to 10.47 mg g^{-1} on a dry leaf basis relative to K326 *per se* (Figure 2.1). K326 222, which is homozygous for EMS-induced mutations in the three most highly expressed *BBL* genes (Lewis et al., 2020), exhibited a mean nicotine level of 2.48 mg g^{-1}. Further reductions in nicotine (0.45 mg g^{-1}) were observed for the high nicotine-converting version of this genotype (K326 222 HIC) (Figure 2.1).

The NIL of K326 homozygous for the recessive alleles at the *Nic1* and *Nic2* loci (K326 *nic1/nic2*) accumulated an average level of 4.04 mg g^{-1} of nicotine in its leaves. Nicotine levels were further reduced when the *nic1* and *nic2* alleles derived from LAFC53 were combined with one of the introgressed *NND* genes (*CYP82E2* or *CYP82E3*) (Figure 2.1). When the known *NDM* gene *CYP82E2,* derived from *N. sylvestris*, was active in the K326 *nic1/nic2* background (K326 *nic1/nic2/Cs*), nicotine content in the leaves was reduced to 0.54 mg g^{-1} (Figure 2.1). Similarly, when the *NND* gene *CYP82E3* derived from *N. tomentosiformis* (Gavilano et al., 2007) was active in the K326 *nic1/nic2* background (K326 *nic1/nic2/Ct*) nicotine content in the leaves was observed to be 0.43 mg g^{-1} (Figure 2.1).

K326 *myc2a*, which is homozygous for the previously described 5 bp deletion in the *Myc2a* gene derived from TI313 (Burner et al., 2022), accumulated 5.0 mg g^{-1} of nicotine in its leaves. When this mutation was transferred into the K326 *nic1/nic2* and K326 *nic1/nic2/Ct* genetic backgrounds, nicotine accumulation was further reduced to 0.91 mg g^{-1} and 0.36 mg g^{-1}, respectively (Figure 2.1).

As previously mentioned, nornicotine levels were dramatically increased in K326 HIC due to the high nicotine conversion rate. K326 *myc2a* was the only other tested genotype that exhibited a significantly ($P < 0.05$) greater level of nornicotine as compared to the control line, with a concentration of 5.13 mg g^{-1} (Figure 2.2). K326 *nic1/nic2*, K326 222 *nic1/nic2*, and K326 *nic1/nic2/myc2a* all exhibited significantly ($P < 0.05$) lower levels of nornicotine compared to the control line, with concentrations between 0.11 and 0.04 mg g^{-1} (Figure 2.2).

For anabasine, K326 HIC was the only tested line that did not exhibit a significantly ($P < 0.05$) lower amount in its leaves relative to K326 *per se* (Figure 2.3). Notably, K326 222 *nic1/nic2*, K326 *nic1/nic2/myc2a*, K326 *nic1/nic2/myc2a/Ct*, K326 *nic1/nic2/Ct*, and K326 *nic1/nic2/Cs* all contained only approximately 0.001 mg g^{-1} of anabasine in their leaves (Figure 2.3). Similar observations were made for anatabine, where K326 HIC exhibited significantly ($P < 0.05$) greater amounts of anatabine compared to K326, while all the other tested lines exhibited significantly ($P < 0.05$) lower amounts of anatabine (Figure 2.4). K326 222 *nic1/nic2* and K326 *nic1/nic2/myc2a* accumulated approximately 0.001 mg g^{-1} of anatabine (Figure 2.4).

The number of days from transplanting to flowering was recorded for the 10 low-nicotine NILs in addition to the K326 control line at the Clayton environment. The average number of days from transplanting to flowering was determined for all plots, and a plot was considered to be flowering when 50% of the plants within the plot were flowering.

The earliest flowering genotype was K326 *nic1/nic2/Cs*, with a mean flowering time of 72.5 days (Figure 2.5). K326 had a mean flowering time of 75.5 days post-transplant (Figure 2.5). Only K326 222 and K326 222 HIC exhibited significantly longer ($P < 0.05$) flowering times compared to the control, with mean flowering times of 82.7 and 84.6 days post-transplant, respectively (Figure 2.5). No line exhibited a mean flowering time significantly shorter ($P < 0.05$) than K326.

Yield data was obtained from the Upper Coastal Plain Research Station in Rocky Mount, North Carolina. All lines yielded between 2055 and 3282 kg ha^{-1}. There is little statistical power in this data, however, due to the data coming from only one location (Figure 2.6). Interestingly, all lines yielded numerically less than K326, other than K326 *myc2a*, which yielded about 200 kg ha^{-1} more than K326 (Figure 2.6).

Resistance of Low Alkaloid NILs to Insects, Bacterial Wilt, & Black Shank

Budworm Data

Budworms were present on all genotypes during the growing season at the Central Crops location, but the severity of the infestation was mild. K326 had the lowest average number of budworms with a mean of 0.5625 budworms per plot (Figure 2.7). Six of the ten tested genotypes had significantly more (P < 0.05) budworms present than that for K326 (Figure 2.7). K326 222 HIC had the greatest average number of budworms with an average of 2.03 budworms per plot (Figure 2.7).

Hornworm Data

Hornworm counts were originally set to start in July at the Central Crops Research Station, but due to a lack of hornworm presence, the first count was not taken until the second week in August. Four out of the ten tested NILs had significantly greater (P < 0.05) hornworm presence compared to K326, and no genotype had significantly less presence (Figure 2.8). The lowest nicotine accumulating line (K326 *nic1/nic2/myc2a/Ct*) was also observed to have the highest hornworm presence of about 1.6 hornworms per plot (Figure 2.8). Interestingly, K326 *nic1/nic2/Cs* was the fourth lowest nicotine accumulating line but had the lowest hornworm presence at about 0.3 hornworms per plot (Figure 2.8).

Aphid Data

Based on the analysis of variance procedure, no significant differences ($P < 0.05$) were observed amongst the tested genotypes for aphid presence (Figure 2.9). All lines had an average number of plants per plot infested with aphids of between 0.18 and 1.3 (Figure 2.9). K326 *nic1/nic2/Ct* had the lowest presence, and K326 *nic1/nic2* had the highest presence (Figure 2.9). This data was collected from only a single location, however, so the statistical power is low for this analysis.

Bacterial Wilt

The ten low nicotine accumulating NILs derived from K326, in addition to K326 as the control, were evaluated for bacterial wilt resistance at the Upper Coastal Plain Research Station during the 2023 growing season. K326 *nic1/nic2/myc2a/Ct* and K326 *nic1/nic2/Cs* were the only lines that exhibited significantly less ($P < 0.05$) resistance as compared to K326 (Figure 3.0). Only K326 222 exhibited significantly greater ($P < 0.05$) resistance to bacterial wilt as compared to K326 *per se* (Figure 3.0).

Black Shank

The same ten NILs and K326 were evaluated for black shank resistance at the Lower Coastal Plain Research Station in Kinston, NC, and the Upper Coastal Plain Research Station in Rocky Mount, NC during the 2023 growing season. Of the ten tested lines, three lines exhibited significantly less ($P < 0.05$) resistance than K326 (Figure 3.1). Consistent with the bacterial wilt resistance analysis, K326 *nic1/nic2/myc2a/Ct* and K326 *nic1/nic2/Cs* were among the three lines that were significantly more susceptible to black shank than K326 (Figure 3.1).

Similar to the bacterial wilt analysis, K326 222 was the only line that exhibited significantly greater ($P < 0.05$) resistance to black shank as compared to the K326 control line (Figure 3.1). Importantly, this line was still significantly more resistant than K326 when tested at the $P < 0.01$ threshold.

Greenhouse Experiment to Identify Differentially Expressed Genes

Alkaloid data were generated for plants for the greenhouse RNAseq experiment. This included data from five K326, four K326 *nic1/nic2*, six K326 *myc2a*, and five K326 *nic1/nic2/myc2a* plants. The K326 genotype exhibited significantly ($P < 0.05$) higher nicotine concentrations compared to the three tested NILs, ranging from 6.548 to 11.838 mg g^{-1} (Table 2.4). The three low-nicotine NILs exhibited nicotine concentrations ranging from 0.072 to 1.607 mg g^{-1} (Table 2.4, Figure 3.2) and the means were not statistically different from each other. Nornicotine concentrations were in the range of 0.11 to 0.245 mg g^{-1} except for plants from the K326 *myc2a* background which had concentrations between 1.077 and 4.495 mg g^{-1} (Table 2.4).

Differentially expressed genes were identified and filtered for all pairwise comparisons amongst K326, K326 *nic1/nic2*, K326 *myc2a*, and K326 *nic1/nic2/myc2a* using the R package DESeq2 with a FDR threshold of <0.05, and log(2) fold change threshold of > 1 for relative expression levels. A Multidimensional Scaling (MDS) plot was created, illustrating expression variation amongst the four genotypes (Figure 3.3). The MDS plot revealed that transcription factor genes at the *Nic1* and *Nic2* loci altered global gene expression in the roots more than *Myc2a* when compared to the control (K326).

Overall, 751 genes exhibited differential gene expression in the comparison between K326 and K326 *nic1/nic2* (Table 2.5). In comparison between K326 and K326 *myc2a*, only 104

DEGs were identified. Interestingly, even though K326 *nic1/nic2/myc2a* exhibited the lowest average nicotine concentrations among the four genotypes, an intermediate number of DEGs (288) were identified in the comparison between this genotype and K326 *per se*. In the K326 and K326 *nic1/nic2* comparison, 435 DEGs were repressed in K326 *nic1/nic2* and 316 DEGs were upregulated (Table 2.5). In the comparison between K326 and K326 *myc2a*, 81 DEGs were repressed in the K326 *myc2a* genotype compared to 23 DEGs that were upregulated (Table 2.5). Lastly, in the comparison between K326 and K326 *nic1/nic2/myc2a*, 183 DEGs were repressed in the K326 *nic1/nic2/myc2a* genotype compared to 105 DEGs that were upregulated (Table 2.5).

Of the genes known to be directly associated with nicotine biosynthesis pathway (Table 2.6), more genes were significantly down-regulated in the roots of K326 *nic1/nic2* than the roots of K326 *myc2a* (Table 2.6). This was the case even though both genotypes reduced nicotine content to a similar degree (Table 2.6). Of the 104 DEGs identified in the K326 *myc2a* versus K326 comparison, three (Nitab4.5_0008037g0020.1, Nitab4.5_0002522g0070.1, Nitab4.5_0000633g0060.1) are known to be related to the nicotine biosynthesis pathway (Table 2.7). Of the 751 DEGs identified in the K326 *nic1/nic2* comparison with K326, 22 were identified as being genes involved in the nicotine biosynthesis pathway including Nitab4.5_0008037g0020.1, and Nitab4.5_0002522g0070.1 which were also identified as being repressed in the K326 *myc2a* genotype. Notably, neither *ERF199* (Nitab4.5_0003090g0030) nor *ERF189* (Nitab4.5_0015055g0010) were identified as being repressed in the K326 *myc2a* versus K326 comparison.

Gene Ontology (GO) Analysis of DEGs

Down-regulated genes identified in the K326 *nic1/nic2* versus K326 comparison were enriched in 113 GO terms, while 20 GO terms were enriched for up-regulated genes in K326 *nic1/nic2*. The GO terms enriched for DEGs downregulated by *nic1/nic2* were associated with a large range of 'biological processes' and 'cellular components,' including 'nucleosome assembly' (GO:0006334), 'nucleosome organization' (GO:0034728), 'gene expression' (GO:0010467), 'protein complex' (GO:0043234), and 'DNA packaging complex' (GO:0044815) (Table 2.8).

Down-regulated genes identified for K326 *myc2a* in comparison with K326 were enriched in 19 GO terms (Table 2.9), and no significant GO terms were enriched for up-regulated genes. All of the GO terms enriched for down-regulated genes were also enriched in the K326 *nic1/nic2* vs K326 comparison. The majority of the down-regulated genes were part of the "cellular component" including 'intracellular membrane-bounded organelle' (GO:0043231), 'membrane-bounded organelle' (GO:0043227), 'intracellular organelle' (GO:0043229), but also included "biosynthetic process" terms such as 'organic substance biosynthetic process' (GO:1901576), 'organonitrogen compound metabolic process' (GO:1901564), and 'organonitrogen compound biosynthetic process' (GO:1901566) (Table 2.9).

Down-regulated genes identified for K326 *nic1/nic2/myc2a* in comparison with K326 were enriched in 38 GO terms, while no significant GO terms were enriched for upregulated genes. All of the GO terms enriched for down-regulated genes were part of the "biosynthetic process" or "molecular function" term group. This included 'NAD biosynthetic process' (GO:0009435), 'pyridine nucleotide biosynthetic process' (GO:0019363), 'pyridine nucleotide metabolic process' (GO:0019362), 'S-adenosylmethionine-dependent methyltransferase activity'

(GO:0008757), ‘antioxidant activity’ (GO:0016209), and ‘oxidoreductase activity’ (GO:0016491) (Table 3.0).

DISCUSSION

The results reported here indicate the potential for lowering tobacco nicotine below previously reported levels using naturally occurring variability. In field testing, all ten of the tested K326-derived NILs exhibited significantly ($P < 0.05$) lower nicotine concentrations than the K326 control line (Figure 2.1). K326 *nic1/nic2/myc2a/Ct* was the only line that contained less than the potential WHO-recommended 0.4 mg g^{-1} of nicotine, however, containing 0.36 mg g^{-1} of nicotine in its leaves on a dry weight basis. This line was significantly more susceptible to both of the soil-borne diseases (bacterial wilt and black shank) compared to the K326 control line, as well as being more susceptible to budworms and hornworms (Figures 2.6, 3.0, 3.1). Five other genotypes exhibited reduced leaf nicotine below 1.0 mg g^{-1} with no significant reductions in yield, although the single-environment field experiment had little statistical power for this trait. K326 *nic1/nic2/Cs* was the only other line that, in addition to K326 *nic1/nic2/myc2a/Ct,* that exhibited significantly ($P < 0.05$) less resistance to both bacterial wilt and black shank compared to K326 (Figures 3.0, 3.1). These findings are consistent with Sui et al. (2023) who suggested *nic1*/*nic2* mutant tobacco plants to have a compromised defense response causing potential increased vulnerability to various stressors. Notably, K326 222 appears to likely have authentic resistance to both soil-borne diseases as evidenced by the significantly lower AUDPC values compared to the K326 control (Tables 3.0, 3.1).

Burner et al. (2022) previously reported that a naturally occurring 5 bp deletion in the *Myc2a* gene in tobacco accession TI 313 has the effect of significantly reducing nicotine levels in tobacco plants. The results from this study provide further evidence for that claim, as a significant reduction in nicotine content was observed in K326 *myc2a* compared to the control (Figure 3.2). Interestingly, this genotype exhibited the highest yield out of all lines, yielding over 200 kg ha^{-1} more than K326 (Figure 2.6). This finding is interesting because Yang et al. (2023) had previously reported increased growth and fitness of *Myc2a/b* co-silenced plants in *N. attenuata*. Broz et al. (2010) also reported that, in other species, co-silenced *Myc2a/b* plants can adjust their metabolic productivity based on the growth/defense status of their neighbors.

Resistance to budworms, hornworms, and aphids varied among the tested lines. However, K326 tended to have less insect presence than the low-nicotine accumulating lines and significantly less budworm and hornworm presence compared to the lowest nicotine accumulating line K326 *nic1/nic2/myc2a/Ct*. This is not surprising considering that, in tobacco plants, nicotine is described as typically being an effective defense mechanism against herbivores (Dawson, 1942). Also, in tobacco plants containing the recessive *nic1* and *nic2* alleles, plants have been suggested to be more susceptible to biotic stresses such as herbivores (Sui et. al, 2024). In our study, not all lines containing the mutant *nic1/nic2* alleles exhibited more susceptibility to either insects or soil-borne diseases, however.

The greenhouse-based RNA sequencing experiment further supports the previous finding that the 5-bp deletion in the *Myc2a* gene significantly reduces nicotine accumulation in tobacco leaves. These levels were further reduced when the mutant *myc2a* allele was combined with recessive *nic1/nic2* alleles. The MDS plot indicated distinct root gene expression patterns, with variability at the *Nic1/Nic2* loci causing more variability in the root transcriptome than genetic variability at the *Myc2a* locus (Figure 3.3). Surprisingly, despite similar reductions in nicotine content, the overlap in DEGs between K326 *nic1/nic2* and K326 *myc2a* was less than expected considering that *Myc2a* was previously reported to regulate *Nic2* (Sui et. al, 2021). The results from our study did not reveal *Nic2* as being repressed in the K326 *myc2a* versus K326 comparison. Interestingly, more down-regulated DEGs were identified in K326 *nic1/nic2* as compared to K326 *nic1/nic2/myc2a* although these genotypes reduced nicotine content to similar levels (Table 2.5).

Approximately 104 DEGs were identified in the K326 *myc2a* versus K326 comparison. Eight of the DEGs repressed in K326 *myc2a* were found to overlap with repressed DEGs identified for K326 *nic1/nic2*, and only eleven of the 104 DEGs overlapped with the DEGs repressed in the K326 *nic1/nic2/myc2a* genotype (Figure 3.4). Approximately 78 DEGs were found to be exclusive to the K326 *myc2a* genotype. Down-regulated genes identified between K326 *myc2a* and K326 were enriched in 19 GO terms that are mainly involved in biological processes and cellular components. The 'cellular components' ontology describes subcellular structures and macromolecular complexes that can be used to annotate cellular locations of gene products (Roncaglia et. al, 2013). All of these 19 GO terms were also enriched in the K326 *nic1/nic2* vs. K326 comparison. Among the enriched GO terms, 'organic substance biosynthetic process' (GO:1901576), 'organonitrogen compound metabolic process' (GO:1901564), and 'organonitrogen compound biosynthetic process' (GO:1901566), are a few terms that indicate that the 5 bp deletion in *Myc2a* as well as the recessive *nic1/nic2* alleles inhibit the formation of organic compounds such as nicotine, nornicotine, anatabine, and anabasine. Although the list is large, it could be beneficial to further investigate the genes induced by *Myc2a* and *Nic1/Nic2* because it may lead to the identification of a potentially currently unknown gene(s) involved in the nicotine biosynthesis pathway. One of the 78 DEGs repressed in K326 *myc2a* may be involved in one of the steps of nicotine biosynthesis.

K326 *nic1/nic2/myc2a* exhibited nicotine concentrations lower than those for K326 *nic1/nic2* and K326 *myc2a* when grown in the greenhouse environment, but exhibited higher concentrations than K326 *nic1/nic2* in the field environment. Regardless, it was interesting that nearly 500 more DEGs were identified in the K326 *nic1/nic2* versus K326 comparison as compared to that for the K326 *nic1/nic2/myc2a* versus K326 comparison. Among the 38 enriched GO terms identified for down-regulated genes identified for the K326 *nic1/nic2/myc2a* versus K326 comparison, 'NAD biosynthetic process' (GO:0009435), 'pyridine nucleotide biosynthetic process' (GO:0019363), 'pyridine nucleotide metabolic process' (GO:0019362), and 'S-adenosylmethionine-dependent methyltransferase activity' (GO:0008757) suggest that some of the genes may be involved in the formation of the pyridine ring, an early step in nicotine biosynthesis, which if inhibited would lower nicotine production.

The research described above provides insight into the advantages that certain genetic modifications can have on lowering nicotine content in tobacco plants and also reports on some possible corresponding modifications of soil-borne disease resistance. The RNA sequencing analysis results further support that Myc TFs such as *Myc2a* are significantly involved in the regulation of nicotine biosynthesis. Future research should focus on examining the list of DEGs for potential identification of candidates for involvement in nicotine biosynthesis. This may be difficult, however, due to many of the identified genes not being previously annotated and their functions remaining unknown.

References

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Available Online. *Retrieved May, 17*, 2018.

Aronesty, E. (2011). Fastq-mcf sequence quality filter, clipping and processor. *Com/P/Ea-Utils/Wiki/Fastqmcf.*

Burner, N., McCauley, A., Pramod, S., Frederick, J., Steede, T., Kernodle, S. P., & Lewis, R. S. (2022). Analyses of diverse low alkaloid tobacco germplasm identify naturally occurring nucleotide variability contributing to reduced leaf nicotine accumulation. *Molecular Breeding, 42*(1), 4.

Dewey, R. E., & Xie, J. (2013). Molecular genetics of alkaloid biosynthesis in *Nicotiana tabacum*. *Phytochemistry, 94*, 10-27.

Du, Z., Zhou, X., Ling, Y., Zhang, Z., & Su, Z. (2010). agriGO: A GO analysis toolkit for the agricultural community. *Nucleic Acids Research, 38*(suppl_2), W64-W70.

Edwards, K. D., Fernandez-Pozo, N., Drake-Stowe, K., Humphry, M., Evans, A. D., Bombarely, A., Allen, F., Hurst, R., White, B., Kernodle, S. P., Bromley, J.R., Sanchez-Tamburrion, J. P., Lewis, R. S., & Mueller, L. A. (2017). A reference genome for *Nicotiana tabacum* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. *BMC Genomics, 18*, 1-14.

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. *Journal of Biological Chemistry, 282*(1), 249-256.

Guo, Y., Hiatt, E., Bonnet, C., Kudithipudi, C., Lewis, R. S., Shi, H., Patra, B., Zhao, X., de Borne, F. D., & Gilles, T. (2021). Molecular regulation and genetic manipulation of alkaloid accumulation in tobacco plants. *Studies in Natural Products Chemistry, 70*, 119-149.

Kajikawa, M., Hirai, N., & Hashimoto, T. (2009). A PIP-family protein is required for biosynthesis of tobacco alkaloids. *Plant Molecular Biology, 69*, 287-298.

Kajikawa, M., Shoji, T., Kato, A., & Hashimoto, T. (2011). Vacuole-localized berberine bridge enzyme-like proteins are required for a late step of nicotine biosynthesis in tobacco. *Plant Physiology*, *155*(4), 2010-2022.

Katoh, A., Shoji, T., & Hashimoto, T. (2007). Molecular cloning of *N*-methylputrescine oxidase from tobacco. *Plant and Cell Physiology, 48*(3), 550-554.

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, *12*(4), 357-360.

Lewis, R. S. (2019). Potential mandated lowering of nicotine levels in cigarettes: A plant perspective. *Nicotine and Tobacco Research*, *21*(7), 991-995.

Lewis, R. S., Drake-Stowe, K. E., Heim, C., Steede, T., Smith, W., & Dewey, R. E. (2020). Genetic and agronomic analysis of tobacco genotypes exhibiting reduced nicotine accumulation due to induced mutations in *berberine bridge like* (BBL) genes. *Frontiers in Plant Science*, *11*, 368.

Lewis, R. S., Lopez, H. O., Bowen, S. W., Andres, K. R., Steede, W. T., & Dewey, R. E. (2015). Transgenic and mutation-based suppression of a *berberine bridge enzyme-like* (BBL) gene family reduces alkaloid content in field-grown tobacco. *PloS One*, *10*(2), e0117273.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics, 25*(16), 2078-2079.

Love, M., Anders, S., & Huber, W. (2014). Differential analysis of count data–the DESeq2 package. *Genome Biol*, *15*(550), 10-1186.

Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics, 26, 841-842.

Roncaglia, P., Martone, M. E., Hill, D. P., Berardini, T. Z., Foulger, R. E., Imam, F. T., Drabkin, H., Mungall, C. J., & Lomax, J. (2013). The gene ontology (GO) cellular component ontology: Integration with SAO (subcellular anatomy ontology) and other recent developments. *Journal of Biomedical Semantics*, *4*, 1-11.

Shoji, T., Inai, K., Yazaki, Y., Sato, Y., Takase, H., Shitan, N., Yazaki, K., Goto, Y., Toyooka, K., & Matsuoka, K. (2009). Multidrug and toxic compound extrusion-type transporters implicated in vacuolar sequestration of nicotine in tobacco roots. *Plant Physiology*, *149*(2), 708-718.

Shoji, T., Kajikawa, M., & Hashimoto, T. (2010). Clustered transcription factor genes regulate nicotine biosynthesis in tobacco. *The Plant Cell*, *22*(10), 3390-3409.

Sui, X., He, X., Song, Z., Gao, Y., Zhao, L., Jiao, F., Kong, G., Li, Y., Han, S., & Wang, B. (2021). The gene *NtMYC2a* acts as a 'master switch'in the regulation of JA-induced nicotine accumulation in tobacco. *Plant Biology, 23*(2), 317-326.

Sui, X., Song, Z., Wang, R., Zhang, H., Tong, Z., Yuan, C., Li, Y., Huang, C., Zhao, L., & Di, Y. (2023). Comparative transcriptome analysis reveals nicotine metabolism is a critical component for enhancing stress response intensity of innate immunity system in tobacco. *Frontiers in Plant Science*, *15*, 1338169.

Wu, W., Tang, X., Yang, C., Liu, H., & Guo, N. (2013). Investigation of ecological factors controlling quality of flue-cured tobacco (*Nicotiana tabacum* L.) using classification methods. *Ecological Infrmatics*, *16*, 53-61.

Tables and Figures:

Table 2.0. Plant Materials

Entry	Genotype	Seed Source
1	K326	Gold Leaf Seeds, 2019
2	K326 HIC	GH21-22-3, X
3	K326 *nic1/nic2*	GH20-254-43, X
4	K326 222	GH20-174-1, X
5	K326 222 HIC	N20-697-7, X
6	K326 222 *nic1/nic2*	N20-362-3, X
7	K326 *myc2a/myc2a*	GH22-160-26, X
8	K326 *nic1/nic2 myc2a/myc2a*	GH22-161-56, X
9	K326 *nic1/nic2 myc2a/myc2a Ct/Ct*	GH22-182-2, X
10	K326 *nic1/nic2 Ct/Ct*	GH21-191-13
11	K326 *nic1/nic2 myc2a/myc2a Cs/Cs*	GH21-190-17

Table 2.1. Forward and reverse primers used to amplify *Myc2a*

Gene	Forward Sequence (5'-3')	Reverse Sequence (5'-3')	Tm^{0}C
Myc2a	GTTTTGGCCCGGAACAACTA	CTGAATAGCACATGAGCCCGA	63

Table 2.2. Forward and reverse primers used for KASP marker assay to detect the presence/absence of *Nic1*

Gene	Forward Sequence (5'-3')	Reverse Sequence (5'-3')	Tm^{0}C
Nic1-1_s9	TCAACATTTTGGGTGAACGA	CGCGATTGCAGACATACCTA	61
Nic1-5	TGGGTGGATGTGGCAATACC	AAAACCCTCTGCAACCCACA	62
Nic1-9	CCCCCTTCGTTGCCTTTTTG	TAAGGAGGTCCTGGGCTACC	62
Nic1-13_s21	GTCATAGGGCCTCTCATCCA	ACCCCAACATGAGGAGACAG	61
Nic1-16	CAGGTGCAAGGATGGAAGGA	AGGCCAATTTATTTGCCGCC	66
Nic1-20	AACAGAGGCGTGGAAGACAG	ACGGACAACGACGATGGAAA	64

Table 2.3. Forward and reverse primers used for KASP marker assay to detect the presence/absence of Nic2

Gene	Forward Sequence (5'-3')	Reverse Sequence (5'-3')	Tm^{0}C
ERF115	AGAGGACGCAGCATTGGCTTAC	TTGCTTGGAAGGCATAGCTCCCG	56
ERF189	GGGCAATGGAAATGAATCTAGC	CTTCCTTCCTTTCACATAG	55.4

Table 2.4. RNAseq samples and their measured alkaloid measured content.

GH#	RNAseq Lib I.D.	Genotype	Nicotine	Nornicotine	Anabasine	Anatabine
			-------------------------------mg g-1---------------------			
158-4	CK4	K326	10.178	0.21	0.03	0.223
158-8	CK1	K326	6.951	0.161	0.028	0.162
158-9	CK2	K326	11.838	0.245	0.046	0.262
158-10	CK3	K326	6.548	0.132	0.026	0.202
158-5	-	K326	9.634	0.218	0.033	0.298
159-7	N2	K326 *nic1/nic2*	0.162	-	-	-
159-9	N3	K326 *nic1/nic2*	0.204	0.205	-	0.347
159-10	N1	K326 *nic1/nic2*	0.138	-	-	-
159-8	-	K326 *nic1/nic2*	0.500	-	-	-
160-1	M2	K326 *myc2a*	0.131	2.313	-	0.14
160-2	M3	K326 *myc2a*	0.167	4.495	0.026	0.259
160-5	M4	K326 *myc2a*	0.111	2.13	0.098	0.15
160-6	M1	K326 *myc2a*	0.487	1.077	-	0.071
159-5	M5	K326 *myc2a*	0.084	-	-	-
160-12	-	K326 *myc2a*	1.607	1.077	-	0.071
161-2	NM3	K326 *nic1/nic2/myc2a*	0.081	-	-	-
161-4	NM4	K326 *nic1/nic2/myc2a*	0.114	-	-	-
161-9	NM1	K326 *nic1/nic2/myc2a*	0.093	-	-	-
161-10	NM2	K326 *nic1/nic2/myc2a*	0.072	0.11	-	-
161-7	-	K326 *nic1/nic2/myc2a*	0.075	-	-	-

Table 2.5. Differentially Expressed Genes (DEGs) for K326 NILs

DEG Pattern	K326 *Nic1/Nic2* Vs. K326	K326 *Myc2a* Vs. K326	K326 *Nic1/Nic2/Myc2a* Vs. K326
Up-regulated	316	23	105
Down-regulated	435	81	183
Total	751	104	288

Table 2.6. Nicotine biosynthesis-related Differentially Expressed Genes (DEGs) identified between K326 and three K326 NILs differing for the presence/absence of *nic1/nic2* and *myc2a*.

Gene I.D.	DEG Foldchange (log2)			Nicotine Related Gene	TAIR annotation description
	K326 nic1/nic2 vs. K326	K326 myc2a vs. K326	K326 nic1/nic2/myc2a vs. K326		
Nitab4.5_0000633g0060.1		1.6	2.1	MATE	MATE efflux family protein
Nitab4.5_0000389g0050.1	2.1		2.4	MATE	MATE efflux family protein
Nitab4.5_0000241g0060.1			1.1	MATE	MATE efflux family protein
Nitab4.5_0002358g0050.1				MPO	Copper amine oxidase family protein
Nitab4.5_0010422g0020.1	1.3			BBL	FAD-binding Berberine family protein
Nitab4.5_0003029g0030.1				BBL	FAD-binding Berberine family protein
Nitab4.5_0002631g0030.1	1.1		1.4	MPO	Copper amine oxidase family protein
Nitab4.5_0006559g0020.1				MATE	MATE efflux family protein
Nitab4.5_0005600g0030.1				MATE	MATE efflux family protein
Nitab4.5_0003090g0030.1	-2.7		-2.9	Nic1	ATERF13, EREBP, ERF13
Nitab4.5_0001029g0080.1			-1.5	BBL	FAD-binding Berberine family protein
Nitab4.5_0000025g0290.1			-1.0	MPO	Copper amine oxidase family protein
Nitab4.5_0015055g0010.1	-8.4		-8.4	Nic2	ATERF13, EREBP, ERF13
Nitab4.5_0000983g0060.1	-1.9		-2.0	MPO1	Copper amine oxidase family protein
Nitab4.5_0000742g0010.1	-2.7		-3.0	QPT2A	QPT: quinolinate phoshoribosyltransferase
Nitab4.5_0000316g0090.1	-2.5		-2.6	BBLc_like	unknown
Nitab4.5_0006307g0010.1	-2.8		-2.8	BBLa	FAD-binding Berberine family protein
Nitab4.5_0000031g0420.1	-3.1		-2.9	MATE2	MATE efflux family protein
Nitab4.5_0000884g0030.1	-3.8		-3.9	MATE1	MATE efflux family protein
Nitab4.5_0002522g0070.1	-3.6	-1.2	-3.6	QS	QS, OLD5, SUFE3: quinolinate synthase
Nitab4.5_0000884g0010.1	-3.7		-4.2	A622	NAD(P)-binding Rossmann-fold superfamily protein
Nitab4.5_0003029g0020.1	-4.6		-4.9	BBLc	unknown
Nitab4.5_0000013g0390.1	-5.4		-4.9	PMT	SPDS1: spermidine synthase 1
Nitab4.5_0003280g0010.1	-3.8		-3.4	ODC 1	unknown
Nitab4.5_0008037g0020.1	-4.5	-1.4	-4.5	QS	QS, OLD5, SUFE3: quinolinate synthase
Nitab4.5_0011232g0010.1	-4.5		-6.4	PMT4	SPDS1: spermidine synthase 1
Nitab4.5_0000031g0370.1	-5.1		-5.0	A622	NAD(P)-binding Rossmann-fold superfamily protein
Nitab4.5_0003284g0040.1	-5.2		-5.3	PMT2	SPDS1: spermidine synthase 1
Nitab4.5_0003284g0010.1	-6.0		-6.3	PMT1/3	SPDS1: spermidine synthase 1
Nitab4.5_0000013g0380.1	-4.8		-5.9	PMT	SPDS1: spermidine synthase 1

*Blue indicates upregulated and red indicates downregulated.

Table 2.7. DEGs identified in K326 *myc2a* and genes involved in nicotine biosynthesis

Genes	TAIR annotation description
Nitab4.5_0000073g0400.1	SUS4, ATSUS4: sucrose synthase 4
Nitab4.5_0001769g0010.1	dehydratase family
Nitab4.5_0005383g0020.1	-
Nitab4.5_0000367g0090.1	-
Nitab4.5_0003248g0010.1	MAF4: K-box region and MADS-box transcription factor family protein
Nitab4.5_0025584g0010.1	emb1187: Protein kinase superfamily protein
Nitab4.5_0003099g0070.1	IPP1: isopentenyl diphosphate isomerase 1
Nitab4.5_0003152g0100.1	-
Nitab4.5_0005471g0020.1	U2 snRNP auxilliary factor, large subunit, splicing factor
Nitab4.5_0000826g0130.1	BG3: beta-1,3-glucanase 3
Nitab4.5_0004165g0050.1	unknown protein similar to AT3G17120.2
Nitab4.5_0010958g0020.1	Mitochondrial substrate carrier family protein
Nitab4.5_0000578g0090.1	MAT3: methionine adenosyltransferase 3
Nitab4.5_0009685g0030.1	unknown protein similar to AT5G16250.1
Nitab4.5_0014883g0010.1	CAC1, CAC1A, BCCP: chloroplastic acetylcoenzyme A carboxylase 1
Nitab4.5_0006629g0020.1	PHR1, AtPHR1: phosphate starvation response 1
Nitab4.5_0003073g0020.1	PHO1;H1: EXS (ERD1/XPR1/SYG1) family protein
Nitab4.5_0004173g0030.1	Lipase/lipooxygenase, PLAT/LH2 family protein
Nitab4.5_0004592g0010.1	alpha/beta-Hydrolases superfamily protein
Nitab4.5_0004338g0100.1	EIF2, AteIF3f, eIF3F: eukaryotic translation initiation factor 2
Nitab4.5_0013215g0010.1	Ribosomal L38e protein family
Nitab4.5_0002692g0060.1	H2B, HTB9: Histone superfamily protein
Nitab4.5_0007578g0010.1	ATOCT3, 3-Oct: organic cation/carnitine transporter 3
Nitab4.5_0010222g0050.1	ACO1, ATACO1: ACC oxidase 1
Nitab4.5_0003402g0040.1	IMK2: inflorescence meristem receptor-like kinase 2
Nitab4.5_0001556g0010.1	GR, EMB2360, ATGR2: glutathione reductase
Nitab4.5_0005052g0020.1	APUM7, PUM7: pumilio 7
Nitab4.5_0001585g0010.1	LTP2, LP2, cdf3: lipid transfer protein 2
Nitab4.5_0008969g0020.1	AGD2: Pyridoxal phosphate (PLP)-dependent transferases superfamily protein
Nitab4.5_0003413g0030.1	Basic-leucine zipper (bZIP) transcription factor family protein
Nitab4.5_0000367g0060.1	LTPG1: glycosylphosphatidylinositol-anchored lipid protein transfer 1
Nitab4.5_0011903g0020.1	bacterial transferase hexapeptide repeat-containing protein
Nitab4.5_0001816g0230.1	-
Nitab4.5_0004173g0010.1	Lipase/lipooxygenase, PLAT/LH2 family protein
Nitab4.5_0012382g0010.1	PAL1, ATPAL1: PHE ammonia lyase 1
Nitab4.5_0000033g0010.1	SWEET11, AtSWEET11: Nodulin MtN3 family protein
Nitab4.5_0009412g0020.1	RPL23AB: ribosomal protein L23AB
Nitab4.5_0002289g0120.1	unknown protein similar to AT5G47400.1
Nitab4.5_0008348g0010.1	B-box type zinc finger protein with CCT domain
Nitab4.5_0009204g0010.1	Eukaryotic translation initiation factor 2 subunit 1
Nitab4.5_0010629g0030.1	zinc knuckle (CCHC-type) family protein
Nitab4.5_0005471g0010.1	ATPRMT11, PRMT11, ATPRMT1B: arginine methyltransferase 11
Nitab4.5_0008566g0010.1	-
Nitab4.5_0003391g0090.1	unknown protein similar to AT2G40316.1
Nitab4.5_0006647g0010.1	NAD(P)-binding Rossmann-fold superfamily protein
Nitab4.5_0005264g0010.1	HEME2: Uroporphyrinogen decarboxylase
Nitab4.5_0002003g0070.1	LBD38: LOB domain-containing protein 38
Nitab4.5_0007722g0050.1	-
Nitab4.5_0008533g0020.1	Auxin efflux carrier family protein
Nitab4.5_0006622g0080.1	-
Nitab4.5_0001769g0100.1	HSP60, HSP60-3B: heat shock protein 60
Nitab4.5_0007021g0010.1	-
Nitab4.5_0006126g0020.1	NAP1;2: nucleosome assembly protein 1;2
Nitab4.5_0010110g0060.1	NTMC2TYPE4, NTMC2T4: Calcium-dependent lipid-binding family protein
Nitab4.5_0004237g0020.1	RPT6A, ATSUG1: regulatory particle triple-A ATPase 6A
Nitab4.5_0002647g0070.1	-
Nitab4.5_0006345g0010.1	-
Nitab4.5_0008566g0040.1	-
Nitab4.5_0019528g0010.1	HOG1, EMB1395, SAHH1, MEE58: S-adenosyl-L-homocysteine hydrolase
Nitab4.5_0007066g0010.1	FtsJ-like methyltransferase family protein

Table 2.7. (Continued**).**

Nitab4.5_0004024g0020.1	-
Nitab4.5_0010958g0010.1	ATSERAT1;1, SAT5, SAT-52, SERAT1;1: serine acetyltransferase 1;1
Nitab4.5_0010111g0010.1	SRZ-22, SRZ22, RSZP22, RSZ22, At-RSZ22: serine/arginine-rich 22
Nitab4.5_0002728g0070.1	Aluminium activated malate transporter family protein
Nitab4.5_0010702g0030.1	GRF2, 14-3-3OMEGA, GF14 OMEGA: general regulatory factor 2
Nitab4.5_0010633g0040.1	Ribosomal S17 family protein
Nitab4.5_0002114g0140.1	Guanylate-binding family protein
Nitab4.5_0015027g0010.1	Endomembrane protein 70 protein family
Nitab4.5_0006242g0030.1	ATCDPK2, CPK11, ATCPK11: calcium-dependent protein kinase 2
Nitab4.5_0011108g0010.1	Glycosyl hydrolase superfamily protein
Nitab4.5_0005042g0010.1	Aluminium activated malate transporter family protein
Nitab4.5_0005105g0020.1	ATC4H, C4H, CYP73A5, REF3: cinnamate-4-hydroxylase
Nitab4.5_0022238g0010.1	SNARE-like superfamily protein
Nitab4.5_0011920g0010.1	ARPN: plantacyanin
Nitab4.5_0003965g0050.1	Translation elongation factor EF1B/ribosomal protein S6 family protein
Nitab4.5_0001277g0060.1	Protein of unknown function (DUF620)
Nitab4.5_0005211g0010.1	-
Nitab4.5_0008565g0020.1	PSF2: PSF2
Nitab4.5_0011990g0010.1	ATDGK2, DGK2: diacylglycerol kinase 2
Nitab4.5_0002662g0020.1	alpha/beta-Hydrolases superfamily protein
Nitab4.5_0000940g0020.1	Leucine-rich repeat protein kinase family protein
Nitab4.5_0001454g0100.1	Mitochondrial glycoprotein family protein
Nitab4.5_0012816g0020.1	MAF4, FCL4: K-box region and MADS-box transcription factor family protein
Nitab4.5_0003152g0070.1	-
Nitab4.5_0008566g0050.1	-
Nitab4.5_0001769g0090.1	I-branching beta-1,6-N-acetylglucosaminyltransferase family protein
Nitab4.5_0004474g0020.1	Mitochondrial glycoprotein family protein
Nitab4.5_0004575g0050.1	SAUR-like auxin-responsive protein family
Nitab4.5_0006365g0080.1	Ribosomal L29e protein family
Nitab4.5_0006342g0050.1	Histone superfamily protein
Nitab4.5_0003476g0040.1	Disease resistance-responsive (dirigent-like protein) family protein
Nitab4.5_0003755g0010.1	transducin family protein / WD-40 repeat family protein
Nitab4.5_0005095g0020.1	HAP6: ribophorin II (RPN2) family protein
Nitab4.5_0004338g0030.1	GATA type zinc finger transcription factor family protein
Nitab4.5_0005471g0040.1	PRL: Minichromosome maintenance (MCM2/3/5) family protein
Nitab4.5_0000143g0590.1	CYP83B1, SUR2, RNT1: cytochrome P450, family 83, subfamily B, polypeptide 1
Nitab4.5_0001506g0100.1	COR47, RD17, AtCOR47: cold-regulated 47
Nitab4.5_0012757g0010.1	Lipase/lipooxygenase, PLAT/LH2 family protein
Nitab4.5_0006365g0010.1	NRPB9A, NRPD9A, NRPE9A: RNA polymerases M/15 Kd subunit
Nitab4.5_0011080g0020.1	Ribosomal L29 family protein
Nitab4.5_0011938g0030.1	AIL6: AINTEGUMENTA-like 6

Nicotine Biosynthesis Related DEGs	**TAIR annotation description**
Nitab4.5_0008037g0020.1	QS, OLD5, SUFE3: quinolinate synthase
Nitab4.5_0002522g0070.1	QS, OLD5, SUFE3: quinolinate synthase
Nitab4.5_0000633g0060.1	MATE efflux family protein

Table 2.8. GO terms enriched for DEGs identified between K326 *nic1/nic1 nic2/nic2* vs. K326

Down-Regulated DEGs in K326 *nic1/nic1 nic2/nic2* vs. K326				
GO term	**Term**	**Term Type**	***p*-value**	**FDR**
GO:0006334	nucleosome assembly	P	9.80E-22	3.30E-19
GO:0065004	protein-DNA complex assembly	P	1.60E-21	3.30E-19
GO:0034728	nucleosome organization	P	9.80E-22	3.30E-19
GO:0031497	chromatin assembly	P	9.80E-22	3.30E-19
GO:0071824	protein-DNA complex subunit organization	P	1.60E-21	3.30E-19
GO:0006333	chromatin assembly or disassembly	P	2.40E-21	4.00E-19
GO:0006323	DNA packaging	P	3.80E-21	5.60E-19
GO:0071103	DNA conformation change	P	8.70E-20	1.10E-17
GO:0006325	chromatin organization	P	4.40E-18	5.00E-16
GO:0070271	protein complex biogenesis	P	1.30E-17	1.20E-15
GO:0006461	protein complex assembly	P	1.30E-17	1.20E-15
GO:0034622	cellular macromolecular complex assembly	P	1.90E-17	1.60E-15
GO:0071822	protein complex subunit organization	P	4.80E-17	3.80E-15
GO:0065003	macromolecular complex assembly	P	7.20E-17	5.20E-15
GO:0051276	chromosome organization	P	1.10E-15	7.30E-14
GO:0006996	organelle organization	P	4.70E-15	3.00E-13
GO:0022607	cellular component assembly	P	5.60E-15	3.30E-13
GO:0043933	macromolecular complex subunit organization	P	1.10E-14	6.10E-13
GO:0044085	cellular component biogenesis	P	9.30E-13	5.00E-11
GO:0016043	cellular component organization	P	8.30E-12	4.20E-10
GO:0071840	cellular component organization or biogenesis	P	1.10E-10	5.50E-09
GO:1901564	organonitrogen compound metabolic process	P	1.60E-10	7.50E-09
GO:1901566	organonitrogen compound biosynthetic process	P	3.80E-10	1.70E-08
GO:0044271	cellular nitrogen compound biosynthetic process	P	1.20E-07	5.10E-06
GO:0009058	biosynthetic process	P	2.30E-07	9.50E-06
GO:1901576	organic substance biosynthetic process	P	2.60E-07	1.00E-05
GO:0009165	nucleotide biosynthetic process	P	2.80E-07	1.00E-05
GO:1901293	nucleoside phosphate biosynthetic process	P	3.00E-07	1.10E-05
GO:0009435	NAD biosynthetic process	P	2.90E-07	1.10E-05
GO:0044249	cellular biosynthetic process	P	4.60E-07	1.60E-05
GO:0006412	translation	P	6.20E-07	2.00E-05
GO:0043043	peptide biosynthetic process	P	7.40E-07	2.40E-05
GO:0043604	amide biosynthetic process	P	7.60E-07	2.40E-05
GO:0006807	nitrogen compound metabolic process	P	1.00E-06	3.00E-05
GO:0006518	peptide metabolic process	P	1.00E-06	3.00E-05
GO:0055086	nucleobase-containing small molecule metabolic process	P	1.20E-06	3.50E-05
GO:0043603	cellular amide metabolic process	P	1.30E-06	3.60E-05
GO:0019674	NAD metabolic process	P	1.80E-06	4.50E-05
GO:0019359	nicotinamide nucleotide biosynthetic process	P	1.80E-06	4.50E-05
GO:0019363	pyridine nucleotide biosynthetic process	P	1.80E-06	4.50E-05
GO:0034641	cellular nitrogen compound metabolic process	P	2.10E-06	5.20E-05

Table 2.8. (Continued).

GO:0009987	cellular process	P	2.70E-06	6.50E-05
GO:0072525	pyridine-containing compound biosynthetic process	P	4.10E-06	9.80E-05
GO:0009117	nucleotide metabolic process	P	5.80E-06	0.00013
GO:0006753	nucleoside phosphate metabolic process	P	6.90E-06	0.00016
GO:0090407	organophosphate biosynthetic process	P	8.30E-06	0.00018
GO:0009108	coenzyme biosynthetic process	P	2.00E-05	0.00043
GO:0006732	coenzyme metabolic process	P	0.0002	0.0042
GO:0051188	cofactor biosynthetic process	P	0.0003	0.0063
GO:0034645	cellular macromolecule biosynthetic process	P	0.00045	0.0091
GO:0019637	organophosphate metabolic process	P	0.00047	0.0092
GO:0009059	macromolecule biosynthetic process	P	0.00047	0.0092
GO:0046496	nicotinamide nucleotide metabolic process	P	0.00059	0.011
GO:0019362	pyridine nucleotide metabolic process	P	0.00059	0.011
GO:0044711	single-organism biosynthetic process	P	0.00069	0.013
GO:0072524	pyridine-containing compound metabolic process	P	0.00069	0.013
GO:0006733	oxidoreduction coenzyme metabolic process	P	0.00077	0.014
GO:0009156	ribonucleoside monophosphate biosynthetic process	P	0.001	0.018
GO:0051186	cofactor metabolic process	P	0.0012	0.02
GO:0009124	nucleoside monophosphate biosynthetic process	P	0.0012	0.02
GO:0010467	gene expression	P	0.0013	0.022
GO:0055085	transmembrane transport	P	0.0015	0.025
GO:0009116	nucleoside metabolic process	P	0.0021	0.034
GO:1901657	glycosyl compound metabolic process	P	0.0021	0.034
GO:0009161	ribonucleoside monophosphate metabolic process	P	0.0031	0.046
GO:0009206	purine ribonucleoside triphosphate biosynthetic process	P	0.0031	0.046
GO:0009201	ribonucleoside triphosphate biosynthetic process	P	0.0031	0.046
GO:0009142	nucleoside triphosphate biosynthetic process	P	0.0031	0.046
GO:0009145	purine nucleoside triphosphate biosynthetic process	P	0.0031	0.046
GO:0009123	nucleoside monophosphate metabolic process	P	0.0033	0.049
GO:0046982	protein heterodimerization activity	F	7.20E-23	3.10E-20
GO:0005198	structural molecule activity	F	1.50E-09	3.20E-07
GO:0003735	structural constituent of ribosome	F	3.60E-09	3.90E-07
GO:0046983	protein dimerization activity	F	2.90E-09	3.90E-07
GO:0008757	S-adenosylmethionine-dependent methyltransferase activity	F	0.00032	0.027
GO:0003677	DNA binding	F	0.0006	0.043
GO:0043228	non-membrane-bounded organelle	C	8.70E-23	4.00E-21
GO:0000786	nucleosome	C	6.40E-23	4.00E-21
GO:0043232	intracellular non-membrane-bounded organelle	C	8.70E-23	4.00E-21
GO:0032993	protein-DNA complex	C	6.40E-23	4.00E-21
GO:0044815	DNA packaging complex	C	1.10E-22	4.10E-21
GO:0000785	chromatin	C	5.80E-22	1.80E-20
GO:0044427	chromosomal part	C	1.80E-21	4.70E-20
GO:0005694	chromosome	C	6.80E-20	1.60E-18
GO:0043229	intracellular organelle	C	1.60E-17	3.00E-16

Table 2.8. (Continued).

GO:0043226	organelle	C	1.60E-17	3.00E-16
GO:0032991	macromolecular complex	C	1.20E-15	2.00E-14
GO:0005622	intracellular	C	2.10E-15	3.10E-14
GO:0044424	intracellular part	C	8.60E-15	1.20E-13
GO:0044464	cell part	C	1.40E-14	1.70E-13
GO:0005623	cell	C	1.40E-14	1.70E-13
GO:0044446	intracellular organelle part	C	5.10E-13	5.50E-12
GO:0044422	organelle part	C	5.10E-13	5.50E-12
GO:0005840	ribosome	C	3.80E-09	3.80E-08
GO:0044444	cytoplasmic part	C	4.40E-09	4.20E-08
GO:0030529	intracellular ribonucleoprotein complex	C	1.30E-08	1.10E-07
GO:1990904	ribonucleoprotein complex	C	1.30E-08	1.10E-07
GO:0005737	cytoplasm	C	1.60E-08	1.30E-07
GO:0043234	protein complex	C	6.50E-08	5.10E-07
GO:0043227	membrane-bounded organelle	C	1.10E-07	8.30E-07
GO:0043231	intracellular membrane-bounded organelle	C	1.10E-07	8.30E-07
GO:0005739	mitochondrion	C	1.80E-06	1.30E-05
GO:0005634	nucleus	C	2.50E-06	1.70E-05
GO:0031967	organelle envelope	C	8.80E-06	5.70E-05
GO:0031975	envelope	C	1.20E-05	7.40E-05
GO:0005740	mitochondrial envelope	C	1.30E-05	7.90E-05
GO:0044429	mitochondrial part	C	2.90E-05	0.00017
GO:0031966	mitochondrial membrane	C	5.80E-05	0.00033
GO:0044455	mitochondrial membrane part	C	9.40E-05	0.00052
GO:0005743	mitochondrial inner membrane	C	0.00038	0.002
GO:0019866	organelle inner membrane	C	0.00041	0.0021
GO:0098798	mitochondrial protein complex	C	0.0006	0.003
GO:0031090	organelle membrane	C	0.0025	0.012
Up-Regulated DEGs in K326 *nic1/nic1 nic2/nic2* vs. K326				
GO term	**Term**	**Term Type**	***p*-value**	**FDR**
GO:2000112	regulation of cellular macromolecule biosynthetic process	P	9.80E-05	0.0038
GO:0031326	regulation of cellular biosynthetic process	P	9.80E-05	0.0038
GO:0019219	regulation of nucleobase compound metabolic process	P	9.70E-05	0.0038
GO:0009889	regulation of biosynthetic process	P	9.80E-05	0.0038
GO:1903506	regulation of nucleic acid-templated transcription	P	8.50E-05	0.0038
GO:2001141	regulation of RNA biosynthetic process	P	8.50E-05	0.0038
GO:0051252	regulation of RNA metabolic process	P	8.70E-05	0.0038
GO:0006355	regulation of transcription, DNA-templated	P	8.50E-05	0.0038
GO:0010556	regulation of macromolecule biosynthetic process	P	9.80E-05	0.0038
GO:0051171	regulation of nitrogen compound metabolic process	P	0.00011	0.0038
GO:0010468	regulation of gene expression	P	0.00013	0.004
GO:0031323	regulation of cellular metabolic process	P	0.00018	0.0047
GO:0080090	regulation of primary metabolic process	P	0.00017	0.0047

Table 2.8. (Continued).

GO:0060255	regulation of macromolecule metabolic process	P	0.00023	0.0056
GO:0019222	regulation of metabolic process	P	0.00025	0.0058
GO:0032774	RNA biosynthetic process	P	0.00073	0.014
GO:0097659	nucleic acid-templated transcription	P	0.00071	0.014
GO:0006351	transcription, DNA-templated	P	0.00071	0.014
GO:0001071	nucleic acid binding transcription factor activity	F	5.20E-05	0.0069
GO:0003700	transcription factor activity, DNA binding	F	5.20E-05	0.0069

Table 2.9. GO terms enriched for DEGs identified between K326 *myc2a* and K326

Down-Regulated DEGs in K326 *myc2a* *v*s. K326				
GO term	**Term**	**Term Type**	***p*-value**	**FDR**
GO:0044249	cellular biosynthetic process	P	0.00066	0.03
GO:1901576	organic substance biosynthetic process	P	0.0007	0.03
GO:1901564	organonitrogen compound metabolic process	P	0.00055	0.03
GO:1901566	organonitrogen compound biosynthetic process	P	0.00068	0.03
GO:0009058	biosynthetic process	P	0.0002	0.03
GO:0044424	intracellular part	C	1.80E-06	5.80E-05
GO:0005622	intracellular	C	2.40E-06	5.80E-05
GO:0005737	cytoplasm	C	1.40E-06	5.80E-05
GO:0044464	cell part	C	6.00E-06	8.70E-05
GO:0005623	cell	C	6.00E-06	8.70E-05
GO:0044444	cytoplasmic part	C	4.00E-05	0.00048
GO:0043229	intracellular organelle	C	9.70E-05	0.00087
GO:0043226	organelle	C	9.70E-05	0.00087
GO:0043231	intracellular membrane-bounded organelle	C	0.001	0.0075
GO:0043227	membrane-bounded organelle	C	0.001	0.0075
GO:0032991	macromolecular complex	C	0.0024	0.015
GO:0043232	intracellular non-membrane-bounded organelle	C	0.0087	0.045
GO:0005840	ribosome	C	0.0077	0.045
GO:0043228	non-membrane-bounded organelle	C	0.0087	0.045

*****Term Type: C= Cellular Component, F= Molecular Function, P= Biological Process

Table 3.0. GO terms enriched for DEGs identified between K326 *nic1*/*nic2*/*myc2a* vs. K326

GO term	Term	Term Type	*p*-value	FDR
	Down-Regulated DEGs in K326 *nic1/nic2/myc2a* vs K326 comparison			
GO:0009435	NAD biosynthetic process	P	9.00E-09	4.30E-06
GO:0019359	nicotinamide nucleotide biosynthetic process	P	5.60E-08	6.60E-06
GO:0019674	NAD metabolic process	P	5.60E-08	6.60E-06
GO:0019363	pyridine nucleotide biosynthetic process	P	5.60E-08	6.60E-06
GO:0072525	pyridine-containing compound biosynthetic process	P	1.30E-07	1.30E-05
GO:0009108	coenzyme biosynthetic process	P	1.60E-06	0.00013
GO:0009165	nucleotide biosynthetic process	P	2.90E-06	0.00018
GO:1901293	nucleoside phosphate biosynthetic process	P	3.10E-06	0.00018
GO:0051188	cofactor biosynthetic process	P	2.20E-05	0.0011
GO:0090407	organophosphate biosynthetic process	P	3.20E-05	0.0015
GO:0044710	single-organism metabolic process	P	6.40E-05	0.0028
GO:0006979	response to oxidative stress	P	8.50E-05	0.0034
GO:0055114	oxidation-reduction process	P	0.00011	0.0038
GO:0006732	coenzyme metabolic process	P	0.00023	0.0077
GO:0044711	single-organism biosynthetic process	P	0.00049	0.015
GO:0009117	nucleotide metabolic process	P	0.00051	0.015
GO:0006753	nucleoside phosphate metabolic process	P	0.00056	0.016
GO:0046496	nicotinamide nucleotide metabolic process	P	0.00071	0.018
GO:0019362	pyridine nucleotide metabolic process	P	0.00071	0.018
GO:0044699	single-organism process	P	0.00081	0.018
GO:0072524	pyridine-containing compound metabolic process	P	0.0008	0.018
GO:0051186	cofactor metabolic process	P	0.00091	0.019
GO:0006733	oxidoreduction coenzyme metabolic process	P	0.00087	0.019
GO:0055086	nucleobase-containing small molecule metabolic process	P	0.0011	0.021
GO:0003824	catalytic activity	F	1.80E-06	0.00045
GO:0016741	transferase activity, transferring one-carbon groups	F	6.70E-05	0.0044
GO:0016746	transferase activity, transferring acyl groups	F	0.00011	0.0044
GO:0008168	methyltransferase activity	F	3.90E-05	0.0044
GO:0016684	oxidoreductase activity, acting on peroxide as acceptor	F	0.00014	0.0044
GO:0004601	peroxidase activity	F	0.00012	0.0044
GO:0005507	copper ion binding	F	0.00011	0.0044
GO:0008757	S-adenosylmethionine methyltransferase activity	F	9.30E-05	0.0044
GO:0016747	transferase activity, transferring acyl groups	F	0.00018	0.0049
GO:0016209	antioxidant activity	F	0.00025	0.0064
GO:0020037	heme binding	F	0.00053	0.012
GO:0016491	oxidoreductase activity	F	0.00061	0.013
GO:0046906	tetrapyrrole binding	F	0.00069	0.013
GO:0050660	flavin adenine dinucleotide binding	F	0.0021	0.038

***Term Type: C= Cellular Component, F= Molecular Function, P= Biological Process

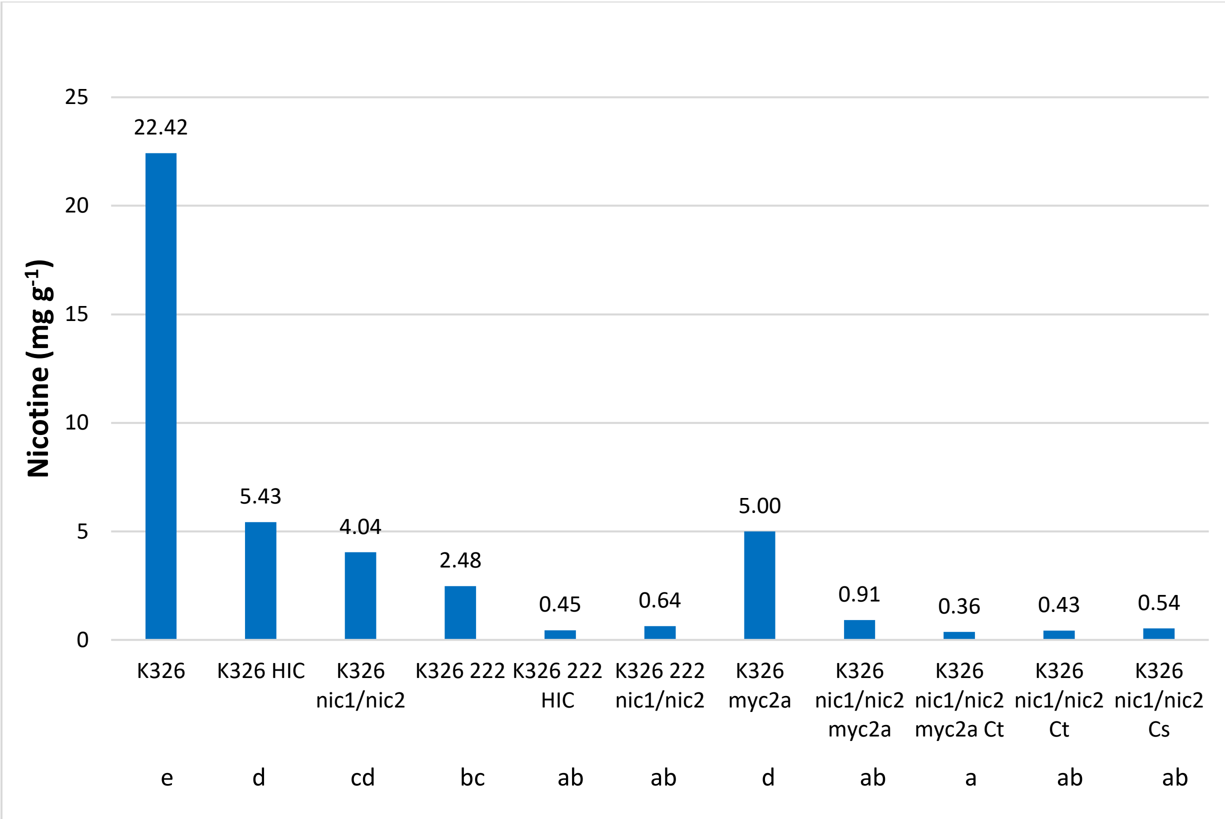


Figure 2.1 Average nicotine accumulation in field comparison of NILs of the K326 genetic background. Numbers above each bar indicate nicotine content in mg g^{-1}. Total nicotine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance.

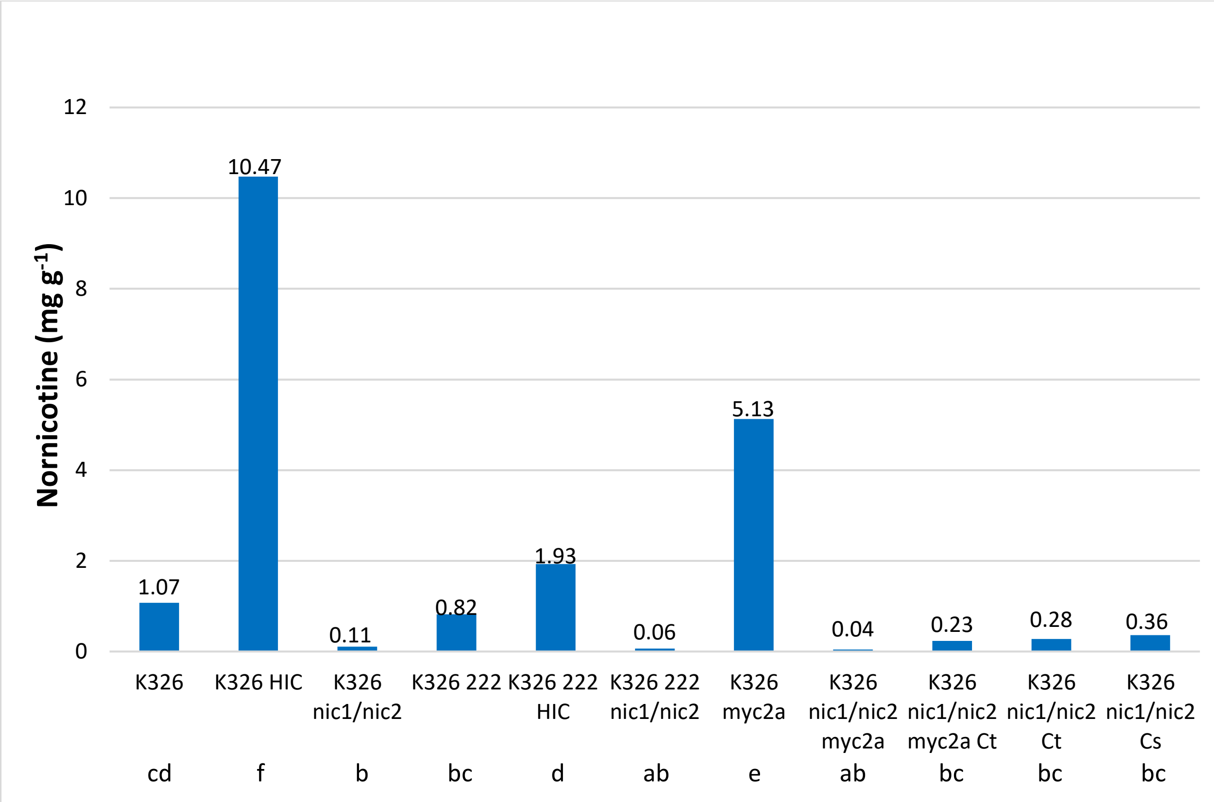


Figure 2.2 Average nornicotine accumulation in field comparison of NILs of the K326 genetic background. Numbers above each bar indicate nornicotine content in mg g^{-1}. Total nornicotine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance.

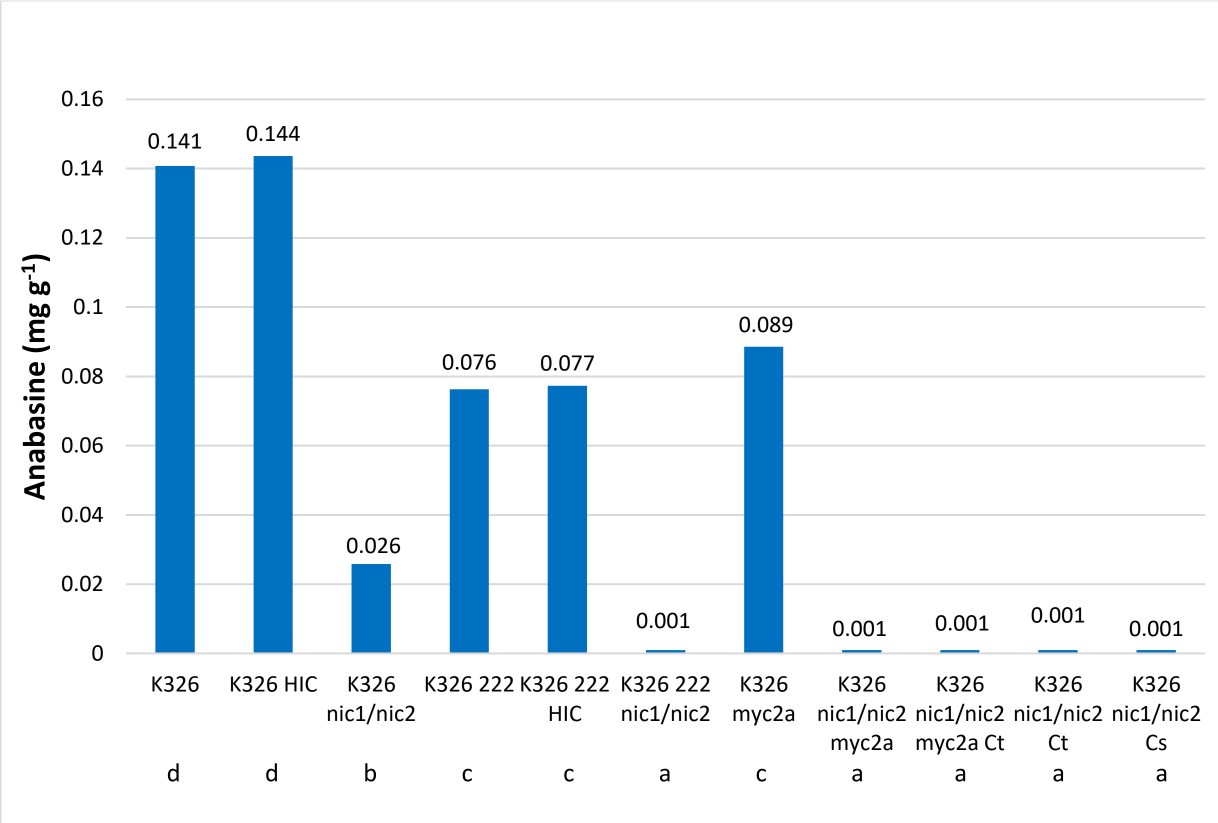


Figure 2.3 Average anabasine accumulation in field comparison of NILs of the K326 genetic background. Numbers above each bar indicate anabasine content in mg g^{-1}. Total anabasine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance.

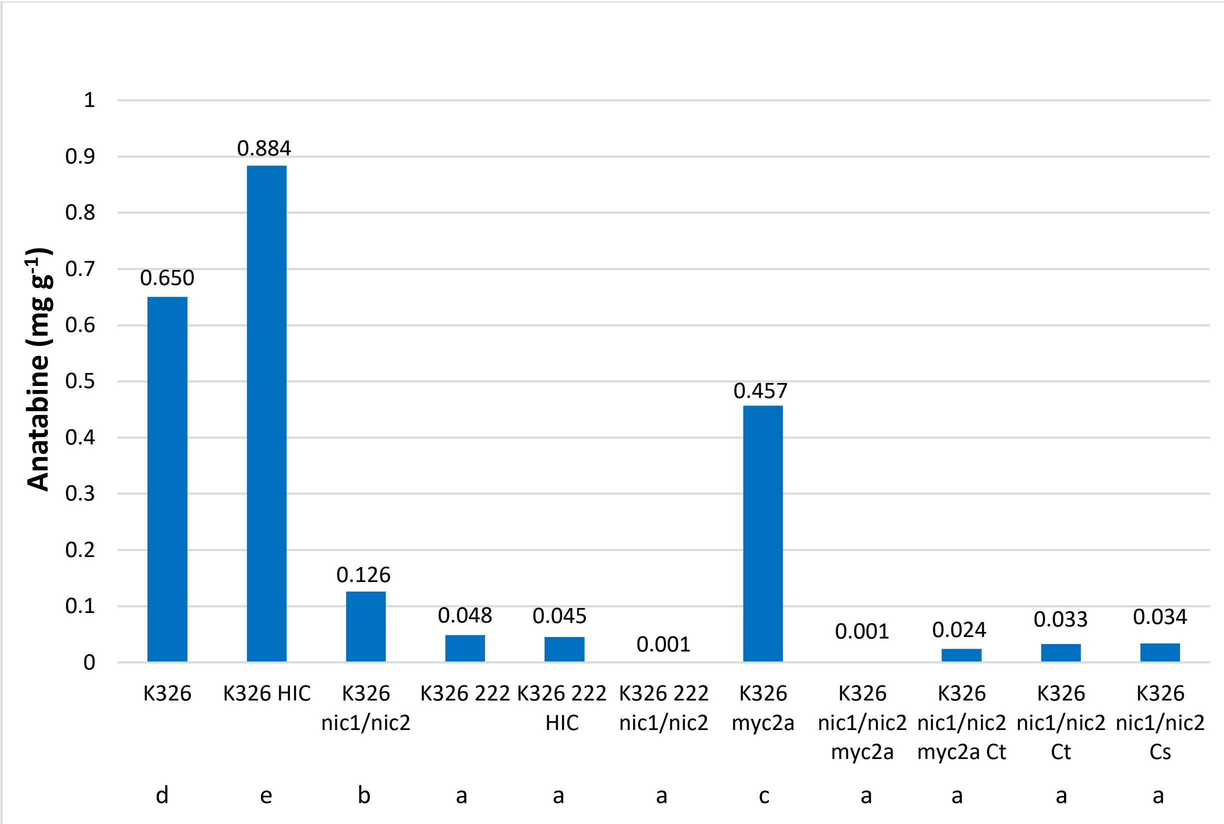


Figure 2.4 Average anatabine accumulation in field comparison of NILs of the K326 genetic background. Numbers above each bar indicate anatabine content in mg g^{-1}. Total anatabine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance.

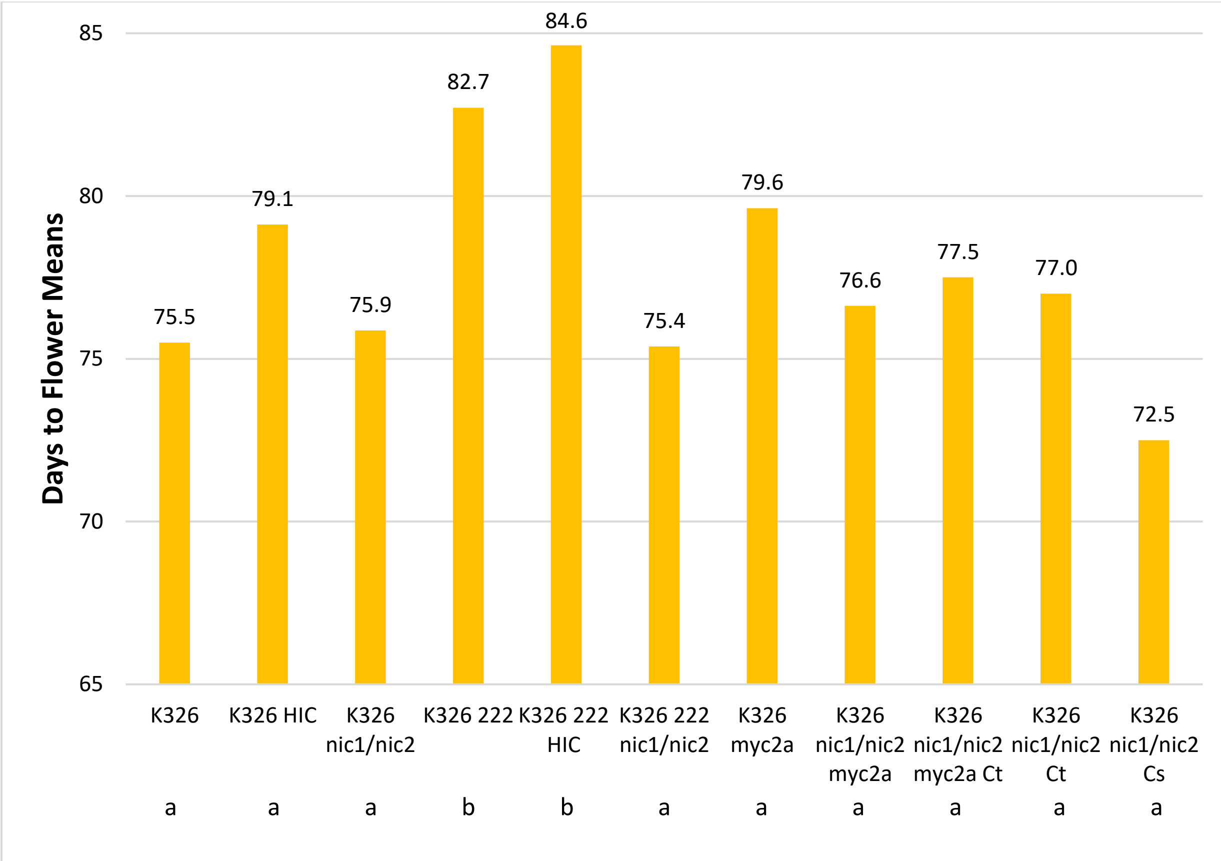


Figure 2.5 Average number of days to flower for NILs of the K326 genetic background evaluated in the Clayton, NC, field experiment. Numbers above each bar indicate the average number of days from transplant to flower. Genotypes with different letters (below genotypic descriptions) exhibited significantly different total numbers of days to flower at the $P < 0.05$ level of significance.

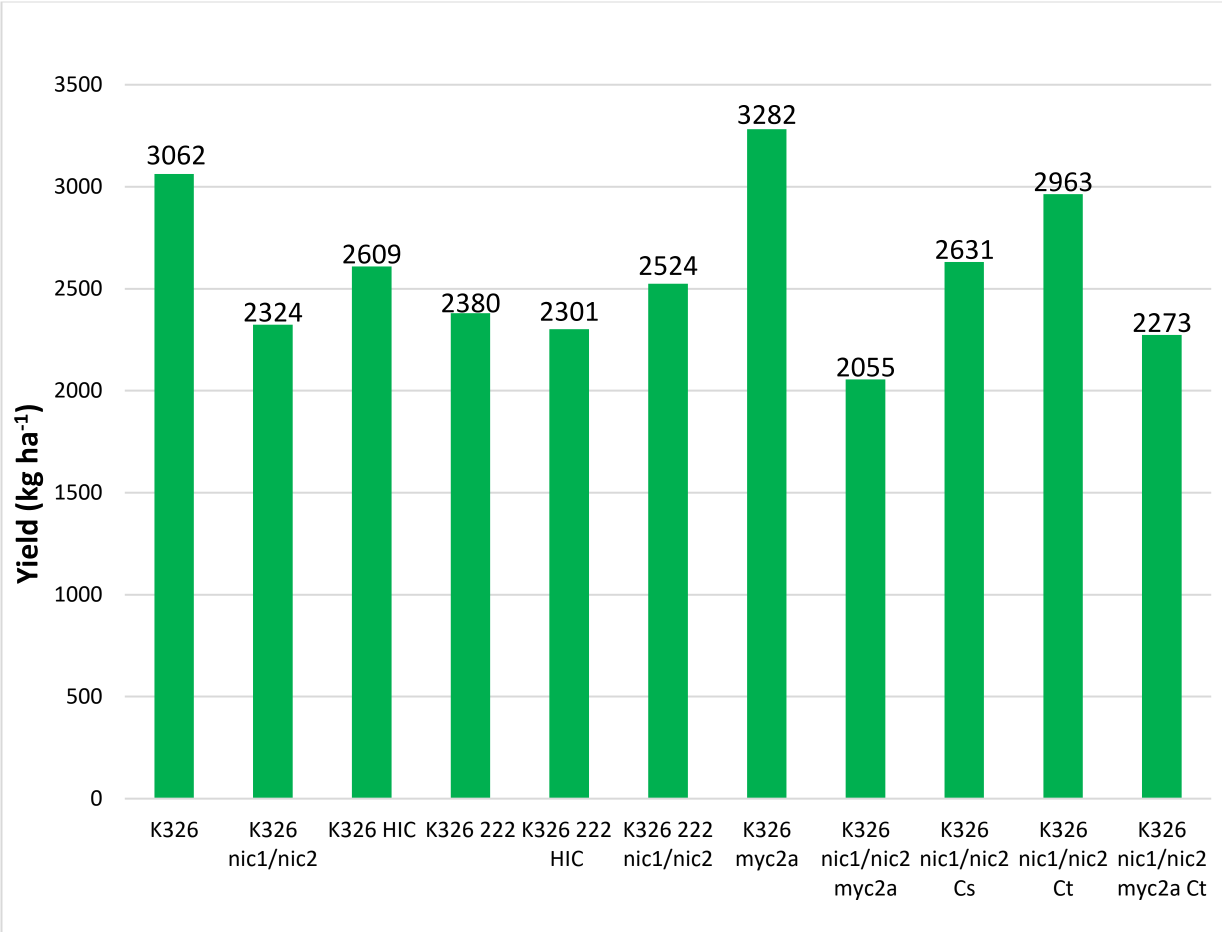


Figure 2.6 Average cured leaf yields for K326 NILs. Numbers above each bar indicate the average yield in kg ha^{-1}. No line was statistically different from another at the *P* < 0.05 level of significance.

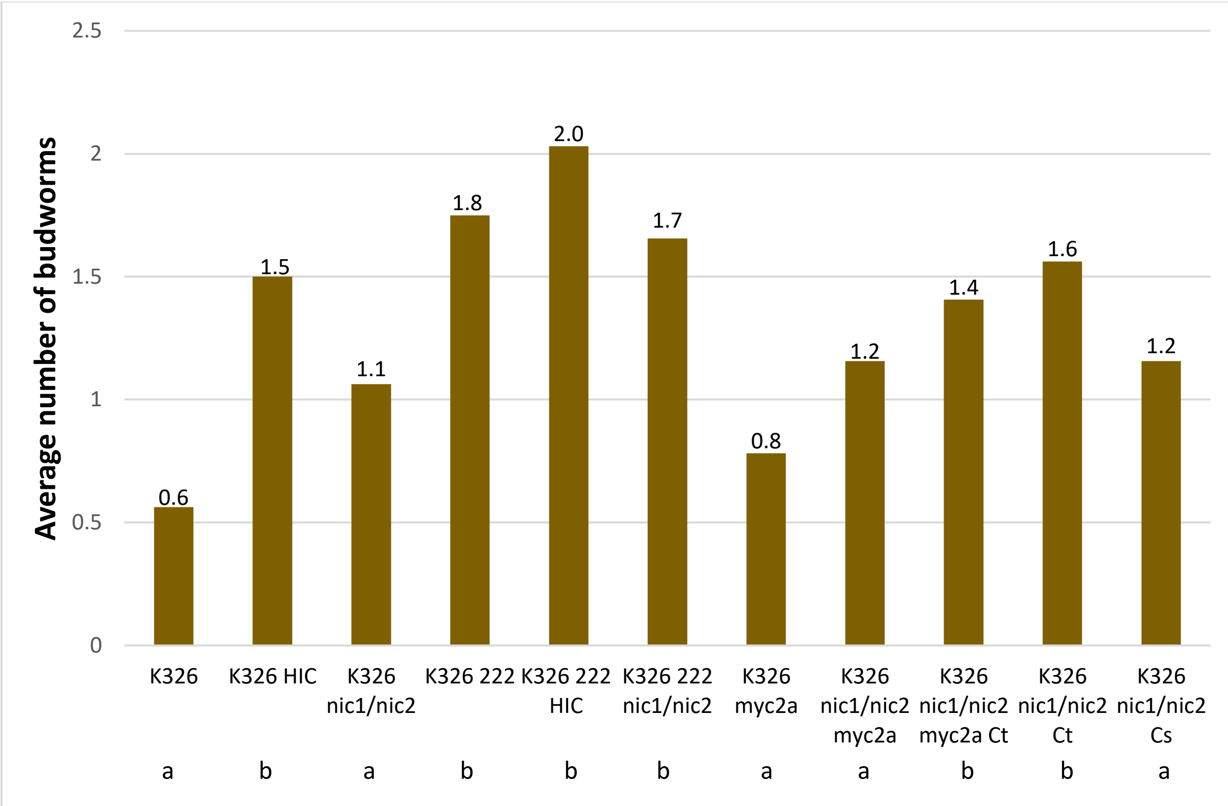


Figure 2.7 Average number of budworms counted on plots of K326 NILs evaluated in the Clayton, NC, field experiment. Numbers above each bar indicate the average number of budworms counted per plot. Genotypes with different letters (below genotypic description) contained statistically different total numbers of budworms at the $P < 0.05$ level of significance.

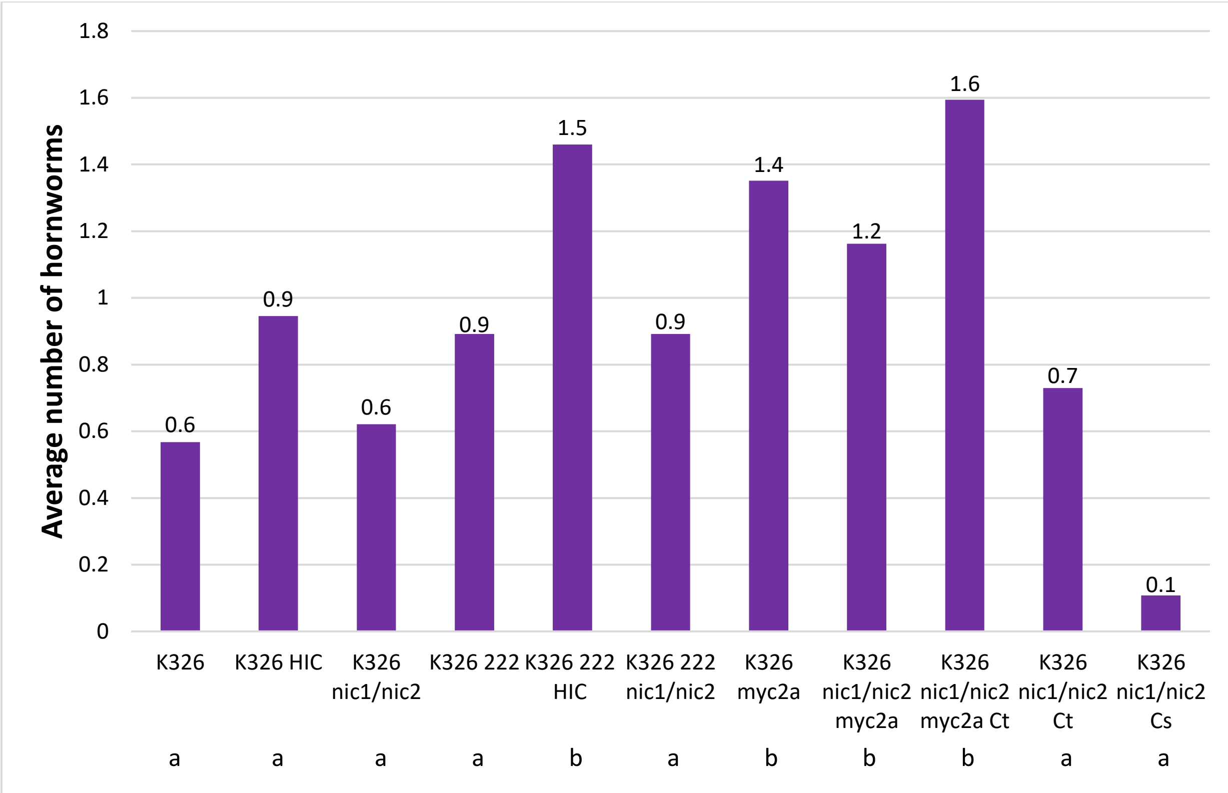


Figure 2.8 Average number of hornworms counted on plots of K326 NILs evaluated in the Clayton, NC, field experiment. Numbers above each bar indicate the average number of hornworms counted per plot. Genotypes with different letters (below genotypic description) contained statistically different total numbers of hornworms at the $P < 0.05$ level of significance.

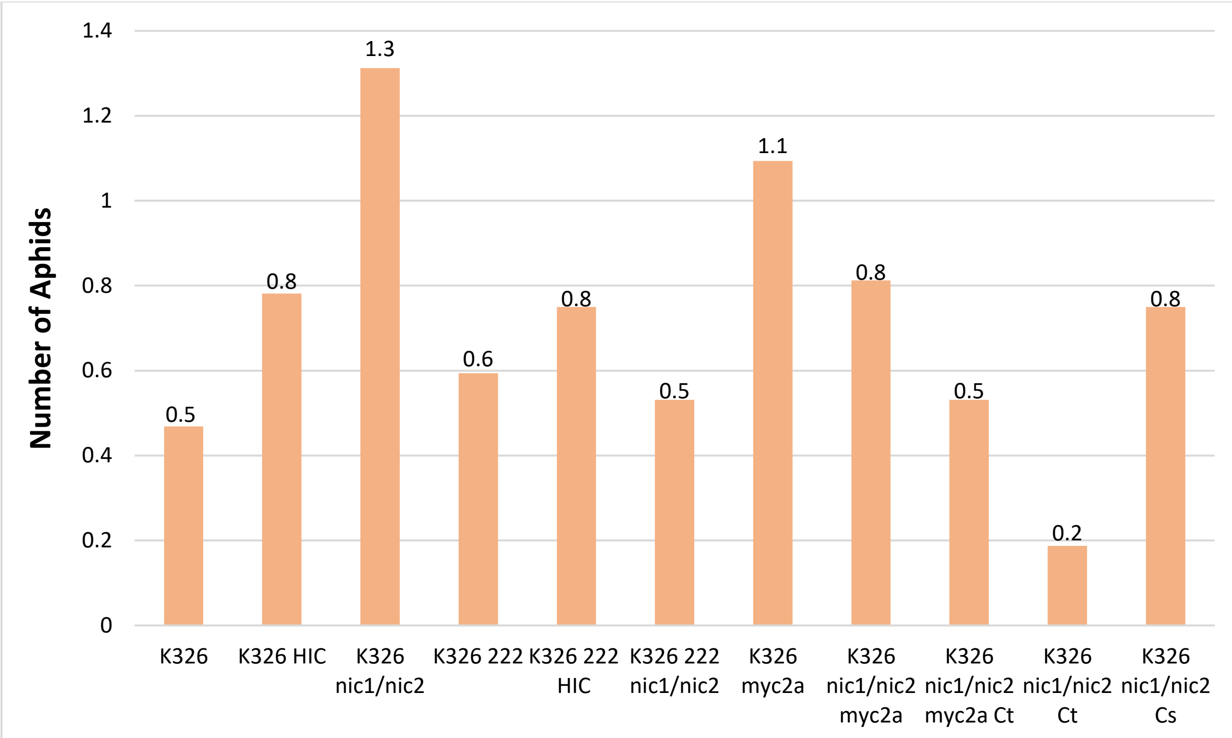


Figure 2.9 Average number of plants infested with aphids in plots of K326 NILs evaluated in the Oxford, NC, field experiment. Numbers above each bar indicate the average number of plants infested with aphids. No line exhibited a statistically different number of infested plants from another at the *P* < 0.05 level of significance.

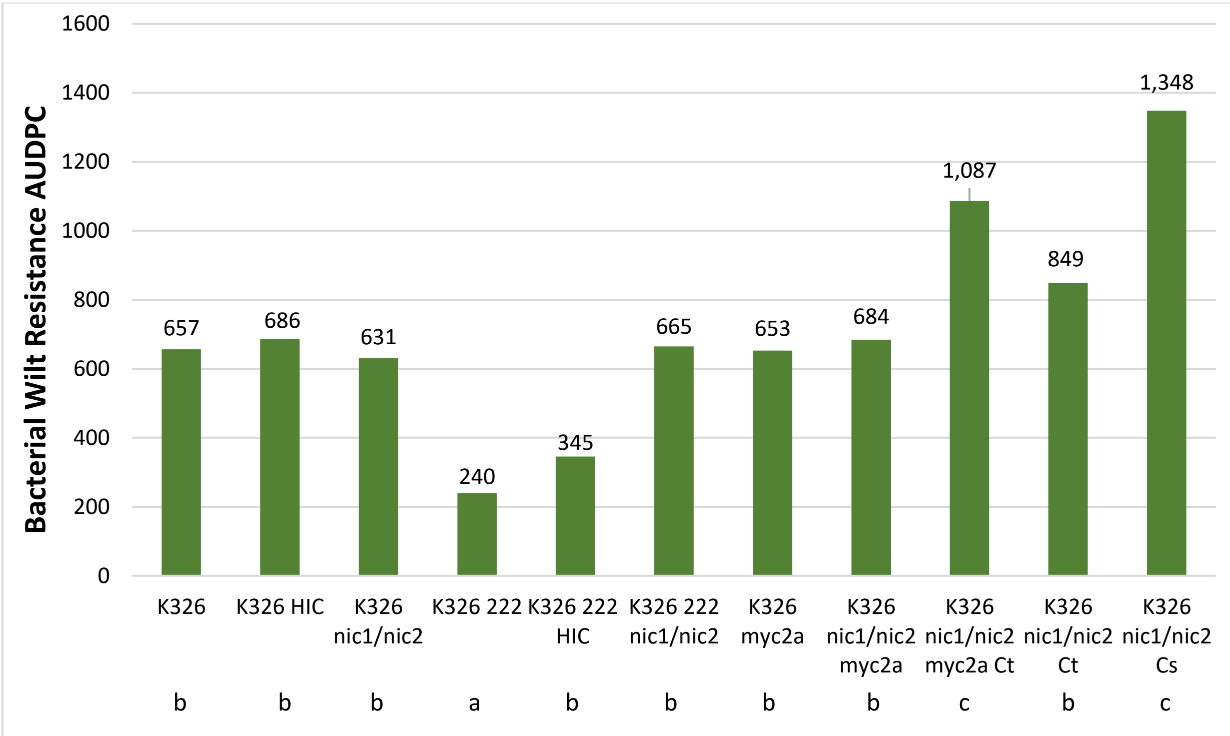


Figure 3.0 Area under disease progress curve (AUDPC) means for bacterial wilt for K326 NILs evaluated in the Rocky Mount, NC, field experiment. Numbers above each bar indicate the AUDPC mean value. Genotypes with different letters (below genotypic description) exhibited statistically different AUDPC values at the $P < 0.05$ level of significance.

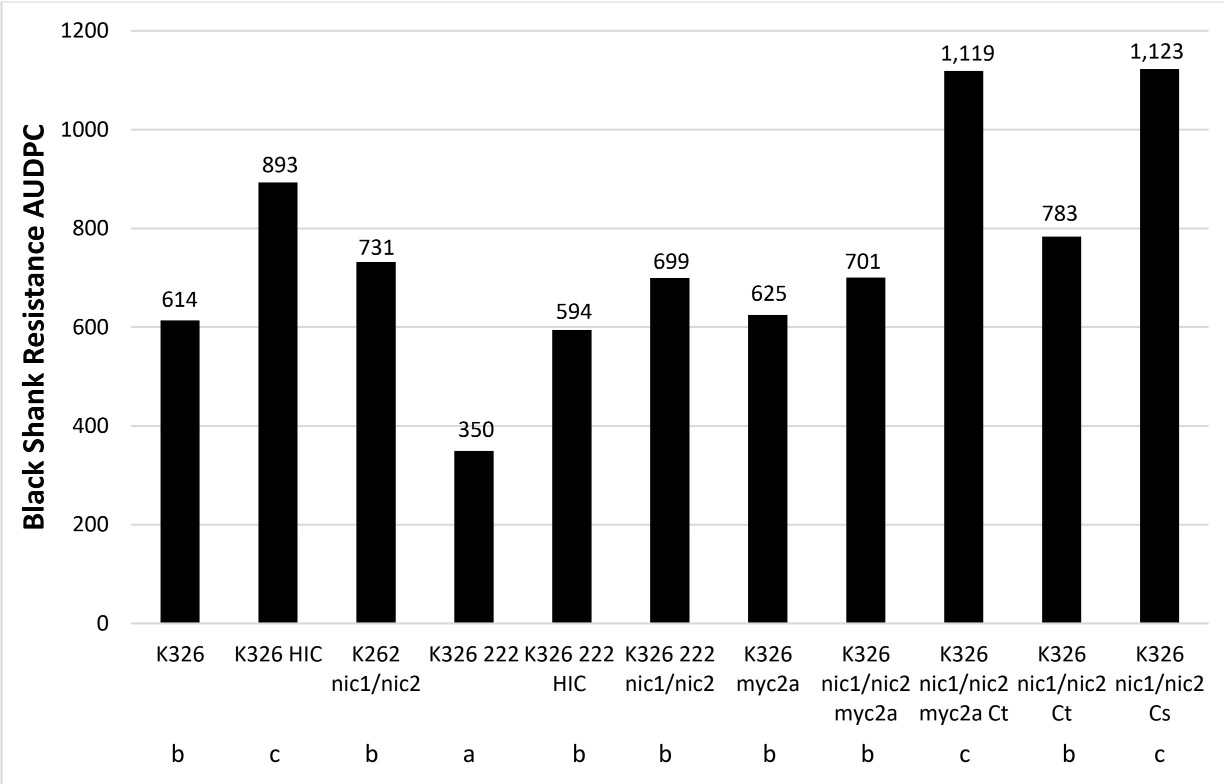


Figure 3.1 Area under disease progress curve (AUDPC) means for black shank for K326 NILs evaluated in the Rocky Mount and Kinston, NC, field experiments. Numbers above each bar indicate the AUDPC mean value. Genotypes with different letters (below genotypic description) exhibited statistically different AUDPC values at the $P < 0.05$ level of significance.

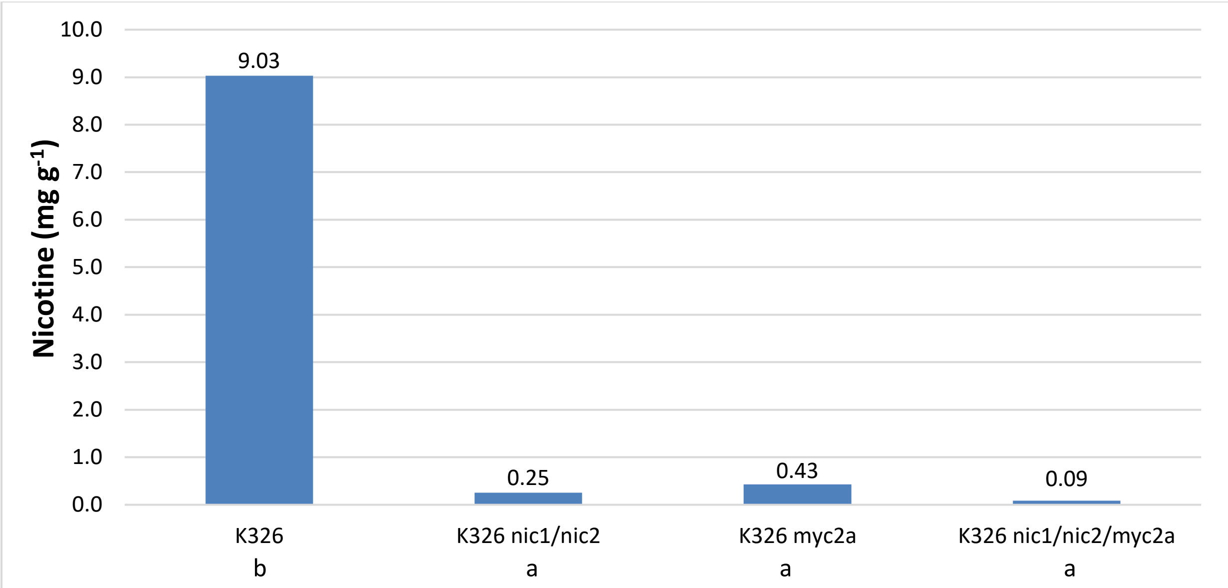


Figure 3.2 Average total nicotine levels for K326 and three K326 NILs differing for allelic variability at the *Nic1*, *Nic2*, and *Myc2a* loci, evaluated in the greenhouse experiment. Numbers above each bar indicate the average nicotine concentrations in mg g^{-1}. Total nicotine means with different letters (below genotypic descriptions) are significantly different from each other at the $P < 0.05$ level of significance.

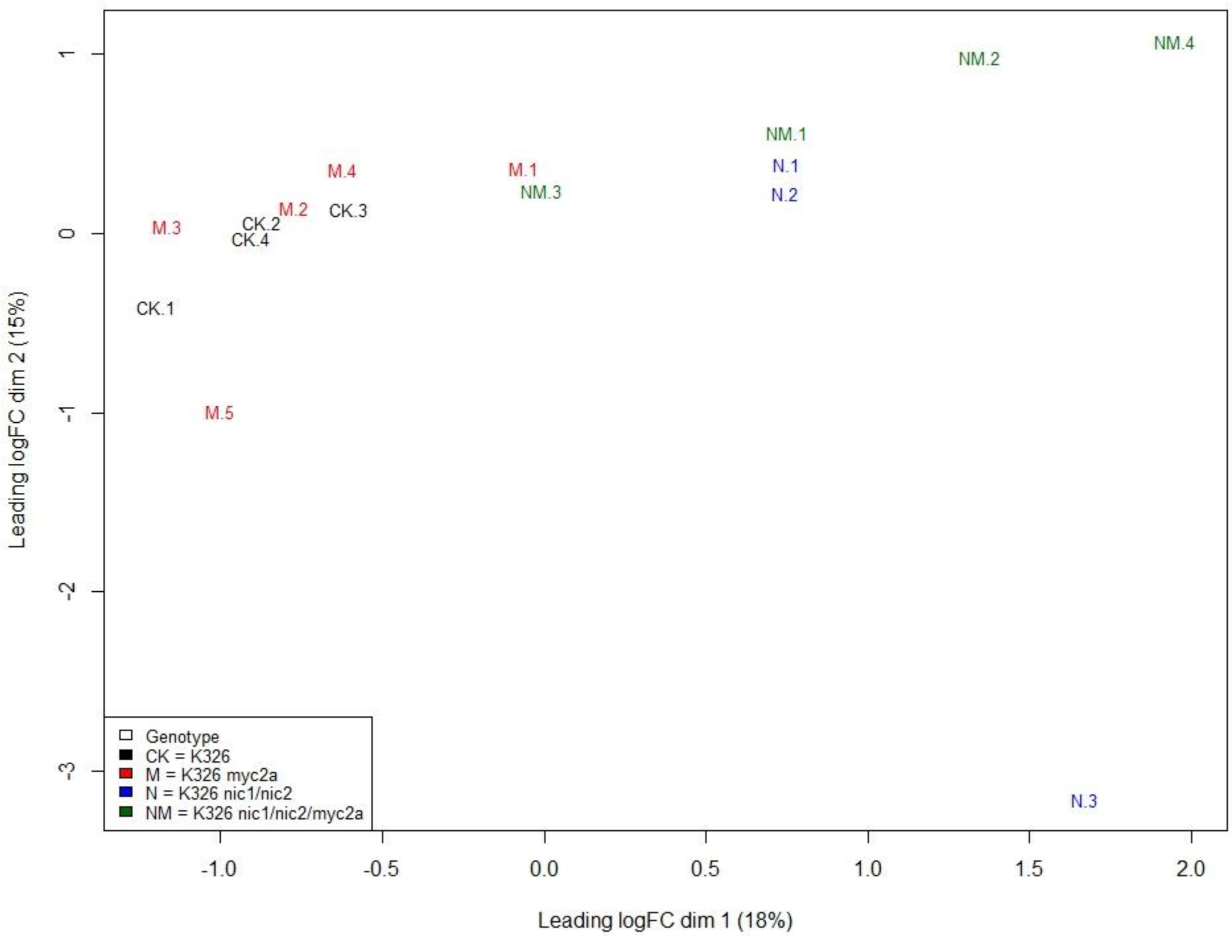


Figure 3.3 Multidimensional Scaling (MDS) Plot for K326 and three K326 NILs differing for allelic variability at the *Nic1*, *Nic2*, and *Myc2a* loci. The distance between the dots represents the similarity of the corresponding samples transcriptional profiles.

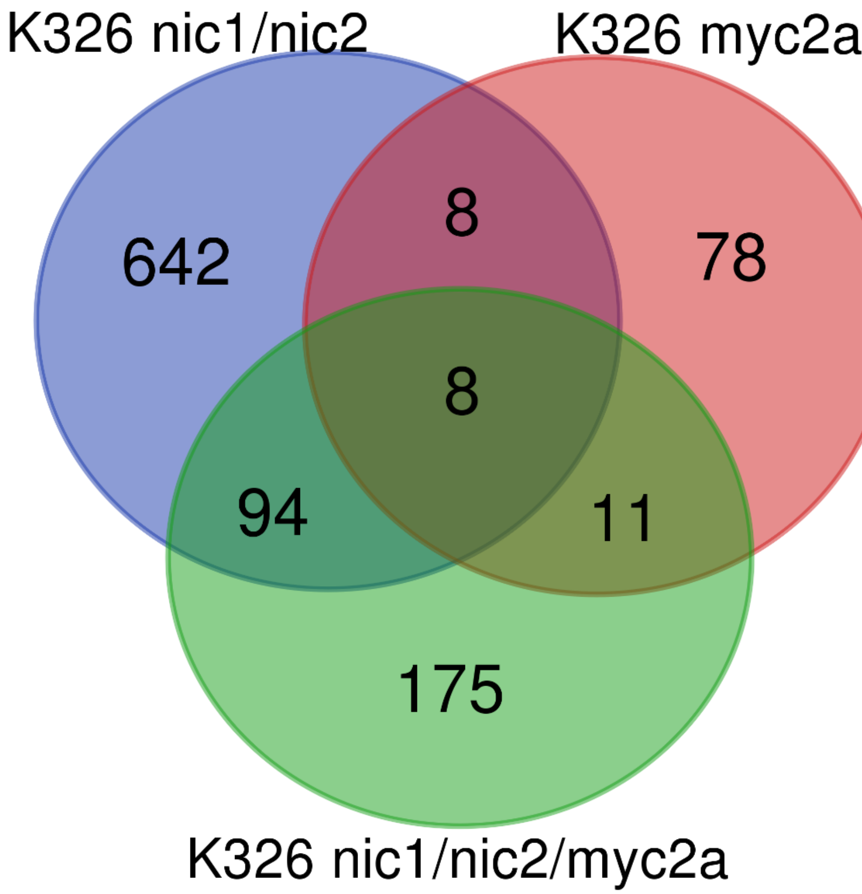


Figure 3.4 Venn diagram of DEGs identified in K326 NILs containing recessive *nic1*/*nic2* alleles and/or mutated *Myc2a* gene.