

## ABSTRACT

BINGWU, WANG. Factors in Nicotine Biosynthesis in Tobacco. (Under the direction of Rongda Qu).

Nicotine is the predominant alkaloid in tobacco leaves and roots. The pathway for its biosynthesis has been mostly elucidated except for the final step of condensation between nicotinic acid derivative and n-methylpyrrolinium cation. Two enzymes, putresine methyltransferase (PMT) and quinolinic acid phosphoribosyltransferase (QPT) acting in this pathway were previously demonstrated to be key enzymes in regulating nicotine synthesis.

In this study, the *PMT* and *QPT* genes were over-expressed individually and simultaneously in a commercial cultivar of *Nicotiana tabacum* in attempt to alter the leaf alkaloid levels. Field trials were conducted to test the effects of the transgene over-expression under field conditions. A total of 101 lines were tested in a preliminary trial in 2007. Several lines with elevated total alkaloid levels from each gene construct were selected as parents for crossing to generate F1 hybrids which would over-express both *PMT* and *QPT* genes. The parent lines and their F1 hybrids were tested in the field in 2008 at two locations in North Carolina. Based on the results obtained in the 2008 trial, lines were chosen and tested again in 2009 at two locations in North Carolina. In the field tests of two successive years of 2008 and 2009, only two *PMT* over-expression lines showed significantly lower total alkaloid levels caused by co-suppression of the *PMT* gene. However, no significant increase in total alkaloids was observed from any of the *PMT* or *QPT* over-expression transgenic plants, or from F1 hybrids derived from crossing between them.

Our results on over-expression and co-suppression of *PMT* and *QPT* genes suggest tight and complicated regulation of the alkaloid biosynthetic pathway in tobacco. Two possible explanations are proposed to explain the observations from this experiment. first, although the reactions that PMT and QPT catalyze were previously demonstrated as being rate limiting steps in each of the two branches of nicotine biosynthesis pathway, they may not be the main limiting steps for overall nicotine synthesis, a role possibly served by nicotine

synthase. Secondly, the key to regulation of PMT and QPT activities may be at translational and/or post-translational levels.

Regulation at the transcriptional level is considered an important mechanism in modulating nicotine biosynthesis. Six transcription factors (TFs) were isolated in this study. Five of them were obtained by yeast one-hybrid screening of a tobacco root cDNA library 30 min after topping with use of the *QPT2* (also known as *TobRD2*) gene promoter as the bait.

Among the six TFs characterized in transgenic over-expression and RNAi plants, four have effects on *QPT* mRNA level and/or nicotine level, and are reported in the dissertation. All four TF genes, named *NtMYC2a*, *NtMYC2b*, *NtETTA* and *NtERF98*, are expressed in root, stem, leaf, and flower organs of tobacco plants and their expression levels are affected in roots by topping, wounding, or methyl jasmonate (MeJA) treatments. *NtMYC2a*, *NtMYC2b* and *NtETTA* mRNA levels are increased by the treatments whereas *NtERF98* is decreased.

Among the four TFs, *NtMYC2a* and *NtMYC2b* are homologs sharing 96% identity in AA sequence and belong to the *bHLH* TF gene family. The change of nicotine level in over-expression and RNAi lines of *NtMYC2a* and *NtMYC2b* genes, and the decreased expression of pathway genes, *PMT* and *QPT*, in RNAi lines suggest that both *NtMYC2a* and *NtMYC2b* are positive regulators in nicotine biosynthesis. *NtMYC2a* seems to play a more important role than *NtMYC2b* in enhancing nicotine synthesis with up to 2.5-fold increase in the constitutive nicotine level in *NtMYC2a* over-expression lines comparing to the 1.5-fold in *NtMYC2b* over-expression lines. However, these two homologs appear to function differently in regulating *PMT* and *QPT* mRNA levels. Over-expression of *NtMYC2b* had little effect on *PMT* and *QPT* mRNA levels whereas over-expression of *NtMYC2a* reduced substantially the expression of these two key structural genes although the nicotine level increased. This phenomenon indicates a complicated regulatory mechanism in which a negative feedback loop and translational or posttranslational regulation of the structural genes may be involved.

*NtERF98* from the *AP2/ERF* gene family has a rather complicated modulating role in nicotine biosynthesis. *NtERF98* is a negative regulator of nicotine synthesis in the sense that it is down-regulated by the three treatments that stimulate nicotine synthesis, and that the nicotine levels increased up to 50% in a majority of its RNAi lines. In addition, there was a

good negative correlation between *NtERF98* and *QPT* mRNA levels in most of the over-expression lines. However, in those lines, while *QPT* was down-regulated, nicotine level was unchanged or moderately increased.

*NtETT<sub>a</sub>*, an *ARF* TF family gene, appears to positively regulate *QPT* expression as shown in the RNAi lines where *QPT* mRNA level was reduced approximately 15 fold, indicating that *NtETT<sub>a</sub>* is required for *QPT* expression. However, over-expression of *NtETT<sub>a</sub>* did not increase *QPT* expression and often moderately reduced its mRNA level, implicating a role of *NtETT<sub>a</sub>* in the complicated tight control of *QPT* expression.

Factors in Nicotine Biosynthesis in Tobacco

by  
Bingwu Wang

A dissertation submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

Crop Science

Raleigh, North Carolina

2011

APPROVED BY:

---

Dr. Rongda Qu  
Chair of Advisory Committee

---

Dr. David Danehower

---

Dr. Ralph Dewey

---

Dr. Ramsey Lewis

---

Dr. Ray Long

---

Dr. Deyu Xie

## **DEDICATION**

This work is dedicated to my parents, parents-in-law, and my wife Yu Zhang.

## **BIOGRAPHY**

Bingwu Wang was born in Jiangxi Province, China on Nov.4, 1976. He entered Agricultural University of Jiangxi at Nanchang in 1992, majoring in Agronomy and obtained his Bachelor of Science degree in 1996. After graduation, Bingwu continued his education and received a Master of Science degree in the College of Life Science at Beijing University, majoring in plant biology, in 1999. With several years of working experience, he realized a strong need for advanced training for research. He applied and was admitted as a graduate student into the Crop Science Department at North Carolina State University under the supervision of Dr. Rongda Qu in 2006. His project concerns the molecular aspects of tobacco breeding, and he completed his Ph.D. study in December, 2010.

## ACKNOWLEDGEMENTS

I could not complete my thesis work without the help from a lot of nice people.

First of all, I would like to give my sincere gratitude to my committee members, especially Dr. Rongda Qu, for his consistent support for my five years graduate study. His inspiration and guidance mean a great deal to me. I would like to thank Dr. Ramsey Lewis for providing me the haploid plants and the expert advice and help for the field tests. I also want to thank him for the comprehensive literature he provided, which built me the basic knowledge of tobacco research on alkaloid. My thanks extend to Dr. David Danehower for his kind help in coordinating the alkaloid quantification and in my field tests. I am grateful to Dr. Deyu Xie for his help in teaching me the data analysis and the experimental design, to Drs. Ray Long and Ralph Dewey for their suggestions, advices and supports to my experiments.

I would like to thank Gary Hamm, Karen Andres, Tyler Steede, Brad Rayfield, and Tomas Moreno for their great help in field tests and alkaloids quantification. Also, I would like to thank Dr. Carole Saravitz, Dr. Janet Shurtleff and all the staff for their help in my phytotron experiments and Dr. Consuelo Arellano for her help on statistics analysis.

I would like to thank Dr. Quanzi Li for providing the vectors and experimental materials, to Shaobin Peng, Wenzheng Li, Jiayan Sun, Kemei Ding, Tye Owens, Albert Buck, and the crews in the research stations for their assistance in my experiments.

The previous and current lab members helped and are helping me a great deal in my experiments, especially Drs. Elumalai Sivamani, Kasi Azhakanandam, Jianli Lu, Minesh Patel, and Wanjun Zhang. I am very thankful to them.

Finally, I have to thank my wife, Yu Zhang. Her love and support make my graduate study at NCSU a wonderful experience in my life.

## TABLE OF CONTENTS

	Page
<b>LIST OF TABLES</b> .....	ix
<b>LIST OF FIGURES</b> .....	x
<b>CHAPTER I Alteration of total alkaloid through over-expression of</b>	
<i>PMT</i> and <i>QPT</i> genes .....	1
<b>INTRODUCTION</b> .....	1
Location of nicotine biosynthesis and its transportation .....	1
Genetic control of the alkaloid accumulation .....	1
Modification of total alkaloid through conventional breeding .....	2
Nicotine biosynthetic pathway .....	3
Key regulatory steps in the nicotine biosynthetic pathway .....	5
Genes involved in nicotine biosynthetic pathway .....	6
Total alkaloid alteration through genetic engineering .....	7
Objectives of this study .....	8
<b>MATERIALS AND METHODS</b> .....	10
Plant materials and transformation .....	10
<i>Agrobacterium</i> strain and vectors .....	11
PCR analysis of the putative transgenic plants .....	12
Southern blot analysis of transgenic plants .....	13
Northern blot analysis of transgenic plants .....	14
Field tests .....	14

## TABLE OF CONTENTS (continued)

	Page
Hydroponic culture .....	15
Quantification of total and individual alkaloids .....	16
Statistical analysis .....	17
<b>RESULTS</b> .....	<b>18</b>
Preliminary field test in 2007 .....	18
Southern analysis of the transgenic plants .....	22
Parental lines and F1 crossing lines used in field tests in 2008 and 2009 .....	25
Northern analysis of parental lines .....	26
Field tests in 2008 and 2009 .....	28
Summary .....	40
<b>DISCUSSION</b> .....	<b>41</b>
<i>Nicotiana africana</i> and midvein culture technique for fast recovery of homozygous transgenic plants .....	41
Multiple transgene copies .....	41
Parental line selection and field tests .....	42
Alteration of total alkaloid through <i>PMT</i> and <i>QPT</i> over-expression .....	43
<b>REFERENCES</b> .....	<b>46</b>
<b>CHAPTER II Identification of transcription factors involved in nicotine     biosynthesis</b> .....	<b>53</b>

**TABLE OF CONTENTS** (continued)

	Page
<b>INTRODUCTION</b> .....	<b>53</b>
Factors affecting nicotine accumulation .....	53
Jasmonic acid (JA) .....	54
Ethylene (ETH) .....	55
Auxin .....	57
Genetic analysis of nicotine synthesis in tobacco .....	57
Transcription factors involved in nicotine biosynthesis .....	58
Objectives of the study .....	60
<b>MATERIALS AND METHODS</b> .....	<b>62</b>
Yeast one-hybrid experiments for cloning transcription factors .....	62
Expression analysis of the isolated transcription factors in tobacco .....	65
Tobacco transformation for over-expression and down-regulation of the isolated transcription factors .....	66
Quantification of major alkaloids .....	68
Phylogenetic tree construction .....	68
<b>RESULTS</b> .....	<b>69</b>
Transcription factor <i>NtMYC2a</i> and <i>NtMYC2b</i> genes .....	71
Transcription factor <i>NtERF98</i> gene .....	78
Transcription factor <i>NtETTA</i> gene .....	84

**TABLE OF CONTENTS** (continued)

	Page
<b>DISCUSSION</b> .....	91
<i>QPT2</i> promoter and yeast one-hybrid screening .....	91
Four transcription factors involved in nicotine biosynthesis were isolated ....	92
Transcription factor <i>NtMYC2a</i> , <i>NtMYC2b</i> are positive regulators in nicotine biosynthesis .....	93
Possible roles of transcription factor <i>NtERF98</i> and <i>NtETT1</i> in nicotine Biosynthesis .....	95
<b>REFERENCES</b> .....	98
<b>APPENDIX</b> .....	108

## LIST OF TABLES

	Page
Table 1.1 Crosses between PMT/K346 and QPT/K346 over-expression lines .....	26
Table 1.2 Crosses between PMT/NCTG61 and QPT/NCTG61 over-expression lines ..	26
ANOVA Table 1 .....	31
ANOVA Table 2 .....	32
ANOVA Table 3 .....	34
ANOVA Table 4 .....	35
ANOVA Table 5 .....	36
ANOVA Table 6 .....	37
ANOVA Table 7 .....	39
ANOVA Table 8 .....	39

## LIST OF FIGURES

	Page
<b>CHAPTER I</b>	
Figure 1.1 Commonly accepted biosynthetic pathway of major alkaloids (from Chintapakorn and Hamill 2003) .....	5
Figure 1.2 <i>PMT</i> and <i>QPT</i> over-expression constructs used for transformation .....	11
Figure 1.3 PTA of Transgenic lines on K346 background in 2007 (Kinston, NC) .....	20
Figure 1.4 PTA of Transgenic lines on NCTG61 background in 2007 (Kinston, NC) ...	21
Figure 1.5 Southern blot analysis of <i>PMT</i> over-expression transformed plants. ....	22
Figure 1.6 Southern blot analysis of <i>PMTox</i> transgenic lines. ....	23
Figure 1.7 Southern blot analysis of <i>QPTox</i> transformed plants. ....	24
Figure 1.8 Southern blot analysis of <i>QPTox</i> lines. ....	25
Figure 1.9 Northern blot analysis of <i>PMT</i> over-expression parental lines .....	27
Figure 1.10 Northern blot analysis of <i>QPT</i> over-expression parental lines .....	28
Figure 1.11 PTA level of transgenic lines in the K346 background tested at Rocky Mount, NC, in 2008 .....	29
Figure 1.12 PTA level of transgenic lines in the NCTG61 background tested in Rocky Mount, NC, in 2008 .....	30
Figure 1.13 PTA of transgenic K346 lines and F1 crossing lines from field test at Oxford, NC, in 2009 .....	30

**LIST OF FIGURES (continued)**

	Page
Figure 1.14 PTA values of transgenic K346 lines and F1 crossing lines in field test at Clayton, NC, in 2009 .....	31
Figure 1.15 PTA level of selected lines in small experiment 1 at Oxford .....	33
Figure 1.16. PTA level of select lines in small experiment 1 at Clayton. ....	33
Figure 1.17 PTA level of select lines in small experiment 2 at Oxford .....	35
Figure 1.18 PTA level of selected lines in small experiment 2 at Clayton .....	36
Figure 1.19 PTA level of selected lines in small experiment 3 at Oxford .....	38
Figure 1.20. PTA level of selected lines in small experiment 3 at Clayton .....	38
 <b>CHAPTER II</b>	
Figure 2.1 Vector map of T-DNA region of pBI121 .....	66
Figure 2.2 Schematic representation of the region of vector pQli used for inverted repeat cloning. ....	67
Figure 2.3 Unrooted phylogenetic tree of NtMYC2a, NtMYC2b, and other 15 bHLH transcription factors .....	72
Figure 2.4 Organ expression pattern of <i>NtMYC2</i> .....	73
Figure 2.5 Time course of <i>NtMYC2</i> expression in root after topping, wounding, or MeJA treatment .....	74
Figure 2.6 Expression of <i>NtMYC2a</i> and <i>NtMYC2b</i> in <i>NtMYC2a</i> or <i>NtMYC2b</i> over expression lines .....	75

## LIST OF FIGURES (continued)

	Page
Figure 2.7. Combined expression of <i>NtMYC2a</i> and <i>NtMYC2b</i> in <i>NtMYC2</i> RNAi lines .....	75
Figure 2.8 Nicotine concentrations of over-expression and RNAi lines of <i>NtMYC2a</i> and <i>NtMYC2b</i> .....	76
Figure 2.9 <i>QPT</i> expression level in <i>NtMYC2 a</i> , or <i>b</i> over-expression and RNAi transgenic lines .....	77
Figure 2.10 <i>PMT</i> expression in <i>NtMYC2a</i> , or <i>b</i> over-expression and RNAi transgenic lines .....	77
Figure 2.11 <i>NtERF98</i> expression in four organs of tobacco plant .....	79
Figure 2.12 Time course of <i>NtERF98</i> expression in root after topping, wounding, or MeJA treatment .....	80
Figure 2.13 <i>NtERF98</i> and <i>QPT</i> expression in ten putative <i>NtERF98</i> over-expression transgenic plants .....	81
Figure 2.14 Nicotine concentrations in ten <i>NtERF98</i> over-expression transformed lines and wild type control .....	81
Figure 2.15 Northern analysis of <i>NtERF98</i> expression in its RNAi tranformed lines ...	82
Figure 2.16 Northern hybridization of <i>QPT</i> expression in the <i>NtERF98</i> RNAi transformed tobacco lines .....	83

**LIST OF FIGURES (continued)**

	Page
Figure 2.17 Nicotine concentrations in <i>NtERF98</i> RNAi transgenic tobacco lines and wild type control .....	83
Figure 2.18 Unrooted phylogenetic tree of <i>NtETT</i> a and ten other related auxin responsive factors .....	85
Figure 2.19 Northern blot analysis of <i>NtETT</i> a expression in tobacco root, stem, leaf and flower .....	86
Figure 2.20 Time course of <i>NtETT</i> a expression response in tobacco root after topping, wounding or MeJA treatment .....	87
Figure 2.21 Expression level of <i>NtETT</i> a in over-expression lines and RNAi lines .....	88
Figure 2.22 <i>QPT</i> expression in <i>NtETT</i> a over-expression and RNAi lines .....	88
Figure 2.23 Nicotine concentrations of <i>NtETT</i> a over-expression and RNAi lines .....	89

## **Chapter 1. Alteration of total alkaloid content in tobacco (*Nicotiana tabacum* L.) through over-expression of *PMT* and *QPT* genes**

### **INTRODUCTION**

Alkaloids are the quintessential natural products identified with tobacco. They greatly affect the quality of commercial tobacco products. There are four major alkaloids found in the leaves of tobacco: nicotine, nornicotine, anabasine, and anatabine. Nicotine is the predominant alkaloid, usually accounting for more than 90% of the total alkaloids in commercial tobacco cultivars (Bush et al. 1999). The measurement of total alkaloids therefore generally reflects the nicotine level in tobacco cultivars unless the conversion of nicotine to nornicotine is substantial. A variety of factors affect alkaloid accumulation including genotype, environment, fertilization, and other agricultural practices. The production of nicotine is stimulated by topping, wounding, and herbivore damage.

#### **Location of nicotine biosynthesis and its transportation**

Nicotine biosynthesis occurs predominantly in the roots of tobacco plants (Dawson, 1941, 1942). The cortex and epidermis of the differentiated region of the root tip is considered the site of nicotine production. The plant then transports the alkaloids through the vascular bundle to the leaves where the alkaloids are then stored in the vacuoles (Shoji et al. 2000a, 2002; Kato et al. 2005). Several transporters may be involved in the translocation process. A transporter gene, named MATE, has recently been cloned and characterized (Morita et al. 2009).

#### **Genetic control of the alkaloid accumulation**

Genetic controls influence the type and level of alkaloids observed in *Nicotiana tabacum*. In one system, two unlinked genetic loci, *Nic1* and *Nic2* (or A and B), have been shown to control the total alkaloid accumulation. *Nic1* and *Nic2* have different dosage effects on alkaloid accumulation. *Nic1* (A) has 2.4-fold greater activity than *Nic2* (B) and their effects are additive (Legg and Collins 1971). Commercial varieties with high total alkaloid content

are considered as homozygous dominant (AABB) at these two loci and low alkaloid level genotypes are considered as homozygous recessive (aabb) at these two loci. Under this genetic model, nine genotypes of tobacco plants with different levels of total alkaloids were developed (Legg et al. 1969, 1971). Very recently, the *Nic2* locus has been molecularly characterized. It is actually a cluster of transcription factor genes from the ethylene responsive factor (*ERF*) family. They regulate nicotine biosynthetic pathway genes (Shoji et al. 2010).

Another system controls the types of alkaloid produced and has a genetic locus which controls the conversion of nicotine to nornicotine. When both alleles of this locus are recessive, the plant contains predominantly nicotine. If one or both alleles are dominant, the plant primarily produces nornicotine. Nornicotine results from demethylation of nicotine (Mann et al. 1964). The gene encoding nicotine demethylase, a cytochrome P450 gene, *CYP82E4*, was cloned and characterized (Siminszky et al. 2005). Suppressing the expression of this gene resulted in a drastic reduction in nornicotine content (Lewis et al. 2008).

In addition to the two genetic systems described above, some minor or quantitative factors are involved in alkaloid synthesis. Therefore, it is possible to produce tobacco lines with varying alkaloid content within the range of the parents' alkaloid levels (Matzinger et al. 1972, 1989).

### **Modification of total alkaloid through conventional breeding**

The low nicotine content trait has been of interest to tobacco breeders. LA Burley 21 is a low total alkaloid line produced by incorporation of a low alkaloid gene(s) from a Cuban cigar variety into Burley 21 through several backcrosses (Legg et al. 1970). It has approximately 0.2% total alkaloids (dry weight) compared to the 3.5% (dry weight) of its parent, Burley 21.

Similarly, Chaplin and Burk (1983) developed some flue-cured tobacco lines with different alkaloid levels by backcrossing. They used NC95, SC58, and Coker 139 as recurrent parent

lines and crossed them with LAFC 53 (a low total alkaloid line in NC95 background) followed by several backcrosses. Five different alkaloid levels were obtained from the NC95 family, six from the SC58 family, and four from the Coker139 family.

However, the breeding of commercial tobacco cultivars is more complicated than simply focusing on the alkaloid content in leaf. It has been found that alkaloid accumulation is genetically linked with other important agronomic traits. For example, a reverse correlation exists between total alkaloids and yield. Selection for increased yield may result in reduced level of total alkaloids (Chaplin and Week 1976). Conventional breeding methods have not been very efficient for breaking this close genetic relationship. It is believed that genetic manipulation at the molecular level can meet this need of tobacco breeders by only modifying alkaloid accumulation without any apparent effect on other traits such as yield.

### **Nicotine biosynthesis pathway**

A great deal of research has been performed to elucidate the pathway of nicotine biosynthesis. Based on the data collected using isotopic feeding experiments, enzyme purification and activity analysis, gene discovery and genetic analysis, a generally accepted pathway has been established although some details still need to be elucidated (Brown and Byerrum 1952; Lamberts and Byerrum 1958; Lamberts et al. 1959; Leete et al. 1964; Yang et al. 1965; Schutte et al. 1966; Frost et al. 1967; Scott and Glynn 1967; Lette 1967; Mizusaki et al. 1967; Mizusaki et al. 1971, 1972, 1973; Liu and Leete 1973; Wagner and Wagner 1985; Feth et al. 1986; Tiburcio and Galston 1986; Hibi et al. 1994; Sinclair et al. 2000; Katoh et al. 2007).

Nicotine has two ring moieties derived from two distinct branches of the pathway (Figure 1.1). One is the pyrrolidine ring, which is formed from arginine or ornithine via putrescine and methylputrescine. The first committed enzyme for this branch is PMT (putrescine methyltransferase). The other moiety is a pyridine ring which originates from nicotinic acid. Nicotinic acid is an intermediate of the pyridine nucleotide cycle, which is regulated by an entry enzyme of the cycle, QPT (quinolinic acid phosphoribosyltransferase). Other enzymes

involved in the pyrrolidine branch route are: ADC (arginine decarboxylase) which catalyzes the decarboxylation of arginine; ODC (ornithine decarboxylase) which catalyzes the decarboxylation of ornithine; MPO (methylputrescine oxidase) which catalyzes the oxidation and deamination of methylputrescine. 4-methylaminobutanal, which comes from deamination of methylputrescine, cyclizes automatically to form an N-methylpyrrolinium cation. In addition, there are eight enzymes involved in the pyridine nucleotide cycle.

PMT catalyzes the methylation of putrescine, which is the product of decarboxylation of ornithine or arginine. The methyl group is believed to be supplied by SAM (S-adenosyl-L-methionine). QPT catalyzes the entry point of the pyridine nucleotide cycle, which serves as a salvage pathway for NAD in the cell.

NS (nicotine synthase), which catalyzes the condensation step between a nicotinic acid derivative and methylpyrrolinium cation, has not been elucidated although two candidates (A622 and NBB1) have been proposed (Hibi et al. 1994, Shoji et al. 2002; Hashimoto and Kato 2007). It has been proposed that a reduction reaction occurs followed by oxidation of nicotinic acid prior to formation of specific pyridine moiety that condenses with the methylpyrrolinium cation to form nicotine (Friesen and Leete 1990). A622 encodes an isoflavone reductase-like protein (Shoji et al. 2002). Down regulation of A622 reduced the capacity of *N. glauca* to produce anatabine suggesting A622 involved in the final step of nicotine synthesis (DeBoer et al. 2009)

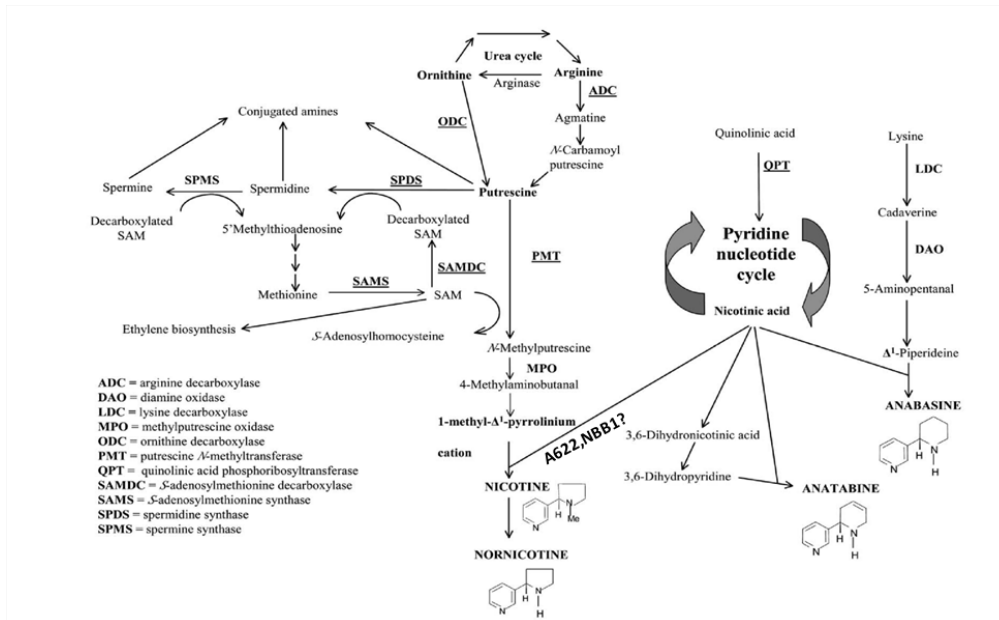


Figure 1.1 Commonly accepted biosynthetic pathway of major alkaloids (from Chintapakorn and Hamill 2003)

### Key regulatory steps in the nicotine biosynthetic pathway

Feth et al. (1986) established tobacco root callus culture in growth medium (with auxin) and nicotine induction medium (with less auxin) to study the regulation of enzyme activity along the nicotine biosynthetic pathway. The results showed that PMT had specific activity of about 0.27 pkat/mg, which is much lower than that of ODC (7.7 pkat/mg) and MPO (5 pkat/mg) measured from tobacco root callus in growth medium. This low activity suggested that PMT is the rate limiting enzyme in this branch pathway to form methylpyrrolidine, and controls the flux of putrescine during nicotine biosynthesis. Moreover, stimulation of PMT activity in the induction medium was rapid and strong, about 30 fold, compared to four-fold for MPO. ODC did not show a significant change in the same treatment (Feth et al. 1986). This result suggested that PMT plays a critical role in regulating of this branch route.

By using the same root culture system, Wagner demonstrated that QPT has a specific activity of 2.6 pkat/mg, which is much lower than those of other eight enzymes that control the

pyridine nucleotide cycle. After induction, four of these enzymes showed constitutive activity, while the other four enzymes showed a two-fold increase. QPT, in contrast, showed four-fold inductions. This result was interpreted as QPT being the most critical enzyme in the branch of the pathway controlling the production of nicotinic acid (Wagner and Wagner 1985).

Genetic and biochemical analysis of the four homozygous tobacco lines, AABB, AAbb, aaBB and aabb, found that the activities of PMT and QPT in root tissues were strictly correlated with leaf nicotine levels, whereas MPO did not show a similar correlation (Saunders et al. 1979).

The ADC protein is present in leaf, stem, flower and root and the level of ADC protein in the root is lower than that of the leaf, stem and flower. This would suggest that ADC is mainly involved in primary metabolism, and might not be strictly regulated enzyme in nicotine biosynthesis (Bortolotti et al. 2004).

### **Genes involved in nicotine biosynthetic pathway**

In the past 15 years, most genes of the nicotine biosynthetic pathway have been identified. The *PMT* gene was cloned in 1994 by subtractive hybridization and has five isoforms in the tobacco genome (Hibi et al. 1994; Riechers and Timko 1999). The *QPT* gene was characterized in 2000 and has two members (Sinclair et al. 2000). *PMT* and *QPT* are only expressed in the root and can be induced by wounding and Jasmonic acid (JA) (Sinclair et al. 2000; Cane et al. 2005). The *MPO* gene was characterized in 2007 and shows an expression pattern similar to *PMT* gene (Kato et al. 2007). The nicotine synthase candidate genes, *A622* and *NBB1*, have also been described and they have expression patterns similar to *PMT* gene (Hibi et al. 1994; Shoji et al. 2002). *A622* is an isoflavone reductase-like gene and might be involved in the final step of nicotine condensation. Both *ADC* and *ODC* genes were also cloned and each has two family members in the genome (Xu et al. 2004).

## **Total alkaloid alteration through genetic engineering**

Based on sequence information obtained for these pathway genes, genetic modification of alkaloid accumulation in tobacco leaf has been conducted. Basically, research efforts have been primarily focused on the *PMT* and *QPT* genes, and transgenic plants were generated for analysis.

Over-expression of the *PMT1a* gene was first performed in *N. sylvestris* under the control of the CaMV 35S promoter. Nicotine levels were increased by about 40% in three transgenic lines with 4-8 fold higher *PMT* transcript levels (Sato et al. 2001). Over-expression of *PMT* was also performed in *N. tabacum* cv. K326 to elevate nicotine content and several transgenic lines with enhanced nicotine level were reported (Hashimoto and Kajikawa 2008).

An antisense strategy to lower nicotine content in *N. tabacum* (Burley 21) by reducing *QPT* expression was reported (Xie et al. 2004). The full-length antisense sequence of the *NtQPT1* cDNA was cloned into the vector under the control of the *NtQPT1* gene promoter.

Transgenic plant lines with nicotine level as low as 4% of the control plants were obtained.

A similar strategy to suppress expression of *PMT* gene was applied in *N. tabacum* (cv. NC95) (Chintapakorn and Hamill 2003). The antisense sequence of the full-length *PMT1a* coding sequence was used in a gene construct under control of the 35S promoter. Nicotine levels in transgenic plants were 2-4 folds lower than that in the control plants. Similarly, co-suppression of *PMT* gene expression was also effective in reducing nicotine content (Sato et al. 2001; Wang et al. 2008). In one co-suppression line, nicotine was only 2%, and *PMT* expression level 16%, of the wild type control (Sato et al. 2001). Concomitant with the reduction of nicotine, anatabine levels were increased (Chintapakorn and Hamill 2003; Wang et al. 2009)

*ODC* and *ADC* genes were also targets for alteration of tobacco alkaloids. A yeast *ODC* gene was over-expressed in root cultures of *N. rustica*. A two-fold increase in accumulation of nicotine was obtained in transformed lines compared to the controls (Hamill et al. 1990).

However, the down regulated ADC activity did not impact on alkaloid level (Chintapakorn and Hamill 2007).

The nicotine demethylase gene can also be used to increase nicotine content when it is silenced to inhibit the conversion of nicotine to nornicotine. The increase is minimal when applied to normal tobacco plants, but can be dramatic when this gene is silenced in strong converter lines (Lewis et al. 2008).

Other pathway genes, such as *A622* and *NBB1*, have also been used to reduce or increase the level of alkaloids in tobacco. Knockdown of *NBB1* reduced the nicotine level to 16% of that of the wild type plant (Hashimoto and Kato 2007).

Finally, over-expression of the allene oxide cyclase gene, which is the gene responsible for the key step in the formation of jasmonate, led to a 4.8-fold higher level of nicotine in transgenic tobacco plants (Jiang et al. 2009).

### **Objectives of this study**

As described above, a number of publications have shown that up or down regulation of the expression level of *PMT* or *QPT* genes can result in the alteration of alkaloid accumulation in the leaves of *N. tabacum* and related species. The data collected in these over-expression cases were obtained from greenhouse-grown plants or root cultures. In the present study, experiments were carried out to over-express *PMT* and *QPT* genes individually and in combination in transgenic tobacco (*N. tabacum*) plants to examine the alkaloid level changes in field-grown transgenic plants. To minimize the possible adverse effect of constitutive expression and to mimic the endogenous genes, both genes were cloned into expression vectors under the *TobRD2* promoter, which is the promoter of the *QPT2* gene.

The objectives of the present experiments were to:

1. Generate transgenic tobacco plants that over-express *PMT* or/and *QPT* genes, either individually or simultaneously, and to determine the effects of over-expression on total alkaloids in field-grown plants, and
2. Gain additional insights into the regulation of alkaloid biosynthesis in tobacco.

## MATERIALS AND METHODS

### Plant materials and transformation

#### *Haploid plants generation*

Haploid tobacco plants were kindly provided by Dr. Ramsey Lewis (Department of Crop Science, NC State University), which were produced by the *Nicotiana africana* technique (Burk et al. 1979). In this technique, the pollen of *Nicotiana africana* was used to pollinate *N. tabacum* cv. K346 or cv. NCTG61 to induce development of haploid seeds. Most of the surviving seedlings of this cross are haploid.

#### *Haploid plant transformation*

Haploid plants were grown in the greenhouse for approximately 2 months. Fully expanded young leaves were collected and sterilized by immersing them in 70% ethanol for 30 sec, then moving them into 0.6% Clorox (The Clorox Company, Oakland, CA) solution, and gently stirring with forceps for 5 min. After that, three successive washes were conducted using sterile water, 5 min per wash. Sterilized leaves were placed onto sterile paper towels and sliced into small pieces for transformation.

Small leaf pieces were mixed with overnight-grown *Agrobacterium tumefaciens* culture (OD<sub>600</sub>=1.0) for 2 min. Excess *Agrobacterium* culture solution was removed using sterile paper towel. Co-cultivation was conducted for 2-3 days on MS medium. MS medium with BA (1 mg/L), IBA (0.1 mg/L) and kanamycin (100 mg/L) was used as the selection medium. Subculturing was performed 2-3 times at 2-wk interval. Non-transformed control plants were also regenerated using the same media and procedure but without antibiotic selection. The rooting medium was MS with 50 mg/L kanamycin and no hormones.

#### *Generation of transgenic double haploid (DH) plants*

Transgenic haploid plants were grown in the greenhouse for approximate 2 months.

Midveins were collected from the middle to bottom leaves just before flowering. Two leaf

midveins were excised from each transgenic haploid plant. Midveins were sterilized as described above, and the same selection and rooting media were used for generation of DH plants. Subculture was performed 3-4 times at 2 weeks interval.

### *T1 and F1 plants generation*

T1 seeds were obtained from self pollination of T0 DH plants and F1 seeds were derived from crosses between *PMT* over-expression lines and *QPT* over-expression DH lines of the same genotype.

### ***Agrobacterium strain and vectors***

*Agrobacterium* strain LBA4404 was used for transformation. Two vectors for over-expression of *PMT* and *QPT* genes were provided by 22<sup>nd</sup> Century, LLC. (Buffalo, NY). For both constructs, the root-specific promoter of *QPT2* (*TobRD2*) gene was used to drive transgene expression.

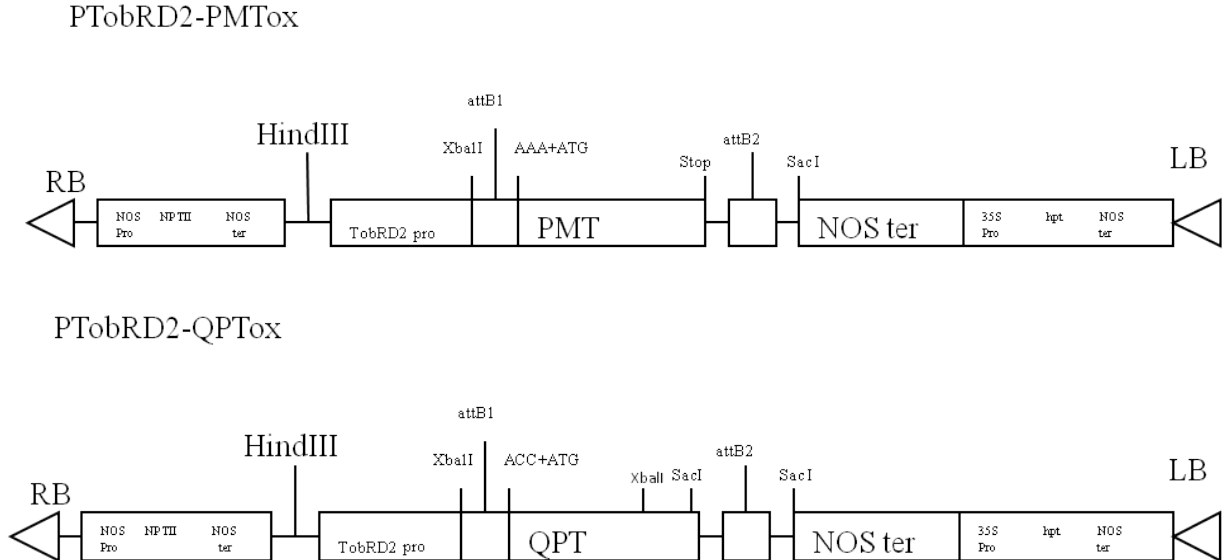


Figure1.2 *PMT* and *QPT* over-expression constructs used for transformation

## PCR analysis of the putative transgenic plants

DNA samples from putative transgenic DH plants were extracted by the CTAB method as follows:

Approximately 0.3g leaf tissue was ground with glass beads by an amalgamator (Ivoclar Vivadent Inc. Amherst, NY). Five hundred  $\mu$ L CTAB buffer [2% CTAB (hexadecyltrimethylammonium bromide), 100 mM Tris HCl (pH=8), 20 mM EDTA, 1.4M NaCl, 0.2%  $\beta$ -mercaptoethanol] was added to each sample. Tubes were incubated at 65 °C for 1 h and 200  $\mu$ L chloroform was added, and vortexed vigorously for 15 sec. The tubes were then centrifuged at 15700 g for 10 min at room temperature. The supernatant was transferred into a new tube. An equal volume of isopropanol was added and mixed well by inversion 7 times. After incubation at room temperature for 10 min, the tubes were centrifuged at 15700 g for 5 min. The resultant pellets were washed by 70% ethanol and 100% ethanol once each, and dried at room temperature. One hundred  $\mu$ L TE buffer (10 mM Tris HCl, pH 7.5, 1 mM EDTA) was added to dissolve the pellet and the DNA concentration was determined by a Nanodrop spectrophotometer (Nanodrop, Wilmington, DE).

Primers used for PCR were

pmtcdsF: 5'-GTATCCAGAAAATTTTCCCTTATCT-3'

nosterR: 5'-ATGATAATCATCGCAAGACCG-3'

qptcdsF: 5'-CCCTGCTTACATCTTGGAGACTAG-3'

nosterR: 5'-ATGATAATCATCGCAAGACCG-3'

PCR program for *PMT* and *QPT* transgenes: 95°C 5min

35 cycles

94°C 1min

55°C 1min

72°C 1min

72°C 10 min

## Southern blot analysis of transgenic plants

Approximately 500 mg young leaf tissue from each plant was used for DNA isolation by using CTAB method as described above. *Hind*III and *Xba*I were used to digest 20 µg genomic DNA from each transformed plant overnight at 37 °C. Genomic DNA from non-transformed cv. K346 was used as a negative control. Digested DNA was separated via electrophoresis on a 0.8% (w/v) agarose gel. The size fractionated DNA was denatured and blotted to Hybond-N+ nylon membrane (GE Healthcare, Buckinghamshire, UK) by capillary transfer overnight as described by Sambrook and Russell (2001). PCR products of *PMT*, *QPT*, and *HPT* genes were generated using the primers listed below, respectively, and each PCR fragment, which is a partial coding sequence of the relevant gene, was used as a template for probe generation to detect the corresponding transgenes.

PMTENDO2F: 5'-TCAGCAACTTATGGGAAGG TTC-3'

PMTENDO2R: 5'-ATCAACTACCACGTCATCGATC-3'

QPTENDO2F: 5'TTGAAACCAGGACAATTGAAGA-3'

QPTENDO2R: 5'CTCCGTATCAAACCTCCCATTG-3'

HPTENDOF: 5'-GATGTTGGCGACCTCGTATT-3'

HPTENDOR: 5'-GATGTAGGAGGGCGTGGATA-3'

The probes were made by using the Prime-It random labeling kit (Stratagene, Santa Clara, CA) according to the manufacturer's manual. Hybridization and washes were performed according to the manual of the MiracleHyb Hybridization instruction manual (Stratagene).

## Northern blot analysis of transgenic plants

Total RNA was extracted using TRIzol Reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol. Approximately 700 mg root tissue or 200 mg leaf tissue were used for RNA isolation. The extracted RNA was dissolved in DEPC- treated water and quantified with the Nanodrop (ND-1000). Ten µg RNA was separated on a 1% agarose gel in MOPS buffer. The gel was stained with EtBr and the image was photographed under UV illumination. Separated RNA on the gel was blotted onto the Hybond-N+ nylon membrane. The primers used in preparing *PMT* and *QPT* probes were:

PMT1F: 5'-TATAGTGGACTCTTCTGATCCC-3'

PMT2R: 5'-CCCTTAAAGACTTGACGACAGT-3'

QPT1F: 5'-GGTTGAGGTTGAAACCAGGA-3'

QPT2R: 5'-CCAGTTTGTCCAATCTTGTG-3'

Pre-hybridization was performed in 1x Denhart's solution at 68 °C for at least 2 hrs and the hybridization was performed at 68 °C for 12 hrs or overnight. Blots were washed twice at 25 °C for 15 min each with 2X SSC and 0.1X SDS and once at 60 °C for 30 min with 0.1X SSC and 0.1X SDS. Washed blots were exposed to a phosphorimager (GE Healthcare) to record the signals.

## Field tests

PCR-positive transgenic DH plants and non-transformed control plants were preliminarily screened at the Lower Coastal Plain Tobacco Research Station (Kinston, NC) in the fall of 2007. There were 30 PMT/K346 lines, 27 QPT/K346 lines, 21 PMT/NCTG61 lines and 23 QPT/NCTG61 T1 lines in the field test. A randomized complete block design (RCBD) was used in the experiment with three replicates for each line and six plants per plot. The plants were transplanted on Sept.14. After five weeks of growth in the field, the plants were topped on Oct. 19. Two weeks after topping (Nov. 2), the leaf samples were harvested from three

representative plants of each plot for total alkaloid quantification. One half of the top leaf was harvested from each of the three plants in each plot, mixed and dried for total alkaloids (TA) analysis.

Southern positive plants with relatively high total alkaloids were used for crossing in an attempt to obtain F1 hybrids with over-expression of both *PMT* and *QPT* genes. Crosses between five *PMT*/K346 lines and five *QPT*/K346 lines, and between five *PMT*/NCTG61 lines and three *QPT*/NCTG61 lines, were performed in the greenhouse in the winter of 2007. A total of 23 transgenic hybrids on K346 genetic background and 15 transgenic hybrids on NCTG61 genetic background were tested at the research stations at Oxford and Rocky Mount research stations, NC, in the 2008 growing season.

Based on the 2008 results, field tests were again carried out at the research stations at Oxford and Clayton, NC, in 2009. Only transgenic lines on K346 genetic background were tested that year. In addition to the two large-scale experiments, three small-scale field experiments were designed, with less number of lines and more replicates, in an attempt to reduce variation caused by the uneven soil conditions in these stations and to increase the sensitivity of the statistical analysis. Some lines with relative high TA levels and their hybrids were tested in those small experiments. All field tests in 2008 and 2009 were designed as RCBD, with 3 replicates and 5 plants per replicate, except for the three small field experiments which had 5 replicates each. Plot was 4 ft wide and 10 ft long.

### **Hydroponic culture**

*PMT* and *QPT* genes are mostly expressed in root tips. For recovery of an intact, clean set of root tissue for gene expression analysis, a hydroponics-based culture experiment was conducted in a growth chamber at NCSU Phytotron. The nutrient solution for the culture contains 0.25 M  $\text{Ca}(\text{NO}_3)_2$ , 0.25 M  $\text{Mg}(\text{NO}_3)_2$ , 0.25 M  $\text{KH}_2\text{PO}_4$ , 0.5 M  $\text{K}_2\text{SO}_4$ , 0.55 g Fe Sequestrene and 110 mL micronutrients (17  $\mu\text{M}$   $\text{H}_3\text{BO}_3$ , 5  $\mu\text{M}$   $\text{MnCl}_2$ , 0.46  $\mu\text{M}$   $\text{ZnSO}_4 \cdot \text{H}_2\text{O}$ ,

0.1  $\mu\text{M}$   $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ , 0.04  $\mu\text{M}$   $\text{H}_2\text{MoO}_4 \cdot \text{H}_2\text{O}$  ) for 100 liter solution in the tank. The nutrient solution was circulated in the tank in which the plants were growing and changed twice a week. The light intensity was  $615 \mu\text{mol m}^{-2}\text{s}^{-1}$  and a light cycle of 16 hrs light and 8 hrs dark was applied. The temperature was maintained at 25 °C.

The lines cultured were: PMT/K346-15, PMT/K346-27, PMT/K346-32, PMT/K346-36, PMT/K346-56; PMT/K346-60, QPT/K346-27, QPT/K346-30, QPT/K346-43, QPT/K346-50, QPT/K346-64, and UT/K346-4.

The plants were topped before flowering (at around 18-leaf stage). Root tissues were collected 6 hrs after topping and frozen in a -80 °C freezer for RNA analysis.

### **Quantification of total and individual alkaloids**

Total alkaloids were determined at Crop Science Tobacco Analytical Services Laboratory at NCSU. Samples of tobacco are dried, ground, and extracted prior to analysis of total alkaloids and reducing sugars using segmented-flow colorimetric methods developed for analysis of tobacco samples as adapted by Skalar Instrument Co (West Chester, PA).

Methods for these analyses are described by Collins et al. (1969) and Davis (1976).

The method used employed an acetic acid/methanol/water extraction and the use of charcoal for decolorization. Determination of total alkaloids was based on the reaction of cyanogen chloride with nicotine alkaloids in the presence of an aromatic amine to form a colored complex which is measured at 460 nm. For the determination of total reducing sugars (without inversion) the sample is dialyzed against a sodium carbonate solution. Copper neocuproin is added to the sample and the solution is heated. The copper neocuproin chelate is reduced in the presence of sugars resulting in a colored complex which is measured at 460 nm.

The analyses utilize an external standard method with a series of calibration standards corresponding to a range of alkaloid concentrations in dried tobacco samples of 0.25 - 6.00% and sugar concentrations of 1.5 – 30.0%. Data are collected on a Skalar Instruments SAN

Plus data system. Routine check samples are used to ensure proper function of the instrumentation and data quality. Our current methods have been amended in 2009 to more closely align with the Coresta (Cooperation Centre for Scientific Research Relative to Tobacco) Recommended Methods.

### **Statistical analysis**

Statistical analysis was performed by using SAS 9.1 software (SAS Institute Inc., Cary, NC).

## RESULTS

Two constructs, pTobRD2-*PMT*Tox and pTobRD2-*QPT*Tox, which over-express *PMT* and *QPT* genes, respectively, under the *QPT2* promoter were used in an attempt to increase the nicotine levels in tobacco. Two genotypes, K346 and NCTG61, were used to generate transgenic haploid tobacco plants. These lines were chosen because K346 is a popular commercial cultivar with good resistance to black shank disease and an intermediate level of TA, and NCTG61 is a widely used parental line for producing hybrid tobacco seeds and has a high level of TA. Plant transformation was mediated using *Agrobacterium tumefaciens*. Leaf midveins collected from transformed haploid plants were cultured on selection medium to generate double haploid plants for recovery of uniform homozygous transgenic plants that would not segregate for transgene copy number in offspring generations. Since both *PMT* and *QPT* are endogenous tobacco genes, PCR primers were designed to amplify the only transgenic versions of these genes. PCR was performed as a preliminary screen for putative double haploid plants. A total of 101 PCR-positive double haploid lines (DH, T<sub>0</sub>) were obtained from the transformation experiments, and their progenies were tested in the field in 2007.

### **Preliminary field test in 2007**

To accelerate the process for identifying transgenic plants with higher TA, a preliminary screening experiment was carried out in the fall of 2007 for T<sub>1</sub> plants of the above 101 transgenic lines. Although the late season planting was not preferred for tobacco growth and the plants had to be topped at an earlier-than-usual stage (5 leaves) to allow the alkaloids to accumulate before the first frost, we believe the effects of transgene over-expression on TA increase could still be revealed. It was decided to select 20% of the lines displaying the highest TA in this test for future analysis. TA was calculated as the percentage of TA (PTA) on leaf dry weight basis.

Figure 1.3 shows PTA values of 30 *PMT*Tox and 27 *QPT*Tox lines on K346 background while Figure 1.4 displays those values of 21 *PMT*Tox and 23 *QPT*Tox lines on NCTG61 background.

The PTA of two *PMTox* lines, PMT/K346-32 and PMT/NCTG61-53 was significantly reduced by 5- and 4-fold, respectively, caused by a phenomenon called “co-suppression” (Napoli et al. 1990).

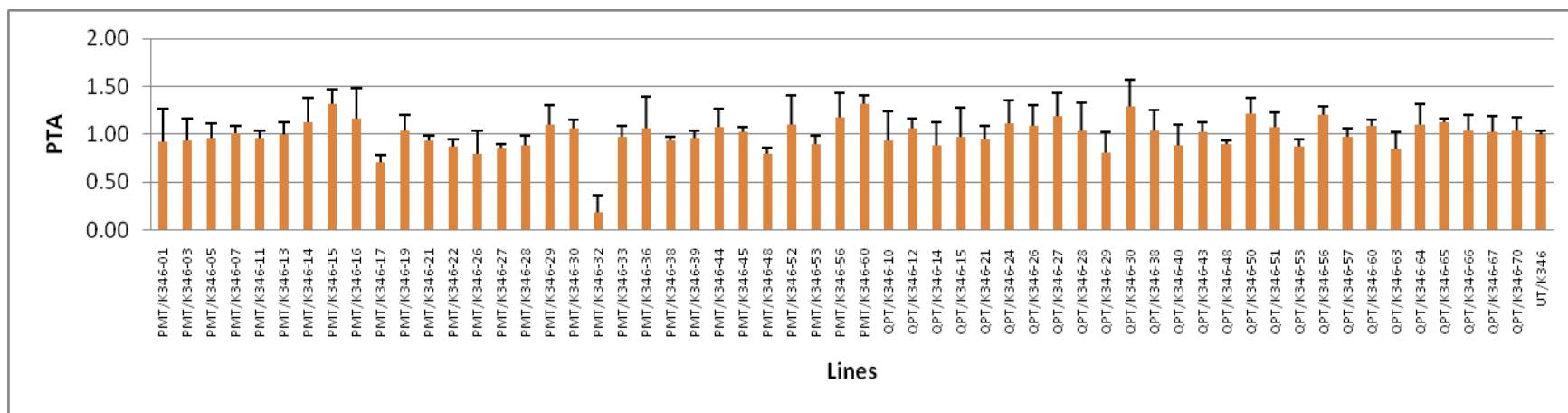


Figure 1.3 PTA of transgenic lines on K346 background in 2007 (Kinston, NC). A total of 30 *PMT*ox lines and 27 *QPT*ox lines were tested. UT/K346 was the non-transformed control. The mean PTA of each line and the standard error are shown.

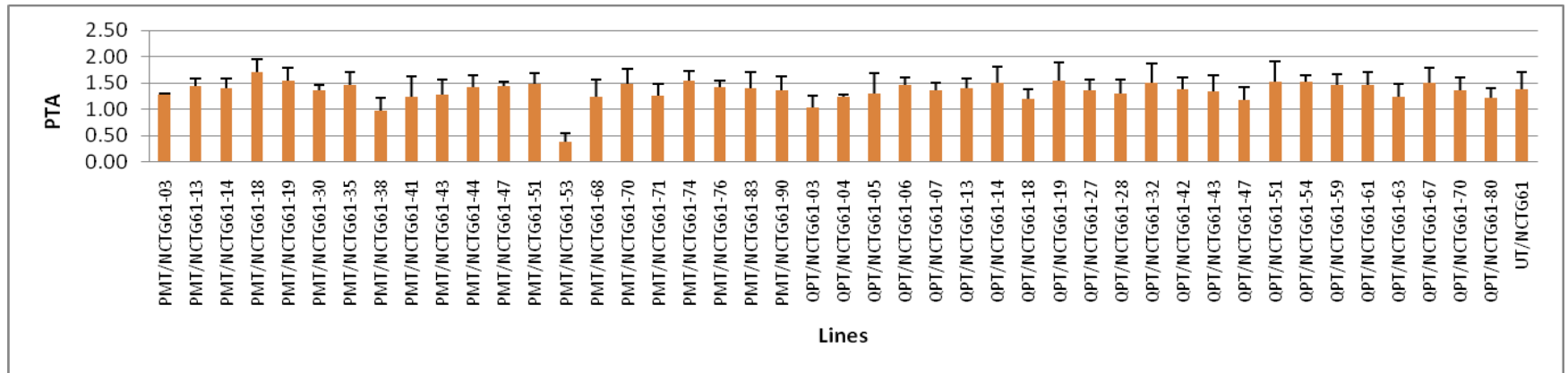


Figure 1.4 PTA of transgenic lines on NCTG61 background in 2007 (Kinston, NC). A total of 21 *PMT*ox lines and 23 *OPT*ox lines were tested. UT/NCTG61 was the non-transformed control. The mean PTA of each line and the standard error are shown.

### Southern analysis of the transgenic plants

The lines with relatively high alkaloid content in the 2007 field test were examined by Southern blot analysis to further confirm that they were indeed transgenic, and to estimate the transgene copy number. *HPT* and *NPTII* genes, selectable markers in the constructs, were used as probes for *PMT*ox and *QPT*ox transgenic plants, respectively.

As shown in Figure 1.5, all of the lines examined in this experiment have the *HPT* transgene. Considering that *Hind*III cleaves only once within the transgene construct, the observation that all the transgenic lines have at least two bands suggest that all plants contain more than one transgene copy. Some plants, such as PMT/K346-32, could have more than 5 copies of the transgene.

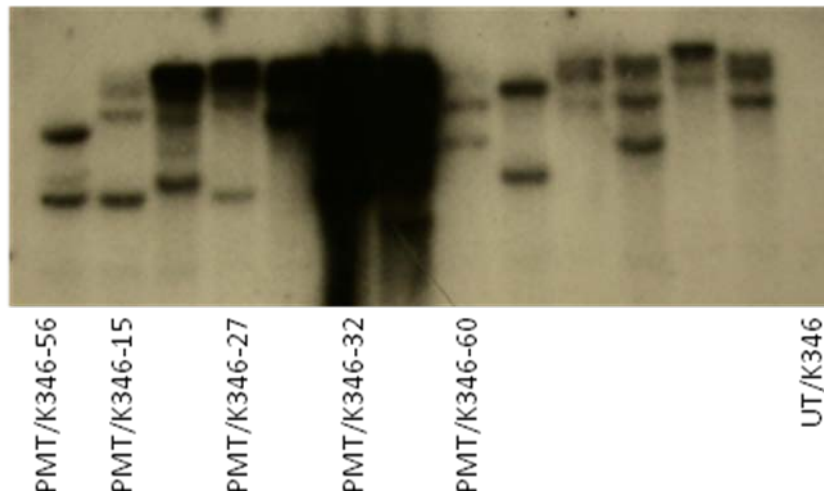


Figure 1.5 Southern blot analysis of *PMT* over-expression transformed plants. Twenty  $\mu$ g genomic DNA from leaf were digested with *Hind*III and hybridized with a  $^{32}$ P –labeled hygromycin phosphotransferase gene (*HPT*) probe. Four *PMT*ox lines, which were chosen for further tests, a co-suppression line and an untransformed K346 plant (UT/K346) are marked.

These transgenic lines were also examined for the presence of the *PMT* transgene. Figure 1.6 shows wild type *N. tabacum* cv. K346 had 5 hybridizing bands. This was in accordance with the previous observation (Riechers and Timko 1999) of 5 hybridization bands in the genomic DNA blot of *N. tabacum* cv. Xanthi. *PMT* is considered to have five members in its gene family. Based on the hybridization patterns and the intensity of the signals, most of the transgenic plants examined in this blot have at least one *PMT* transgene copy. The five *PMT*/K346 transgenic lines that were chosen for further studies based on the Southern and TA results are indicated in the two figures.

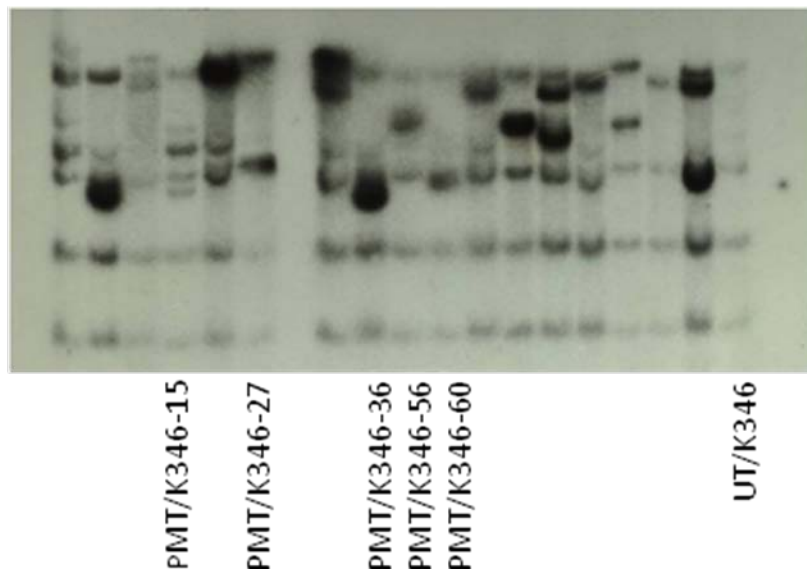


Figure 1.6 Southern blot analysis of *PMTox* transgenic lines. Genomic DNA of *PMT*/K346 plants were digested with *Xba*I and hybridized with  $^{32}\text{P}$  –labeled *PMT* gene probe. UT/K346: untransformed K346 plants as a control.

Southern blot analysis was also performed on select PCR-positive *QPTox* lines with *HPT* as well as *QPT* coding regions as probes. Figure 1.7 shows genomic DNA of *QPT*/K346 lines digested with *Hind*III and probed with a *HPT* coding sequence. This blot reveals that most of the tested plant lines are transgenic with one to three hybridization signal bands. Figure 1.8

shows that there were four major bands for *QPT* gene in the untransformed NCTG61 plants. This hybridization pattern is different from the observation of Sinclair et al. (2000) who found only two bands in the Southern analysis of the *QPT* gene in *N. tabacum* cv. NC95 with *Hind*III digestion. The discrepancy is most likely caused by the restriction enzymes used: *QPT* genes may have two more *Xba*I sites within the genes than *Hind*III sites. The five transgenic QPT/K346 lines chosen for further studies are marked in Figure 1.8 (and four of them are marked in Figure 1.7).

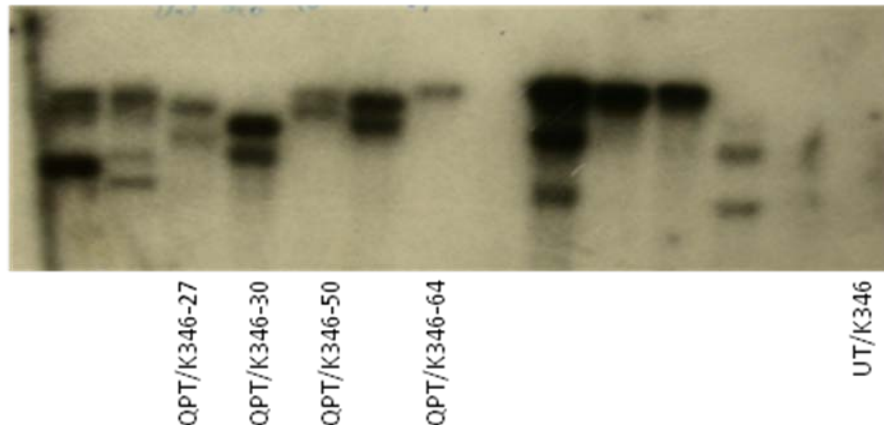


Figure 1.7 Southern blot analysis of *QPT*ox transformed plants. Genomic DNA from QPT/K346 lines was digested with *Hind*III and the blot was hybridized with probes of <sup>32</sup>P – labeled hygromycin phosphotransferase (*HPT*) coding sequences. Four lines of *QPT* over-expression were indicated. UT/K346: untransformed K346 plants as a negative control.

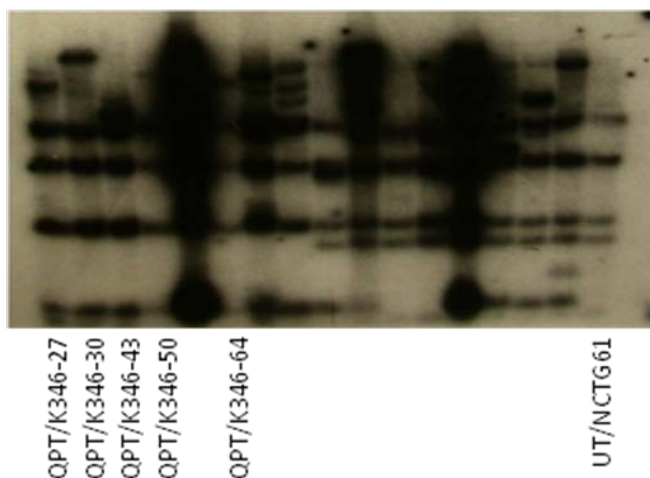


Figure 1.8 Southern blot analysis of *QPT*ox lines. Genomic DNA of QPT/K346 lines was digested with *Xba*I and probed with *QPT* coding sequence. UT/NCTG61: untransformed control of NCTG61.

#### **Parental lines and F1 crossing lines used in field tests in 2008 and 2009**

We hypothesized that transgenic plants that over-expressed both *PMT* and *QPT* genes may have even higher TA. To obtain such plants, we crossed the *PMT*ox lines with the *QPT*ox lines of the same cultivar. Parental lines with relative high TA in 2007 field tests and low transgene copy number were chosen. Based on these criteria, five PMT/K346, five QPT/K346, five PMT/ NCTG61, and three QPT/NCTG61 lines were chosen as parents for the crossings. Tables 1.1 and 1.2 show the lines used for the crossings and the names of the derived F1 lines. There were 23 F1 lines on K346 background and 15 such lines on NCTG61 background.

<b>PMT</b> QPT	PMT/K346-15	PMT/K346-27	PMT/K346-56	PMT/K346-60	PMT/K346-36
QPT/K346-64	PQKC-11	PQKC-12	PQKC-14	PQKC-15	PQKC-19
QPT/K346-30	PQKC-21	PQKC-22	PQKC-24	PQKC-25	PQKC-29
QPT/K346-50	PQKC-31	PQKC-32	PQKC-34	PQKC-35	PQKC-39
QPT/K346-27	PQKC-41	PQKC-42	PQKC-44	PQKC-45	PQKC-49
QPT/K346-43			PQKC-74	PQKC-75	PQKC-79

Table 1.1 Crosses between PMT/K346 and QPT/K346 over-expression lines

<b>PMT</b> QPT	PMT/NCTG61-18	PMT/NCTG61-19	PMT/NCTG61-70	PMT/NCTG61-51	PMT/NCTG61-83
QPT/NCTG61-67	PQNC-31	PQNC-32	PQNC-33	PQNC-35	PQNC-36
QPT/NCTG61-63	PQNC-71	PQNC-72	PQNC-73	PQNC-75	PQNC-76
QPT/NCTG61-04	PQNC-81	PQNC-82	PQNC-83	PQNC-85	PQNC-86

Table 1.2 Crosses between PMT/NCTG61 and QPT/NCTG61 over-expression lines

### Northern analysis of parental lines

Previous reports indicated that *QPT* gene expression is significantly enhanced 6 hrs after topping (Xie et al. 2004). To evaluate the levels of *PMT* and *QPT* transcript accumulation in transgenic plants, selected parental lines were grown hydroponically in a growth chamber. Root samples were collected 6 hrs after topping when the plants were about to flower.

Northern analysis (Figure 1.9) shows that PMT/K346-15 and PMT/K346-56 had substantially higher levels of *PMT* gene expression than the untransformed control plants.

PMT/K346-32 showed reduced level of *PMT* gene transcript, suggesting that it is a co-suppression line, most likely caused by the multiple transgene insertion events as shown in Figure 1.5. The low *PMT* expression level is correlated with the greatly reduced TA level of this transgenic line (Figure 1.3). For *QPTox* lines, there were no apparent increases in *QPT* gene expression among these lines (Figure 1.10).

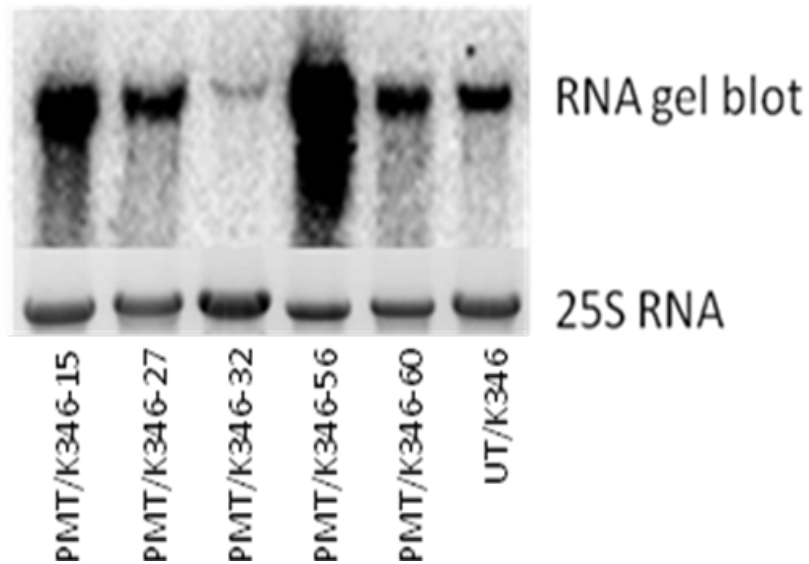


Figure 1.9 Northern blot analysis of *PMT* over-expression parental lines. Ten  $\mu\text{g}$  of RNA from roots of transgenic plants was separated on an 1% MOPS gel. The blot was hybridized with *PMT* gene probe. UT/K346 is an untransformed negative control.

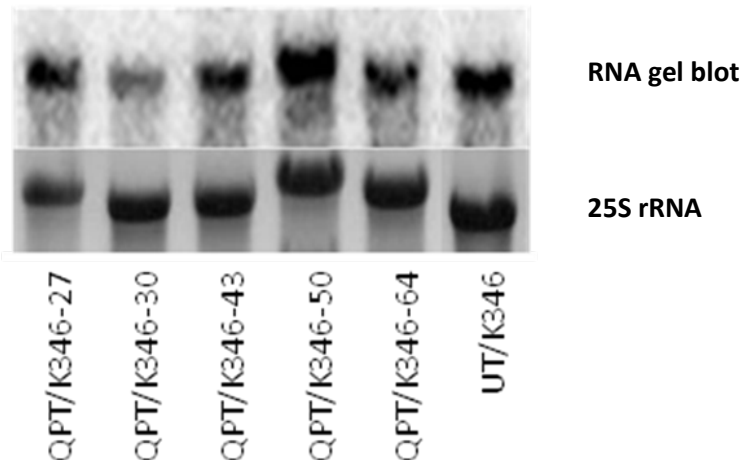


Figure 1.10 Northern blot analysis of *QPT* over-expression parental lines. Ten µg of RNA from roots of transgenic plants was separated on an 1% MOPS gel. The blot was hybridized with a *QPT* gene probe. UT/K346 is an untransformed negative control.

#### Field tests in 2008 and 2009

A total of 58 transgenic lines were tested in 2008 at both Rocky Mount and Oxford research stations. However, the experiment at Oxford was lost due to the decomposition of the leaves. Thus, only data from Rocky Mount is presented here and analyzed (Figures 1.11 and 1.12). PMT/K346-32 and PMT/NCTG61-53, the two co-suppression lines, had significantly lower TA content. The TA levels in these two lines were approximately 12 fold less than the control plants. The line with the highest mean increase in TA level was PQKC-45 with a TA level approximately 14% higher than the control.

In 2009, the field test of transgenic lines in the K346 background was carried out again at Oxford and Clayton research stations. Lines in the NCTG61 background were not tested because the field test data from 2008 failed to show any promising results for TA accumulation. Although the TA contents in the tests from Clayton were only about half of that observed from the Oxford test, probably reflecting the differences in soil, weather and

management between the two locations, the trends were very similar as shown in Figures 1.13 and 1.14: only the PMT/K346-32 showed greatly reduced TA levels (approximately 11% of the wild type).

Because the three field experiments of the transgenic lines in the K346 background were conducted in the same RCBD design in 2008 and 2009, these data were combined for ANOVA analysis. The co-suppression line PMT/K346-32 was not included due to its extreme low TA (11.5% of wild type), which might skew the results of the statistical analysis.

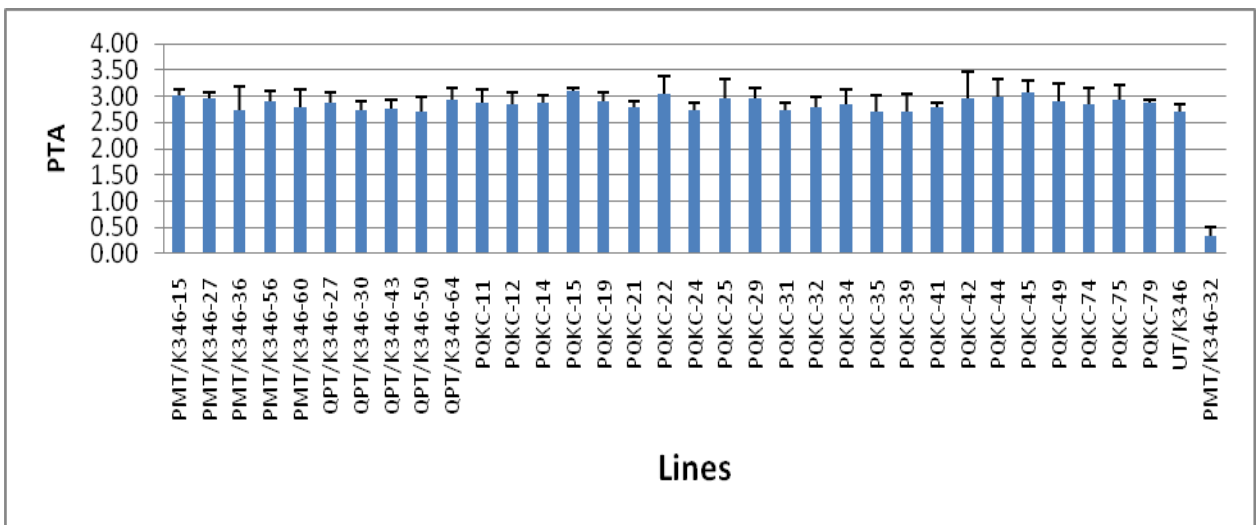


Figure 1.11 PTA levels of transgenic lines in the K346 background tested at Rocky Mount, NC, in 2008. A total of 35 lines were assayed. PQKC: F1 crossing lines. UT/K346: control (untransformed K346).

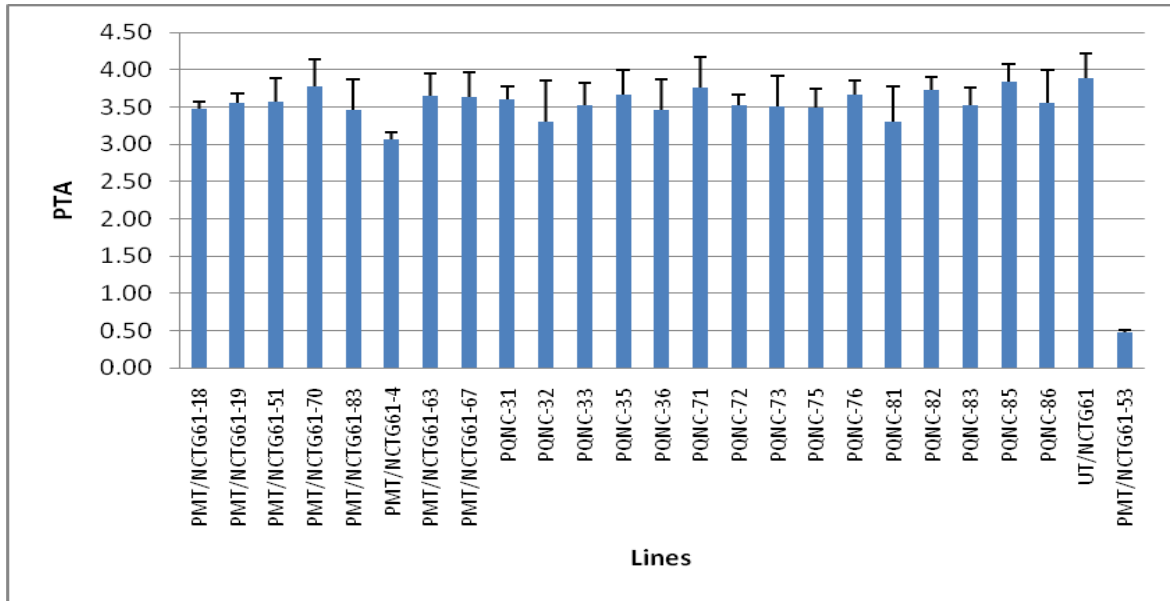


Figure 1.12 PTA levels of transgenic lines in the NCTG61 background tested at Rocky Mount, NC, in 2008. A total of 25 lines were tested. PQNC: F1 crossing lines. UT/NCTG61: control (untransformed NCTG61 plant).

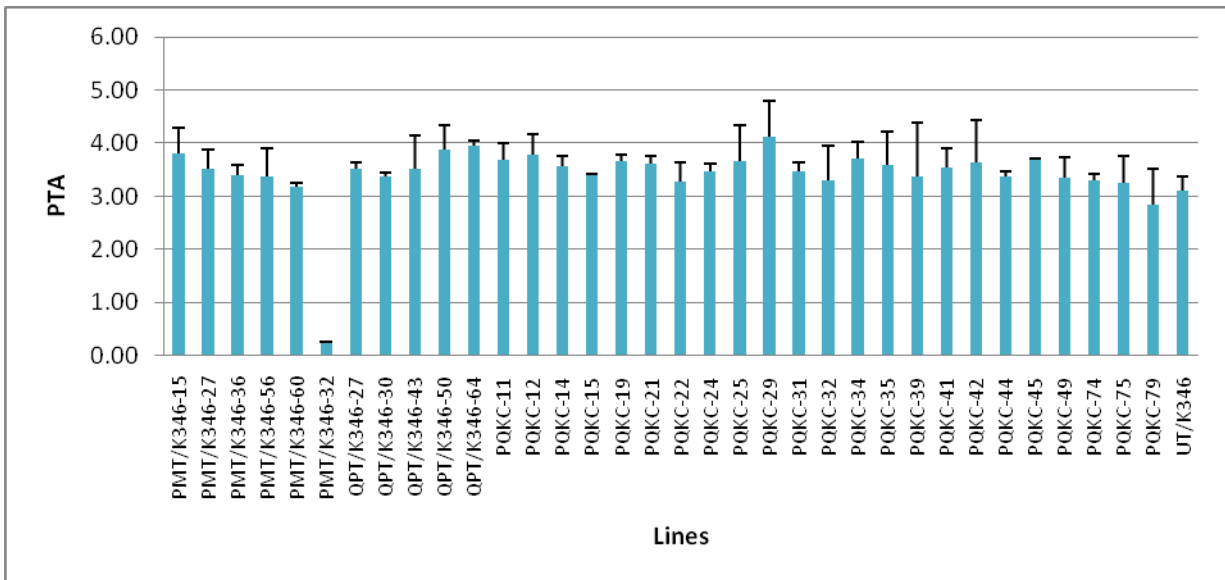


Figure 1.13 PTA of transgenic K346 lines and F1 crossing lines from the field test at Oxford, NC, in 2009.

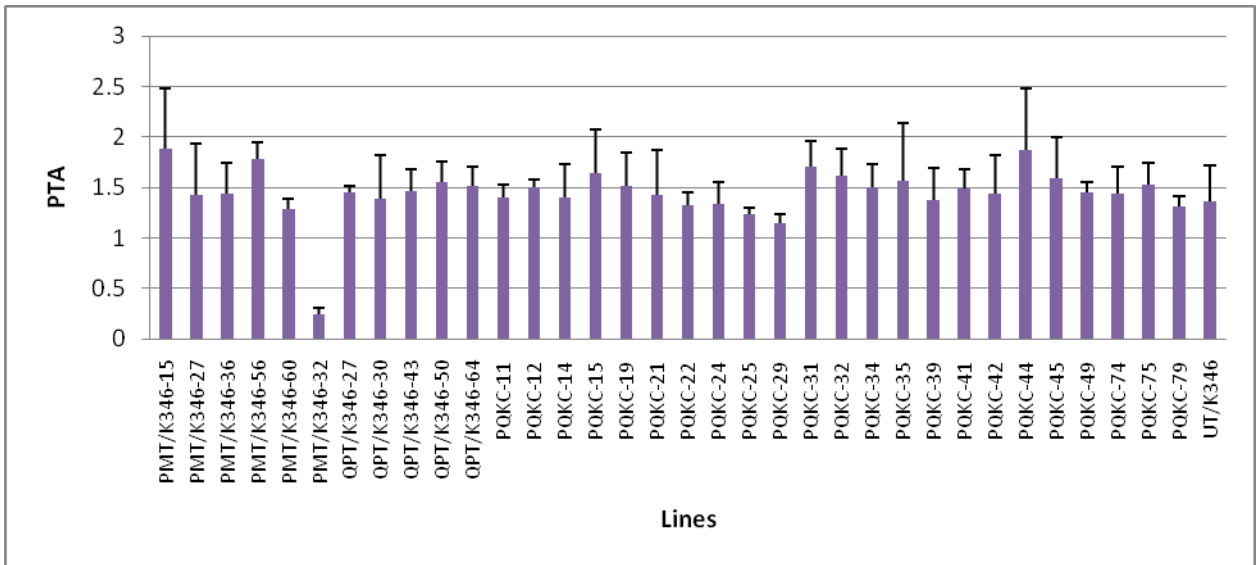


Figure 1.14 PTA values of transgenic K346 lines and F1 crossing lines in the field test at Clayton, NC, in 2009.

ANOVA Table 1 showed the results of this analysis: no significant difference in TA accumulation was observed among the parental transgenic and hybrid lines ( $P > 0.05$ ) compared to wild type control.

ANOVA Table 1:

Source of variation	DF	Sum of Squares	Mean Square	F value	Pr >F
Genotype	33	4.1639052	0.1261789	1.02	0.4484
Replicate	2	0.0049705	0.0024852	0.02	0.9802
Environment	2	219.0024084	109.5012042	881.93	<.0001
Error	268	33.2750839	0.1241608		
Total	305	256.4463679			

In another analysis only comparing the co-suppression line PMT/K346-32 and the control, significant difference of TA level ( $P < 0.0001$ ) was observed as indicated in ANOVA Table 2.

ANOVA Table 2:

Source of variation	DF	Sum of Squares	Mean Square	F value	Pr >F
Genotype	1	20.14959012	20.14959012	80.98	<.0001
Replicate	2	0.10789165	0.05394582	0.22	0.8082
Environment	2	2.60284369	1.30142185	5.23	0.0233
Error	12	2.98579427	0.24881619		
Total	17	25.84611972			

Meanwhile, in an attempt to reduce the effects of the non-uniformity of the soil and to increase the sensitivity of the statistical analysis, three small-scale experiments were designed in 2009 for selected lines that showed significantly lower, or relatively higher TA accumulation in the 2008 field tests. Plants of non-transformed K346, and sometimes the commercial cultivar K346 and low TA mutant line LAFC-53 were also included in these tests for comparison.

Each small-scale experiment had 5-7 entries and 5 replicates, and conducted at both Oxford and Clayton locations. PMT/K346-32 again displayed a much lower TA levels, comparable to the mutant line LAFC-53. Both had significantly lower TA levels when compared to the untransformed K346 control and the K346 plant from commercial seed source. Other lines did not show any significant change from controls in TA content (Figures 1.15-1.20).

Small experiment 1:

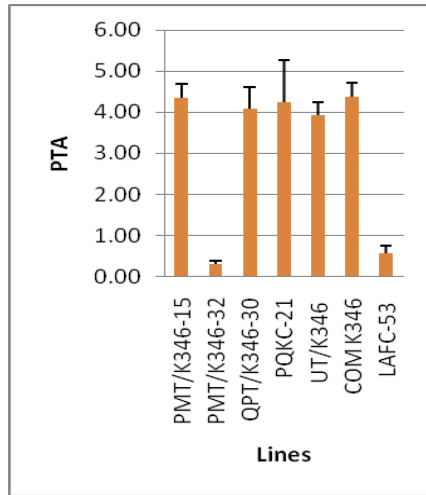


Figure 1.15 PTA levels of selected lines in small experiment 1 at Oxford. L AFC-53: low alkaloid mutant line; COMK346: commercial cultivar K346; UT/K346: untransformed negative control; PQKC-21: crossing line between PMT/K346-15 and QPT/K346 -30; PMT/K346-32: co-suppression line.

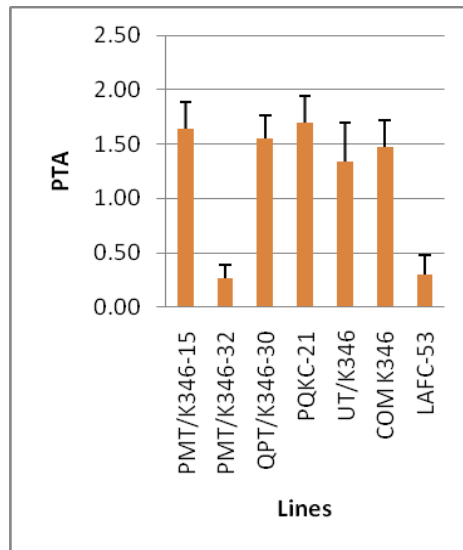


Figure 1.16 PTA levels of selected lines in small experiment 1 at Clayton. L AFC-53: low alkaloid mutant line; COMK346: commercial cultivar K346; UT/K346: untransformed negative control; PQKC-21: F1 crossing line between PMT/K346-15 and QPT/K346-30; PMT/K346-32: *PMT* co-suppression line.

Data from the same small experiment in two locations were also combined for ANOVA analysis. In small experiment 1, the subset data of PMT/K346-15, QPT/K346-30, their hybrid PQKC-21 and UT/K346, COM K346, and subset data of PMT/K346-32, LAFC-53 and UT/K346 were analyzed separately. ANOVA Table 3 ( $P > 0.05$ ) shows that no significant change of TA level was found among lines of PMT/K346-15, QPT/K346-30, their hybrid PQKC-21, UT/K346 and COM K346. The P value in ANOVA Table 4 indicates that there is significant difference in TA level among the lines of PMT/K346-32, LAFC-53 and UT/K346. LSD analysis further shows that co-suppression line PMT/K346-32 exhibited no difference with the mutant line LAFC-53, but significantly lower ( $P < 0.0001$ ) than the control line.

ANOVA Table 3:

Source of variation	DF	Sum of Squares	Mean Square	F value	Pr >F
Genotype	4	0.88630800	0.22157700	1.30	0.2864
Replicate	4	1.58896800	0.39724200	2.33	0.0724
Environment	1	88.31205000	88.31205000	518.18	<.0001
Error	40	6.81713200	0.17042830		
Total	49	97.60445800			

ANOVA Table 4:

Source of variation	DF	Sum of Squares	Mean Square	F value	Pr >F
Genotype	2	34.72928667	17.36464333	34.31	<.0001
Replicate	4	0.19928000	0.04982000	0.10	0.9818
Environment	1	6.96972000	6.96972000	13.77	0.0012
Error	22	11.13346000	0.50606636		
Total	29	53.03174667			

Small experiment 2:

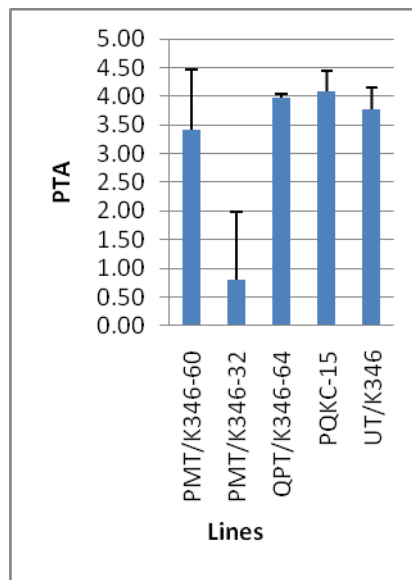


Figure 1.17 PTA levels of selected lines in small experiment 2 at Oxford. UT/K346: untransformed negative control; PQKC-15: F1 crossing line between PMT/K346-60 and QPT/K346-64; PMT/K346-32: *PMT* co-suppression line.

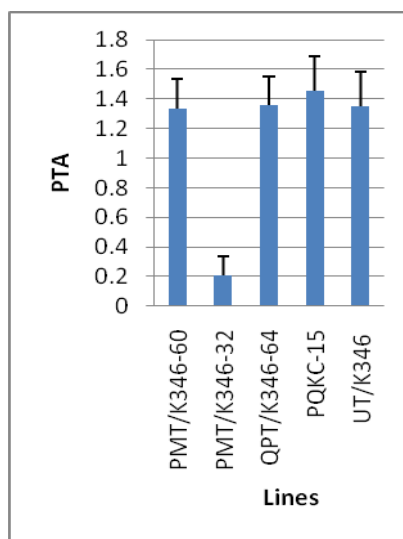


Figure 1.18 PTA levels of selected lines in small experiment 2 at Clayton. UT/K346: untransformed negative control; PQKC-15: crossing line between PMT/K346-60 and QPT/K346-64; PMT/K346-32: *PMT* co-suppression line.

The P value in ANOVA Table 5 shows no significant change of TA level was found among the parental, hybrid and control lines.

ANOVA Table 5:

Source of variation	DF	Sum of Squares	Mean Square	F value	Pr >F
Genotype	3	0.80009000	0.26669667	1.54	0.2245
Replicate	4	1.24571000	0.31142750	1.80	0.1550
Environment	1	59.53600000	59.53600000	343.17	<.0001
Error	31	5.37811000	0.17348742		
Total	39	66.95991000			

ANOVA Table 6 again shows a significant difference ( $P < 0.001$ ) in TA level between control line UT/K346 and the co-suppression line PMT/K346-32.

ANOVA Table 6:

Source of variation	DF	Sum of Squares	Mean Square	F value	Pr >F
Genotype	1	21.30048000	21.30048000	27.63	0.0002
Replicate	4	0.74357000	0.18589250	0.24	0.9099
Environment	1	11.34018000	11.34018000	14.71	0.0021
Error	13	10.02159000	0.77089154		
Total	19	43.40582000			

Small experiment 3:

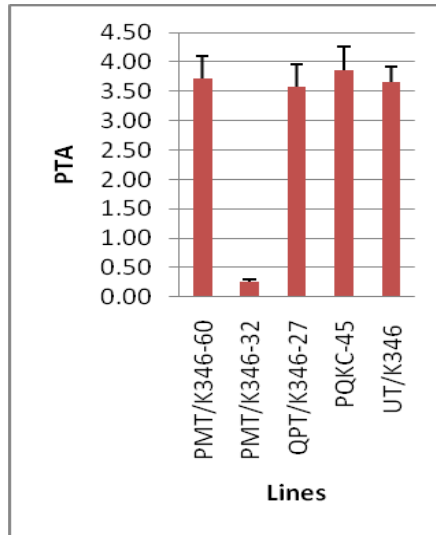


Figure 1.19 PTA levels of selected lines in small experiment 3 at Oxford. UT/K346: untransformed negative control; PQKC-45: F1 crossing line between PMT/K346-60 and QPT/K346-27; PMT/K346-32: *PMT* co-suppression line.

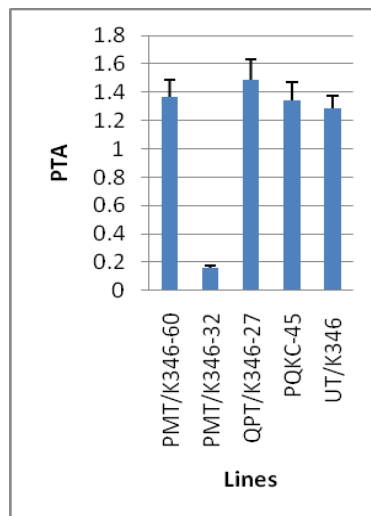


Figure 1.20 PTA levels of selected lines in small experiment 3 at Clayton. UT/K346: untransformed negative control; PQKC-45: crossing line between PMT/K346-60 and QPT/K346 -27; PMT/K346-32: *PMT* co-suppression line.

Lines of PMT/K346-60, QPT/K346-30 and their hybrid PQKC-45 did not exhibit any significant change in TA level compared to the control line in this experiment as indicated in the ANOVA Table 7.

ANOVA Table 7:

Source of variation	DF	Sum of Squares	Mean Square	F value	Pr >F
Genotype	3	0.08136750	0.02712250	0.43	0.7358
Replicate	4	0.54586000	0.13646500	2.14	0.0991
Environment	1	54.59232250	54.59232250	857.22	<.0001
Error	31	1.97424750	0.06368540		
Total	39	57.19379750			

ANOVA Table 8 shows again a significant difference ( $P < 0.0001$ ) in TA level between the co-suppression line PMT/K346-32 and the control line UT/K346.

ANOVA Table 8:

Source of variation	DF	Sum of Squares	Mean Square	F value	Pr >F
Genotype	1	25.67378000	25.67378000	49.96	<.0001
Replicate	4	0.08103000	0.02025750	0.04	0.9967
Environment	1	7.66322000	7.66322000	14.91	0.0020
Error	13	6.68115000	0.51393462		
Total	19	40.09918000			

## Summary

In this study, *PMT* over-expression was observed in two of the five lines examined. The *PMT* mRNA level of line PMT/K346-15 was about 4 fold higher than the control plant in roots 6 hrs after topping (Figure 1.9). This line showed about a 30% increase in TA level in the 2007 field test, 12% increases at Rocky Mount in 2008 (Figure 1.11), and 22% and 38% increases at Oxford and Clayton, respectively, in 2009 (Figures 1.13, 1.14). In small experiment 1, this line shows 23% increase of TA level in Clayton. However, the increase in TA of this line was not statistically significant in any of these tests, even when in the “small” experiment with five replicates. Moreover, the TA of the F1 crossing lines, PQKC-11, 21, 31, and 41, with PMT/K346-15 as a parent were not increased significantly in the large and small experiments (Figures 1.11, 1.13, 1.14, 1.15, 1.16) either. Another *PMT*ox line, PMT/K346-56, over-expressed *PMT* by about 8 fold (Figure 1.9). However, the TA levels of this line and the hybrids with this line as a parent (PQKC14, 24, 34, 44, and 74) never stood out in the tests covering three locations and two years (Figures 1.11, 1.13, 1.14). Among the five *QPT*ox lines examined by Northern blot analysis, only a slight increase (no more than 50%) of mRNA was observed after topping when compared to the wild type control. None of these lines or their F1 offspring lines had significantly higher TA levels.

## DISCUSSION

### ***Nicotiana africana* and midvein culture technique for fast recovery of homozygous transgenic plants**

The *Nicotiana africana* technique (Burk et al. 1979) used in the transformation of tobacco plants allowed easy recovery of transgenic haploid plants. After being pollinated with pollen grains from *Nicotiana africana*, most survived plants are haploids. By using leaf midvein culture of transgenic haploid plants, double haploid plants were generated. One advantage of using midvein culturing is that all the double haploid plants are homozygous transgenic plants without further segregation. This greatly shortens the time to obtain homozygous transgenic plants, which otherwise requires a long time of self-pollination and segregation to recover. In this study, around 50% of the plants generated from midvein culture were double haploids. Leaf midveins were collected just before flowering. The developmental stage of the midvein may be critical in generating double haploid plants. Young midvein from top leaves and very old ones from bottom leaf were not recommended for double haploid production. The lamina of leaf might be the most undesirable above-ground tissue for generation of double haploid plants. In this study, only one out of about 360 plants derived from leaf lamina culture was true double haploid plant. In my experience, the midvein from the leaves next to the bottom of the plants was superior in generating double haploid plants.

### **Multiple transgene copies**

Southern blotting data showed that most of the transgenic plants had multiple transgene copies. This result was unexpected because *Agrobacterium*-mediated transformation usually results in low copy numbers of the transgene. In this study, only one transgenic line shows one band signal in Southern analysis. The reason why this happened might be because the concentration of antibiotic (kanamycin) in the selection medium was high (100 mg/L in this study) so that only transgenic plants with multiple transgenes (including selectable marker gene which provides the resistance to the antibiotic) were selected.

## **Parental line selection and field tests**

Field test is the core value of this project. Nicotine increase by over-expression of *PMT* gene has been reported (Sato et al. 2001; Hashimoto and Kajikawa 2008), but all the research was conducted under greenhouse conditions with limited number of pot-grown plants. Whether similar results can be obtained in field-grown plants is quite important in determining the value of the approach in future applications. To achieve this, large scale field experiments are necessary.

In our approach, instead of using construct containing both *PMT* and *QPT* transgenes, we attempted to over-express both genes in F1 hybrids through crossing between selected transgenic double haploid lines. We expected a variety of nicotine levels by combining different expression levels of these two genes.

To accelerate the process, the preliminary screening experiment to select certain lines for crossing was carried out in 2007 at Kinston from September to November which was not the optimum growth season for tobacco, and the plants grew only about 48 days compared to the normal growth period of 140 days for tobacco during growth season. However, the data obtained still provided meaningful information on the effect of the transgenes. Based on the total alkaloid data and the Southern analysis, we selected lines with relatively high total alkaloid levels and fewer transgene copy number to serve as parental lines. Plants with fewer transgene copy number are preferred because the transgene expression may be less prone to gene silencing. In both 2008 and 2009, the experiments were conducted in two locations in North Carolina. The field test in 2009 was narrowed down to only those transgenic lines on K346 background because the transgenic lines on NCTG61 background showed little changes on total alkaloid level, probably due to the cultivar's already high TA content. In 2009, additional small-scale experiments were performed to include only a few promising lines and to increase replicates in an attempt to increase the sensitivity in detection of the difference between transgenic plants and the control plants, which may be masked in the large-scale field test mostly due to the non-uniformity of the soil conditions.

### **Alteration of total alkaloid through *PMT* and *QPT* over-expression**

Our data show that *PMT* mRNA levels can be enhanced by over-expressing *PMT* as a transgene in *N. tabacum* using the *QPT2* promoter after topping. However, a substantial increase in *PMT* mRNA, as observed in PMT/K346-15 and -56, had little effect on TA accumulation. Furthermore, we examined five *QPTox* lines using its own promoter, and none of them had substantial increases in *QPT* mRNA accumulation after topping (less than 50%).

In the literature, there are no reports describing an attempt to increase nicotine levels by over-expression of a *QPT* gene. In a publication reporting *PMT* over-expression, the authors used the CaMV 35S promoter to over-express the *PMT* gene in *N. sylvestris* plants. While the *PMT* mRNA level was raised by four-eight folds, a 40% increase in nicotine content was observed (Sato et al. 2001). However, when their results were carefully examined, we found that the differences between their systems and ours were not only the promoter and the species employed in the experiments (*N. sylvestris* is a diploid and considered a progenitor of *N. tabacum*, which is an allotetraploid). Apparently, topping or other ways to induce nicotine synthesis was not utilized in that report. The basal nicotine content in their non-transgenic plants was extremely low, ~ 0.06% (3.6 n mol/mg) of the dry leaf weight, and only ~0.09% (5.6 n mol/mg) in the transgenics, approximately 20-40 fold lower than the TA levels of the field-grown *N. tabacum* plants observed in our experiments. Whether the increase margin (40%) observed in their experiments would be observed in *N. tabacum* which has much higher constitutive nicotine level, and/or after topping (which they did not perform) is very questionable. Moreover, their experiments were performed in a greenhouse and ours was conducted in the field which typically yields higher alkaloid accumulation. Thus, the results from the two experiments may not be directly comparable, and the increase they observed may not translate to a similar increase in field-grown plants of *N. tabacum*.

Although no significant TA increases were observed among the *PMTox* and *QPTox* lines, we did identify two *PMTox* lines (PMT/K346-32 and PMT/NCTG61-53) that had much lower TA levels than the control. Northern blot analysis of one such line, PMT/K346-32, suggested

that they are co-suppression lines. Southern analysis of this line (when probed with the *HPT* transgene within the construct) indicated that it might have more than 5 transgene copies. The high transgene copy number, and their configuration in the genome, could be the reason for the observed co-suppression phenotype. The *PMT* co-suppression lines reveal that when the *PMT* mRNA levels were down-regulated, the TA content was reduced substantially. Although we did not observe co-suppression among the *QPTox* lines, there is a report that reduced *QPT* mRNA and TA lines were obtained when an anti-sense *QPT* gene was introduced into tobacco (Xie et al., 2004). In our experiments, the TA level of line *PMT/K346-32* was stably low and only 10-15% of that of the control plants in all the 2008 and 2009 field tests.

Our results on over-expression and co-suppression of *PMT* and *QPT* genes suggest a tight and complicated regulation of the alkaloid biosynthetic pathway in tobacco. In our experiments, both constructs used the *QPT2* promoter to drive over-expression of the *PMT* or *QPT* transgene, but the results appeared to be different. We did see substantially increased accumulation of *PMT* mRNA and believe that the *QPT2* promoter functions in tobacco and could transcribe a gene it drives to a high level. Although the number of the transgenic lines examined was limited, the fact that the *QPT* mRNA did not increase to the level as high as the *PMT*'s, when similar constructs were used, seems to indicate a possible control point of *QPT* gene expression regulation at the RNA level. This point is further reinforced by our observation in Chapter 2.

The published reports (Chintapakorn and Hamill 2003; Xie et al. 2004) and our results show that suppression of *PMT* or *QPT* expression resulted in dramatically reduced nicotine and TA accumulation. However, our experiments strongly suggest that over-expression of *PMT* or/and *QPT* does not lead to significant increases in TA levels in transgenic tobacco plants even when the mRNA is substantially increased as in the case of *PMT*. While we believe *QPT* expression is more regulated at the RNA level, we propose two hypotheses to explain what happened in the *PMT* over-expression lines. The first hypothesis is that the main limiting step in nicotine biosynthesis may not occur at the steps catalyzed by *PMT* or *QPT*,

but may instead occur at the last condensation step of the pathway catalyzed by nicotine synthase. In this case, over-expression of *PMT* mRNA and the increased enzyme activities would not lead to higher nicotine and TA accumulation, and may result in over-accumulation of methylputrescine and/or other intermediates. This hypothesis is experimentally testable by measuring the *PMT* protein accumulation or its enzyme activity as well as the accumulation of methylputrescine and other intermediates under certain conditions. The second hypothesis proposes the possibility that there is a tight control either at the translational level or at the post-translational level. Thus, although *PMT* mRNA accumulation is increased in the over-expression lines as shown by Northern blot analysis, the *PMT* enzyme activity may not be, either by reduced protein synthesis, enhanced protein turnover, or by a feedback inhibition to the *PMT* enzyme activity by a downstream product in this pathway. This hypothesis is also experimentally testable by measuring the *PMT* enzyme activities in roots of both transgenic and control plants. Our transcription factor studies seem to support a tight control mechanism at the translational or post-translational level (please see Chapter 2).

Overall, our results, using over 100 transgenic plants in three year's field tests, suggest that over-expressing *PMT* and/or *QPT* genes does not appear to be an efficient way to elevate nicotine and TA levels in tobacco.

## REFERENCES

- Bortolotti C, Cordeiro A, Alcazar R, Borrell A, Culiañez-Macià<sup>2</sup> FA, Tiburcio AF, Altabella T(2004). Localization of arginine decarboxylase in tobacco plants. *Physiol Plant* 120:84–92.
- Brown SA, Byerrum RU (1952).The origin of the methyl carbon of nicotine formed by *Nicotiana rustica* L. *J Amer Chem Soc* 74(6): 1523-1526.
- Burk LG, Gerstel DU, Wernsman EA (1979). Maternal haploids of *Nicotiana tabacum* L. from seed. *Science* 206: 585.
- Burton LL, Dewey LJ, Byerrum RU (1958). Glutamate as a Precursor for the Pyrrolidine Ring of Nicotine. *The Journal of Biological Chemistry* 233(4): 939-943.
- Burton LL, Dewey LJ, Byerrum RU (1959). Ornithine as a Precursor for the Pyrrolidine Ring of Nicotine. *Biochimica et Biophysica Acta* 33: 22-26.
- Bush L, Hempfling WP, Burton H (1999). Biosynthesis of nicotine and related compounds. Analytical determination of nicotine and related compounds and their metabolites. Chapter 2:13-44.
- Cane K, Mayer M, Lidgett A, Michael A, Hamill J (2005). Molecular analysis of alkaloid metabolism in AABB v.aabb genotype *Nicotiana tabacum* in response to wounding of aerial tissues and methyljasmonate treatment of cultured root. *Functional Plant Biology* 32: 305-320.
- Chaplin JF, Weeks WW (1976). Association between percent total alkaloids and other traits in flue-cured tobacco.*Crop Sci* 16: 416-418.
- Chaplin JF and Burk LG (1983). Agronomic, Chemical, and Smoke Characteristics of Flue-cured Tobacco Lines with Different Levels of Total Alkaloids *Agronomy Journal* 75: 133-136

Chintapakorn Y, Hamill JD (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 Mol Biol* 53: 87-105.

Chintapakorn Y, Hamill JD (2007). Antisense-mediated reduction in ADC activity causes minor alterations in the alkaloid profile of cultured hairy roots and regenerated transgenic plants of *Nicotiana tabacum*. *Phytochemistry* 68: 2465-2479.

Collins PF, Sarji NM, Williams JF (1969). Determination of Nicotine Alkaloids in Tobacco Using the Autoanalyzer. *Tob. Sci.* 13:79-81

Davis RE (1976). A combined automated procedure for the determination of reducing sugar and nicotine alkaloids in tobacco products. *Tobacco Science* 20:139-144.

Dawson RF (1941). The localization of the nicotine synthetic mechanism in the tobacco plant. *Science* 94 (2443):396-397.

Dawson RF (1942). Nicotine synthesis in excised tobacco root. *American Journal of Botany* 29(10):813-815.

Dawson RF, Christman DR, D'Adamo A, Solt ML, Wolf AP (1960). The biosynthesis of nicotine from isotopically labeled nicotinic acids. *J Amer Chem Soc* 82(10): 2628-2633.

DeBoer KD, Lye JC, Aitken CD, Su AK-K, Hamill JD (2009). The A622 gene in *Nicotiana glauca* (Tree tobacco): evidence for a functional role in pyridine alkaloid synthesis. *Plant Mol. Biol.* 69:299-312.

Feth F, Wagner R, Wagner KG (1986). Regulation in tobacco callus of enzyme activities of the nicotine pathway. *Planta* 168: 402-407.

Friesen JB, Leete E (1990). Nicotine synthase -- an enzyme from *Nicotiana* species which catalyzes the formation of (S)-Nicotine from nicotinic acid and 1-methyl-pyrrolinium chloride. *Tetrahedron Letters* 31(44): 6295-6298.

Frost GM, Yang KS, Waller GR (1967). Nicotinamide adenine dinucleotide as a precursor of nicotine in *Nicotiana rustica* L. The Journal of Biological Chemistry 242(No.5): 887-888.

Hamill JD, Robins RJ, Parr AJ, Evans DM, Furze JM, Rhodesn MJC (1990). Over-expression of a yeast ornithine decarboxylase gene in roots of transgenic *Nicotiana rustica* can lead to enhanced nicotine accumulation. Plant Mol Biol **15**: 27-38.

Hashimoto T, Kato A (2007). Reducing levels of nicotinic alkaloids in plants. United States Patent Application US 2007/0240728 A1.

Hashimoto T, Kajikawa M (2008). Increasing levels of nicotinic alkaloids in plants. United States Patent Application US 2008/ 0120737A1.

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

Jiang K, Pi Y, Hou R, Jiang L, Sun X, Tang K (2009). Promotion of nicotine biosynthesis in transgenic tobacco by overexpressing allene oxide cyclase from *Hyoscyamus niger*. Planta 229: 1057-1063.

Katoh A, Ohki H, Inai K, Hashimoto T (2005). Molecular regulation of nicotine biosynthesis. Plant Biotechnology 22: 389-392.

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

Lamberts BL, Byerrum, RV (1958). Glutamate as a precursor for the pyrrolidine ring of nicotine. J. Biol. Chem. 233: 933-939.

Lamberts BL, Dewey LJ, Byerrum RV (1959). Ornithine as a precursor for the pyrrolidine ring of nicotine. Biochim. Biophys Acta 33: 22-26.

Leete E, Gros EG, Gibertson TJ (1964). Biosynthesis of the Pyrrolidine Ring of Nicotine: Feeding Experiments with N<sup>15</sup>- Labeled Ornithine -2-C<sup>14</sup>. Tetrahedron Letters 5(11): 587-592.

Leete E (1967). Biosynthesis of the Nicotiana alkaloids XI . Investigation of tautomerism in N-methyl-pyrrolinium chloride and its incorporation into nicotine. J Amer Chem Soc 89: 7081-7084.

Legg P D, Chaplin JF, Collins GB (1969). Inheritance of percent total alkaloids in *Nicotiana tabacum* L. J. Hered 60: 213-217.

Legg PD, Collins GB, Litton CC (1970). Registration of LA Burley 21 tobacco germplasm. Crop Sci 10: 212.

Legg PD, Collins GB (1971). Inheritance of percent total alkaloids in *Nicotiana tabacum* L. II. Genetic effects of two loci in Burley 21 x LA Burley 21 populations. Can. J. Genet. Cytol 13: 287-291.

Lewis RS, Jack AM, Morris JW, Robert VJ, Gavilano LB, Siminszky B, Bush LP, Hayes AJ, Dewey RE (2008). RNA interference (RNAi)-induced suppression of nicotine demethylase activity reduces levels of a key carcinogen in cured tobacco leaves. Plant Biotechnology Journal 6: 346-354.

Liu YY, Leete E (1973). Metabolism of [2-3H]- and [6-3H]-nicotinic acid in intact *Nicotiana tabacum* plants. Phytochemistry 12: 593-596.

Mann TJ, Weybrew JA, Matzinger DF, Hall JL (1964). Inheritance of the conversion of nicotine to nornicotine in varieties of *Nicotiana tabacum* L. and related amphiploids. Crop Sci. 4:349-353.

Matzinger DF, Mann TJ, Cockerham CC (1972). Recurrent family selection and correlated response in *Nicotiana tabacum* L. I. 'Dixie Bright 244' x 'Coker 139'. Crop Sci. 12: 40-43.

- Matzinger DF, Wernsman EA, Weeks WW (1989). Restricted index selection for total alkaloids and yield in tobacco. *Crop Sci.* 29: 74-77.
- Mizusaki S, Kisaki T, Tamaki E (1967). Phytochemical studies on the tobacco alkaloids XII. Identification of r-methylaminobutyraldehyde and its precursor role in nicotine biosynthesis. *Plant Physiol.* 43: 93-98.
- Mizusaki S, Tanabi Y, Noguchi M, Tamaki E (1971). Phytochemical studies on tobacco alkaloids XIV. The occurrence and properties of putrescine N-methyltransferase in tobacco roots. *Plant & Cell Physiol.* 12: 633-640.
- Mizusaki S, Tanabe Y, Noguchi M, Tamaki E (1972). N-methylputrescine oxidase from tobacco roots. *Phytochemistry* 11: 2757-2762.
- Mizusaki S, Tanabi 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 & Cell Physiol* 14: 103-110.
- Morita M, Shitan N, Sawada K, Van Montagu MCE, Inze D, Rischer H, Goossens A, Oksman-Caldentey K-M, Moriyama Y, Yazaki K (2009). Vacuolar transport of nicotine is mediated by a multidrug and toxic compound extrusion (MATE) transporter in *Nicotiana tabacum*. *Proc Natl Acad Sci (USA)* 106: 2447-2452.
- Napoli C, Lamieux C, Jorgensen R (1990). Introduction of a chimera chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *Plant Cell* 2: 279-289.
- Riechers DE, Timko MP (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 Mol Biol* 41: 387-401.
- Sambrook J, Russell DW (2001). *Molecular cloning, a laboratory manual* (3rd ed) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

- Sato F, Hashimoto T, Hachiya A, Tamura K, Choi K, Morishige T, Fujimoto H and Yamada Y (2001). Metabolic engineering of plant alkaloid biosynthesis. *Proc Natl Acad Sci (USA)* 98:367-372.
- Saunders JW, Bush LP (1979). Nicotine biosynthetic enzyme activities in *Nicotiana tabacum* L. genotypes with different alkaloid levels. *Plant Physiology* 64: 236-240.
- Schutte HE, Maier M, Mothes K (1966). Methylputrescine as a possible precursor of nicotine in *Nicotiana rustica*. *Acta Biochim. Polon.* 13: 401.
- Scott TA, Glynn JP (1967). The incorporation of [2,3,7-<sup>14</sup>C] nicotinic acid into nicotine by *Nicotiana tabacum*. *Phytochemistry* 6: 505-510.
- Shoji T, Yamada Y, Hashimoto T (2000). Jasmonate induction of putrescine N-methyltransferase genes in the roots of *Nicotiana glauca*. *Plant & Cell Physiol* 41: 831-839.
- Shoji T, Winz R, Iwase T, Nakajima K, Yamada Y, and Hashimoto T (2002). Expression Patterns of two isoflavone reductase-like genes and their possible roles in secondary metabolism in tobacco. *Plant Mol Biol* 50: 427-440.
- Shoji T, Kajikawa M, Hashimoto T (2010). Clustered transcription factor genes regulate nicotine biosynthesis in tobacco. *The Plant Cell* 22: 3390-3409.
- Siminszky B, Gavilano L, Bowen SW, Dewey RE (2005). Conversion of nicotine to nornicotine in *Nicotiana glauca* is mediated by CYP82E4, a cytochrome P450 monooxygenase. *Proc Natl Acad Sci (USA)* 102: 14919-14924.
- Sinclair SJ, Murphy KJ, Birch CD, Hamill JD (2000). Molecular characterization of quinolinate phosphoribosyltransferase (QPRTase) in *Nicotiana glauca*. *Plant Mol Biol* 44: 603-617.
- Tiburcio AF, Galston AW (1986). Arginine decarboxylase as the source of putrescine for tobacco alkaloids. *Phytochemistry* 25(1): 107-110.

Wagner R, Wagner KG (1985). The pyridine-nucleotide cycle in tobacco. Enzyme activities for the de-novo synthesis of NAD. *Planta* 165: 532-537.

Wagner R, Feth F, Wagner KG (1986). Regulation in tobacco callus of enzyme activities of the nicotine pathway. II. The pyridine-nucleotide cycle. *Planta* 168: 408-413.

Wagner R, Feth F, Wagner KG (1986). The regulation of enzyme activities of the nicotine pathway in tobacco. *Physiol Plant* 68: 667-672.

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. *J Biosci* 33: 177-184.

Wang P, Zeng J., Liang Z, Miao Z, Sun X, Tang K (2009). Silencing of PMT expression caused a surge of anatabine accumulation in tobacco. *Mol Biol Rep* 36: 2285-2289.

Xie J, Song W, Maksymowicz W, Jin W, Cheah K, Chen W, Carnes C, Ke J, Conkling MA (2004). Biotechnology: A tool for reduced-risk tobacco products -- the nicotine experience from test tube to cigarette pack. *Rec Adv Tob Sci* 30: 17-38.

Xu BF, Sheehan MJ, and Timko MP (2004). Differential induction of ornithine decarboxylase (ODC) gene family members in transgenic tobacco (*Nicotiana tabacum* L. cv. Bright Yellow 2) cell suspensions by methyl-jasmonate treatment. *Plant Growth Regul* 44: 101-116.

Yang KS, Gholson RK, Waller GR (1965). Studies on nicotine biosynthesis. *J Amer Chem Soc* 87: 4184-4188.

## **Chapter 2. Identification of transcription factors involved in nicotine biosynthesis**

### **INTRODUCTION**

Nicotine is found predominantly in the genus *Nicotiana* and in low quantities in some other species of the family *Solanaceae* (Sheen 1988). Nicotine is the most abundant alkaloid in commercial tobacco (*N. tabacum* L.) cultivars (Siminszky et al. 2005), and naturally plays an important role in tobacco resistance to insect herbivores (Steppuhn et al. 2004). The location for nicotine formation in tobacco is the root tip. Nicotine is then transported to the leaf and stored in leaf vacuole by a multidrug and toxic compound extrusion (MATE) transporter (Morita et al. 2009; Shoji et al. 2009). The pathway for nicotine biosynthesis involves the convergence of two biosynthetically distinct branches. The enzymes and the genes encoding these enzymes involved in the pathway have been identified except for the enzyme/gene that catalyzes the final condensation step of nicotinic acid and methy-pyrrolinium cation to form nicotine, generally referred to as nicotine synthase. The enzyme A622, a reductase of the PIP family, is considered a good candidate for the nicotine synthase (Hibi et al. 1994; Kajikawa et al. 2009). As stated in Chapter 1, PMT and QPT, the first committing enzymes in each branch of the nicotine biosynthetic pathway, are believed to be the key enzymes in the regulation of nicotine production (Feth et al. 1986; Wagner et al. 1986; Katoh et al. 2005, 2007). A few minor alkaloids, including nornicotine, anabasine, and anatabine, are also synthesized in this pathway, with nornicotine being directly converted from nicotine by nicotine demethylase whose corresponding genes have been cloned (Siminszky et al. 2005).

#### **Factors affecting nicotine accumulation**

Biotic and abiotic stresses, such as mechanical injury, insect attack, and herbivore damage, can significantly increase the nicotine accumulation in the leaf (Baldwin 1988, 1989; Baldwin et al. 1994, 1997; Shi et al. 2006). Moreover, nicotine accumulation in tobacco plants is greatly affected by agricultural practice, such as water availability, nitrogen supply,

topping, and suckering (topping is decapitation of the apical meristem at an early stage of flowering and suckering is removing the axillary buds of plants activated by topping ) (Bush and Saunders 1977; Wang et al. 2008). Furthermore, nicotine accumulation is dependent on the developmental stage of the root. Only young, growing root tips produce nicotine, whereas old and fully mature regions of the root do not (Solt 1957). Nicotine synthesis and accumulation is mediated by endogenous phytohormone changes, which affect expression of the genes involved in nicotine synthesis. So far, jasmonic acid, ethylene, and auxin have been shown to affect nicotine synthesis.

### **Jasmonic acid (JA)**

Jasmonic acid (JA) is derived from linolenic acid. JA and its derivative, methyl jasmonate (MeJA), are natural plant growth regulators commonly present in the plant kingdom (Creelman and Mullet 1995). JA is believed to be generated in wounded tissue and act as the transmissible signal to mediate the wounding responses of plant including the production of nicotine in tobacco roots (Howe 2004; Cane et al. 2005; Katoh et al. 2005; Koo and Howe 2009).

A JA signaling pathway has been proposed. In a plant cell with low JA level (un-induced), the JA responsive genes are repressed by JASMONATE ZIM-DOMAIN (JAZ) proteins through suppressing the activity of the MYC2 transcription factor that is required to activate the downstream genes of JA response. When induced, COI1, which is an F-box protein and serves as the JA receptor (Yan et al. 2009), binds to the most bioactive form of JA, the JA isoleucine conjugate (JA-Ile), to form an active complex. This complex binds JAZ proteins and leads to its ubiquitination and degradation through the 26S proteasome pathway. The MYC2 released from the suppression of JAZ then activates the downstream genes (Chini et al. 2007; Staswick 2007; Yan et al. 2009). As part of the complex negative and positive regulatory feedback loops in JA signaling (Kazan and Manners 2008), JAZ expression is up-

regulated by MeJA, and MYC2 appears to up-regulate the expression of most members of its family (Chini et al. 2007).

Nicotine production in tobacco is stimulated by JA treatment. Major nicotine biosynthetic pathway genes, including *PMT*, *QPT*, *ODC*, *ADC*, *MPO*, *A622* and *NBB1*, have been shown to be up-regulated to various extents by MeJA (Imanishi et al. 1998; Shoji et al. 2000a, 2002; Sachan and Falcone 2002; Goossens et al. 2003; Xu et al. 2004; Xu and Timko 2004; Cane et al. 2005; Katoh et al. 2007). Stimulation of nicotine production by MeJA or wounding in tobacco is mediated by the JA signaling pathway similar to that in *Arabidopsis*. Both NtCOI1 expression and NtJAZ protein degradation are required for nicotine biosynthesis (Paschold et al. 2007; Shoji et al. 2008).

An important component of the JA signal pathway is the protein MYC2, which is a basic helix –loop-helix leucine zipper transcription factor. It was first identified from the *Arabidopsis* mutant *JIN1*, which encodes AtMYC2 (Lorenzo et al. 2004). AtMYC2a was shown to be up-regulated by JA signal and to have negative feedback regulation on its own expression (Lorenzo et al. 2004; Dombrecht et al. 2007). In tomato, two MYC transcription factors, JAMYC2 and JAMYC10, were characterized and have strong homology with AtMYC2 (Boter et al. 2004). The binding motif of MYC2 transcription factor was identified and named G-box. G-box has a hexamer consensus core sequences as CACGTG (Sessa et al. 1995; Ouwerkerk and Memelink 1999). Some G-box like motifs were also found in certain genes' promoters, such as AACGTG and CACCTG (Wang et al. 2002; Guerineau et al. 2003; Boter et al. 2004). The *PMT* and *NtODC1* genes of the nicotine biosynthetic pathway have G-box in their promoter sequences (Xu and Timko 2004; Xu et al. 2004; Timko et al. 2010).

### **Ethylene (ETH)**

Ethylene is a gaseous plant hormone. It affects many plant developmental processes including germination, flower and leaf senescence and fruit ripening (Guo and Ecker 2004).

The ethylene signaling pathway has attracted intensive research interest, and remarkable progress has been made in elucidating this pathway (Stepanova and Alonso 2009). Ethylene was shown to have a negative effect on nicotine biosynthesis (Wang et al. 1994; Winz and Baldwin 2001). The transcription of the pathway genes, *PMT*, *ODC* and *A622*, seemed not to be affected by ethylene treatment individually, but ethylene significantly suppressed their expression induced by JA and is considered to antagonize JA signaling (Shoji et al. 2000b).

Ethylene response element binding factors (ERFs, previously called EREBPs) are important components of the ethylene signaling pathway (Guo and Ecker 2004). ERFs belong to a large *AP2* family of transcription factors with conserved AP2/ERF domain. Transcription factors possessing only one conserved AP2/ERF domain forms a subfamily designated as *ERF* (Magnani et al. 2004). The conserved DNA binding domain contains 58 or 59 amino acids

(Okamuro et al. 1997). The first ERFs were identified from tobacco and called NtEREBP 1-4 (Ohme-Takagi and Shinshi 1995). The core sequence of ERE from the promoters of PR (Pathogenesis-related) genes of tobacco was identified to be the GCC- box (Ohme-Takagi and Shinshi 1995; Shinshi et al. 1995). The motif AGCCGCC of GCC- box was shown to be necessary for the binding of ERFs in *Arabidopsis* and tobacco (Hao et al. 1998). Tobacco *PMT* gene of the nicotine biosynthetic pathway has a GCC-box (TGTGCCC) (Xu and Timko 2004; Timko et al. 2010).

ERFs bind to the GCC-box to regulate expression of stress or ethylene responsive genes. Some ERFs show positive effects as activators, while other ERFs have negative effects as repressors (McGrath et al. 2005). Among the five AtERFs identified from *Arabidopsis*, three of them act as transcription activators and the other two function as transcription repressors (Fujimoto et al. 2000). The repressing feature was also found in NtERF3 (Ohta et al. 2000).

Some ERFs have been shown to be involved in the regulation of secondary metabolism. Both ORCA2 and ORCA3 are ERF transcription factors identified in *Catharanthus roseus*.

ORCA2 was shown to induce expression of *STR*, a secondary metabolite biosynthetic gene. ORCA3 is also an ERF transcription factor and functions to activate the expression of several genes in the TIA (Terpenoid indole alkaloid) biosynthetic pathway (Menke et al. 1999; van der Fits and Memelink 2000).

## **Auxin**

Auxin has a negative effect on nicotine biosynthesis. Reducing the concentration of the naphthaleneacetic acid (NAA) in culture medium increased the enzyme activity of PMT and subsequent nicotine level of rootderived callus of *N. tabacum* (Feth et al. 1986). Nicotine level was decreased by the addition of 3-indoleacetic acid (IAA) (Yasumatsu 1967). Removal of indolebutyric acid (IBA) from the medium of tobacco root cultures led to an increase in gene transcription of *PMT*, *QPT*, *ADC*, and *ODC* (Hibi et al. 1994; Reed and Jelesko 2004). Application of NAA to the cut surface immediately after topping inhibited JA increase in the damaged stem tissue and roots, and resulted in decreased nicotine accumulation in tobacco leaf. Moreover, application of an auxin transport inhibitor below the shoot apex resulted in increased nicotine level in the plants (Shi et al. 2006). In addition, it was recently shown that, in *Nicotiana benthamiana* Domin, an auxin responsive factor, NbARF1, negatively regulated nicotine synthesis (Todd et al. 2010). All these strongly suggest a critical negative regulation role of auxin in nicotine biosynthesis.

## **Genetic analysis of nicotine synthesis in tobacco**

Genetic analysis of tobacco mutants exhibiting variable levels of nicotine revealed two non-linked loci named as *NIC1* and *NIC2* (or A and B). These two loci work synergistically to control the accumulation of alkaloid and *NIC1* has a more profound (2.4 fold) effect than *NIC2* (Legg et al. 1969; Legg and Collins 1971). The activity of the three enzymes of the nicotine biosynthesis pathway, PMT, QPT and MPO, were shown to be controlled under this system (Saunders and Bush 1979). These two loci were also demonstrated to regulate mRNA

accumulation of *PMT*, *QPT*, *A622*, *ODC* and *ADC* genes (Hibi et al. 1994; Reed and Jelesko 2004; Kidd et al. 2006). Because most of the pathway genes of nicotine biosynthesis have been characterized and many pathway genes are regulated by the *NIC* loci, it has been believed that these loci possibly encode transcription factor(s) although it cannot be excluded that one pathway enzyme has an effect on transcription of the other pathway gene and/or its enzyme activity. Very recently, the *NIC2* locus was identified to include a cluster of at least seven *ERF* transcription factor genes (Shoji et al. 2010).

### **Transcription factors involved in nicotine biosynthesis**

Major steps in nicotine biosynthetic pathway have been elucidated in the past two decades or so. However, studies on the regulation of this pathway at the transcription level have only begun recently. Transcription factors are the key players in regulating the transcription of pathway genes. A tobacco transcription factor database has been established based on *in silico* analysis of genomic sequencing data. More than 2500 transcription factors from 64 families were identified (Rushton et al. 2008). However, which transcription factors are involved in nicotine biosynthesis and how those transcription factors function together to regulate the nicotine biosynthetic pathway is still far from clear.

Two JA inducible AP2/ERF transcription factors genes, *NtORC1* and *NtJAP1*, were identified from a group of genes showing induction of expression after MeJA elicitation of *N. tabacum* BY-2 protoplasts. Both transcription factors up-regulated *PMT* gene expression. *NtORC1* seemed to have a stronger induction effect than *NtJAP1* on the *PMT* promoter and they function synergistically (De Sutter et al. 2005).

In a novel functional genomics approach, Todd et al. (2010) employed virus-induced gene silencing (VIGS) to screen for TFs that affect nicotine levels in the leaves of *Nicotiana benthamiana*. In a cDNA library enriched with JA responsive genes and containing 1898 unigenes, 69 putative TFs were identified and screened for their effects on nicotine level in leaves when they were over-expressed individually through the tobacco rattle virus (TRV)

vector to induce gene silencing. Among them, six affected nicotine level: one belonging to the *ERF* family (*NbERF1*), three to the *bHLH* family (*NbbHLH1*, 2, 3), one from the *ARF* family (*NbARF1*), and one in the homeobox family (*NbHB1*). Silencing of *NbERF1* gene by VIGS resulted in a decrease of nicotine level in the infected plants when treated with MeJA, indicating it is a positive regulator in nicotine biosynthesis. Suppression of the *NbHB1* gene caused much lower nicotine concentration in MeJA-treated plants, indicating that it also regulates nicotine synthesis positively. Meanwhile, silencing of *NbARF1* with the VIGS approach led to an increase of nicotine level in untreated plants, suggesting it is a suppressor of nicotine biosynthesis. Three basic helix-loop-helix (*bHLH*) transcription factor genes (*NbbHLH1*, *NbbHLH2* and *NbbHLH3*) were found to be involved in nicotine biosynthesis, with *NbbHLH1* and *NbbHLH2* being positive regulators and *NbbHLH3* a negative one. Over-expression and RNAi transgenic plants of *NbbHLH1* and *NbbHLH2* further confirmed this observation. Interestingly, while all eight major genes involved in nicotine biosynthesis were down-regulated in RNAi transgenic plants of each of the two genes, only three such genes (including *PMT*) were up-regulated in over-expressing transgenic plants and five others were still down-regulated (including *QPT*), suggesting a complex regulation network. Sequence analysis placed *NbbHLH1* and 2 genes in subgroup N of the *bHLH* family, to which *MYC2* genes also belong. It is not clear which subgroup *NbbHLH3* belongs to since no full-length cDNA was recovered.

Extensive studies on the effect of TFs in nicotine biosynthesis are also being carried out in Timko's laboratory (Timko et al. 2010). By using yeast one-hybrid screening with the *PMT* promoter sequence as a bait, several *ERF* genes were identified, all of which belong to the Group IX subfamily, which also includes *NtORC1* and *NtJAP1* as described above. In addition, three *MYC2* genes in *Nicotiana tabacum* were characterized and were designated as *NtMYC2a/b/c*. NtMYC2a and b formed a nuclear complex with the NtJAZ1 repressor. RNAi knockdown of expression of *NtMYC2a* and *b* decreased the transcription levels of *PMT1a* and other nicotine biosynthetic pathway genes, whereas over-expression of each gene had no

effect on *PMT1a* expression in the presence or absence of exogenously added MeJA. A detailed report of this research is yet to be published.

The *NIC2* locus in the tobacco genome has recently been molecularly characterized, and shown to contain a cluster of at least 7 *ERF* genes which positively regulate expression of nicotine biosynthetic pathway genes (Shoji et al. 2010). The recessive *nic2* locus is actually a deletion mutation with a missing DNA fragment covering this cluster of transcription factors, as shown by Southern analysis and PCR in the genome of *nic2* mutant plants. The seven ethylene responsive factor genes showed considerably different transcription level in wild type (*NIC1/NIC2*) plants compared to the double mutants (*nic1/nic2*). Protein sequence analysis of these highly homologous ERFs suggested that they belong to clade 2 of Group IXa subfamily of the *AP2* TF gene family. The seven *ERFs* had substantial expression in roots, which were induced by MeJA. But they were also expressed in other organs. These *ERF* genes were not equally efficient in promoting pathway gene expression and nicotine level, ranging from significant to negligible, suggesting functional redundancy and divergence developed during evolution of the duplicated genes. The most efficient one, *ERF189*, was able to enhance major pathway gene expression and tobacco alkaloids by 2-3 fold in transgenic hairy roots on the wild-type background. Suppression of the *ERF189* gene resulted in decreases in nicotine and transcription levels of pathway genes in transgenic hairy root lines. The molecular nature of the *NIC1* locus remains to be elucidated.

### **Objectives of the study**

This work started in mid 2008 when little was known regarding the transcription factors that regulate nicotine biosynthesis. Previous publications showed that expression of the pathway genes of nicotine biosynthesis was coordinated to some extent under MeJA or auxin treatment. Those findings led to the hypothesis that major or master transcription factor(s) exist to co-regulate these pathway genes. At the time, alteration of alkaloids appeared to be unachievable by simply over-expressing key pathway gene(s), such as *PMT* and *QPT*,

especially when significant increases are expected. It was hypothesized that transcription factor(s) could provide an alternative to achieving that goal. In this study, the objectives were to:

1. Use yeast one-hybrid screening to identify transcription factors involved in *QPT* gene transcription regulation in tobacco;
2. Gain insight into the mechanism of co-regulation of genes involved in the nicotine biosynthesis pathway;
3. Explore ways to up- or down-regulate nicotine level by adjusting TF expression.

## MATERIALS AND METHODS

### Yeast one-hybrid experiments for cloning transcription factors

#### *Bait vector construction*

A yeast one-hybrid system, Matchmaker™ One-Hybrid Library Construction and Screening Kit (Clontech, Mountain View, CA), was employed to screen for transcription factors that bind to the promoter of the tobacco *QPT2* gene (United States Patent 5837876). The promoter, 1034 bp in length, was cleaved from pTobRD2-PMT<sub>OX</sub> construct provided by the 22<sup>nd</sup> Century LLC., and inserted upstream of the GAL4 minimal promoter in vector pHIS2.1 to form the bait construct pTobHis. The insert was first verified by *Bam*HI digestion and the orientation of the promoter was confirmed by PCR using the following primers designed from sequence of the *QPT2* promoter and *His3* gene:

Tob1F: 5'-ACATCTGTAACCGGAACAGCAC-3'

His1R: 5'-GGTCGTCTATGTGTAAGTCACC-3'

Yeast strain Y187 was transformed with the bait construct pTobHis and 100 µL of 1/15 dilution of the original transformants was plated on SD/-Trp and SD/-Trp/-His media, respectively. The plates were titrated with a series of 3-AT (3-amino-1, 2, 4-triazole) (0, 5, 10, 15, 20, 30, 50, 75, 100 mM) to optimize the concentration that suppresses the basal expression of the bait construct. The plates were cultured at 30 °C for five days.

#### *cDNA library construction*

Total RNA was extracted from roots of greenhouse grown tobacco cv. NC95 0.5 hr after topping as described in Chapter 1. An oligo(dT) primer, modified by fusion with a 25-mer sequence, called CDS III, and the SMART™ III primer from the kit were used for synthesis of the first strand cDNA so to flank the synthesized cDNA with the two primer sequences. Long distance PCR (LD-PCR) was performed to amplify the synthesized cDNA using the Advantage 2 PCR kit (Clontech) with the CDS III and SMART™ III primers based on the

manufacturer's instruction. The PCR product was examined on agarose gel and the double-stranded cDNA was purified with CHROMA SPIN TE-400 columns from the kit and concentrated.

#### *Yeast transformation and positive colony selection*

Yeast competent cells were prepared and transformed with the bait vector pTobHis, the linearized prey vector pGADT7-Rec2, and the ds cDNA (~4 µg) following the manufacturer's instructions. The homologous recombination between the prey vector and the ds cDNA took place at the CDS III and SMART<sup>TM</sup> III sites of the linearized prey vector inside the yeast cells. The transformed cells were propagated, collected, and spread on DDO medium (SD/-Leu/-Trp) to estimate the screening efficiency of the co-transformation, and on TDO medium (SD/-Leu/-Trp/-His) to identify colonies positive in the one-hybrid selection.

#### *Characterization of positive colonies from yeast one-hybrid selection*

The insert length of the positive colonies was evaluated by a standard 30-cycle PCR (Bioneer, Alameda, CA) with 1 µL overnight yeast culture and the following two primers provided by the kit:

5' PCR Primer: 5'-TTCCACCCAAGCAGTGGTATCAACGCAGAGTGG-3'

3'PCR Primer: 5'-GTATCGATGCCCACCCTCTAGAGGCCGAGGCGGCCGAC-3'

The PCR products were examined on agarose gels, and all the prey constructs with cDNA inserts longer than 500 bp were subjected to plasmid isolation with QIAprep Spin Miniprep kit (Qiagen, Valencia, CA). The plasmids isolated from yeast colonies were subsequently transformed into *E. coli* DH5α for propagation. The prey plasmids were isolated again from *E. coli* and the cDNA inserts were subjected to sequence analysis using the T7 primer and the 3'PCR primer. Any yeast colony which contained more than one prey constructs was re-streaked on SD/-Leu for 2-3 times until the PCR result yields a single amplified fragment.

The sequences of the cDNA inserts were used for BLAST analysis with the NCBI GenBank database. A total of five transcription factors were identified from about 100 positive colonies. However, none of these five transcription factor cDNAs was in full length but they all had poly(A) tails.

*Recovery of full-length cDNAs of the TFs*

RNA was isolated from root tissue of greenhouse grown NC95 tobacco plants 0.5 hr after topping. The GeneRacer kit (Invitrogen, Carlsbad, CA) was used to obtain the missing 5' sequences of the TF cDNAs following the manufacturer's manual. One or two gene specific primers were designed based on the sequences near the 5' termini of the partial cDNAs of the six isolated TF genes for PCR together with the GeneRacer<sup>TM</sup> 5' primer provided by the kit. The sequences of these gene specific primers are:

NtGRAS GSP1: 5'-ATAGTCGTTTCTCCACTGCTGCCAATC -3'

NtGRAS GSP2: 5'-CACTGCATTAAACTCTACTGCAATGTTG -3'

NtERF GSP1: 5'-CTATCTCCGACTTCTGGTCTTCCTCT-3'

NtERF GSP2: 5'-CCACGGTCTCTGCCTTATTCCTCTGTA-3'

NtMYC2a GSP1: 5'-ACACATTTGGTACAACAGCTCTAAGTGC-3'

NtMYC2a GSP2: 5'-TGCAATTGCATCACCAAGAAGTGATGCT-3'

NtMYC2b GSP: 5'-CGGGGAGTTGGTGTAGTAG-3'

NtARF GSP1: 5'- CCTTTTGTGTCTCCCTTCCTACTGATG-3'

NtARF GSP2: 5'-CTAAGTTTTGAGAGCACTGGGTCCCAAG-3'

NtWRKY GSP1: 5'-TTGGTCACGAACGTACAATTCCTAGTG-3'

NtWRKY GSP2: 5'-TGCTGTATTTGTTGAGACTGAGTCTGAG-3'

Later transgenic studies indicated that the GRAS and WRKY TFs isolated from this project had no effect on nicotine synthesis, so my report will focus on the other four TFs from now on. After the full-length cDNAs were recovered, primers were designed to obtain the full-length coding sequences of these four transcription factors by PCR:

NtERFFL1F:5'-TCTAGAGGATCCCGGGATGTGTGGAGGTGCCATAATCC-3'

NtERFFL1R:5'GCGCCCGGGTTCAGTAAACAGCTGCTGCTGC-3'

NtARFFL1F:5'-GGATCCATGATGTGTGGACTTATTGATC-3'

NtARFFL1R:5'-CCCGGGCTACAAAGCAATATCAAGAATC-3'

NtMYC2a 1F:5'-GCGGTCTAGACAGATCTGAATTGATTTGTCT-3'

NtMYC2a 1R:5'-GCGGTCTAGAACATTATTCAGAGCTCACTATG-3'

NtMYC2b 1F:5'-GCGTCTAGAATGACGGACTATAGAATACCA-3'

NtMYC2b 1R:5'-GCGTCTAGATCATCGCGATTCAGCAATTCT-3'

All PCR reactions were performed using the high-fidelity Taq DNA polymerase (Phusion, Finnzymes, Espoo, Finland). The PCR products were cloned into pCRBluntII-TOPO or pCR4-TOPO vector (Invitrogen) for sequence analysis by a commercial provider.

### **Expression analysis of the isolated transcription factors in tobacco**

Tobacco plants (cv. NC 95), grown in the greenhouse for about two months until just before flowering, were subjected to gene expression analysis in various organs as well as in roots after topping, wounding, or MeJA (Sigma, St. Louis, MO) treatments. Topping was performed by cutting off the shoot apex of plants immediately below the bud. Wounding was performed by cutting leaves on the plants with scissors: five cuts per leaf on top three fully-grown leaves. For the MeJA treatment, 50  $\mu$ M solution (about 5 mL) was sprayed on all the plant leaves.

Root tissues were collected at the time points of 0 h, 0.25 h, 0.5 h, 1 h, 2 h, 4 h, and 6 h after treatments. RNA samples were isolated and Northern blot analysis was performed as described in Chapter 1. The primers designed from the coding sequences used to generate the probes for Northern hybridization were:

NtERFNOR F: 5'-ACACTGCACTAGCACCATCCC-3'

NtERFNOR R: 5'-CTGCATTGTACTACGTACTACC-3'

NtMYC2a/bNOR F: 5'-GAAGTAACGGATACTGAATGG-3'

NtMYC2a/bNOR R: 5'-ATCCTTGTGTTTGCTGAGAAT-3'

NtARFNOR F: 5'-CTGCCTATAGCCAACTGTTG-3'

NtARFNOR R: 5'-AAGCTGCTGGATACAGGAGC-3'

### Tobacco transformation for over-expression and down-regulation of the isolated transcription factors

pBI121 was used as the backbone vector to make the over-expression and RNAi gene constructs. The coding sequences of the isolated transcription factors were cleaved from their cloning vectors (pCRBluntII-TOPO or pCR4-TOPO) and inserted into pBI121 at the place of the *GUS* gene to be under control of the constitutive CaMV 35S promoter (Fig. 2.1).

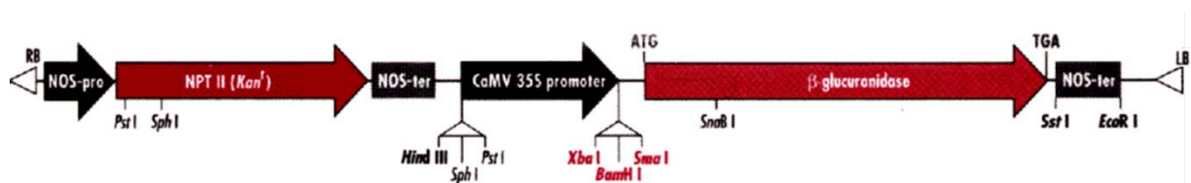


Figure 2.1 vector map of T-DNA region of pBI121.

The RNAi technique was used to down-regulate the individual TFs in transgenic tobacco plants. For RNAi vector construction, the partial coding sequence of each TF was obtained by PCR with the following pairs of the primers:

NtERFRNAiF: 5'-GACTGAGCTCTCTAGAGTGGAGGTGCCATAATCCCCGA-3'

NtERFRNAiR: 5'-GACTCCCGGGGATATCCGGTCTCTGCCTTATTCCTCTGTA-3'

NtMYC2RNAiF: 5'- GGGGAGCTCTCTAGAGCTGCAACAGCGACTCCAGA -3'

NtMYC2RNAiR: 5'-ATT CCCGGGGTTCGACCCGTTAACAAACGATTGAGTC-3'

NtARFRNAiF: 5'-GGGGAGCTCGGATCCGATGGGATTGCAGTATCAGAC-3'

NtARFRNAiR: 5'-GGGACTAGTGTTCGACGAGTACTTGGATTGCAATGAC-3'

The PCR fragment was first cloned into pQLi (kindly provided by Dr. Quanzi Li, Dept. of Forestry and Environmental Resource, NCSU) to replace the red and blue regions (as shown in Figure 2.2) to create inverted repeat, and then the whole cassette was inserted into pBI121 to replace the *GUS* gene. All the resultant over-expression and RNAi vectors were confirmed by appropriate restriction digestions and sequencing analysis.

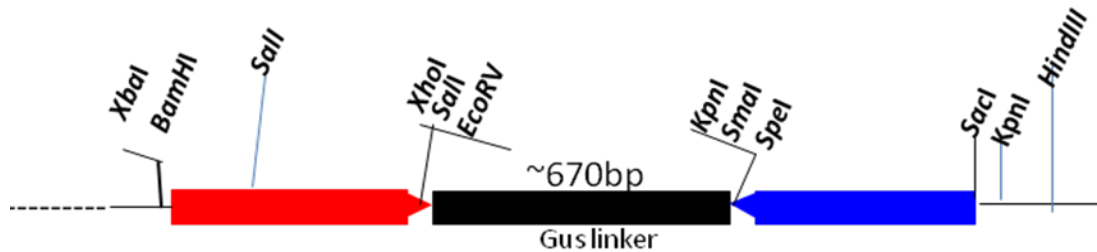


Figure 2.2 Schematic representation of the region of vector pQLi used for inverted repeat cloning. *Gus* linker: partial sequence from *GUS* gene.

*Agrobacterium*-mediated tobacco transformation was performed as described in Chapter 1. Putative transgenic tobacco plants (cv. NC 95) were grown in the greenhouse for about two months and root tissues were collected just before flowering. RNA extraction from root tissues and Northern blot analysis of the transformed plants were performed as described in Chapter 1. The primers used to generate probes were the same as described above.

## Quantification of major alkaloids

Nicotine levels in dried leaves of the transgenic plants were kindly quantified by gas chromatography (GC, Coresta recommended method N<sup>o</sup> 62) in Dr. D. Danehower's laboratory (Crop Science Dept., NCSU). Each sample was prepared by placing 0.2000 ± 0.0010 g of dried ground tobacco leaves into a 50 mL Erlenmeyer flask. Two mL of 2N NaOH solution was added to each flask and swirled to moisten the tobacco. After 15 min of rest, 10 mL of methyl tertiary butyl ether (MTBE) containing 0.1062 g/mL of quinoline was added to the flask. The flasks were placed on a shaker for 2.5 hrs. After shaking the flasks were allowed to sit overnight to separate. Approximately 1 mL of the top MTBE layer was transferred into a vial. GC analysis was conducted using a split injection (40:1) on an Agilent HP 6890 GC-FID (Agilent Technologies, Santa Clara, CA) using a 30 meter DB-5MS column (0.53 mm ID and 1.5 µm film thickness). The carrier gas was helium at a linear velocity of approximately 38 cm/sec. The injector and detector were both set at 250 °C. The analysis consists of a temperature program from 110 °C initially held for 0.5 min followed by a ramp to 280 °C at a rate of 25 °C/min where the final temperature was held for 20 min. Data were collected and analyzed using Agilent Chemstation software. A multi point internal standard calibration table was constructed for each compound. The curves for each compound are as follows:

$$\text{Nicotine: } Y = 2.32779e-1 * x + 8.05332e-3$$

$$\text{Nornicotine: } Y = 2.26220e-1 * x - 3.49890e-3$$

$$\text{Anabasine: } Y = 2.23584e-1 * x + 2.27888e-4$$

$$\text{Anatabine: } Y = 1.33963e-1 * x - 3.08881e-3$$

## Phylogenetic tree construction

Mega 4.1 software (Biodesign Institute, Tempe, AZ) was used to generate the phylogenetic trees for the isolated TFs.

## RESULTS

The yeast one-hybrid technique was used in this study to identify the transcription factors. This technique includes three important components: bait construct, prey vector, and cDNA. The *QPT2* promoter region was inserted upstream the GAL4 minimal promoter in the bait vector, which drives a histidine synthesis gene *HIS3*. Thus, if the yeast cell contains the bait vector and a prey vector expressing a TF-GAL4 AD (activation domain) fusion protein that binds to the *QPT2* promoter, the *HIS3* gene will express and the cell will be able to grow on a screening medium which lacks histidine. The bait construct was confirmed by restriction digestion and PCR for the insertion of the *QPT2* promoter and its orientation. It was then tested for leaking expression of *HIS3* (or basal expression because of endogenous yeast transcription factors). 3-amino-1, 2, 4-triazole (3-AT) is a competitive inhibitor of the reporter gene *HIS3* product, a histidine biosynthetic enzyme. A titration experiment of 3-AT was conducted to optimize 3-AT concentration in the yeast culture medium to minimize potential false positives in the screening experiments. No colony grew on the medium (SD/-His/-Trp containing a series of 3-AT from 5 mM to 75 mM) while many colonies grew on the control plate of SD/-Trp. This experiment was repeated and the same result was obtained. Thus, it was not necessary to add 3-AT to the screening medium to suppress the basal expression of *HIS3* gene and the background growth.

Total RNA was extracted from the root tissue collected 30 min after topping and used to make a cDNA library for the yeast one-hybrid screening. Many of the nicotine biosynthetic pathway genes are induced several hrs after topping and it was expected that the transcription factor genes would be induced earlier. A time of 30 min after topping was estimated to be appropriate to capture a “snapshot” of the expression of these transcription factors. The quality of the cDNA was examined by PCR.

A total of three screening experiments were performed by transforming yeast competent cells with the bait vector, the prey vector, and the cDNA collection (homologous recombination would take place between the prey vector and the cDNA inside the yeast cell so the cDNA-

GAL4 AD fusion gene will express from the prey vector), and approximately 1.6 million yeast colonies were screened. After seven days of incubation, actively-growing yeast colonies were selected for colony PCR to screen for prey plasmid which has the cDNA insert longer than 500 base pair. All yeast colonies which contained more than one prey plasmid (more than one amplified DNA fragment on agarose gels) were subjected to successive re-streaking on SD/-Leu medium until only one prey plasmid was left in the colony. The isolated prey plasmid was used to transform *E. coli* (strain DH5 $\alpha$ ) for plasmid propagation. Approximately 100 yeast positive colonies were finally isolated and sequenced, among which five were putative transcription factors as identified based on the BLAST analysis with the NCBI GenBank database. These five transcription factors were shown to belong to five TF families: *GRAS*, *AP2/ERF*, *bHLH*, *ARF* and *WRKY*.

All the cloned TF cDNAs were partial in length with their 5' sequences missing. To obtain full-length cDNAs for these TFs, the 5' RACE (Rapid Amplification of cDNA Ends) technique was performed, and high-fidelity Taq DNA polymerase was used in all the PCR reactions. For each TF, at least five randomly picked colonies were subjected to sequence analysis. During the process, another *bHLH* gene with high homology to the cloned one was identified and cloned. The sequences and expression patterns of these TF genes were characterized. To test whether these transcription factors have effects on nicotine biosynthesis pathway, transgenic over-expression and RNAi lines of all the six transcription factors were produced. Transformed plants were grown in the greenhouse for about two months. Total RNA was extracted from root tissue and subjected to Northern blot analysis of pathway gene expression and nicotine concentration. Because the TFs from *GRAS* and *WRKY* families did not show any apparent effect on *QPT* gene expression, they are not reported in this dissertation.

### **Transcription factor *NtMYC2a* and *NtMYC2b* genes**

Two full-length *bHLH* transcription factor cDNAs were cloned with 2214 and 2391bp in length and encoding 659 and 658 AA, respectively (Appendix). They share 96% identity at the cDNA sequence level. BLAST analysis to the NCBI database showed that these two MYC transcription factors are 99.9% (1979/1980 nt) and 100% (1977/1977 nt) identical to the coding sequences named *NtMYC2a* (GenBank No. GQ859160) and *NtMYC2b* (GenBank No. GQ859161) in the database. The single nucleotide difference in *NtMYC2a* might be caused by sequencing error or PCR error or by the cultivar difference. So the two TFs are named *NtMYC2a* and *NtMYC2b*.

A phylogenetic tree was made by using the maximum parsimony method in MEGA4.1 ([www.megasoftware.net](http://www.megasoftware.net)) based on the AA sequence to infer the evolutionary relationship of both *NtMYC2a* and *NtMYC2b* with other *bHLH* transcription factors. Several closely related transcription factors from the BLAST result were also included. The results show that *NTMYC2a* and *NTMYC2b* are closely related to each other, and most closely related to *NbbHLH2* and *StMYC10*, and slightly more distantly to *NtMYC1a*, *NbbHLH1* and *StMYC2* as reported in *N. benthamiana* and tomato (Todd et al. 2010; Boter et al. 2004).

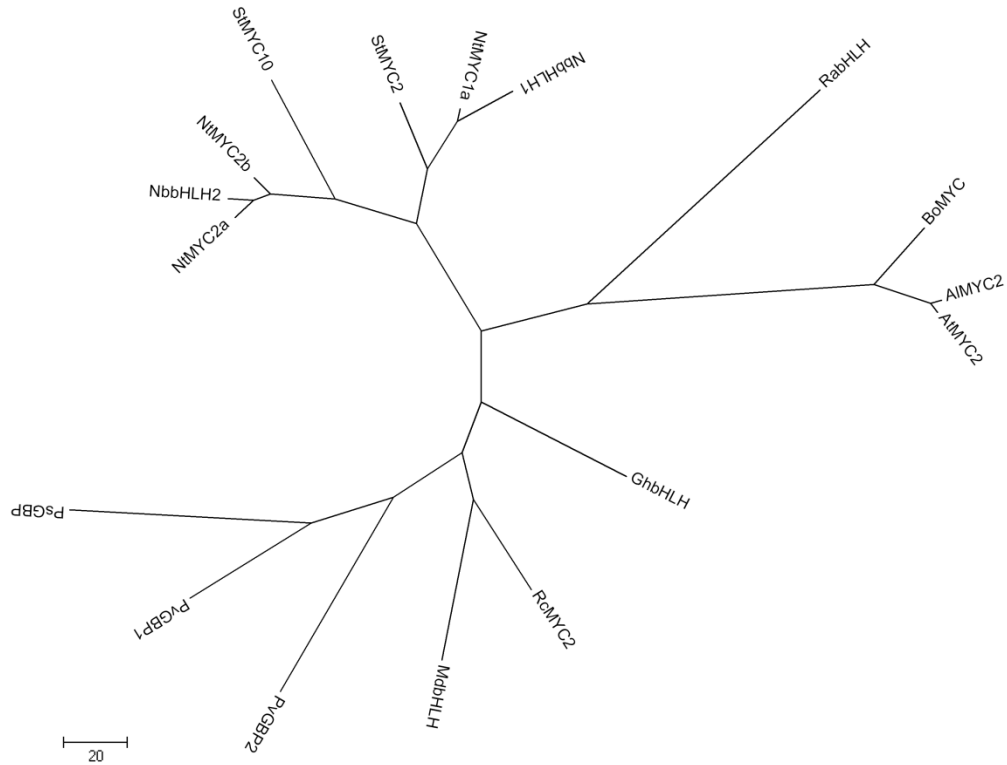


Figure 2.3 Unrooted phylogenetic tree of *NtMYC2a*, *NtMYC2b* and 15 other *bHLH* transcription factors

#### *NtMYC2* Expression pattern in tobacco

Northern blot analysis (Figure 2.4) shows *NtMYC2* genes expressed in tobacco root, stem, leaf and flower. Due to the high homology between the two *MYC2* genes, the probe for all the Northern blot analysis of *MYC2* is a partial coding sequence of 505 bp from *MYC2b*, and cannot distinguish the two genes and thus they were analyzed together. *MYC2* expresses in all the four organs examined with the highest expression in stem. Expression in all major organs may indicate the genes also function in metabolic pathways other than nicotine biosynthesis.

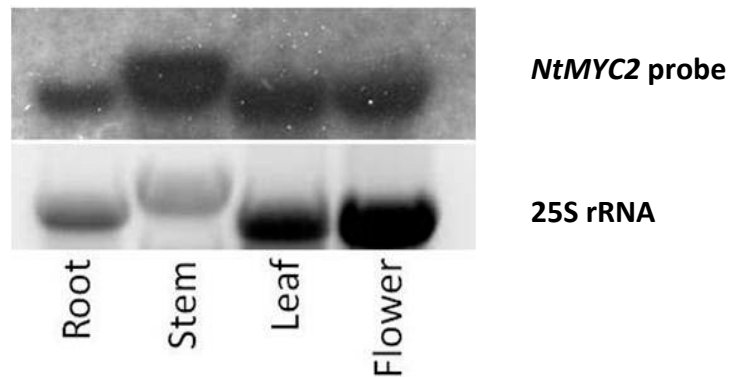


Figure 2.4 Organ expression pattern of *NtMYC2*. Northern blot hybridization of total RNA isolated from various organs of fully grown plant and probed with *NtMYC2b* PCR fragment probes. The 25S rRNA stained with EtBr in gel was also shown as a loading reference.

*NtMYC2 expression patterns in tobacco root after topping, wounding, and MeJA treatment*

Expression of the *NtMYC2* genes in root after topping, wounding, or MeJA treatment was investigated. As shown in Figure 2.5, *NtMYC2* expression was induced in root by topping, wounding, or MeJA treatment. Compared to the control, the induced expression of *NtMYC2* seems to be biphasic: the expression increased 0.25 h after the treatments, declined slightly afterwards, and increased again by 4 or 6 hrs, with wounding having the strongest induction effect among the three treatments.

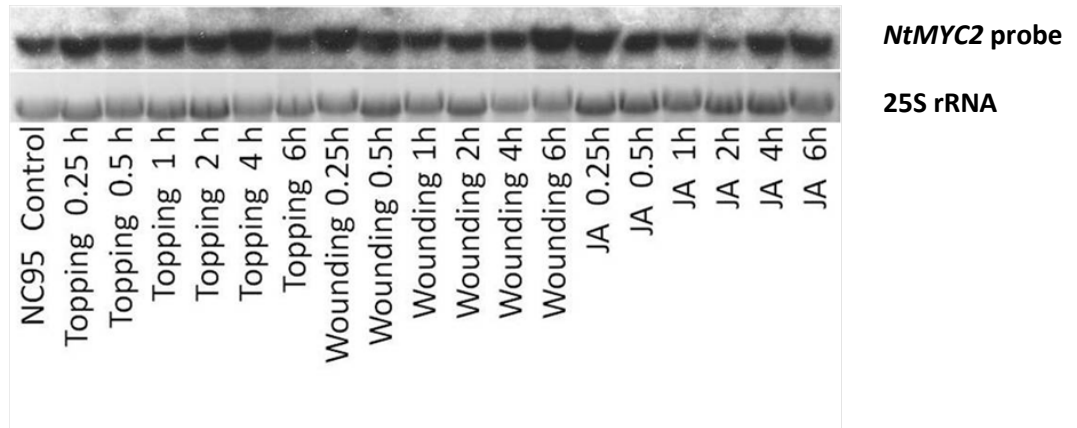


Figure 2.5 *NtMYC2* expression in tobacco root after topping, wounding, or MeJA treatment. Northern hybridization shows a time course induction pattern of *NtMYC2* after each treatment. The probe is from *NtMYC2b*. The 25S rRNA stained in gel with EtBr is shown as a loading reference.

#### *Analysis of NtMYC2a and NtMYC2b transgenic lines*

Seven *NtMYC2a* and 10 *NtMYC2b* over-expression lines and nine *NtMYC2* RNAi transgenic lines, in which both genes are expected to be suppressed, were generated to test the effect of these two TF genes on nicotine biosynthesis. The 336 bp fragment used for RNAi construct was from *NtMYC2b* coding sequence, which shares 94% identity with that region of *NtMYC2a* gene.

Figure 2.6 shows the expression levels of *NtMYC2a* and *NtMYC2b* in transgenic plants over-expressing *NtMYC2a* or *NtMYC2b* gene. Compared to the wild type and vector control, two lines of *NtMYC2a* (AOE-3, and 6) and seven lines of *NtMYC2b* (BOE-7, 10, 11, 13, 14, 16, and 17) were obvious over-expression lines.

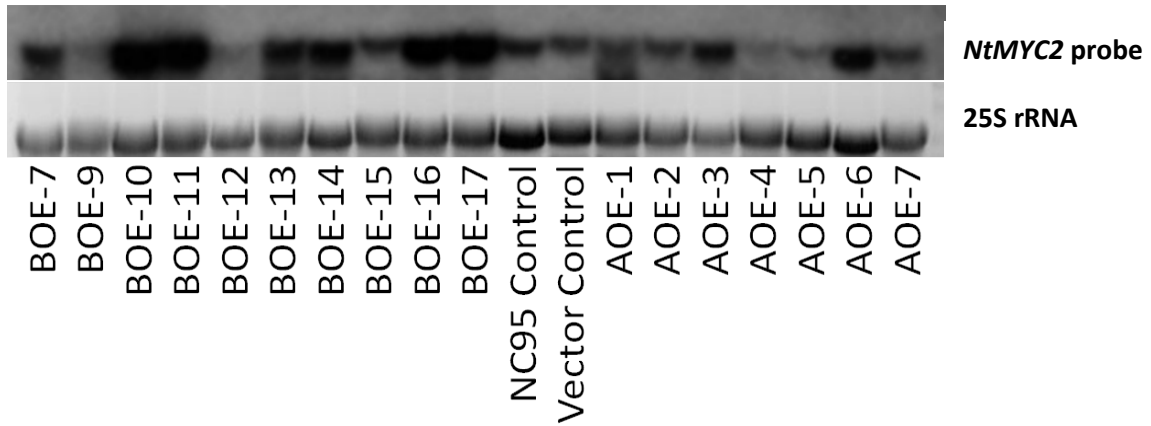


Figure 2.6 Expression of *NtMYC2a* and *NtMYC2b* in *NtMYC2a* or *NtMYC2b* over-expression lines. Northern hybridization shows *NtMYC2* mRNA level in *NtMYC2a* or *NtMYC2b* over-expression lines. The bottom panel shows the ribosomal RNA in gel stained with EtBr as an RNA loading reference. The wild-type and vector control are also shown.

Figure 2.7 shows the effects of the RNAi construct on *NtMYC2a/b* expression. Expression level of *NtMYC2* in three RNAi lines (RNAi-1, 2, 3) was greatly reduced. All other lines did not show substantial changes in *NtMYC2* expression.

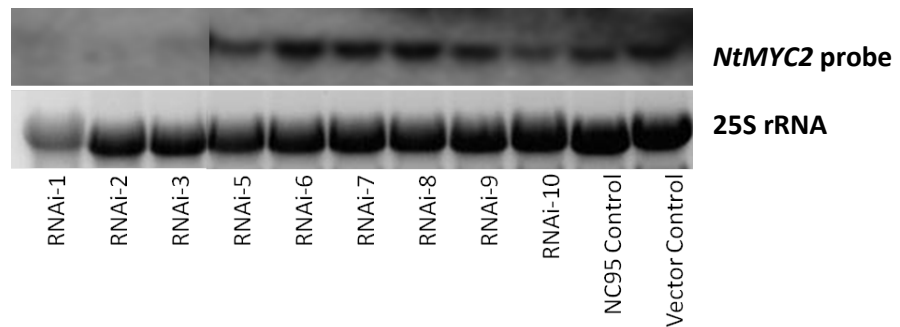


Figure 2.7 Combined expression of *NtMYC2a* and *NtMYC2b* in *NtMYC2* RNAi lines. Top panel: Northern hybridization of *NtMYC2* mRNA level in *NtMYC2* RNAi lines. The bottom panel shows ribosomal RNA in gel stained with EtBr as an RNA loading reference. The wild-type and vector control are also shown.

Based on the Northern analysis, nicotine levels of two *NtMYC2a* over-expression and seven *NtMYC2b* over-expression lines with high transgene expression and three RNAi lines with much lower expression of *NtMYC2* were quantified. Figure 2.8 shows that three *NtMYC2b* over-expression lines (BOE-10, 16, and 17) and two *NtMYC2a* over-expression lines (AOE-3 and AOE-6) had higher nicotine level than the controls (41% to 149% higher than vector control). However, other lines with higher expression level of *NtMYC2b*, such as BOE-7, 11, 13, and 14 didn't have higher nicotine level. The three RNAi lines (RNAi-1, 2 and 3) showed much lower nicotine level (about five-fold less) than the controls.

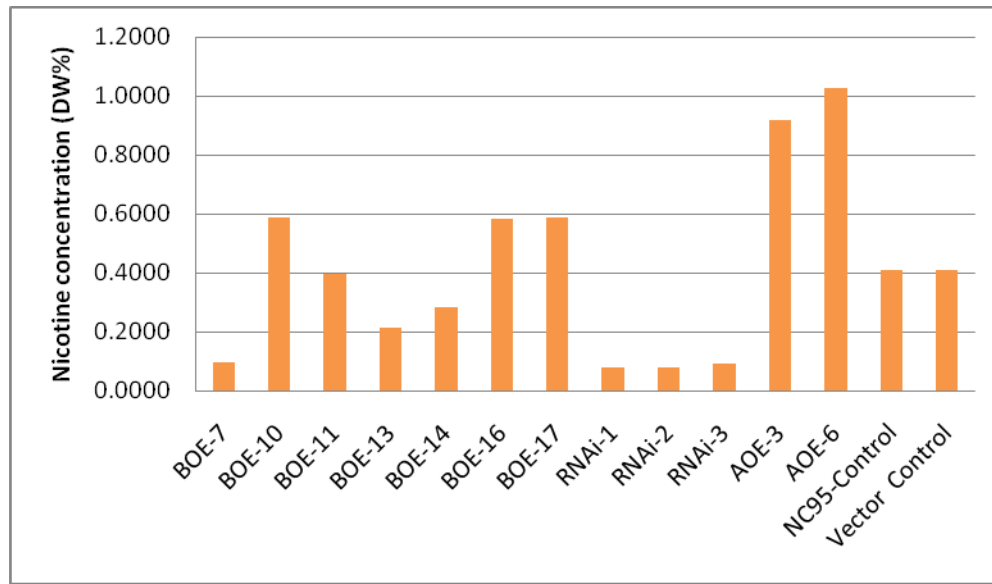


Figure 2.8 nicotine concentration of over-expression and RNAi lines of *NtMYC2a* and *NtMYC2b*. DW: dry weight

Because over-expression and knock down of the *NtMYC2* genes altered both the nicotine and total alkaloid levels (data not shown), selected over-expression and RNAi lines were subjected to Northern blot hybridization analysis to evaluate the effects on *PMT* and *QPT* gene expression. Figures 2.9 and 2.10 show the expression of *QPT* and *PMT* genes as affected by the two transcription factors. Over-expression of *NtMYC2b* did not change much

of the mRNA levels of *PMT* and *QPT*. However, the expression of these two key structural genes were remarkably reduced by over-expression of *NtMYC2a* and RNAi of both *NtMYC2* genes.

Surprisingly, *QPT* and *PMT* expression levels in *NtMYC2a* over-expression lines (AOE-3 and AOE-4) were even lower than in the RNAi lines.

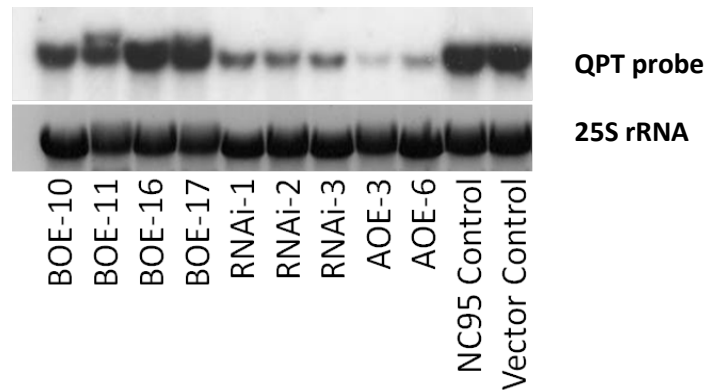


Figure 2.9 *QPT* expression level in *NtMYC2 a*, or *b* over-expression and RNAi transgenic tobacco lines. Northern blot hybridization uses partial *QPT* coding sequence as probes. The rRNA stained with EtBr in gel is shown at the lower panel as an RNA loading reference.

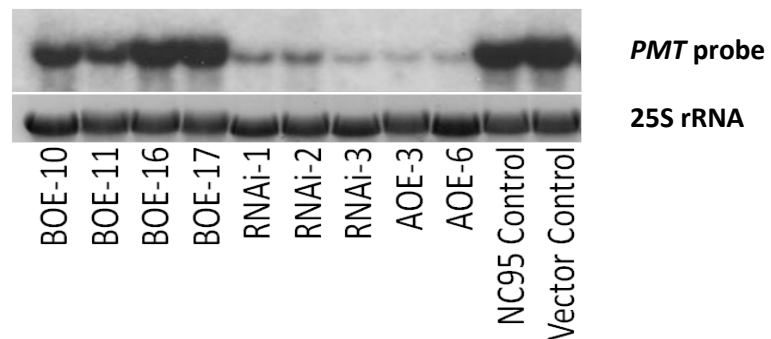


Figure 2.10 *PMT* expression in *NtMYC2a*, or *b* over-expression and RNAi transgenic lines. Northern blot hybridization using partial *PMT* coding sequence as probes. The rRNA stained with EtBr in gel is shown in the lower panel as an RNA loading reference.

From the data on over-expression and RNAi lines, it is evident that *NtMYC2a* and/or *NtMYC2b* genes modulate nicotine biosynthetic pathway gene expression and the nicotine level. It is more remarkable that in the three RNAi lines (RNAi-1, -2, and -3) which had *NtMYC2* expression substantially decreased, both *PMT* and *QPT* mRNA levels were reduced by approximately ten- fold, and the nicotine level decreased about five-fold. More interestingly, although *NtMYC2a* and *NtMYC2b* are highly homologous, their functions appear to be diverse: Over-expression of *NtMYC2a* led to greatly reduced *PMT* and *QPT* mRNA levels yet the highest nicotine level (more than two-fold of the controls' and around 1% leaf dry weight) whereas over-expression of *NtMYC2b* caused little change in *PMT* and *QPT* mRNA levels among the four lines analyzed when nicotine concentration had a moderate increase (nearly 50%) in three out of the four lines. Moreover, it is surprising to see that plants having the highest nicotine level (AOE-3 and -6) were the ones with the lowest *PMT* and *QPT* mRNA levels, and plants with the lowest level of nicotine (RNAi-1, -2, and -3) had slightly more (but much reduced comparing the controls) *PMT* and *QPT* mRNA. The data suggest that the nicotine level is not necessarily associated with *PMT* and *QPT* mRNA levels. It is needed to measure the nicotine level and relevant gene expression of these transgenic plants after topping (or other treatments) to further evaluate the impact of the *MYC2* genes on the gene expression and nicotine accumulation in this pathway. When doing so, it would be more helpful to perform qRT-PCR to distinguish the expression of the two *MYC2* genes to gain more insight in their distinct functions.

### **Transcription factor *NtERF98* gene**

The full-length *ERF* transcription factor cDNA was 1019 bp in length and encodes a protein of 257 AA. In BLAST analysis against the tobacco transcription factor database (TOBFAC, <http://compsysbio.achs.virginia.edu/tobfac/>), this *ERF* transcription factor gene showed 99.9% identity (1001/1002 nt) with the genomic sequence of a gene named *NtERF98* (GenBank No. [AJ299252](#)) in the database, suggesting that they are the same gene. This gene

belongs to subfamily VII of the *AP2/ERF* family according to the phylogenetic tree in the TOBFAC database. The role this gene plays in the nicotine biosynthesis has not been reported.

*Expression pattern of NtERF98 in different organs of plant*

The expression pattern of *NtERF98* in root, stem, leaf, and flower in wild type NC95 tobacco plants were examined. Figure 2.11 shows that *NtERF98* was mainly expressed in root, stem, and flower, and had little expression, if any, in leaf.

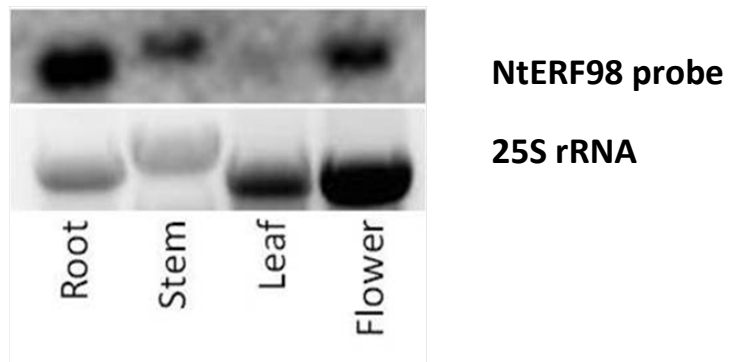


Figure 2.11 *NtERF98* expression in four organs of tobacco plant. Northern blot hybridization of total RNA isolated from various organs of fully grown plant was probed with a partial *NtERF98* coding sequence. The 25S rRNA stained with EtBr in gel was also shown as a loading reference.

*Expression pattern of NtERF98 under treatments of topping, wounding, and MeJA*

Because nicotine accumulation is induced by topping, wounding, and JA treatment, these three treatments were applied individually in this study to test whether *NtERF98* expression was also affected by these treatments. Root tissues were collected 0.25 h, 0.5 h, 1 h, 2 h, 4 h and 6 h after treatments. Figure 2.12 shows the gene expression pattern in roots of the treated tobacco plants.

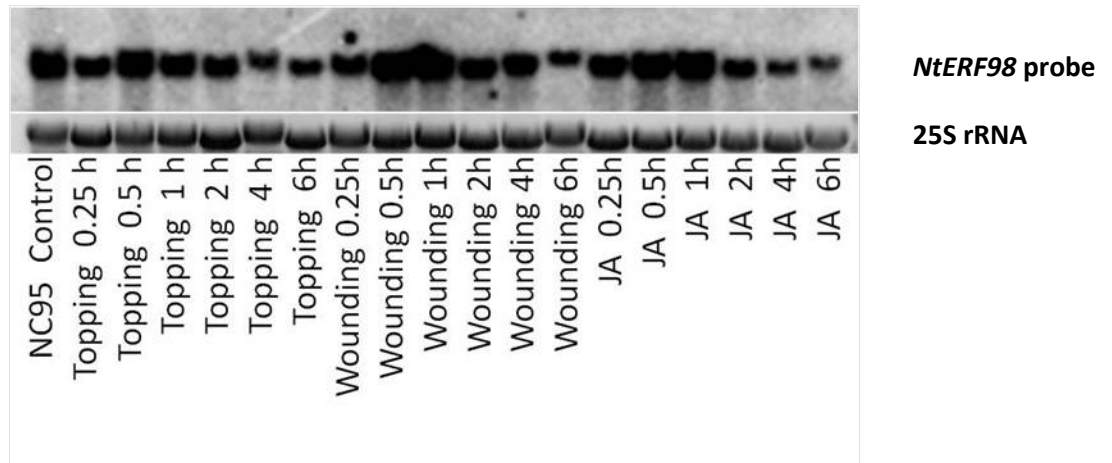


Figure 2.12 Time course of *NtERF98* expression in tobacco roots after topping, wounding, or MeJA treatment. Northern blot was hybridized with *NtERF98* coding sequence as a probe, and 25S rRNA with EtBr staining as a loading reference.

*NtERF98* expression under these three treatments had a similar biphasic decline pattern. The steady-state *NtERF98* mRNA level was reduced within 15 min after initiation of each of these treatments. From 0.25 h to 0.5 h, the expression level increased up to the basal level of the wild type control in all three treatments. After 0.5 h, the mRNA levels were in decline again in the topping and wounding treatments. In MeJA treatment, the expression level reached its peak at 1 h, and declined thereafter.

#### *Analysis of transgenic overexpression lines*

Figure 2.13 shows the Northern analysis of *NtERF98* and *QPT* in *NtERF98* putative over-expression lines. Six out of 10 lines showed much higher expression level of *NtERF98*. Among the six *NtERF98* over-expression lines, five had reduced *QPT* expression. OE-19 and OE-25 showed similar pattern of expression of these two genes as the control plant, which had very low expression of *NtERF98* and high expression of *QPT*. A negative correlation seems to exist between the expression pattern of *NtERF98* and *QPT* in most of the transgenic plants that over-expressed *NtERF98*.

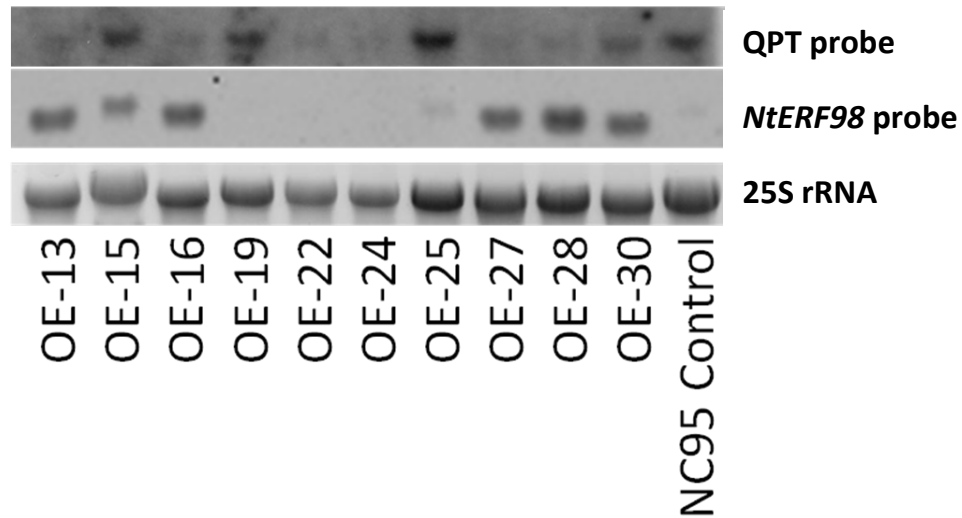


Figure 2.13 *NtERF98* and *QPT* expression in ten putative *NtERF98* over-expression transgenic plants. Northern hybridization of *QPT* (top) and *NtERR98* mRNA (middle) is shown. 25S rRNA stained with EtBr is also shown as a loading reference (bottom).

Leaf nicotine concentration of all the transformed lines were determined. Figure 2.14 shows that the nicotine levels in the tranformed over-expression lines of *NtERF98* increased except for the OE-15 and OE-16 lines, which exhibited slight reductions. No apparent correlation was found between the nicotine concentration and the *NtERF98* or *QPT* expression level.

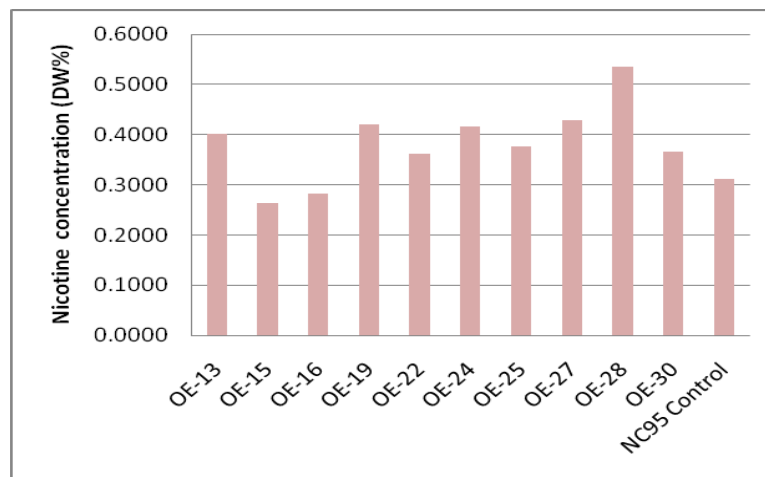


Figure 2.14 Nicotine concentration in ten *NtERF98* over-expression transformed tobacco lines and wild type control. DW: dry weight.

*Analysis of NtERF98 transgenic RNAi lines*

A PCR fragment of 309 bp from *NtERF98* coding sequence was used to make the RNAi construct. Northern analysis was performed to evaluate whether the knockdown of *NtERF98* has an effect on the *QPT* expression. Figure 2.15 shows that, out of 10 transformed lines, six had the *NtERF98* expression completely repressed (RNAi-6 through -12), and the other four had slightly or moderately reduced expression.

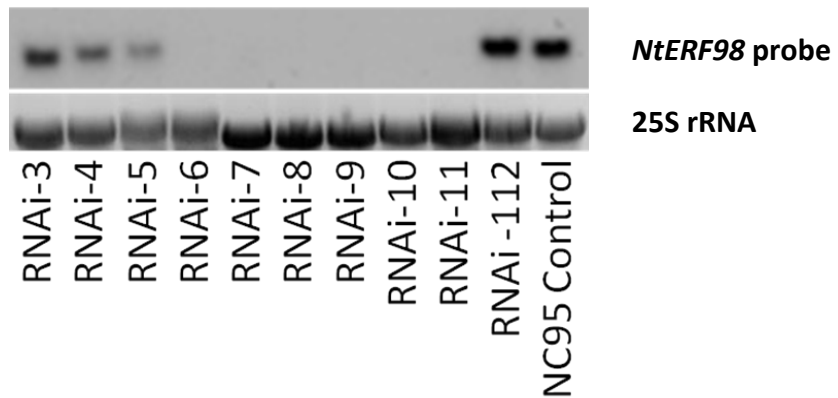


Figure 2.15 Northern analysis of *NtERF98* expression in its RNAi transgenic lines. The 25S rRNA stained with EtBr is shown as a loading reference.

Northern analysis of *QPT* gene expression of the *NtERF98* RNAi lines was performed. Figure 2.16 shows the *QPT* expression in the RNAi plants when compared to a control plant. Various degree of reduction in *QPT* expression was observed among the transgenic plants. However, when nicotine concentration was examined, seven out of ten RNAi plants had moderate increase in nicotine level (30-60% higher) with the other three having levels similar to that of the non-transgenic control (Figure 2.17).

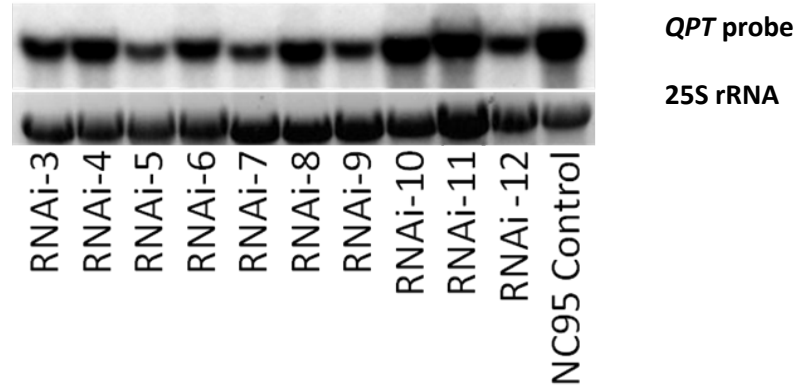


Figure 2.16 Northern hybridization of *QPT* expression in the *NtERF98* RNAi transformed tobacco lines. The 25S rRNA stained with EtBr is shown in the lower panel as a loading reference.

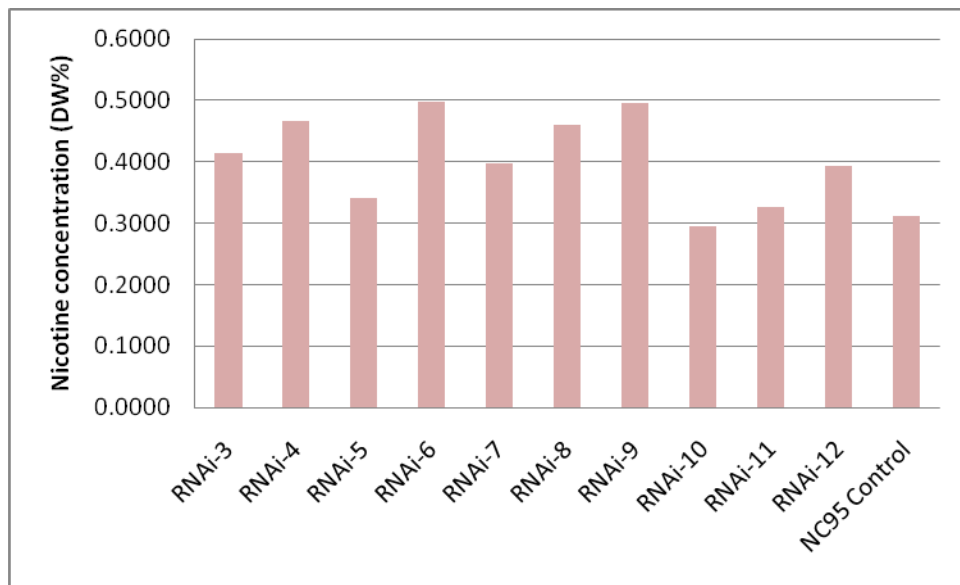


Figure 2.17 Nicotine concentration in *NtERF98* RNAi transgenic tobacco lines and wild type control. DW: dry weight.

Overall, *NtERF98* appears to play a role in modulating *QPT* expression. It is intriguing that both over-expression and suppression of its expression generally led to a reduction of the *QPT* mRNA level and to a slight to moderate increase of nicotine concentration. The observation underscores the complex nature of the regulation of the nicotine biosynthesis

pathway. Investigation of the relevant gene expression and nicotine level of the transgenic plants after topping or wounding treatment may provide more insight into the role of *NtERF98* in nicotine biosynthesis.

### **Transcription factor *NtETTA* gene**

The full length cDNA sequence of the transcription factor *NtETTA* gene was 2429 bp in length and encodes a protein of 739 AA. BLAST analysis against NCBI GenBank database showed that the gene belongs to the *ARF* (auxin responsive factor) gene family, and shares 99.8% identity (1008/1010) of nucleotides with an EST sequence named *NtETTA* gene (GenBank No. DQ340256) in the database, and thus the cloned TF factor gene is called *NtETTA*.

To infer the evolutionary relationship of the *NtETTA* with its homologues in other plant species, a phylogenetic tree was constructed by using the maximum parsimony method in MEGA4.1 based on the amino acid sequences of the *NtETTA* and other proteins that show the highest homology with *NtETTA* in the NCBI BLAST hits. The *NbARF1* from *Nicotiana benthamiana* was included because it was shown to be an ARF TF involved in nicotine biosynthesis (Todd et al. 2010). Figure 2.18 shows that the *NtETTA* is most closely related to *NtETTb* (98% identity, GenBank No. DQ340257) of tobacco, and is only distantly related to the reported *NbARF1* (50% identity).

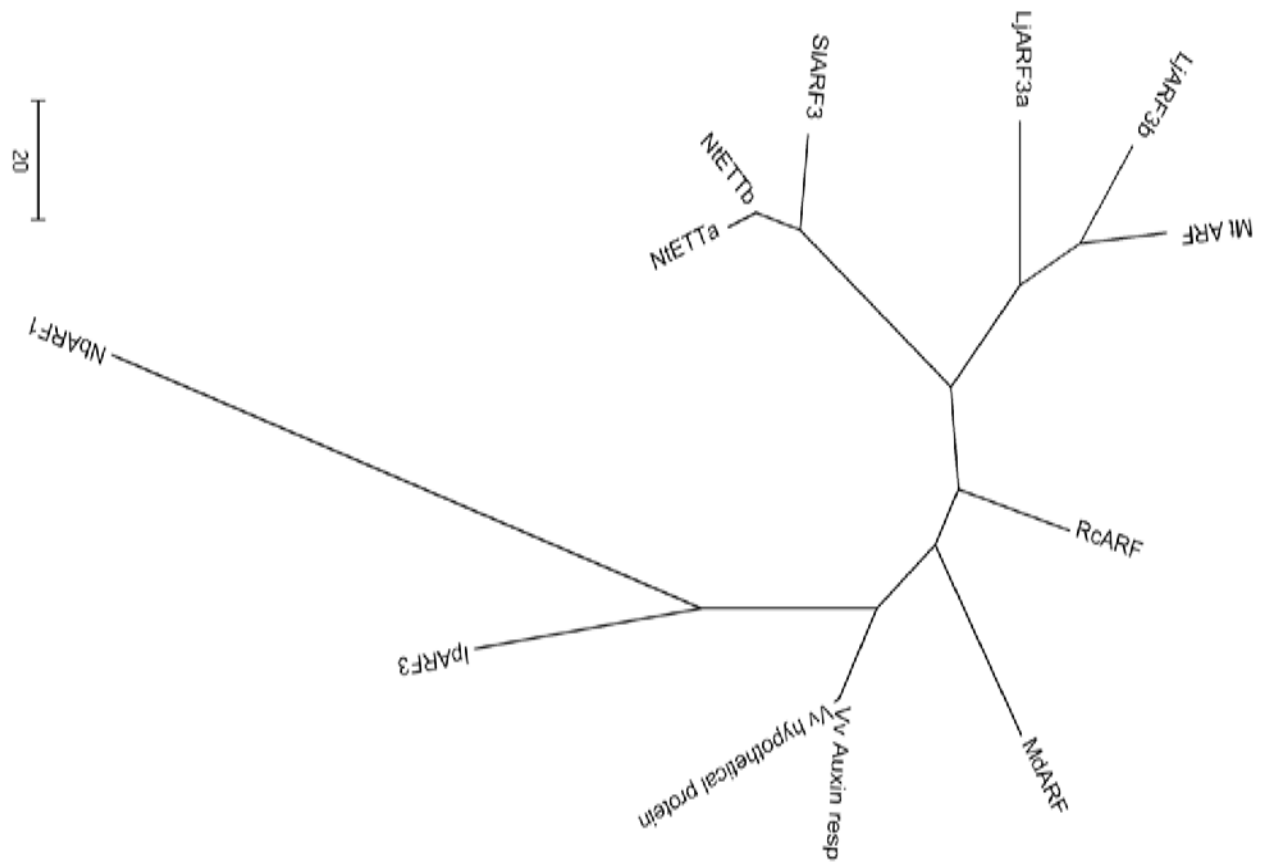


Figure 2.18 Unrooted phylogenetic tree of NtETTA and ten other related auxin responsive factors.

*Expression pattern of NtETTA in tobacco plant*

The expression of the *NtETTA* gene in mature tobacco plant was examined by Northern analysis. Figure 2.19 shows that it has higher expression in stem and low expression in root, leaf, and flower.

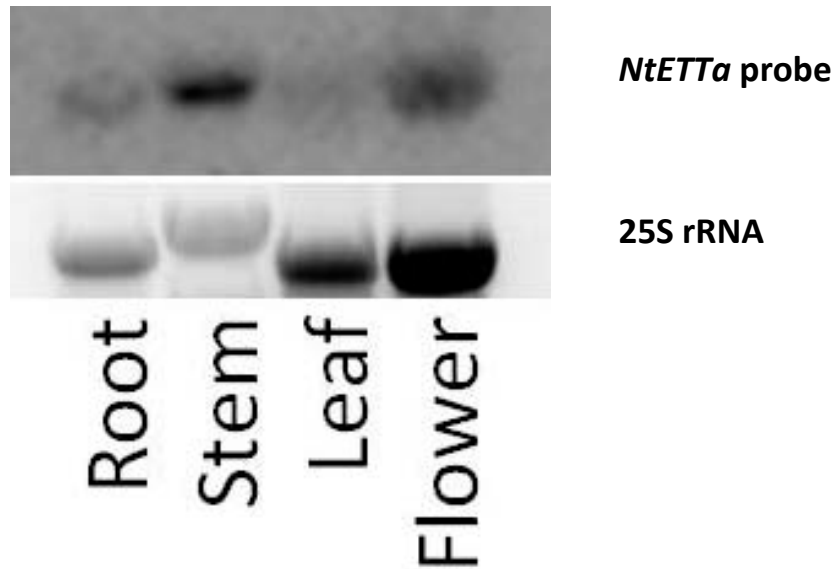


Figure 2.19 Northern blot analysis of *NtETTα* expressions in tobacco root, stem, leaf, and flower. The 25S rRNA stained with EtBr on the gel is shown as a loading reference.

*NtETTα expression in tobacco root after topping, wounding, or MeJA treatment*

The three treatments were also applied to test the responses of the *NtETTα* gene expression in tobacco root. Figure 2.20 shows a clear up-regulation of *NtETTα* gene expression 0.25 h after any of these treatments with mRNA level increased more by topping and wounding. However, unlike the *MYC2* genes, no clear biphasic increase pattern was observed for *NtETTα* gene expression. MeJA seems to be the least effective treatment with respect to *NtETTα* induction.

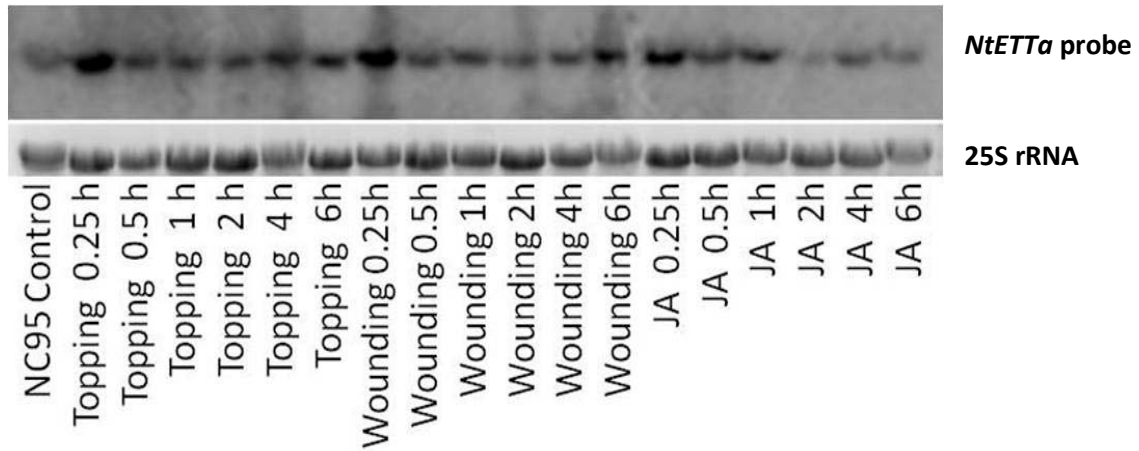


Figure 2.20 Time course of *NtETTA* expression response in tobacco root after topping, wounding, or MeJA treatment. Shown are the Northern blot hybridization of *NtETTA* (top panel) and the 25S rRNA stained with EtBr as a loading reference (bottom panel).

*Analysis of transgenic plants that over- or under-express NtETTA gene*

Four *NtETTA* over-expression and three RNAi transgenic tobacco lines were obtained. Northern blot analysis was undertaken to characterize these lines. Figure 2.21 shows that all the four over-expression lines had higher *NtETTA* expression levels with OE-6 and OE-7 being much higher than the controls, and three RNAi lines showed almost no detectable expression.

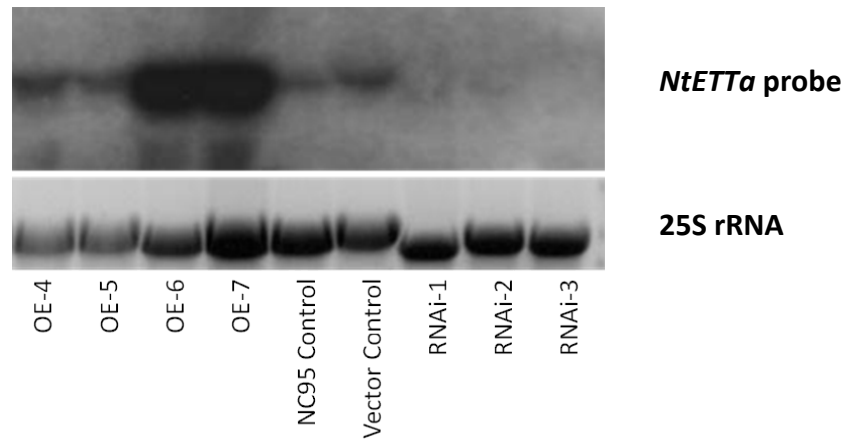


Figure 2.21 . Expression level of *NtETTA* in over-expression lines and RNAi lines. Shown are Northern hybridization of *NtETTA* (top panel) and the 25S rRNA stained with EtBr as a loading reference (bottom panel).

To test the effect of *NtETTA* on *QPT* gene expression, Northern analysis of these over-expression and RNAi lines was performed. Figure 2.22 shows that both *NtETTA* over-expression lines and RNAi lines had lower *QPT* expression levels when compared to the controls. Two RNAi lines (RNAi-1 and RNAi-2) exhibited the lowest *QPT* expression level.

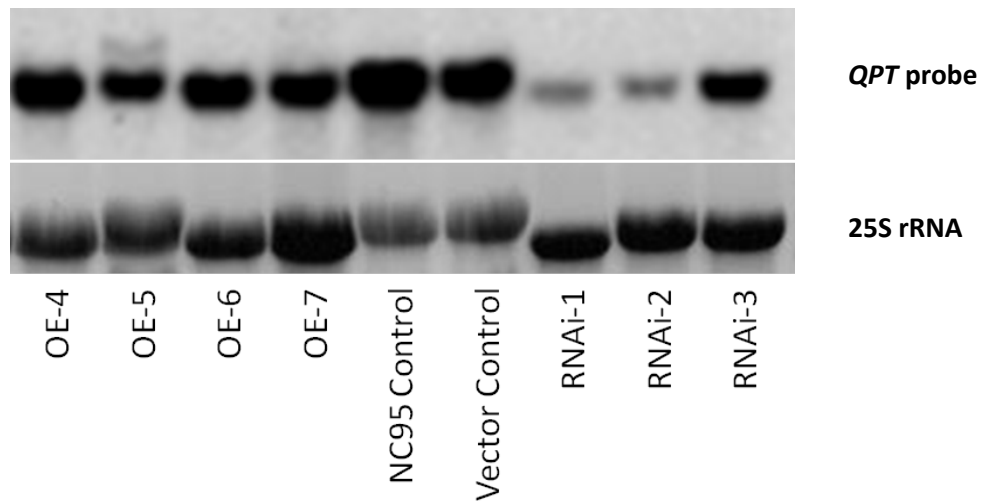


Figure 2.22 *QPT* expressions in *NtETTA* over-expression and RNAi lines. Shown are Northern blot hybridization of *QPT* (top panel) and the 25S rRNA stained with EtBr as a loading reference (bottom panel).

For all those lines, nicotine concentration in the leaf was determined. Figure 2.23 shows that OE-4, -5 and RNAi-2, -3 had approximately 25% lower nicotine level than the controls while the nicotine concentration in OE-6 was slightly higher (25%). Taken together, it can probably be concluded that there were no significant changes in nicotine level in these lines.

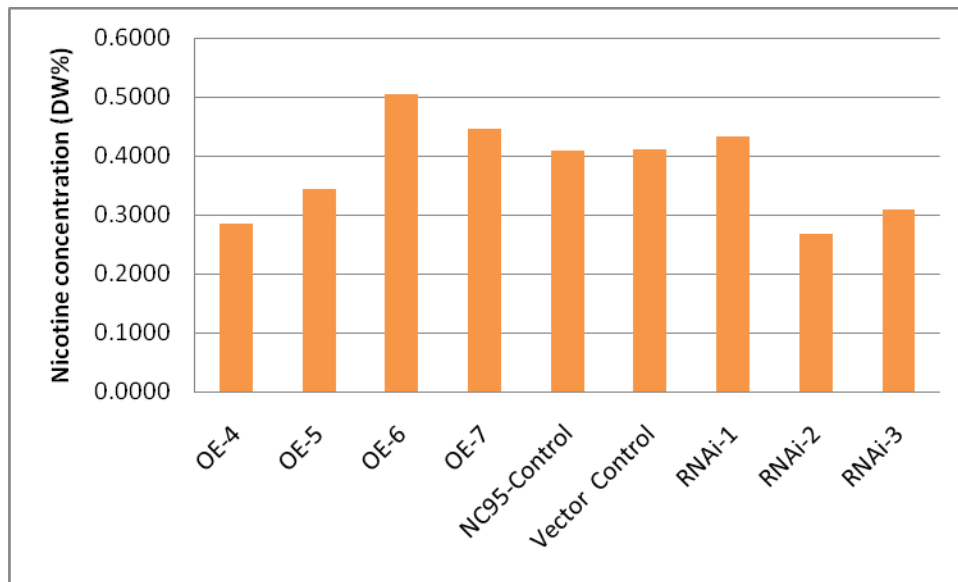


Figure 2.23 Nicotine concentrations of *NtETTA* over-expression and RNAi lines. The line numbers are the same as in the previous figures. DW: dry weight.

The expression of the *ARF* transcription factor gene isolated in this experiment, *NtETTA*, was quickly induced by topping, wounding, and MeJA treatment. It affects *QPT* expression as revealed in either over-expression or RNAi transgenic lines. It was most explicitly shown in RNAi-1 and -2 lines: suppression of *NtETTA* expression could lead to severe reduction of *QPT* mRNA level. However, nicotine level had little changes among the transgenic lines. This again points to the complicated control of nicotine biosynthesis, and to our previous observation that nicotine level is not correlated to *QPT* (and probably *PMT* too) mRNA level.

It remains to be seen whether stimulation like topping, wounding, or MeJA treatment would make the transgenic plants perform more differently than the control plants, in terms of *PMT* and *QPT* mRNA levels and the nicotine accumulation, so the role of *NtETTα* in nicotine biosynthesis pathway could be further elucidated.

## DISCUSSION

### ***QPT2* promoter and yeast one-hybrid screening**

Among the major nicotine synthesis-related genes, only the promoter sequence of *NtPMT1a* was analyzed in detail, and three basic transcription regulatory motifs were identified. They are a G-box (GCACGTTG, -103 to -96 bp from transcription initiation site), a GCC-like box (TGCGCCC, -62 to -56 bp) and an AT rich region in between (-80 to -69 bp, 92.8% A and T bases). These elements were demonstrated to be important in regulation of nicotine accumulation under JA treatment, and were named GAG motif (Timko et al. 2010). Although the GAG motif was found in all *NtPMT* gene promoters, a BLAST search could not identify such a motif in other gene promoters in public available database (Timko et al. 2010), suggesting that the motif is a critical element required to coordinate expression of the *NtPMT* family members. It also implies that the expression of other nicotine synthesis pathway genes may not be regulated in the same way although their expression has to be somewhat coordinated to produce nicotine efficiently. The tobacco *QPT2* gene promoter was used in this study as a bait to isolate transcription factors that bind to the *QPT2* promoter and regulate its gene expression. The disadvantage is that the *QPT2* gene promoter has not previously been analyzed and characterized in detail. A sequence scanning analysis of the promoter found that a “G-box” sequence (AACGTG) lies at -205 bp upstream of the translational start site ATG as predicted by the web-based software (plantCARE, <http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). Whether this sequence functions as a G-box, which is the binding site of MYC2 TF, remains to be experimentally demonstrated.

In the yeast one-hybrid system, a library of proteins fused with the activation domain of GAL4 are screened by their binding to the bait promoter sequence and activates the reporter gene expression (*His3* in this case). As recommended in the user manual, at least three tandem copies of the target sequence (bait) should be included in the bait vector upstream of the reporter gene to increase the chance for binding. Usually, cis-elements less than 20 bp are

used to make the tandem copies and the total bait is around 100 bp long. In this study, because the *QPT* promoter region was not well characterized, the 1kb *QPT2* gene promoter was used as a bait. A total of around 100 positive colonies were isolated and their sequences were determined and analyzed. Except for the isolated five TFs and a couple of histone-like proteins (which bind to DNA unspecifically), most of the clones are unrelated false positives. The overall poor efficiency of our system maybe mostly due to the length of the *QPT2* promoter sequence (> 1 kb) and the lack of the tandem repeats. In spite of these disadvantages, three out of five TF genes isolated from the system have roles in regulating genes in the nicotine biosynthesis pathway.

#### **Four transcription factors involved in nicotine biosynthesis were isolated**

A total of four TFs were isolated from this project, which are involved in the nicotine biosynthesis pathway. They modulate the *QPT* mRNA level, and affect nicotine concentration. Their expression in roots is responsive to topping, wounding, and MeJA treatments in a similar manner: their expression levels change as quickly as within 15 min upon the treatments. *NtMYC2a and 2b* are positively induced by the treatments and showed two induction peaks within the first six hours after treatments whereas *NtERF98* mRNA level was negatively regulated and exhibited biphasic decreases within six hours. *NtETTα* expression was also induced by the treatments within 15 min but did not show a clear biphasic pattern. The long-distance movement signal that causes such rapid changes (within 15 min) in gene expression level in roots is still unknown. Neither JA nor auxin seems to move that fast within plants (Hertel and Flory 1968; Baldwin et al. 1997; Shi et al. 2006). From another perspective, wounding can induce production of smRNA. Recent research shows that silenced RdR1 (RNA directed RNA polymerase) in *Nicotiana attenuate* made the plants susceptible to herbivores (Pandey et al. 2008). This result suggests that the smRNA is involved in the wounding-induced response of tobacco plant and may be a good fast-moving signal candidate.

The fact that these TF genes express in all the tissues tested is not at all surprising. All these TFs are likely involved in the JA signaling pathway, which regulates not only nicotine (or other alkaloids) biosynthesis, but also many other physiological and developmental processes, such as root growth, fertility, resistance to diseases and even responses to abiotic stress like drought (Kazan and Manners 2008). Interestingly, despite their potential multiple roles in plant growth, transgenic plants with constitutive over-expression or down-regulation of each of these four TF genes all looked normal when growing in greenhouse.

### **Transcription factors *NtMYC2a*, *NtMYC2b* are positive regulators in nicotine biosynthesis**

MYC2 has been considered the core TF in JA signaling pathway by regulating a cascade of transcription factors in plant responses to JA. It is known that MYC2 binds to the G-box motif of a promoter (Dombrecht et al. 2007). Because all the *PMT* gene promoters contain the G-box (Timko et al. 2010) and the *QPT* promoter has a “G-box” sequence, it is likely that MYC2 directly regulates expression of these two key genes. In two over-expression lines of *NtMYC2a*, the constitutive nicotine level (untreated) was more than doubled attaining about 1% of leaf dry weight, which is very high. *NtMYC2b* also facilitated approximately 50% increase in nicotine level in three over-expression lines. On the other hand, three RNAi lines, in which it is likely that both *NtMYC2a* and *b* were down-regulated (the probe used in Northern blot analysis cannot distinguish the two), had approximately a five-fold decrease in nicotine level accompanied by approximate 10-fold decrease in mRNA levels of *PMT* and *QPT* genes. All indicate a positive regulator role of *NtMYC2a* and *b* in nicotine biosynthesis.

However, what puzzled us was the dramatic decrease of both *PMT* and *QPT* mRNA levels in the *NtMYC2a* over-expression lines, which may indicate a negative feedback loop, when the nicotine level reaches a certain threshold level, to prevent an uncontrollable runaway situation (Kazan and Manners, 2008). Alternatively, although the mRNAs of *PMT* and *QPT* were down-regulated by an unknown mechanism related to *NtMYC2a* over-expression, their protein levels were increased or their enzymes were more active to enhance nicotine

production. The latter hypothesis is testable by measuring the enzyme activities or the protein concentration in the transgenic plants. In addition, our results also reveal a somewhat diverse function between the 2a and 2b isoforms, because *PMT* and *QPT* mRNA levels were not affected much in the *NtMYC2b* over-expression lines. The phenomenon may be explained by a negative regulatory loop found in *Arabidopsis*: AtMYC2 was demonstrated to up-regulate repressor JAZ protein, which in turn binds, and represses the activity of, AtMYC2 (Staswick 2007). Similarly, high levels of *NtMYC2b* wouldn't be able to activate the *PMT* and *QPT* expression if it up-regulates *NtJAZ* expression. The hypothesis can be tested by measuring *NtJAZ* expression in *NtMYC2b* over-expression lines and by treating the over-expression lines with MeJA and measuring the expression of the *PMT* and *QPT* genes. Overall, *NtMYC2b* seems to be the "weaker" isoform between the two because, in its over-expression lines, its mRNA levels increased approximately 10-fold and the nicotine level was only enhanced by about 50%. In the future, use of qRT-PCR instead of Northern analysis will provide more insights into the two *MYC2* genes and distinguish better between their expression and functions. Moreover, because *MYC2* plays a key role in the expression of JA responsive genes, it is necessary to evaluate nicotine level and relevant gene expression with topping, wounding, or MeJA treatments in the over-expression and RNAi transgenic plants.

The TF genes, *NbbHLH1* and *NbbHLH2*, from the same subgroup of the *bHLH* family were recently isolated and characterized in *N. benthamiana* (Todd et al. 2010). *NbbHLH1* shares 70% AA homology with the *NtMYC2b* while *NbbHLH2* has 96% AA identity with *NtMYC2a* gene. *NbbHLH1* and 2 bind at the G-box sequence of the *NbPMT* gene promoter, and positively regulate nicotine biosynthesis in transgenic tobacco plants. However, Todd et al. (2010) were also puzzled by the observation of little change or even reduction of mRNA levels of most nicotine biosynthesis genes, including *QPT*, in *NbbHLH1* and *NbbHLH2* over-expression transgenic tobacco plants. This report verifies that our observation was not an exception or experimental error, and points to a more complicated regulatory network in nicotine biosynthesis pathway. Timko et al. (2010) reported isolation of three *NtMYC2* isoform genes, but the detailed description of the work has not yet been published.

## Possible roles of transcription factors NtERF98 and NtETTa in nicotine biosynthesis

Ethylene and JA signaling interaction is rather complex and could be both synergistic and antagonistic (Kazan and Manners, 2008). Both repressor- and activator-type ethylene reaction factors (ERFs) have been reported (McGrath et al. 2005). In tobacco, ethylene is shown to have negative effects on nicotine biosynthesis (Wang et al. 1994; Shoji et al. 2000b; Winz and Baldwin 2001). Our research revealed that NtERF98 has a rather complicated role in nicotine biosynthesis. *NtERF98* is a negative regulator of nicotine synthesis in the sense that its expression is down-regulated by all three treatments that stimulate nicotine synthesis, and that the nicotine levels increased by up to 50% in a majority of its RNAi lines. In addition, there was a good negative correlation between *NtERF98* and *QPT* mRNA levels in most of the over-expression lines. However, in those lines, while *QPT* was down-regulated, nicotine level was unchanged or increased only moderately. Moreover, in the RNAi lines where nicotine levels were generally increased, the *QPT* mRNA levels were reduced slightly. Overall *NtERF98* may have a modifying function on nicotine level but may be more responsible for modulating *QPT* mRNA level. This TF may be involved in the tight control of *QPT* mRNA level as reported in Chapter 1 and in the observation in this chapter that over-expression of *NtMYC2a* led to a higher nicotine level but lower *QPT* mRNA. It would be informative to see whether *NtMYC2a* regulates *NtERF98* expression because it was demonstrated that *AtMYC2* does regulate expression of some *ERF* genes in *Arabidopsis* (Dombrecht et al. 2007). The hypothesis can be tested in both *NtMYC2a* over-expression and RNAi lines.

All the *ERF* transcription factor genes reportedly involved in nicotine biosynthesis other than *NtERF98* belong to subfamily IX of *AP2/ERF* family, such as *NtORC1*, *NtJAP* (De Satter et al. 2005), and *NtERF189*, *115*, *17,221*, *104*, *179*, and *168* (Shoji et al. 2010). Most of them have positive regulatory roles in nicotine production, although a few have only minor effects. Another *ERF* gene, *NbERF1*, was recently shown to be a positive regulator as well (Todd et al. 2010). Suppression of this gene expression by VIGS lowered nicotine level when the plants were treated with MeJA. *NbERF1* shares 82% AA identity with *NtORC1* and should

also belong to subfamily IX. *NtERF98* is the first repressor-type *ERF* isolated in nicotine biosynthesis regulation. *NtERF98* belongs to subfamily VII and does not share high homology with those in subfamily IX.

The Auxin and JA signaling pathways are closely interlinked. It has been shown that one phytohormone activates biosynthesis genes of the other and vice versa. Moreover, at least two ARFs are required for JA biosynthesis and plant fertility (for review, see Kazan and Manners, 2008). Recently it was reported that the *JAZ1* repressor gene is activated by both JA and auxin (Grunewald et al. 2009). In our experiment, the cloned ARF TF gene, *NtETT**a*, appears to positively regulate *QPT* expression as shown in the RNAi lines where *QPT* mRNA was reduced approximately 15 fold, indicating that *NtETT**a* is required for *QPT* expression. However, over-expression of *NtETT**a* did not increase *QPT* expression and often moderately reduced its mRNA level, implicating a role of *NtETT**a* in the complicated tight control of *QPT* mRNA level. Recently, an auxin responsive transcription factor, *NbARF1*, was reported as a negative regulator in nicotine synthesis. Suppression of *NbARF1* by VIGS significantly enhanced nicotine level in untreated plants (Todd et al. 2010). *NtETT**a* is an auxin responsive transcription factor but it acts as a positive regulator on *QPT* expression although it appears to have minor effects on nicotine level. The phylogenetic tree analysis indicates a distant relationship between *NtETT**a* and *NbARF1*.

In summary, we have cloned four TF genes in this study. They are *NtMYC2a* and *NtMYC2b* from the *bHLH* family, *NtERF98* from the *AP2/ERF* family, and *NtETT**a* from the *ARF* family. They are all involved in regulation of root *QPT* mRNA level and/or leaf nicotine level, with *NtMYC2a* and *NtMYC2b* having a more positive effects, and *NtERF98* and *NtETT**a* being rather more complicated modulators. *NtMYC2a* appears to play a more important role in regulating nicotine synthesis: Over-expression increases constitutive nicotine level by up to 2.5- fold whereas down-regulation of the gene (together with *NtMYC2b*) reduces it by five fold. Future studies of the transgenic over-expression and RNAi lines with MeJA treatment are necessary to provide more insights into the functions and mechanisms of these TFs in nicotine biosynthesis. Together with TFs isolated from other

laboratories, more than a dozen TFs have been reportedly having roles in affecting nicotine biosynthesis, indicating the complexity of the pathway regulation. Our research also revealed that higher nicotine synthesis is not always associated with higher mRNA levels of a key pathway gene, *QPT*, pointing to a negative feedback loop and/or possible translational and/or posttranslational control in the pathway. In addition, involvement of smRNA in modulating the pathway gene mRNA levels cannot be excluded. From all of our transgenic plants reported in this chapter and the first chapter, we never observed a plant with substantial increase in *QPT* mRNA level, implicating an extremely tight control of this transcript as pointed in Chapter 1. It also shows differential control between *PMT* and *QPT*, the two key structural genes in nicotine biosynthesis. Overall, the research suggests a very complicated network that regulates nicotine synthesis pathway genes and the nicotine accumulation.

## REFERENCES

- Arimura G, Kost C, Boland W (2005). Herbivore –induced, indirect plant defences. *Biochimica et Biophysica Acta* 1734: 91-111.
- Baldwin IT (1988). The alkaloidal responses of wild tobacco to real and simulated herbivory. *Oecologia* 77: 378-381.
- Baldwin IT (1989). Mechanism of damage-induced alkaloid production in wild tobacco. *Journal of Chemical Ecology* 15: 1661-1680.
- Baldwin IT, Schmelz EA, Ohnmeiss TE (1994). Wound-induced changes in root and shoot jasmonic acid pools correlate with induced nicotine synthesis in *Nicotiana sylvestris* spegazzini and comes. *Journal of Chemical Ecology* 20: 2139-2157.
- Baldwin IT, Zhang ZP, Diab N, Ohnmeis ET, McCloud ES, Lynds GY, Schmelz (1997). Quantification, correlations and manipulations of wound-induced changes in jasmonic acid and nicotine in *Nicotiana sylvestris*. *Planta* 201: 397-404
- Boter M, Ruiz-Rivero O, Abdeen A, Prat Salome (2004). Conserved MYC transcription factors play a key role in jasmonate signaling both in tomato and Arabidopsis. *Gene and Dev* 18: 1577-1591.
- Bush LP, Saunders JL (1977). Accumulation, manipulation and regulation of nicotine level in tobacco. *Proceedings of American Chemical Society Symposium*. 173: 388-425.
- Cane KA, Mayer M, Lidgett AJ, Michael AJ, Hamill JD (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: 305-320.

Chen G, Hu Z, Grierson D (2008). Differential regulation of tomato ethylene responsive factor Le ERF3b, a putative repressor, and the activator Pti4 in ripening mutants and in response to environmental stresses. *Journal of Plant Physiol* 165: 662-670.

Chini A, Fonseca S, Fernandez G, Adie B, Chico JM, Lorenzo O, Garcia-Casado G, Lopez-Vidriero , Lozano FM, Ponce MR, Micol JL, Solano R (2007). The JAZ family of repressors is the missing link in jasmonate signaling. *Nature* 448: 666-671.

Creelman RA, Mullet JE (1995). Jasmonic acid distribution and action in plants: regulation during development and response to biotic and abiotic stress. *Proc Natl Acad Sci USA* 92(10): 4114-4119.

De Sutter VD, Vanderhaeghen R, Tilleman S, Lammertyn F, Vanhoutte I, Karimi M, Inze D, Goossens A, Hilson P (2005). Exploration of jasmonate signaling via automated and standardized transient expression assays in tobacco cells. *Plant J* 44: 1065-1076.

Dombrecht B, Xue GP, Sprague SJ, Kirkegaard JA, Ross JJ, Reid JB, Fitt GP, Sewelam N, Schenk PM, Manners JM, Kazan K (2007). MYC2 differentially modulates diverse jasmonate-dependent functions in *Arabidopsis*. *Plant Cell* 19: 2225-2245.

Feth F, Wagner R, Wagner KG (1986). Regulation in tobacco callus of enzyme activities of the nicotine pathway. *Planta* 168: 402-407.

Fujimoto SY, Ohta M, Usui A, Shinshi H, Ohme-Takagi M (2000). *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. *Plant Cell* 12: 393-404.

Goossens A, Hakkinen ST, Laakso I, Seppanen-Laakso T, Biondi S, Sutter VD, Lammertyn F, Nuutila AM, Soderlun H, Zabeau M, Inze D, Okasma-Galdentey KM (2003). A functional genomics approach toward the understanding of secondary metabolism in plant cells. *Proc Natl Acad Sci USA* 100: 8595-8600.

- Grunewald W, Vanhome B, Pauwels L, Plovie E, Inze D, Gheysen G, Goossens A(2009). Expression of the Arabidopsis jasmonate signalling repressor *JAZ1/TIFY10A* is stimulated by auxin. *EMBO Rep.* 10: 923-928.
- Guerineau F, Benjdia M, Zhou DX (2003). A jasmonate-responsive element within the *A. thaliana* *vsp1* promoter. *Journal of Experimental Botany* 54: 1153-1162.
- Guo H, Ecker JR (2004). The ethylene signaling pathway: new insights. *Current Opinion in Plant Biology* 7: 40-49.
- Guo ZE, Chen XJ, Wu XL, Ling JQ, Xu P (2004). Over-expression of the AP2/EREBP transcription factor OPBP1 enhances disease resistance and salt tolerance in tobacco. *Plant Mol Biol* 55: 607-618.
- Hao D, Ohme-Takagi M, Sarai A (1998). Unique mode of GCC box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF domain) in plant. *The Journal Biological Chemistry* 273: 26857-26861.
- Hertel R, Flory R (1968). Auxin movement in corn coleoptiles. *Planta* 82: 123-144.
- Hibi N, Higashiguchi S, Hashimoto T, Yamada Y (1994). Gene expression in tobacco low-nicotine mutants. *Plant Cell* 6: 723-735.
- Howe GA (2004). Jasmonates as signals in the wound response. *Journal of Plant Growth Regulation* 23: 223-237.
- Imanishi S, Hashizume K, Nakakita M, Kojima H, Matsubayashi Y, Hashimoto T, Sakagami Y, Yamada Y, Nakamura K (1998). Differential induction by methyl jasmonate of genes encoding ornithine decarboxylase and other enzymes involved in nicotine biosynthesis in tobacco cell cultures. *Plant Mol Biol* 38: 1101-1111.

- Okamuro JK, Caster B, Villarroel R, Montagu MV, Jofuku D (1997). The AP2 domain of APETALA2 defines a large new family of DNA binding protein in Arabidopsis. *Proc Natl Acad Sci USA* 94: 7076-7081.
- Kajikawa M, Hirai N, Hashimoto T (2009). A PIP family protein is required for biosynthesis of tobacco alkaloids. *Plant Mol Biol* 69: 287-298.
- Katoh A, Ohki H, Hashimoto T (2005). Molecular regulation of nicotine biosynthesis. *Plant Biotechnology* 22: 389-392.
- Katoh A, Shoji T, Hashimoto T (2007). Molecular cloning of N-methylputrescine oxidase from tobacco. *Plant Cell Physiol* 48: 550-554.
- Kazan K, Manners JM (2008). Jasmonate signaling: toward an integrated view. *Plant Physiol* 146: 1459-1468.
- Kidd SK, Melillo AA, Lu RH, Reed DG, Kuno N, Uchida K, Furuya M, Jelesko JG (2006). The A and B loci in tobacco regulate a network of stress response genes, few of which are associated with nicotine biosynthesis. *Plant Mol Biol* 60: 699-716.
- Koo AJK, Howe GA (2009). The wound hormone jasmonate. *Phytochemistry* 70: 1571-1580.
- Koyama T, Okada T, Kitajima S, Ohme-Takagi M, Shinshi H, Sato F (2003). Isolation of tobacco ubiquitin- conjugating enzyme c DNA in a yeast two-hybrid system with tobacco ERF3 as bait and its characterization of specific interaction. *Journal of Experimental Botany* 54: 1175-1181.
- Legg P D, Chaplin JF, Collins GB (1969). Inheritance of percent total alkaloids in *Nicotiana tabacum* L. *J. Hered.* 60: 213-217.
- Legg PD, Collins GB, Litton CC (1970). Registration of LA Burley 21 tobacco germplasm. *Crop Sci* 10: 212.

Legg PD, Collins GB (1971). Inheritance of percent total alkaloids in *Nicotiana tabacum* L. II. Genetic effects of two loci in Burley 21 x LA Burley 21 populations. *Can J Genet Cytol* 13: 287-291.

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998). Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell* 10: 1391-1406.

Lorenzo O, Chico JM, Sanchez-Serrano JJ, Solano R (2004). Jasmonate-insensitive 1 encodes a MYC transcription factor essential to discriminate between different jasmonate-regulated defense responses in *Arabidopsis*. *Plant Cell* 16: 1938-1950.

Lorenzo O, Piqueras R, Sanchez-Serrano JJ, Solano R (2003). Ethylene response factor 1 integrates signals from ethylene and jasmonate pathways in plant defense. *Plant Cell* 15: 165-178.

Magnani E, Sjolander K, Hake S (2004). From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. *Plant Cell* 16: 2265-2277.

McGrath KC, Dombrecht B, Manners JM (2005). Repressor- and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of *Arabidopsis* transcription factor gene expression. *Plant Physiol* 139: 949-959.

Menke FLH, Champion A, Kijne JW, Memelink J (1999). A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene *Str* interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. *EMBO J* 18: 4455-4463.

Morita M, Shitan N, Sawada K, Van Montagu MCE, Inze D, Rischer H, Goossens A, Oksman-Caldentey K-M, Moriyama Y, Yazaki K (2009). Vacuolar transport of nicotine is

mediated by a multidrug and toxic compound extrusion (MATE) transporter in *Nicotiana tabacum*. *Proc Natl Acad Sci USA* 106: 2447-2452.

Ohme-Takagi M, Shinshi H (1995). Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell* 7: 173-182.

Ohta M, Matsui K, Hiratsu K, Shinshi H, Ohme-Takagi M (2001). Repression domains of class II ERF transcriptional Repressors share an essential motif for active repression. *Plant Cell* 13: 1959-1968.

Ohta M, Ohme-Takagi M, Shinshi H (2000). Three ethylene-responsive transcription factors in tobacco with distinct transactivation functions. *Plant J* 22: 29-38.

Ouwerkerk PBF Memelink J (1999). A G-box element from the *Catharanthus roseus* strictosidine synthase (Str) gene promoter confers seed- specific expression in transgenic tobacco plants. *Mol Gen Genet* 261: 635-643.

Pandey SP, Shahi P, Gase K, Baldwin IT (2008). Herbivory-induced change in the small-RNA transcriptome and phytohormone signaling in *Nicotiana attenuate*. *Proc Natl Acad Sci USA* 105:4559-4564

Paschold A, Halitschke R, Baldwin IT (2007). Co (i)-ordinating defenses: NaCOI1 mediates herbivore induced resistance in *Nicotiana attenuate* and reveals the role of herbivore movement in avoiding defenses. *Plant J* 51: 79-91.

Reed DG, Jelesko JG (2004). The A and B loci of *Nicotiana tabacum* have non-equivalent effects on the mRNA levels of four alkaloid biosynthesis genes. *Plant Sci* 167: 1123-1130.

Rushton PJ, Bokowiec MT, Han S, Zhang H, Brannock JF, Chen X, Laudeman TW, Timko MP (2008). Tobacco Transcription factors: novel insights into transcriptional regulation in the solanaceae. *Plant Physiol* 147: 280-295.

- Sachan N, Falcone DL (2002). Wound-induced gene expression of putrescine N – methyltransferase in leaves of *Nicotiana tabacum*. *Phytochemistry* 61: 797-805.
- Saunders JW, Bush LP (1979). Nicotine biosynthetic enzyme activities in *Nicotiana tabacum* L. genotypes with different alkaloid levels. *Plant Physiol* 64: 236-240.
- Sessa G, Meller Y, Fluhr R (1995). A GCC element and a G-box motif participate in ethylene-induced expression of the PRB-1b gene. *Plant Mol Biol* 28: 145-153.
- Sheen SJ (1988). Detection of nicotine in foods and plant materials. *J Food Sci* 53(5): 1572-1573.
- Shinshi H (2008). Ethylene-regulated transcription and crosstalk with jasmonic acid. *Plant Sci* 175: 18-23.
- Shinshi H, Usami S, Ohme-Takagi M (1995). Identification of an ethylene-responsive region in the promoter of a tobacco class I chitinase gene. *Plant Mol Biol* 27: 923-932.
- Shi Q, Li C, Zhang F (2006). Nicotine synthesis in *Nicotiana tabacum* L. induced by mechanical wounding is regulated by auxin. *J Exp Bot* 57: 2899-2907.
- Shoji T, Yamada Y, Hashimoto T (2000a). Jasmonate induction of putrescine N-methyltransferase genes in the root of *Nicotiana glauca*. *Plant Cell Physiol* 41: 831-839.
- Shoji T, Nakajima K, Hashimoto T (2000b). Ethylene suppresses jasmonate –induced gene expression in nicotine biosynthesis. *Plant Cell Physiol* 41: 1072-1076.
- Shoji T, Winz R, Iwase T, Nakajima K, Yamada Y, Hashimoto T (2002). Expression patterns of two tobacco isoflavone reductase-like genes and their possible roles in secondary metabolism in tobacco. *Plant Mol Biol* 50: 427-440.
- Shoji T, Ogawa T, Hashimoto T (2008). Jasmonate- induced nicotine formation in tobacco is mediated by tobacco COI1 and JAZ genes. *Plant Cell Physiol* 49: 1003-1012.

Shoji T, Inai K, Yazaki Y, et al. (2009). Multidrug and toxic compound extrusion-type transporters implicated in vacuolar sequestration of nicotine in tobacco roots. *Plant Physiol* 149: 708-718.

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

Siminszky B, Gavilano L, Bowen SW, Dewey RE (2005). Conversion of nicotine to nornicotine in *Nicotiana tabacum* is mediated by CYP 82E4, a cytochrome P450 monooxygenase. *Proc Natl Acad Sci USA* 102: 14919-14924.

Solano R, Stepanova A, Chao Q, Ecker JR (1998). Nuclear events in ethylene signaling: a transcriptional cascade mediated by ethylene-insensitive 3 and ethylene-response-factor1. *Genes & Dev* 12: 3703-3714.

Solt ML (1957). Nicotine production and growth of excised tobacco root cultures. *Plant Physiol* 32: 480-484.

Staswick PE (2007). JAZing up jasmonate signaling. *Trends in Plant Sci* 13: 66-71.

Stepanova AN, Alonso JM (2009). Ethylene signaling and response: where different regulatory modules meet. *Curr Opin in Plant Biol* 12: 548-555.

Steppuhn A, Gase K, Krock B, Halitschke R, Baldwin IT (2004). Nicotine's defensive function in nature. *PLoS Biology* 2: 1074-1080

Timko MP, Rushton PJ, Bokowiec MT, Zhang H (2010). Functional genomic approaches to harm reduction in tobacco products. *Recent Advances in Tobacco Science* 36: 25-39.

Todd AT, Liu E, Polvi SL, Pammatt RT, Page JE (2010). Functional genomics screen identifies diverse transcription factors that regulate alkaloid biosynthesis in *Nicotiana benthamiana*. *Plant J* 62: 589-600.

- Van der Fits L, Memelink J (2000). ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. *Science* 289: 295-297.
- Wagner R, Feth F, Wagner KG (1986). The regulation of enzyme activities of the nicotine pathway in tobacco. *Physiol Plant* 68: 667-672.
- Wang HL, Ho HH, Su JC (1994). Effect of ethylene on adventitious root formation and nicotine level of tobacco callus tissues. *Bot Bull Acad Sin* 35: 217-222.
- Wang SJ, Lan YI, Chen SF, Chen YM, Yeh KW (2002). Wound-response regulation of the sweet potato sporamin gene promoter region. *Plant Mol Biol* 48: 223-231.
- Wang SS, Shi QM, Li WQ, Niu JF, Li CJ, Zhang FS (2008). Nicotine concentration in leaves of flue-cured tobacco plants as affected by removal of the shoot apex and lateral buds. *Journal of Integrative Plant Biol* 50: 958-964.
- Winz RA, Baldwin IT (2001). Molecular interactions between the specialist herbivore *Manduca sexta* (Lepidoptera, Sphingidae) and its natural host *Nicotiana attenuata*. IV. Insect-induced ethylene reduces jasmonate-induced nicotine accumulation by regulating putrescine N-methyltransferase transcripts. *Plant Physiol* 125: 2189-2202.
- Xu B, Sheehan MJ, Timko MP (2004). Differential induction of ornithine decarboxylase (ODC) gene family members in transgenic tobacco (*Nicotiana tabacum* L. cv. Bright Yellow 2) cell suspensions by methyl-jasmonate treatment. *Plant Growth Regulation* 44: 101-106.
- Xu B, Timko MP (2004). Methyl jasmonate induced expression of the tobacco putrescine N-methyltransferase genes requires both G-box and GCC-motif elements. *Plant Mol Biol* 55: 743-761.
- Yan J, Zhang C, Gu M, Bai Z, Zhang W, Qi T, Cheng Z, Peng W, Luo H, Nan F, Wang Z, Xie D (2009). The *Arabidopsis* coronation insensitive 1 protein is a jasmonate receptor. *The Plant Cell* 21: 2220-2236.

Yasumatsu N (1967). Studies on the chemical regulation of alkaloid biosynthesis in tobacco plants. *Agricultural and Biological Chemistry* 31: 1441-1447.

## APPENDIX

**Appendix 1: *NtMYC2a* cDNA (full-length, 2214 bp)**

CACACACTCTCTCCATTTTCACTCACTCCTTATCACCAAACAATTCTTGGGTGTTT  
GAATATATACCCGAAATAATTTCTCTCTGTATCAAGAATCAAACAGATCTGAAT  
TGATTTGTCTGTTTTTTTTTCTTGATTTTGTATATGGAATGACGGATTATAGAAT  
ACCAACGATGACTAATATATGGAGCAATACTACATCCGATGATAATATGATGGA  
AGCTTTTTTATCTTCTGATCCGTCGTCGTTTTGGCCCGGAACAACACTACTACACCAA  
CTCCCCGGAGTTCAGTTTCTCCAGCGCCGGCGCCGGTGACGGGGATTGCCGGAG  
ACCCATTAAAGTCTATGCCATATTTCAACCAAGAGTCACTGCAACAGCGACTCCA  
GACTTTAATCGATGGGGCTCGCAAAGGGTGGACGTATGCCATATTTTGGCAATCG  
TCTGTTGTGGATTTTCGCGAGCCCCTCGGTTTTGGGGTGGGGAGATGGGTATTATA  
AAGGTGAAGAAGATAAAAATAAGCGTAAAACGGCGTCGTTTTCGCCTGACTTTA  
TCACGGAACAAGCACACCGGAAAAAGGTTCTCCGGGAGCTGAATTCTTTAATTT  
CCGGCACACAAACCGGTGGTGAAAATGATGCTGTAGATGAAGAAGTAACTGATA  
CTGAATGGTTTTTTCTGATTTCCATGACACAATCGTTTGTTAACGGAAGCGGGCT  
TCCGGGCCTGGCGATGTATAGTTCAAGCCCGATTTGGGTTACTGGAACAGAGAG  
ATTAGCTGTTTCTCACTGTGAACGGGCCCGACAGGCCCAAGGTTTCGGGCTTCAG  
ACTATTGTTTGTATTCCTTCAGCTAATGGTGTGTTGAGCTCGGGTCAACTGAGTT  
GATATTCCAGACTGCTGATTTAATGAACAAGGTTAAAGTTTTGTTTAATTTAAT  
ATTGATATGGGTGCGACTACGGGCTCAGGATCGGGCTCATGTGCTATTCAGGCCG  
AGCCCGATCCTTCAGCCCTTTGGCTGACTGATCCGGCTTCTTCAGTTGTGGAAGT  
CAAGGATTCGTCGAATACAGTTCCTTCAAGGAATACCAGTAAGCAACTTGTGTTT  
GGAAATGAGAATTCTGAAAATGGTAATCAAATCTCAGCAAACACAAGGATTT  
TTCCTAGGGAGTTGAATTTTTCCGAATATGGATTTGATGGAAGTAATACTCGGT  
ATGGAAATGGGAATGCGAATTCTTCGCGTTCTTGCAAGCCTGAGTCTGGTGAAT  
CTTGAATTTTGGTGATAGTACTAAGAGGAGTGCTTGCAGTGCAAATGGGAGCTTG  
TTTTCGGGCCAATCACAGTTCGGGCCCGGGCCTGCGGAGGAGAACAAGAACAAG  
AACAAGAAAAGGTCACCTGCATCAAGAGGAAGCAACGATGAAGGAATCCTTTCA  
TTTGTTCGGGTGTGATTTTGCCAAGTTCAAACACGGGGAAGTCCGGTGGAGGTG  
GCGATTCGGATCAATCAGATCTCGAGGCTTCGGTGGTGAAGGAGGCGGATAGTA  
GTAGAGTTGTAGACCCCGAGAAGAAGCCGAGGAAACGAGGGAGGAAACCGGCT  
AACGGGAGAGAGGAGCCATTGAATCATGTGGAGGCAGAGAGACAAAGGAGGGA  
GAAATTGAATCAAAGATTCTATGCACTTAGAGCTGTTGTACCAAATGTGTCAAAA  
ATGGATAAAGCATCACTTCTTGGTGATGCAATTGCATTTATCAATGAGTTGAAAT  
CAAAGTTCAGAATTCTGACTCAGATAAAGAGGACTTGAGGAACCAAATCGAAT  
CTTTAAGGAATGAATTAGCCAACAAGGGATCAAACACTATACCGGTCCTCCCCCGT

CAAATCAAGAACTCAAGATTGTAGATATGGACATCGACGTTAAGGTGATCGGAT  
GGGATGCTATGATTCGTATAACAATCTAATAAAAAGAACCATCCAGCCGCGAGGT  
TAATGACCGCTCTCATGGAATTGGACTTAGATGTGCACCATGCTAGTGTTCAGT  
TGTC AACGAGTTGATGATCCAACAAGCGACTGTGAAAATGGGAAGCCGGCTTTA  
CACGCAAGAACA ACTTCGGATATCATTGACATCCAGAATTGCTGAATCGCGATG  
AAGAGAAATACAGTAAATGGAAATTATCATAGTGAGCTCTGAATAATGTTATCT  
TTCATTGAGCTATTTTAAGAGAATTTCTCCT  
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

**NtMYC2a amino acid sequence (659 AA)**

MTDYRIPTMTNIWSNTTSDDNMMEAFLSSDPSSFWPGTTTT  
PTPRSSVSPAPAPVTGIAGDPLKSMPYFNQESLQQRLQTLID  
GARKGWTYAIFWQSSVVDFA SPV LGWGDGYYKGEEDKNK  
RKTASFSPDFITEQAHRKKVLRELSLISGTQTGGENDAVDE  
EVTDTEWFFLISMTQSFVNGSGLPGLAMYSSSPIWVTGTERL  
AVSHCERARQAQGFGLQTIVCIPSANGVVELGSTELIFQTAD  
LMNKVKVLFNFNIDMGATTGSGSGSCAIQAEPDPSALWLT  
PASSVVEVKDSSNTVPSRNTSKQLVFGNENSENGNQNSQQT  
QGFFTRELNFSEYGF DGSNTRYGNGNANSSRSCKPESGEILN  
FGDSTKRSAC SANGSLFSGQSQFGPGPAEENKNKNKKRSPA  
SRGSNDEGILSFVSGVILPSSNTGKSGGGGSDSDLEASVV  
KEADSSRVVDPEKKPRKRGRKPANGREEPLNHVEAERQRRE  
KLNQRFYALRAVVPNVSKMDKASLLGD AIAFINELKSKVQN  
SDSDKEDLRNQIESLRNELANKGSNYTGPPPSNQELKIVDM  
DIDVKVIGWDAMIRIQSNKKNHPAARLMTALMELDLDVHH  
ASVSVVNELMIQQATVKMGSRLYTQEQLRISLTSRIAESR

**Appendix 2: *NtMYC2b* cDNA (full-length, 2391bp)**

GTAACAAACCCTCTCCATTTTCACTCACTCCAAAAA ACTTTCTCTCTATTTTTTC  
TCTCTGTATCAAGAATCAAACAGATCTGAATTGATTTGGGAGTTTTTTTTCTTCTT  
GTTTTTGTATATGGAATGACGGACTATAGAATACCAACGATGACTAATATATGG  
AGCAATACAACATCCGACGATAACATGATGGAAGCTTTTTTATCTTCTGATCCGT  
CGTCGTTTTGGGCCGGAACAAATACACCAACTCCACGGAGTTCAGTTTCTCCGGC  
GCCGGCGCCGGTGACGGGGATTGCCGGAGACCCATTAAGTCGATGCCGTATTT  
CAACCAAGAGTCGCTGCAACAGCGACTCCAGACGTTAATCGACGGGGCTCGCGA  
AGCGTGGACTTACGCCATATTCTGGCAATCGTCTGTTGTGGATTTTCGTGAGCCCC  
TCGGTGTGGGGTGGGGAGATGGATATTATAAAGGAGAAGAAGACAAGAATAA  
GCGTAAAACGGCGGCGTTTTTCGCCTGATTTTATTACGGAGCAAGAACACCGGAA  
AAAAGTTCTCCGGGAGCTGAATTCTTTAATTTCCGGCACACAAACTGGTGGTGAA  
AATGATGCTGTAGATGAAGAAGTAACGGATACTGAATGGTTTTTTCTGATTTCAA  
TGACTCAATCGTTTGTAAACGGAAGCGGGCTTCCGGGCCTGGCTATGTACAGCTC  
AAGCCCGATTTGGGTTACTGGAAGAGAAAGATTAGCTGCTTCTCACTGTGAACG  
GGCCCGACAGGCCCAAGGTTTCGGGCTTCAGACTATGGTTTGTATTCCTTCAGCT  
AATGGTGTGTTGAGCTCGGGTCAACTGAGTTGATATTCCAGAGCGCTGATTTAA  
TGAACAAGGTTAAAATCTTGTTTGATTTTAATATTGATATGGGCGCGACTACGGG  
CTCAGGTTCCGGGCTCATGTGCTATTCAGGCTGAGCCCGATCCTTCAACCCTTTGG  
CTTACGGATCCACCTTCCTCAGTTGTGGAAGTCAAGGATTCGTCGAATACAGTTC  
CTTCAAGTAATAGTAGTAAGCAACTTGTGTTTGGAAATGAGAATTCTGAAAATGT  
TAATCAA AATTCTCAGCAAACACAAGGATTTTTCACTAGGGAGTTGAATTTTTCC  
GAATATGGATTTGATGGAAGTAATACTAGGAGTGGAATGGGAATGTGAATTCT  
TCGCGTTCTTGCAAGCCTGAGTCTGGCGAAATCTTGAATTTTGGTGATAGTACTA  
AGAGAAATGCTTCAAGTGCAAATGGGAGCTTGTTTTTCGGGCCAATCGCAGTTTCG  
GTCCCGGCCTGCGGAGGAGAACAAGAACAAGAACAAGAAAAGGTCACCTGCA  
TCAAGAGGAAGCAATGAAGAAGGAATGCTTTCATTTGTTTCGGGTGTGATCCTTGC  
CAAGTTCAAACACGGGGAAGTCCGGTGGAGGTGGCGATTTCGGATCATT CAGATC  
TCGAGGCTTCGGTGGTGAAGGAGGCGGATAGTAGTAGAGTTGTAGACCCCGAGA  
AGAGGCCGAGGAAACGAGGAAGGAAACCGGCTAACGGGAGAGAGGAGCCATTG  
AATCATGTGGAGGCAGAGAGGCAAAGGAGGGAGAAATTGAATCAAAGATTCTA  
TGCACTTAGAGCTGTTGTACCAAATGTGTCAAAAATGGATAAAGCATCACTTCTT  
GGTGATGCAATTGCATTTATCAATGAGTTGAAATCAAAGGTT CAGAATTCTGACT  
CAGATAAAGATGAGTTGAGGAACCAAATTGAATCTTTAAGGAATGAATTAGCCA  
ACAAGGGATCAA ACTATACCGGTCCTCCACCGCCAAATCAAGATCTCAAGATTG  
TAGATATGGATATCGACGTTAAAGTCATCGGATGGGATGCTATGATTCGTATACA  
ATCTAATAAAAAGAACCATCCAGCCGCGAGGTTAATGGCCGCTCTCATGGAATT

GGACTTAGATGTGCACCATGCTAGTGTTCAGTTGTCAACGAGTTGATGATCCAA  
CAAGCGACAGTGAAAATGGGGAGCCGGCTTTACACGCAAGAGCAGCTTCGGATA  
TCATTGACATCCAGAATTGCTGAATCGCGATGAAGAGAAATACAGTAAATGGAA  
ATTATTAGTGAGCTCTGAATAATGTTATCTTTCATTGAGCTATTTTAAGAGAATTT  
CTCCTATAGTTAGATCTTGAGATTAAGGCTACTTAAAAGTGGAAAGTTGATTGAG  
CTTTCCTCTTAGTTTTTTGGGTATTTTTCAACTTTTATATCTAGTTTGTTTTCCACA  
TTTTCTGTACATATAATGTGAAACCAATACTAGATCTCAAGATCTGGTTTTTAGTT  
CTGTAATTAGAAATAAATATGCAGCTTCATCTTTTTCTGTAAAAAAAAAAAAAAAA  
AAAAAAAAAAAA

**NtMYC2b amino acid sequence (658 AA)**

MTDYRIPTMTNIWSNTTSDDNMMEAFSSDPSSFWAGTNTPT  
TPRSSVSPAPAPVTGIAGDPLKSMFYFNQESLQQRLQTLIDG  
AREAWTYAIFWQSSVDFVSPSVLGGWGDGYKGEEDKNKR  
KTAAFSPDFITEQHRKKVLRELSLISGTQTGGENDAVDEE  
VTDTEWFFLISMTQSFVNGSGLPGLAMYSSSPIWVTGRERL  
AASHCERARQAQGFGLQTMVCIPSANGVVELGSTELIFQSA  
DLMNKVKILDFDNIDMGATTGSGSGSCAIQAEPDPSTLWLT  
DPPSSVVEVKDSSNTVPSSNSSKQLVFGNENSENVNQNSQQ  
TQGFFTRELNFSEYGFDSNTRSGNGNVNSSRSCKPESGEIL  
NFGDSTKRNASSANGSLFSGQSQFGPGPAEENKNKNKKRSP  
ASRGSNEEGMLSFVSGVILPSSNTGKSGGGGDSHSDLEAS  
VVKEADSSRVVDPEKRPRKRGRKPANGREEPLNHVEAERQR  
REKLNQRFYALRAVVPNVSKMDKASLLGDAIAFINELKSKV  
QNSDSDKDELRNQIESLRNELANKGSNYTGPPPPNQDLKIVD  
MDIDVKVIGWDAMIRIQSNKKNHPAARLMAALMELDLVDH  
HASVSVVNELMIQQATVKMGSRLYTQEQLRISLTSRIAESR

**Appendix 3: *NtERF98* cDNA (full-length, 1019 bp)**

CACCGTCTCTTTCCATTTCTTTCTCTTAAAAGAAAAACATCTCAATAACAAAA  
GAAAAATGTGTGGAGGTGCCATAATCCCCGACTATGAACCCGTCGGAAACCGCT  
GCCGGAAAATCACTGCTAGTGACCTCTGGGCTGAGCTTGACCCTATCTCCGACTT  
CTGGTCTTCCTCTTCCTCTTCCTCCTCCATTGCCGGCAAATCTGATTCCGTTTCAGT  
CGCTAACCCACTCCTACAATAAGCCTCAGAAATCAGATTCCGGCAAACCTTAATC  
AACTCGAAAAAGGTACAATAAGTGTGAAGGTTGAGAAGGAGAGCAGTGGCCCA  
AGGGCGAGGAAGAACAATAACAGAGGAATAAGGCAGAGACCGTGGGGAAAATG  
GGCTGCTGAGATACGTGATCCTCAGAAAGGCGTCCGCGTGTGGTTAGGTACATTC  
AACACGGCTGAGGAAGCTGCCAGGGCATATGACGAGGCTGCAAAGCGAATCCG  
CGGTGACAAGGCTAAGCTCAACTTTCCAGAGCCACCTTCGCCACCAGCCAAGCG  
ACACTGCACTAGCACCATCCCTGATCAGCCCACACGTTCTGACTTAATGTCTCAG  
AAACCGGCCTCAATAATGTTGAACTATGGATATGAAAACCAAACACCCTACTAC  
CCCATGGAAATGCCCGCTGCTGAGGATCCTCAACATCATGATTATGAGCTCAAG  
GAGCAGATTTCCAACCTTGGAGTCATTCCTGGATTTAGAGCCAGACTCAGGGATCG  
TCGATTCTGACCCCTCAATATTTTTCTGATGGATGACTTTGCTGCAACTCAGCAG  
CAGCAGCTGTTTTACTGAACACTGTAAAAATTATCATATACTACTAGTTAATTTT  
ATCCTAAGTTGTTGGTGTGCGTTTTCTGATGAGTGACTAGTTAGCTTTTGGTAGT  
ACGTAGTACAATGCAGAAAGTACATAACAATAAAGTTGCGTGCCTTTGCATGC  
AATTTGTAATATTAATGTCATGTTGTTTTGTGCTGTTAAAAAAAAAAAAAAAAAA  
AA

***NtERF98* amino acid sequence (257 AA)**

MCGGAIIPDYEPVGNRCRKITASDLWAELDPISDFWSSSSSS  
SSIAGKSDSVQSLTHSYNKPQKSDSGKLNQLEKGTISVKVEK  
ESSGPRARKKNKYRQIRQRPWGKWA AEIRD PQKGV RVWLGT  
FNTAEEA ARAYDEAAKRIRGDKAKLNFPEPPSPPAKRHCTS  
TIPDQPTRSDLMSQKPASIMLNYGYENQTPYYPMEMPAED  
PQHHDYELKEQISNLESFLDLEPDSGIVDS DPLNIFLMDDFA  
ATQQQQQLFY

**Appendix 4: *NtETTA* cDNA (full-length, 2429bp)**

AGCAAAAGGGTTTGAAGATGATGTGTGGACTTATTGATCTAAATACTGTGGATA  
ACGATGACGTCCGAGAAGAAACGACGGCGCCGGTGTACCAGCGTCATCGTCGA  
CGGCGTCTGGATGTTCCGATTTGACGTCGTCATCTCTGCCGGCGATGGCATCGGT  
TTGTCTGGAGCTGTGGCATGCGTGTGCTGGACCGTTGATTTCTCTGCCGAAGAAA  
GGAAGTGCTGTTGTGTACCTACCTCAAGGTCACCTTGAACATCTCTCTGAGTACC  
CGCCATAGCCTATAACCTCCCTCCTCACGTTTTTTGTGCGGTCTAGACGTGAA  
GCTACAAGCGGATGCGGCGAGTGATGAGGTCTATGCACAAGTCTCACTGGTTCC  
AGACAATCAGATTGAGCAGAAATGGAGGGATGGAGACATTGATGCAGATACTG  
AAGAGGAGGAAATAGAAGGTGCTGGAAAATCAACAACACCACACATGTTCTGC  
AAGACTCTCACTGCTTCGGATACCAGCACTCATGGCGGTTTTTTCTGTCCCTCGCC  
GGGCTGCAGAAGATTGCTTTCCTCCATTGGATTACAGACAACAGCGGCCCTCAC  
AGGAGCTGGTAGCCAAAGATCTACATGGTATCGAGTGGAAATTTCCGGCATATCT  
ATCGTGGTCAGCCACGAAGGCATCTGCTCACTACAGGATGGAGTGCGTTTTGTAA  
ACAGGAAGAAGCTTGTTTCTGGTGACGCTGTGCTTTTCTTAAGGACTGCTGATGG  
AGAACTTAGGCTAGGGGTGAGACGAGCTGCCCAAGCTAAAACATGTTCAAATTA  
TCTAGCTGCCTATAGCCAACCTGTTGAATGTCAGTGGTATTGTGGATGTGGTTAAG  
GCCATATCTAGCACAAATGCCTTCAGTATCTGTTATAACCCGAGGGCTAGCTCAT  
CAGGCTTCATTTTACCTTACCACAAATTCTCAAAGACTCTTGCACATCCCTTTTCA  
GCTGGAATGAGATTTAAGATGCGTGTGCGAAACAGAAGATGCAGCTGAACAAAGG  
TTCACTGGACTTGTTGTAGGAGTCAGCGATGTAGATCCAGTTCGCTGGCCTGGTT  
CTAAATGGAGGTGCCTATTGGTTCAGGTGGGATGATCTTGATGTTTCTCGGCATAA  
TAGGGTTTACCGTGGGAAATTGAGCCATCTGGTTCAGCTCCTGTATCCAGCAGC  
TTGGTGATGCCTTCTGCGAAGAGGACCAGGGTTGGCTTCCAATTACAAAGGCCG  
ATTTTCCAATTCCTAGAGATGGGATTGCAGTATCAGACTTTGGGGAATCTTCTAG  
GTTCCAGAAGGTCTTGCAAGGTCAAGAAATTTTGGGGATTAGTTCTCCTTTTGTG  
GGTTTTGATGCTCACAGTCCTCGTACAGCGGGGATAAGATGCTTTCCTGGTTTTC  
CTAGTTCTGGGGCATCTAGATTGGGAAACAGCATCAGAACCTGCTTGGTGACA  
CAGACAAGTCCCCTGAAAGCATTGGCTTTAGTGATTCTTCTCGATAACAATAAGGT  
CTTGCAAGGTCAAGAACTTTTTCAACCCCTCCTTATGGGAGAGGTCATGCAGGT  
AGCCTAATGCAGGAAAAAAGTAGAACTGGTATTATCGTCGGTATTCAGGTTCCA  
AGCCACGTAAACAGGTGGTCTGCTCCAAATCAGGGTAATCGCAGTCATTGCAAT  
CCAAGTACTCTTGTCCCAGCATCATCACCTCCTTCTGTGCTCAGCTTTCAGCCTCC  
CAGGTCTCCAGCATCAAAATTCAGGCTATGTTCAATCATAAACATGGGAAGCTT  
GAGACTGCTACCCAGGCTTTGGATATGTCTGAGAGCTGTAGTAGGCATCTCGCAT  
CTGGCTCACATGCCGAGGACATCAGTAGGAAGGGAGACACAAAAGGAATCAGTT  
CTTTTAGTTTCTTAAAGGAGCAAAAGCAAACAGGAATTCATATCTTTCTCTTGG

GACCCAGTCGTCTCAAAACTTAGTTTCCATGTGTAAAACCAGTTGCAGGATCTTT  
GGATTCCCCTTGACCGAGAGTAAAATAAATGCAGCTAGAGCGGAGAATCCTGCC  
GAGGCTGTATATTCACATGGTCTAGAAACAACATTTCTGCCTTCCAGTGATGGAA  
AGTTGCAGCCAGGGCCACCATTGATGACTAATGTTGTGGGAACAAACTTTACTA  
AAGTAAATGACCTCTATGCTGCAAGAGATGTGATTCTTGATATTGCTTTGTAGCA  
AGTATTTGTTGTGAAGTCATGAGCATATGTAAACTGAAGGATGTGTGAGCAGTAT  
TATTGATTCTTAGATTTTAGTTGGCTGATTAGTTTTGGCCAATGAACGCAAGCAT  
GTAGTTGCCAGTACAATGCTTATCCTGAGATGAGTATTGAGAGTTTTTATTGTAA  
GGAACACAGTGAAGATTAGTATTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
A

**NtETa amino acid sequence (739 AA)**

MMCGLIDLNTVDNDDVGEETTAPVSPASSSTASGCSDLTSSS  
LPAMASVCLELWHACAGPLISLPKKGSVVYLPQGHLEHLS  
EYPPIAYNLPPHVFCRVVDVKLQADAASDEVYAQVSLVPDN  
QIEQKWRDGDIDADTEEEEIEGAGKSTTPHMFCKTLTASDTS  
THGGFSVPRRAEDCFPLDYRQQRPSQELVAKDLHGIEWK  
FRHIYRGQPRRHLLTTGWSAFVNRKKLVSGDAVLFLRTADG  
ELRLGVRRAAQAKTCSNYLAAYSQLLNVSGIVDVVKAISST  
NAFSICYNPRASSSGFILPYHKFSKTLAHPFSAGMRFKMRVE  
TEDAAEQRF TGLVVGVS DVDPVRWPGSKWRCLLVRWDDL  
VSRHNRVSPWEIEPSGSAPVSSSLVMPSAKRTRVGFPIKAD  
FPIPRDGIAVSDFGESSRFQKVLQGQEILGISSPFVGFDAHSP  
RTAGIRCFPGFPSSGASRLGNSIRTL LGD TDKSPESIGFSDSS  
RYNKVLQGQETFS TPPYGRGHAGSLMQEKSRTGIIVGIQVPS  
HVNRWSAPNQGNRSHCNPSTLVPASSPPSVLSFQPPRSPASK  
FQAMFNHKGKLETATQALDMESECSRHLASGSHAEDISRK  
GDTKGISSFSFLKEQKQTGISYLSLGTQSSQNLVSMCKTSCR  
IFGFPLTESKINAARAENPAEAVYSHGLETTFLPSSDGKLQP  
GPPLMTNVVGTNFTKVNDLYAARDVILDIAL