Abstract

Identification, Characterization, and Abiotic Stress Analysis of microRNAs in *Nicotiana tabacum*

by Taylor P. Frazier

July, 2010

Director: Dr. Baohong Zhang

DEPARTMENT OF BIOLOGY

microRNAs (miRNAs) are non-coding RNAs with short sequences that negatively regulate gene expression at the post-transcriptional levels by either binding to mRNAs for degradation, or by inhibiting protein translation. miRNAs are highly evolutionarily conserved, from lower mosses to higher flowering plants, and have been shown to play an important role in plants by regulating growth and development, developmental timing, hormone signaling, organogenesis, and response to environmental stresses. Based on the conservation of mature miRNA sequences, computational methods have been used to predict thousands of miRNAs in numerous plant species, such as soybean, maize, cotton, potato, rice, apple, and switchgrass. In this project, using tobacco as my model organism, I first employed genome survey sequence analysis to identify conserved miRNAs in tobacco and then investigated their expression profiles during different growth conditions. My results identified 259 potentially conserved miRNAs in tobacco, belonging to 65 miRNA families. In addition, I also discovered antisense miRNAs as well as miRNA clusters in tobacco and predicted putative target genes for the newly identified tobacco miRNAs. Eleven of these miRNAs are highly expressed in young tobacco seedlings as well as under different environmental stress conditions. The expression profiles of tobacco miRNAs were significantly affected by Titanium Dioxide nanoparticles, salt, and drought

stresses, in a dosage-dependent manner. Some miRNAs, for example miR395, exhibited a change in expression of one thousand fold after exposure to stress conditions. Abiotic stresses also affected the expression of two stress-related genes, alcohol dehydrogenase and alcohol peroxidase. Given the results of this project, I believe that miRNAs may play an important role in tobacco growth and development and that miRNAs may function in tobacco tolerance to environmental stresses.

Identification, Characterization, and Abiotic Stress Analysis of microRNAs in *Nicotiana tabacum*

A Thesis

Presented To

The Faculty of the Department of Biology

East Carolina University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Biology

by

Taylor P. Frazier

July, 2010

IDENTIFICATION, CHARACTERIZATION, AND ABIOTIC STRESS ANALYSIS OF MICRORNAS IN NICOTIANA TABACUM

by

Taylor P. Frazier

APPROVED BY:

DIRECTOR OF THESIS:_____
Baohong Zhang, Ph.D

COMMITTEE MEMBER:_____
Carol Goodwillie, Ph.D

COMMITTEE MEMBER:_____
Xiaoping Pan, Ph.D

COMMITTEE MEMBER:_____
Anne Spuches, Ph.D

CHAIR OF THE DEPARTMENT OF BIOLOGY:_____
Jeff McKinnon, Ph.D

DEAN OF THE GRADUATE SCHOOL:_____
Paul J. Gemperline, PhD

ACKNOWLEDGEMENTS

I would like to thank Dr. Baohong Zhang for giving me the opportunity to work in his lab. I have greatly appreciated his mentorship, encouragement, and support. I would also like to thank the other members of my thesis committee, Dr. Carol Goodwillie, Dr. Xiaoping Pan, and Dr. Anne Spuches, for their guidance and feedback.

Additionally, I would like to thank my colleagues in the Zhang Lab, Maitri Shah and Lindsey Fix, for their assistance and support. I would like to thank Fuliang Xie, Caitlin Burklew, Cornelius Flowers, Michael Bossé, and Andrew Freistaeder for their help with the computational portion of this project. I would also like to thank the rest of the members of the Zhang lab, as well as the members of the Farwell and Pan Labs, for continued encouragement and support.

I would like to thank Dr. Ramsey Lewis of the Department of Crop Science at North Carolina State University for kindly providing the tobacco seeds used in this project. I would also like to thank Julie Marik of the Greenhouse Facility at East Carolina University for her care of the tobacco plants.

Finally, I would like to thank the Department of Biology at East Carolina University for financial support. This project was supported by ECU New Faculty Startup Funds allotted to Dr. Baohong Zhang as well as a Science and Engineering Grant from DuPont.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABA	Abscisic Acid
ADH	Alcohol dehydrogenase
AGO	ARGONAUTE
Al_2O_3	Aluminum Oxide
APX	Alchohol peroxidase
AP2	APETALA2-like
BLAST	Basic Local Alignment Search Tool
C	Celsius
CBC	Nuclear cap-binding complex
cDNA	complementary DNA
C_T	Cycle threshold
DCL1	DICER-LIKE1
DDL	DAWDLE
DNA	Deoxyribonucleic acid
dsRNA	Double stranded RNA
EF1α	Elongation Factor 1 α
EST	Expressed sequence tag
GSS	Genome survey sequence
HEN1	HUA ENHANCER1
HYL1	HYPONASTIC LEAVES1
INR	Transcription Initiator
kg	Kilogram
MFE	Minimal folding free energy
MFEI	Minimal folding free energy index

miRNA	microRNA
miRNA*	microRNA complementary sequence
mRNA	messenger RNA
MS	Murashige and Skoog
MYB	Myeloblast
NaCl	Sodium Chloride
nano-TiO_2	TiO_2 nanoparticles
NCBI	National Center for Biotechnology Information
nt	nucleotide
PEG	Polyethyleneglycol
pre-miRNA	Precursor microRNA
pri-miRNA	Primary microRNA
pol II	RNA Polymerase II
qRT-PCR	Quantitative Real Time PCR
RNA	Ribonucleic acid
rRNA	ribosomal RNA
RISC	RNA Induced Silencing Complex
RT-PCR	Reverse transcription PCR
SPL	SQUAMOSA promoter binding protein-like
SE	SERRATE
TCP	TEOSINTE BRANCHED1, CYCLOIDEA, PCF
TiO_2	Titanium Dioxide
TIR1	TRANSPORT INHIBITOR RESPONSE1
tRNA	transfer RNA
µg	Microgram

Zn	Zinc
ZnO	Zinc Oxide

CHAPTER 1: INTRODUCTION

microRNAs (miRNAs) are a newly recognized class of endogenous gene regulators that negatively control gene expression at the post-transcriptional levels by binding to messenger RNAs (mRNAs) and either targeting them for degradation, or inhibiting protein translation (Bartel 2004; Dugas and Bartel 2004). The first plant miRNA was discovered in *Arabidopsis thaliana* in 2002 (Park et al. 2002; Reinhart et al. 2002) and since then, computational and experimental methods have identified thousands of miRNAs in a wide range of plant species (Raymond et al. 2005; Zhang et al. 2005). miRNAs have been shown to be highly evolutionarily conserved, from lower mosses to higher flowering eudicots (Zhang et al. 2005). Due to their function in gene regulation, miRNAs have been shown to play an important role in a variety of plant metabolic and biological processes including organ maturation (Juarez et al. 2004; Mlotshwa et al. 2006), signal transduction (Vernoux and Benfey 2005), hormone signaling (Achard et al. 2004; Eckardt, 2005), phase change from vegetative to reproductive growth (Aukerman and Sakai 2003; Lauter et al. 2005), and response to environmental stresses (Sunkar and Zhu 2004; Lu et al. 2005; Zhang and Anderson 2006).

Generally speaking, plant miRNA genes are located at intergenic regions within the genome, and are believed to be transcribed by RNA Polymerase II (pol II) (Xie et al. 2010). This is supported by the presence of certain miRNA gene core promoter elements, such as the TATA box and transcription initiator (INR) elements, that are characteristic of pol II transcripts (Xie et al. 2005; Zhou et al. 2007). miRNAs are often found in the intergenic genomic region, although there have been several reports of miRNAs located in intron segements within the protein-coding genes (Zhao et al. 2010). Long primary miRNA (pri-miRNA) transcripts contain a 5’ cap as well as a 3’ poly-A tail and fold into a stem-loop hairpin structure (Xie et al. 2010). Some pri-

miRNAs also contain introns and subsequently enter the splicing pathway before processing (Aukerman and Sakai 2003). Usually, one pri-miRNA transcript gives rise to one mature miRNA. However, several miRNAs can also be transcribed into one long polycistronic miRNA that can give rise to different mature miRNAs, and recent studies have shown that some plant miRNAs are clustered together within genomes (Zhang et al. 2006a; Zhang et al. 2008; Merchan et al. 2009).

Pri-miRNA processing is a crucial step in mature miRNA biogenesis. While the precise mechanism remains unknown, several important key players have emerged that help piece together a framework for how pri-miRNAs are processed. Since plant miRNAs contain a 5' cap and 3' poly-A tail, they are believed to be processed almost immediately after leaving the pol II complex and before translation can begin (Xie et al. 2010). Recent studies have demonstrated that the nuclear cap-binding complex (CBC), known to be involved in premature mRNA processing, plays an important part in ensuring miRNAs enter into the miRNA biogenesis pathway (Gregory et al. 2008; Kim et al. 2008). This complex has been reported to interact with SE, a C2H2-type zinc-finger domain-containing protein, and possibly mediate pri-miRNA splicing (Yang et al. 2006a; Laubinger et al. 2008).

Once pri-miRNAs enter the miRNA biogenesis pathway, they are then processed by the multi-domain RNase III-like enzyme DICER-LIKE1 (DCL1) into shorter precursor miRNA (pre-miRNA) sequences (Kurihara and Watanabe 2004; Olmedo and Guzman 2008). DCL1 contains a N-terminal helicase domain, a mid-protein PAZ domain, and two RNase III domains, along with a dsRNA-binding motif at the C-terminal end (MacRae et al. 2006). Starting at the N-terminal helical end, DCL1 measures dsRNA molecules and cleaves them into smaller RNAs, the length of which is determined by the distance between the mid-protein PAZ domain and the

RNase III catalytic domains (MacRae et al. 2006). Additional proteins, such as HYPONASTIC LEAVES1 (HYL1), DAWDLE (DDL), and SE, are also necessary for pri-miRNA processing, and have been shown to interact with DCL1 (Han et al. 2004; Kurihara et al. 2006; Lobbes et al. 2006; Morris et al. 2006).

In plants, DCL1 also functions in the processing of pri-miRNAs to miRNA precursors (pre-miRNAs). This pre-miRNA is further processed by DCL1, and other associated proteins, to form a miRNA:miRNA* complex (Jones-Rhoades et al. 2006; Zhang et al. 2006b). Each miRNA in the miRNA:miRNA* complex, as a result of DCL1 cleavage, contains a 3' 2 nt overhang that is recognized by HUA ENHANCER1 (HEN1) (Dezulian et al. 2006; Yang et al. 2006b). HEN1, a small RNA methyltransferase, deposits a methyl group onto each of the terminal 3' nucleotide ends in the miRNA:miRNA* complex (Yang et al. 2006b). This is believed to contribute to the stability of the small RNAs and prevent further cellular modifications (Xie et al. 2010).

Finally, one of the strands in the miRNA:miRNA* complex, often times the miRNA strand, is selected for entry into the ARGONAUTE (AGO)-containing effecter complex (Bartel 2004). This complex is also known as the RNA-induced silencing complex (RISC), and targets sequence-specific mRNAs by one of two mechanisms: 1) perfect or nearly perfect miRNAs bind to mRNAs and promote mRNA transcript cleavage or 2) miRNAs that are complementary to mRNAs and do not induce transcript cleavage, bind and prevent protein translation (Bartel 2004; Vaucheret et al. 2006). In plants, the majority of miRNA-mediated gene regulation results in mRNA cleavage (Bartel 2004). However, some reports have identified cases of plant miRNAs, such as miR172, inhibiting protein translation (Aukerman and Sakai 2003; Chen 2004).

The majority of plant miRNAs have been shown to target transcription factors, elements that control gene expression during plant growth and development (Rhoades et al. 2002; Zhang

et al. 2008). miR156 targets SQUAMOSA promoter binding protein-like protein (SPL) transcription factors that function in leaf development and developmental timing (Mallory et al. 2004; Gandikota et al. 2007). miR172 controls flower timing and development by regulating members of the APETALA2 (AP2)-like proteins (Aukerman and Sakai 2003; Chen 2004). miR166 has been shown to target class III HD-ZIP genes that function in vascular patterning and leaf polarity (Emery et al. 2003; Williams et al. 2005). miR319, or miRJAW, aids in leaf development by targeting TEOSINTE BRANCHED1, CYCLOIDEA, and PCF (TCP) transcription factors (Palatnik et al. 2003). Recent studies have shown that mutations in the complementary binding site of miRJAW to MYB33, a transcription factor, results in upwardly curled leaves (Palatnik et al. 2003).

Plants are sessile organisms and have developed complex gene regulatory networks to combat environmental stresses. Recent studies are showing that more and more miRNAs are involved in plant tolerance to abiotic stresses such as drought, salinity, cold, heavy metal, and nutrient deprivation (Jones-Rhoades and Bartel 2004; Sunkar 2010). For example, miR398 has been shown to target superoxide dismutase genes, and its expression is influenced after plant exposure to heavy metals (Sunkar et al. 2006), high salt (Jia et al. 2009), and water deficiency (Trindade et al. 2010). Expression of miR169 is induced under high salinity, and expression of miR395 is up-regulated under sulfur and inorganic phosphate starvation (Jones-Rhoades and Bartel 2004; Kawashima et al. 2009; Zhao et al. 2009). Therefore, miRNAs play a key role in regulating plant response to environmental conditions.

Although much research has been conducted on plant miRNAs, these studies have mainly focused on model organisms such as *Arabidopsis* and rice (Zhang et al. 2008). Tobacco (*Nicotiana tabacum*) is an important agricultural and economic crop, and is the most commonly

cultivated plant in the *Nicotiana* genus. Tobacco is grown in more than 100 countries around the globe. The tobacco plant is commonly used for its high nicotine content; in the mature plant, 64% of the nicotine content can be found in the leaves, 18% in the stem, 13% in the roots, and 5% in the flowers. As the tobacco plant matures, the nicotine content inside the plant increases, explaining why tobacco farmers harvest the leaves in the later stages of growth. Nicotine is the main carcinogen found in cigarettes, cigars, snuff, and chewing tobacco and almost all tobacco plants grown commercially are used to make these tobacco products.

Since tobacco is a non-food crop, it is currently being pursued as a potential biofuel and a system to produce recombinant proteins for the pharmaceutical industry (McCabe et al. 2008; Andrianov et al. 2010). Some varieties of tobacco can generate large amounts of biomass. For example, the Maryland Mammoth variety can grow to over 3 m in height and prior to flowering, can produce 1-2 kg of leaf biomass per plant (McCabe et al. 2008). When grow for biofuel purposes, tobacco plants can produce 170 tons of green biomass per hectare and are an ideal candidate because multiple harvests can happen per year (Andrianov et al. 2010).

Despite its agricultural and economic significance, a list of miRNAs in tobacco has never been generated and therefore, miRNA functions in tobacco remain poorly understood. In this project, ***I hypothesize that conserved miRNAs exist in tobacco and that the expression levels of tobacco miRNAs are altered after exposure to heavy metal oxide, salt, and drought stresses.***

To test this hypothesis, this project will identify conserved tobacco miRNAs and investigate changes in miRNA expression profiles under environmental stress with the following specific objectives:

1. Identify *Nicotiana tabacum* miRNAs using a homology-based computational approach (Chapter 2)

miRNAs that have been identified in *Arabidopsis*, poplar, and rice were used to identify conserved miRNAs in tobacco using Genome Survey Sequence (GSS) analysis. Potential target genes for the newly identified tobacco miRNAs were also predicted using conservation of miRNA target sites across species.

2. Analyze miRNA expression profiles in response to heavy metal oxide stress (Chapter 3)

Stem-loop RT-PCR and TaqMan qRT-PCR Assays were used to determine miRNA expression profiles in three week old tobacco seedlings after exposure to increasing concentrations of Titanium Dioxide nanoparticles.

3. Analyze miRNA expression profiles in response salt and drought stresses (Chapter 4)

Stem-loop RT-PCR and TaqMan qRT-PCR Assays were used to determine miRNA expression profiles in 3 week old tobacco seedlings after exposure to different concentrations of NaCl and PEG.

References

Achard P, Herr A, Baulcombe DC. Harberd NP (2004) Modulation of floral development by a gibberellin-regulated microRNA. Development 131: 3357-3365

Andrianov V, Borisjuk N, Pogrebnyak N, Brinker A, Dixon J, Spitsin S, Flynn J, Matyszczuk P, Andryszak K, Laurelli M, Golovkin M, Koprowski H (2010) Tobacco as a production platform for biofuel: overexpression of *Arabidopsis DGAT* and *LEC2* genes increases accumulation and shifts the composition of lipids in green biomass. Plant Biotech J 8: 277-287

Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes. Plant Cell 15: 2730-2741

Bartel DP (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 116: 281-297

Chen XM (2004) A microRNA as a translational repressor of *APETALA2* in *Arabidopsis* flower development. Science 303: 2022-2025

Dezulian T, Remmert M, Palatnik JF, Weigel D, Huson DH (2006) Identification of plant microRNA homologs. Bioinforma 22: 359-360

Dugas DV, Bartel B (2004) MicroRNA regulation of gene expression in plants. Curr Opin Plant Biol 7: 512-520

Eckardt NA (2005) MicroRNAs regulate auxin homeostasis and plant development. Plant Cell 17: 1335-1338

Emery JF, Floyd SK, Alvarez J, Eshed Y, Hawker NP, Izhaki A, Baum SF, Bowman JL (2003) Radial patterning of *Arabidopsis* shoots by class IIIHD-ZIP and KANADI genes. Curr Biol 13: 1768-1774

Gandikota M, Birkenbihl RP, Hohmann S, Cardon GH, Saedler H, Huijser P (2007) The miRNA156/157 recognition element in the 3' UTR of the *Arabidopsis* SBP box gene *SPL3* prevents early flowering by translational inhibition in seedlings. Plant J 49: 683-693

Gregory BD, O'Malley RC, Lister R, Urich MA, Tonti-Filippini J, Chen H, Millar AH, Ecker JR (2008) A Link between RNA Metabolism and Silencing Affecting *Arabidopsis* Development. Dev Cell 14: 854-866

Han MH, Goud S, Song L, Fedoroff N (2004) The *Arabidopsis* double-stranded RNA-binding protein *HYL1* plays a role in microRNA-mediated gene regulation. Proc Natl Acad Sci 101: 1093-1098

Jia XY, Wang W, Ren L, Chen Q, Mendu V, Willcut B, Dinkins R, Tang X, Tang G (2009) Differential and dynamic regulation of *miR398* in response to ABA and salt stress in *Populus tremula* and *Arabidopsis thaliana*. Plant Mol Biol 71: 51-59

Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell 14: 787-799

Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs and their regulatory roles in plants. Annu Rev Plant Biol 57: 19-53

Juarez MT, Kui JS, Thomas J, Heller BA, Timmermans MCP (2004) microRNA-mediated repression of *rolled leaf1* specifies maize leaf polarity. Nature 428: 84-88

Kawashima CG, Yoshimoto N, Maruyama-Nakashita A, Tsuchiya YN, Saito K, Takahashi H, Dalmay T (2009) Sulphur starvation induces the expression of microRNA-395 and one of its target genes but in different cell types. Plant J 57: 313-321

Kim S, Yang JY, Xu J, Jang IC, Prigge MJ, Chua NH (2008) Two cap-binding proteins CBP20 and CBP80 are involved in processing primary microRNAs. Plant Cell Phys 49: 1634-1644

Kurihara Y, Watanabe Y (2004) *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. Proc Nat Acad Sci 101: 12753-12758

Kurihara Y, Takashi Y, Watanabe Y (2006) The interaction between *DCL1* and *HYL1* is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. RNA 12: 206-212

Laubinger S, Sachsenberg T, Zeller G, Busch W, Lohmann J, Ratsch G, Weigel D (2008) Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in *Arabidopsis thaliana*. Proc Nat Acad Sci 105: 8795-8800

Lauter N, Kampani A, Carlson S, Goebel M, Moose SP (2005) microRNA172 down-regulates *glossy15* to promote vegetative phase change in maize. Proc Nat Acad Sci 102: 9412-9417

Lobbes D, Rallapalli G, Schmidt DD, Martin C, Clarke J (2006) SERRATE: a new player on the plant microRNA scene. Embo Rep 7: 1052-1058

Lu SF, Sun YH, Shi R, Clark C, Li LG, Chiang VL (2005) Novel and mechanical stress-responsive microRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. Plant Cell 17: 2186-2203

MacRae IJ, Zhou K, Li F, Repic A, Brooks A, Cande WZ, Adams PD, Doudna JA (2006) Structural Basis for Double-Stranded RNA Processing by Dicer. Science 311: 195-198

Mallory AC, Reinhart BJ, Jones-Rhoades MW, Tang GL, Zamore PD, Barton MK, and Bartel DP (2004) MicroRNA control of *PHABULOSA* in leaf development: importance of pairing to the microRNA 5 ' region. EMBO J 23: 3356-3364

McCabe M, Klaas M, Gonzalez-Rabade N, Poage M, Badillo-Corona JA, Zhou F, Karcher D, Bock R, Gray J, Dix P (2008) Plastid transformation of high-biomass tobacco variety Maryland Mammoth for production of human immunodeficiency virus type 1 (HIV-1) p24 antigen. Plant Biotech J 6: 914-929

Merchan F, Boualem A, Crespi M, Frugier F (2009) Plant polycistronic precursors containing non-homologous microRNAs target transcripts encoding functionally related proteins. Gen Biol 10: R136

Mlotshwa S, Yang ZY, Kim YJ, Chen XM (2006) Floral patterning defects induced by *Arabidopsis APETALA2* and *microRNA172* expression in *Nicotiana benthamiana*. Plant Mol Biol 61: 781-793

Morris ER, Chevalier D, Walker JC (2006) *DAWDLE*, a forkhead-associated domain gene, regulates multiple aspects of plant development. Plant Phys 141: 932-941

Olmedo G, Guzman P (2008) Processing precursors with RNase III in plants. Plant Sci 175: 741-746

Palatnik JF, Allen E, Wu XL, Schommer C, Schwab R, Carrington JC, Weigel D (2003) Control of leaf morphogenesis by microRNAs. Nature 425: 257-263

Park W, Li JJ, Song RT, Messing J, Chen XM (2002) CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. Curr Biol 12: 1484-1495

Raymond CK, Roberts BS, Garrett-Engele P, Lim LP, Johnson JM (2005) Simple, quantitative primer-extension PCR assay for direct monitoring of microRNAs and short-interfering RNAs. RNA 11: 1737-1744

Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP (2002) MicroRNAs in plants. Gen Dev 16: 1616-1626.

Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of plant microRNA targets. Cell 110: 513-520

Sunkar R, Zhu JK (2004) Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. Plant Cell 16: 2001-2019

Sunkar R, Kapoor A, Zhu JK (2006) Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in *Arabidopsis* is mediated by downregulation of *miR398* and important for oxidative stress tolerance. Plant Cell 18: 2051-2065

Sunkar R (2010) MicroRNAs with macro-effects on plant stress responses. Sem Cell Dev Biol In Press: 10.1016/j.semcdb.2010.1004.1001.

Trindade I, Capitao C, Dalmay T, Fevereiro MP, dos Santos DM (2010) *miR398* and *miR408* are up-regulated in response to water deficit in *Medicago truncatula*. Planta 231: 705-716

Vaucheret H, Mallory AC, Bartel DP (2006) AGO1 homeostasis entails coexpression of *MIR168* and *AGO1* and preferential stabilization of miR168 by AGO1. Mol Cell 22: 129-136

Vernoux T, Benfey PN (2005) Signals that regulate stem cell activity during plant development. Curr Opin Gen Dev 15: 388-394

Williams L, Grigg SP, Xie MT, Christensen S, Fletcher JC (2005) Regulation of *Arabidopsis* shoot apical meristem and lateral organ formation by microRNA *miR166g* and its AtHD-ZIP target genes. Development 132: 3657-3668

Xie Z, Khanna K, Ruan S (2010) Expression of microRNAs and its regulation in plants. Sem Cell Dev Biol In press: doi:10.1016/j.semcdb.2010.03.012

Xie ZX, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC (2005) Expression of *Arabidopsis* miRNA genes. Plant Phys 138: 2145-2154

Yang L, Liu ZQ, Lu F, Dong AW, Huang H. (2006a) SERRATE is a novel nuclear regulator in primary microRNA processing in *Arabidopsis*. Plant J 47: 841-850

Yang ZY, Ebright YW, Yu B, Chen XM (2006b) HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2 ' OH of the 3 ' terminal nucleotide. Nuc Acids Res 34: 667-675

Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA (2005) Identification and characterization of new plant microRNAs using EST analysis. Cell Res 15: 336-360

Zhang BH, Anderson TA (2006) MicroRNAs: New players in plant responses to agrochemicals and environmental stress. 232nd Amer Chem Soc (ACS) Nat Meet, San Francisco, CA, USA, Sep 10-14, 2006

Zhang BH, Pan XP, Anderson TA (2006a) Identification of 188 conserved maize microRNAs and their targets. FEBS Lett 580: 3753-3762

Zhang, BH, Pan XP, Wang QL, Cobb GP, Anderson TA (2006b) Computational identification of microRNAs and their targets. Comput Biol Chem 30: 395-407

Zhang, BH, Pan XP, Stellwag EJ (2008) Identification of soybean microRNAs and their targets. Planta 229: 161-182

Zhao B, L Ge, Liang R, Li W, Ruan K, Lin H, Jin Y (2009) Members of miR-169 family are induced by high salinity and transiently inhibit the NF-YA transcription factor. BMC Mol Biol 10: 29

Zhao C-Z, Xia H, Frazier TP, Yao Y-Y, Bi Y-P, Li A-Q, Li M-J, Li C-S, Zhang B, Wang X-J (2010) Deep sequencing identifies novel and conserved microRNAs in peanuts (*Arachis hypogaea L.*). BMC Plant Biol 10: 3

Zhou XF, Ruan JH, Wang GD, Zhang WX (2007) Characterization and identification of microRNA core promoters in four model species. Plos Comput Biol 3: 412-423

CHAPTER 2: Computational identification of tobacco microRNAs and their target genes

Abstract:

microRNAs (miRNAs) are a recently discovered class of small (~21 nt) endogenous gene regulators that have been shown to play an important role in plant growth and development by aiding in organ maturation, hormone signaling, tissue differentiation, and plant tolerance to environmental stress. Since a list of miRNAs has never been generated for tobacco, we employed genome survey sequence (GSS) analysis to computationally identify 259 potentially conserved tobacco miRNAs, belonging to 65 families, and validated 11 of these miRNAs using qRT-PCR. The 65 miRNA families were dramatically different in size. miRNA precursor (pre-miRNA) sequence analysis showed that tobacco pre-miRNAs greatly varied from 45 nt to 635 nt in length with an average of 141 ± 108 nt. We were also able to determine the presence of antisense miRNAs as well as miRNA clusters in tobacco. Using previously established protocols, a total of 1,225 potential target genes were predicted for the newly identified tobacco miRNAs. These target genes include transcription factors, DNA replication proteins, metabolic enzymes, as well as other gene targets necessary for proper plant maturation. The results of this study show that conserved miRNAs exist in tobacco and suggest that these miRNAs may play an important role in tobacco growth and development.

Introduction:

microRNAs (miRNAs) are a newly discovered class of small (~21 nt) endogenous regulator molecules that negatively control gene expression at the post-transcriptional levels by either targeting messenger RNAs (mRNAs) for degradation or by inhibiting protein translation (Bartel 2004). In plants, the majority of miRNA-mediated gene regulation results in the cleavage of mRNA transcripts (Rhoades et al. 2002; Bartel 2004). Due to their role in gene regulation,

miRNAs have been shown to take part in a wide variety of plant metabolic and biological processes and play an important function in leaf development (Palatnik et al. 2003), stem and root growth (Guo et al. 2005; Williams et al. 2005), organ maturation (Aukerman and Sakai 2003), signal transduction (Achard et al. 2004), and tolerance to environmental stresses (Sunkar et al. 2006; Zhao et al. 2009; Sunkar 2010).

The first plant miRNA was discovered in *Arabidopsis* in 2002 (Park et al. 2002; Reinhart et al. 2002) and since then, hundreds of miRNAs have been identified in a wide range of plant species using computational and experimental methods. Currently, 2,043 miRNAs have been verified in 29 different plants and have been deposited in the publicly available miRNA database miRBase (Griffiths-Jones et al. 2008). Of these 2,043 miRNAs, the majority have been found in plant species whose genomes have already been sequenced including 190 from *Arabidopsis*, 230 from *Physcomitrella patens*, 108 from *Medicago truncatula*, 234 from poplar, 140 from grapevine, 414 from rice, 140 from sorghum, and 109 from maize (Griffiths-Jones et al. 2008). Since miRNAs are vital to plant growth and development, they have been shown to be highly evolutionarily conserved from lower mosses to higher flowering plants (Zhang et al. 2006a). This provides a power tool for using comparative genomics approaches to identify miRNAs in new plant species.

Comparative genome-based computational approaches have been used to identify hundreds of miRNAs in a variety of plants including cotton (Qiu et al. 2007; Zhang et al. 2007), soybean (Zhang et al. 2008a), maize (Zhang et al. 2006b), rice (Lu et al. 2005), wheat (Yin and Shen 2010), tomato (Yin et al. 2008; Zhang et al. 2008b), potato (Guo et al. 2007; Zhang et al. 2009), oil rape (Xie et al. 2007), citrus (Song et al. 2009), lettuce (Han et al. 2010), apple (Gleave et al. 2008), *Brachypodium distachyon* (Unver and Budak 2009), and *Populus euphratica* (Li et al.

2009). This method has been preferred over experimental approaches, such as direct cloning and deep sequencing, for two reasons. First, no specific software is required and all sequences used are readily accessible from the NCBI GenBank Databases. Therefore, computational approaches are inexpensive compared to experimental methods. Second, this method relies on the expressed sequence tags (EST) and genome survey sequences (GSS) databases. For this reason, computational approaches provide an ideal way for identifying miRNAs in plants whose genomes have not been fully sequenced.

In this paper, we employed GSS analysis to identify conserved miRNAs in tobacco (*Nicotiana tabacum*). Tobacco is an important economic and agricultural crop in the United States as well as around the world. It has also recently been investigated as a potential biofuel crop (Andrianov et al. 2010). The complete tobacco genome remains unknown; however, there are approximately 1.4 million GSS sequences for tobacco deposited in the GSS database in NCBI GenBank. Consequently, this makes tobacco an ideal candidate for using computational approaches to identify conserved miRNAs. Despite much research being dedicated to this crop, a list of miRNAs in tobacco has never been generated. In this study, we compiled a list of 349 reference miRNAs from *Arabidopsis*, poplar, and rice and BLASTn searched these against the 1,420,579 GSS sequences for tobacco. We were able to successfully identify over 250 conserved miRNAs, belonging to 65 miRNA families, as well as determine the presence of antisense miRNAs and miRNA clusters in tobacco. Using a similar method, we also identified potential target genes of the newly identified miRNAs. Given the abundance of conserved miRNAs in tobacco, we believe that miRNAs may play an important role in tobacco growth and development.

Materials and methods:

miRNA reference set and tobacco GSS sequences

In order to identify potentially conserved miRNAs in tobacco, a total of 801 mature miRNA sequences were obtained for three different plant species from the publicly available miRNA database miRBase, version 14.0 (Griffiths-Jones et al. 2008). This list included 187 miRNAs from *Arabidopsis*, 237 from poplar, and 377 from rice. After removing repeated sequences, a total number of 349 unique mature miRNAs were used to probe for potential miRNAs in tobacco. We chose to computationally identify conserved miRNAs in tobacco using the GSS database because there are significantly more GSS sequences (1,420,579) available for tobacco than EST sequences (317,272).

Identifying potential tobacco miRNAs using GSS analysis

Potential tobacco miRNAs were identified according to previously established methods (Zhang et al. 2005). Briefly, the reference set of 349 mature miRNA sequences was BLASTn searched against the GSS sequence database for tobacco in order to identify potential tobacco miRNA homologues. Results of these BLASTn searches were then subjected to secondary structure analysis using mFOLD version 3.2 (Mathews et al. 1999; Zuker 2003). A result was considered a potential miRNA if it met the following criteria: (i) there were no more than 4 nucleotides substituted between the query miRNA sequence and the BLASTn result sequence; (ii) the potential miRNA was part of a longer precursor sequence that folded into a stem-loop hairpin secondary structure; (iii) the potential miRNA was located in one arm of the hairpin; (iv) there were no more than seven nucleotides mismatched in the hairpin between the potential miRNA and its opposite miRNA* sequence; (v) there were no loops or breaks in the miRNA or miRNA* sequences; and (vi) the potential miRNA precursor sequence had a high negative

minimal folding energy (MFE) and a high MFE index (MFEI) value. GSS sequences that met these criteria were recorded along with the potential precursor sequence, the nucleotide position of the potential miRNA within the precursor sequence, the length of the precursor sequence including the total number of As, Cs, Gs, and Ts, the arm of the hairpin in which the potential miRNA was located (5' or 3'), the MFE, and the MFEI value. The initial list of potential tobacco miRNAs was further evaluated by removing repeated sequences and analyzing the remaining sequences for the presence of miRNA clusters and antisense miRNAs.

Verification of tobacco miRNAs using quantitative real time PCR (qRT-PCR)

Young leaf samples were collected from tobacco plants growing in the East Carolina University Greenhouse and immediately frozen in liquid nitrogen. Total RNA was extracted from the leaves using the mirVana^{TM} miRNA Isolation Kit (Ambion, Austin, TX, USA) according to the manufacturer's protocol. Using a Nanodrop ND-1000 (Nanodrop technologies, Wilmington, DE, USA), the total RNA was assessed for quantity and quality and then stored at -80°C until future use. Applied Biosystems TaqMan® microRNA assays were used to verify the expression of 11 identified tobacco miRNAs: nta-miR156, nta-miR159, nta-miR162, nta-miR167, nta-miR169, nta-miR172, nta-miR393, nta-miR395, nta-miR396, nta-miR398, and nta-miR399. First, reverse transcription PCR (RT-PCR) was performed using the TaqMan microRNA Reverse Transcription Kit and miRNA-specific stem-loop primers included in the kit. Briefly, 1 µg of total RNA was used to generate a single-stranded miRNA cDNA for each of the miRNAs listed above. RT-PCR conditions were followed according to the manufacturer's protocol. Second, quantitative real time PCR (qRT-PCR) and miRNA specific primers were used to verify the expression of the 11 selected miRNAs. The reactions were run on an Applied

Biosystems 7300 Sequence Detection System (Foster City, CA, USA) according to the manufacturer's protocol. Three biological replicates along with 2 technique replicates were conducted for each miRNA. Following qRT-PCR, differences in miRNA expression levels were assessed by analyzing the mean C_T values.

Prediction of potential tobacco miRNA target genes

We used a BLASTn search method, similar to that described above for predicting tobacco miRNA homologues, to predict potential target genes of tobacco miRNAs. First, all of the potential tobacco miRNAs were BLASTn searched against the Genbank protein-coding gene databases and tobacco EST sequences. When developing the list of potential target genes, the total number of mismatched nucleotides between the miRNAs and their potential targets were considered as well as alignment structures. In order to eliminate false positives, the conservation of the miRNA target site in alternate plant species was analyzed. The following criteria were used to determine complementary sites between predicted tobacco miRNAs and potential mRNA targets: (i) There were no more than 4 mismatches between predicted tobacco miRNAs and their potential mRNA targets; (ii) Of these 4 mismatches, no mismatches were allowed between positions 10 and 11, no more than 1 mismatch was allowed between positions 1 and 9, and no more than 2 mismatches were allowed at any other positions; and (iii) No gaps were allowed between the potential tobacco miRNAs and the predicted mRNA targets. Due to the limited number of protein-coding sequences available for tobacco, we used BLASTn searches against the protein-coding databases of other plant species as well as BLASTn searches against the tobacco EST database. In the event that a potential mRNA target was found in tobacco ESTs, we performed a homology search against the protein-coding databases of other plant species in order

to identify potential target genes. This was conducted based on the degree of similarity of protein-coding genes among plant species.

Results:

Identification of tobacco miRNAs

In order to identify potentially conserved miRNAs in tobacco, a reference set of 349 previously known miRNAs from *Arabidopsis*, poplar, and rice was BLASTn searched against the 1,420,579 sequences available for tobacco in the NCBI GSS database. Using mFOLD version 3.2, secondary structure analysis of the results identified 259 potentially conserved miRNAs in tobacco (Table 2-1, Table 2-2, Figure 1). Therefore, an estimated 0.0182% of tobacco GSS sequences contain potential miRNAs.

The 259 potential tobacco miRNAs belong to 65 families. Of these 65 families, 6 major families were identified that contained 10 or more members (Figure 2-2). These families included miR156 with 15 members, miR166 with 10 members, miR169 with 47 members, miR170 with 10 members, miR172 with 26 members, and miR399 with 17 members. All other miRNA families contained fewer than 10 members and most only contained 1 or 2 miRNAs per family.

Characteristics of the potentially conserved miRNAs in tobacco varied between families (Table 2-1, Table 2-2). The majority of identified potential tobacco miRNAs (72.6%) were 21 nt in length followed by 20 nt (17%), 22 nt (6.2%), 19 nt (3.1%), and 18 nt (1.1%), respectively. A total number of 129 (49.8%) miRNAs were predicted to be found on the 5’ arm of the pre-miRNA stem-hairpin loop. In contrast, 130 miRNAs (50.2%) were predicted to be found on the 3’ arm of the pre-miRNA stem-hairpin loop. Potential tobacco pre-miRNA sequences also showed great variability (Table 2-2 and Table 2-3). The average length of a potential tobacco

pre-miRNA sequence was 141 ± 108nt; however, the majority of potential tobacco pre-miRNA sequences were only 70-100 nt in length (Figure 2-3). nta-miR1435b exhibited the shortest precursor length of 45 nt whereas nta-miR845a exhibited the longest precursor length of 635 nt. The nucleotide composition of the newly identified potential tobacco miRNA precursor sequences had an average G+C content of 39.4 ± 7.5% and an average A+U content of 60.5 ± 7.5%. The average A/U nucleotide ratio of the potential tobacco miRNA precursor sequences was 0.912 ± 0.261. The average minimal folding free energy (MFE) of the putative tobacco pre-miRNAs was -45.93 ± 26.19 kcal/mol.

The stem-loop hairpin structure is not unique to miRNAs. Therefore, other measures must be used to further validate that a result could potentially be a miRNA. Minimal folding free energy index (MFEI) is a criterion that has been established to eliminate false positives and distinguish miRNAs from other types of RNA molecules (Zhang et al. 2006c). Potential miRNA sequences are more likely to be miRNAs if they have a MFEI value of 0.85 or greater (Zhang et al. 2006c). In this study, we found that the 259 identified potential tobacco miRNAs exhibited a wide range of MFEI values with the lowest being 0.28 and the highest being 1.71. The average MFEI value of the potential tobacco miRNAs was 0.91 ± 0.23. This suggests that the majority of the 259 potential tobacco miRNAs are more likely to be miRNAs than any other type of RNA molecule that can form the same stem-loop hairpin structure, such as tRNA.

Quantitative real time PCR validation of tobacco miRNAs

qRT-PCR has become a sensitive and reliable method to determine the presence of, as well as quantitatively assess the expression levels of, miRNAs (Chen et al. 2005). Applied Biosystems TaqMan® microRNA assays, using miRNA specific stem-loop primers, were used

to verify the existence of 11 conserved tobacco miRNAs that were identified by GSS analysis. Using these assays, we were able to confirm *in vivo* the expression of nta-miR156, nta-miR159, nta-miR162, nta-miR167, nta-miR169, nta-miR172, nta-miR393, nta-miR395, nta-miR396, nta-miR398, and nta-miR399 in young tobacco leaves (Figure 2-4a). By analyzing differences in the mean C_T values for each miRNA (Figure 2-4b) , we were able to determine that nta-miR167 and nta-miR396 are expressed at higher levels, as compared to the other 9 miRNAs in this study, in young tobacco leaves. nta-miR169 and nta-miR399 exhibited the lowest levels of expressions.

Antisense miRNAs in tobacco

miRNAs have been shown to be transcribed from both sense and antisense DNA strands located at the same genomic miRNA loci in animal (Stark et al. 2008). These precursor sequences are processed and give rise to functional mature miRNAs (Stark et al. 2008). However, there have not been many reports of antisense miRNAs in plants. The first identification of a plant miRNA, transcribed from both sense and antisense DNA strands, was reported in soybeans in 2008 (Zhang et al. 2008a). In this experiment, we identified 12 pairs of sense and anti-sense strand tobacco miRNAs belonging to 3 different miRNA families (Table 2-4). One pair, nta-miR827c and nta-miR827b, belonged to the miR827 family while 2 pairs (nta-miR169l and nta-miR169ar, and nta-miR169as and nta-miR169at) belonged to the miR169 family. The miR172 family exhibited the greatest number of sense and antisense miRNA pairs with 9 including nta-miR172w and nta-miR172q, nta-miR172x and nta-miR172i, nta-miR172k and nta-miR172r, nta-miR172h and nta-miR172s, nta-miR172e and nta-miR172m, nta-miR172a and nta-miR172l, nta-miR172n and nta-miR172d, nta-miR172o and nta-miR172v, and nta-miR172p and nta-miR172y. miRNAs participating in the sense/antisense pairs had the same pre-

miRNA precursor length; however, the MFE, nucleotide composition, and secondary hairpin structure of these pre-miRNAs differed within the sense/antisense pair (Figure 2-5).

miRNA clusters in tobacco

Recent evidence has shown that polycistronic miRNAs, or several different mature miRNAs derived from the same pri-miRNA, exist in plants (Merchan et al. 2009). Evidence of miRNA clusters has also been reported in several plant species including *Arabidopsis* (Jones-Rhoades and Bartel 2004), moss (Talmor-Neiman et al. 2006), cotton (Zhang et al. 2007), and soybean (Zhang et al. 2008a). In this study, we identified 8 potential miRNA clusters involving 3 different miRNA families in tobacco. The miR156 family had two clusters including the nta-miR156a-miR156b-miR156c cluster (Figure 2-6) and the nta-miR156o-miR156m cluster. The miR169 family had 5 clusters including the nta-miR169k-miR169l-miR169m cluster, the nta-miR169af-miR169ai cluster, the nta-miR169n-miR169an cluster, the nta-miR169ag-miR169ah-miR169ak cluster, and the nta-miR169af-miR169an cluster. The miR172 family had one cluster that contained two pairs of sense/antisense miRNAs (Table 2-4). This cluster contained the nta-miR172e and nta-miR172m pair as well as the nta-miR172a and nta-miR172l pair.

Prediction of potential tobacco miRNA targets

The 259 potential tobacco miRNAs were BLASTn searched against tobacco EST sequences, as well as the plant protein-coding databases, in NCBI GenBank in order to identify potential target genes. A total of 1,225 putative target genes, belonging to a variety of gene families that partake in various biological and physiological functions, were identified in a wide range of plant species. These target genes include transcription factors, DNA replication proteins, genes

involved in cellular metabolism, and genes involved in plant response to environmental stresses (Table 2-5). For most miRNAs, more than one potential target gene was predicted. The potential tobacco miRNAs were found to either perfectly or near-perfectly complement their predicted target genes (Figure 2-7).

Discussion:

Computational methods have been successfully used to predict hundreds of miRNAs in a wide variety of plant species. In this study, we employed GSS analysis to identify 259 potentially conserved miRNAs in tobacco belonging to 65 different families. Therefore, 0.0182% of tobacco GSS sequences contain one or more miRNAs. This number is similar to our previous report (Zhang et al. 2008a) and has the potential to increase as new conserved miRNAs as well as tobacco-specific miRNAs are identified.

The newly identified potential tobacco miRNAs exhibited a wide range of characteristics between different families and even among members of the same family. For example, while only one member was identified for a majority of the miRNA families, a total number of 47 potential miRNAs were identified for the miR169 family and 15 miRNAs were identified for the miR156 family. Some miRNA family numbers in tobacco are larger than the miRNA numbers reported for the same families in different plant species. miR169, with 14 members, is the largest miRNA family in *Arabidopsis* and has been shown to be expressed mainly in young seedlings as opposed to mature tissues (Li et al. 2010). miR156, with 7 members, is also one of the more predominant miRNA families in *Arabidopsis* (Griffiths-Jones et al. 2008) and has been shown to function in developmental timing, as well as the phase change from vegetative to reproductive growth (Wu and Poethig 2006). The GSS database in NCBI GenBank is a collection of nucleotide sequences that are derived from genome sequencing and is different from the EST

database, which is based on mRNA transcripts (Benson et al. 2008). Therefore, the larger number of identified potential tobacco miRNAs could be explained by using a more comprehensive collection of sequences that better represents the tobacco genome, as opposed to searching the developmental timing-based ESTs.

The majority of the potential tobacco miRNAs (72.6%) were found to be 21 nt in length. This is similar to conserved miRNAs predicted for soybean (Zhang et al. 2008a), maize (Zhang et al. 2006b), and grass (Unver and Budak 2009). Potential tobacco miRNA precursor sequences, however, exhibited more diversity. Most animal miRNAs have precursor lengths of ~70-100 nt (Ambros 2004), however, plant miRNA precursor sequences have been shown to be highly variable (Zhang et al. 2006b). Potential tobacco pre-miRNAs ranged from 45 nt to 635 nt in length with an average of 141 ± 108 nt. The majority (50.6%) of tobacco miRNA precursor sequences are 70-110 nt in length, consistent with other reports for *Arabidopsis*, rice, cotton, maize, and soybean (Sunkar et al. 2005; Zhang et al. 2006b; Zhang et al. 2007; Zhang et al. 2008a). Precursor sequences with less than 70 nt in length comprised 11.2% of the total number of precursor sequences while those with lengths greater than 110 nt comprised 38.2% of the total. The length of the potential tobacco pre-miRNAs varied greatly among members of the same family. This suggests that members of the same family can have differing expression patterns, such as in a spatiotemporal or tissue specific manner (Zhang et al. 2008a).

miRNAs are located within one arm of the stem-loop hairpin structure. In this study, we found that 129 tobacco miRNAs (49.8%) resided in the 5’ arm of the stem-loop hairpin structure while 130 (50.2%) resided in the 3’ arm. These results indicate that there is no preference as to in which arm of the hairpin structure a tobacco miRNA will be located.

miRNA precursor sequences have a higher negative minimal folding free energy (MFE) compared to other RNAs as they form a stable stem-loop hairpin structure (Bonnet et al. 2004). However, other RNAs, such as tRNAs and rRNAs, can also form the hairpin structure. Therefore, prediction of potential miRNAs cannot be based on MFE alone. A new criterion, minimal folding free energy index (MFEI), distinguishes miRNAs from other types of RNA molecules if they have an MFEI value of 0.85 or greater (Zhang et al. 2006c). We found that the potential tobacco miRNAs had MFEI values of 0.28 to 1.71 with an average of 0.91 ± 0.23. This value is significantly higher than the MEFI values of other RNAs such as tRNAs (0.64), rRNAs (0.59), and mRNAs (0.62-0.66) (Zhang et al. 2006c). Therefore, the newly predicted potential tobacco miRNAs are more likely to be miRNAs than any other type of RNA molecule.

qRT-PCR was used to confirm the expression of 11 potential tobacco miRNAs. Using this method, we were able to validate the existence of nta-miR156, nta-miR159, nta-miR162, nta-miR167, nta-miR169, nta-miR172, nta-miR393, nta-miR395, nta-miR396, nta-miR398, and nta-miR399 in young tobacco leaves. These miRNAs were differentially expressed in comparison to the reference gene, EF1α, which suggests a certain role for these miRNAs in tobacco development. In young tobacco leaves, nta-miR169 and nta-miR396 exhibited higher levels of expression compared to the other 9 miRNAs. This indicates that these two miRNAs have a function in leaf development. miR396 has been shown to target growth regulating factors, ultimately contributing to proper leaf development (Liu et al. 2009). miR169, although known to be stress-induced under high salt conditions (Zhao et al. 2009), was found to be expressed in young tobacco leaves. Given the large number (47) of potential miR169 genes in tobacco, we believe that miR169 may play other roles in plant development, aside from conferring tolerance to environmental stresses. miRNAs have been shown to have variable expression patterns with

regard to tissue differentiation and developmental stages (Zhang et al. 2006d). Therefore more studies will need to be done to determine the specific expression patterns of tobacco miRNAs.

Recent studies have shown that miRNAs can be transcribed from either the sense or antisense strand of DNA at miRNA genomic loci (Stark et al. 2008). We found 12 pairs of sense/antisense miRNAs in tobacco. The majority of pairs (9) belonged to the miR172 family. To our knowledge, this is the first report of miR172 sense/antisense pairs in plants. Antisense miRNAs in tobacco differed from their sense partners by 1-3 nucleotides, similar to what has been reported for animals (Stark et al. 2008) and other plant species (Zhang et al. 2008a; Zhou et al. 2009). For the miR172 sense/antisense pairs, the change in nucleotides between partners typically occurred at position 1, position 18, and position 21 of the mature miRNAs. This indicates that these changes could alter miRNA biogenesis or that members of a sense/antisense pair could target different genes involved in various developmental processes (Zhang et al. 2006a; Zhang et al. 2008a). In all cases, the pre-miRNA sequences also differed within members of a sense/antisense pair. Each pair member contains the same number of nucleotides in the pre-miRNA; however, the nucleotide composition and secondary hairpin structure of the pre-miRNA varies between members. Since the complete tobacco genome remains unknown, it is impossible to tell on which DNA strand these miRNAs reside.

In some cases, miRNAs are clustered together in plant genomes, where one polycistronic mRNA transcript can give rise to multiple mature miRNAs of the same family (Merchan et al. 2009). miRNA clusters have been reported in soybeans (Zhang et al. 2008a), maize (Zhang et al. 2006b), switchgrass (Xie et al. 2010), rice (Cui et al. 2009), and *Arabidopsis* (Jones-Rhoades and Bartel 2004). We identified 8 miRNA clusters in tobacco belonging to three families, miR172 (1), miR156 (2), and miR169 (5). Due to the origin of GSS sequences, 3 of the identified

miR169 clusters may overlap to form one cluster containing 4 miR169 family members. miR169 has been shown to play a role in plant tolerance to environmental stress (Zhao et al. 2009) and clusters of this miRNA have previously been identified in soybean (Zhang et al. 2008a), cotton (Zhang et al. 2007) and rice (Zhao et al. 2009). Therefore, miR169 and these newly identified miR169 clusters in tobacco may have similar evolutionary origins and ultimately, may play the same role in tobacco development as in other plant species.

Plant miRNAs are a perfect or near-perfect match to their target mRNAs and help regulate post-transcriptional gene expression by binding to mRNAs and promoting mRNA degradation, or binding and inhibiting protein translation (Bartel 2004). Most plant miRNAs that bind to mRNAs lead to transcript cleavage (Bartel 2004), however, some miRNAs have been shown to inhibit protein translation in plants (Aukerman and Sakai 2003). Based on the complementarities between miRNAs and their target mRNAs, we employed BLASTn searches to identify potential target genes in tobacco, as well as in other plant species. Next, we performed a homology-based search to determine protein identity. This is because the complete tobacco genome and proteome are unknown and there is a high degree of conservation of mature miRNA sequences as well as miRNA target sites among the plant kingdom (Zhang et al. 2006b). We were able to identify 1,225 potential target genes for the 259 newly identified tobacco miRNAs. miRNAs have been shown to target transcription factors as a means of regulating plant growth and development (Jones-Rhoades and Bartel 2004). We predicted that miR156 in tobacco targets the SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) transcription factor, a protein that is involved in regulating developmental timing. miR156 has also been shown to target SPL in *Arabidopsis* (Jones-Rhoades et al. 2006) and is predicted to target SPL in other plant species (Zhang et al. 2006e). miR166 targets the class III HD-ZIP transcription factors that aid in the

establishment of leaf polarity (Emery et al. 2003). We predicted that the potential tobacco miR166 family will also target class III HD-ZIP transcription factors, thereby regulating leaf development in tobacco. Additionally, we predicted that the potential tobacco miR172 will target APETALA2-like (AP2) proteins. This is similar to other reports that have shown that miR172 targets AP2 and AP2-like genes, ultimately promoting floral organ maturation (Aukerman and Sakai 2003). Almost all of the 259 potential tobacco miRNAs were predicted to target more than 10 target genes. Interestingly, a recent estimate has stated that miRNAs have approximately 100 target sites within the protein coding genes (Brennecke et al. 2005). Additionally, miRNAs are thought to target more than 30% of protein-coding genes in humans and this number is expected to rise as more miRNAs are discovered (Lewis et al. 2005). Therefore, tobacco miRNAs may target multiple genes that are involved in a wide variety of biological and metabolic pathways. While we have used a homology-based approach to predict potential tobacco miRNA target genes, these target predictions need to be confirmed using experimental methods such as northern blotting and 5' Rapid Amplification of cDNA Ends (RACE).

Conclusion:

In this study, we BLASTn searched a reference set of 349 conserved miRNAs from *Arabidopsis*, rice, and poplar against the 1,420,579 sequences deposited for tobacco in the GSS database. We were able to computationally identify 259 potential tobacco miRNAs, belonging to 65 families, and using qRT-PCR, we confirmed *in vivo* the expression of 11 of these predicted miRNAs in young tobacco leaves. The majority of the predicted tobacco miRNAs were 21 nt in length and they were found equally on both the 5' and 3' arms of the stem-loop hairpin secondary structure. The tobacco pre-miRNAs varied in length from 45 to 635 nt with an average of 141 ± 108 nt. Using GSS analysis, we were able to identify 12 sense/antisense pairs of

miRNAs in tobacco and determined the presence of 8 miRNA clusters. We also predicted 1,225 potential target genes for the newly identified tobacco miRNAs that included transcription factors, DNA replication proteins, and metabolic enzymes. With these results, we are able to conclude that conserved miRNAs exist in tobacco and we believe that these miRNAs may play an important role in tobacco growth and development.

References

Achard P, Herr A, Baulcombe DC, Harberd NP (2004) Modulation of floral development by a gibberellin-regulated microRNA. Development 131: 3357-3365

Ambros V (2004) The functions of animal microRNAs. Nature 431: 350-355

Andrianov V, Borisjuk N, Pogrebnyak N, Brinker A, Dixon J, Spitsin S, Flynn J, Matyszczuk P, Andryszak K, Laurelli M, Golovkin M, Koprowski H (2010) Tobacco as a production platform for biofuel: overexpression of *Arabidopsis DGAT* and *LEC2* genes increases accumulation and shifts the composition of lipids in green biomass. Plant Biotech J 8: 277-287

Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes. Plant Cell 15: 2730-2741

Bartel DP (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 116: 281-297

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2008) GenBank. Nucleic Acids Res 36: D25-30

Bonnet E, Wuyts J, Rouze P, Van de Peer Y (2004) Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. Bioinformatics 20: 2911-2917

Brennecke J, Stark A, Russell RB, Cohen SM (2005) Principles of microRNA-target recognition. PLoS Biol 3: 404-418

Chen CF, Ridzon DA, Broomer AJ, Zhou ZH, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, Lao KQ, Livak KJ, Guegler KJ (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res 33: e179

Cui X, Xu SM, Mu DS, Yang ZM (2009) Genomic analysis of rice microRNA promoters and clusters. Gene 431: 61-66

Emery JF, Floyd SK, Alvarez J, Eshed Y, Hawker NP, Izhaki A, Baum SF, Bowman JL (2003) Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes. Curr Biol 13: 1768-1774

Gleave AP, Ampomah-Dwamena C, Berthold S, Dejnoprat S, Karunairetnam S, Nain B, Wang Y-Y, Crowhurst RN, MacDiarmid RM (2008) Identification and characterisation of primary microRNAs from apple (*Malus domestica* cv. Royal Gala) expressed sequence tags. Tree Genet Genomes 4: 343-358

Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. Nucleic Acids Res 36: D154-D158

Guo HS, Xie Q, Fei JF, Chua NH (2005) MicroRNA directs mRNA cleavage of the transcription factor *NAC1* to downregulate auxin signals for *Arabidopsis* lateral root development. Plant Cell 17: 1376-1386

Guo Q, Xiang AL, Yang Q, Yang ZM (2007) Bioinformatic identification of microRNAs and their target genes from *Solanum tuberosum* expressed sequence tags. Chin Science Bull 52: 2380-2389

Han Y, Zhu B, Luan F, Zhu H, Shao Y, Chen A, Lu C, Luo Y (2010) Conserved miRNAs and their targets identified in lettuce (*Lactuca*) by EST analysis. Gene 463:1-7

Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Molecular Cell 14: 787-799

Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs and their regulatory roles in plants. Ann Rev Plant Biol 57: 19-53

Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120: 15-20

Li B, Yin W, Xia X (2009) Identification of microRNAs and their targets from *Populus euphratica*. Biochem Biophys Res Comm 388: 272-277

Li Y, Fu Y, Ji L, Wu C, Zheng C (2010) Characterization and expression analysis of the *Arabidopsis* mir169 family. Plant Science 178: 271-280

Liu D, Song Y, Chen Z, Yu D (2009) Ectopic expression of miR396 suppresses *GRF* target gene expression and alters leaf growth in *Arabidopsis*. Physiol Plant 136: 223-236

Lu YZ, Yan DW, Lu YT (2005) Identification of microRNAs from rice. Funct Plant Biol 32: 963-971

Mathews DH, Sabina J, Zuker M, Turner DH (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Biol 288: 911-940

Merchan F, Boualem A, Crespi M, Frugier F (2009) Plant polycistronic precursors containing non-homologous microRNAs target transcripts encoding functionally related proteins. Genome Biol 10: R136

Palatnik JF, Allen E, Wu XL, Schommer C, Schwab R, Carrington JC, Weigel D (2003) Control of leaf morphogenesis by microRNAs. Nat 425: 257-263

Park W, Li JJ, Song RT, Messing J, Chen XM (2002) CARPEL FACTORY, a dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. Curr Biol 12: 1484-1495

Qiu CX, Xie FL, Zhu YY, Guo K, Huang SQ, Nie L, Yang ZM (2007) Computational identification of microRNAs and their targets in *Gossypium hirsutum* expressed sequence tags. Gene 395: 49-61

Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP (2002) MicroRNAs in plants. Genes Dev 16: 1616-1626

Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of plant microRNA targets. Cell 110: 513-520

Song C, Fang J, Li X, Liu H, Chao CT (2009) Identification and characterization of 27 conserved microRNAs in citrus. Planta 230: 671-685

Stark A, Bushati N, Jan CH, Kheradpour P, Hodges E, Brennecke J, Bartel DP, Cohen SM, Kellis M (2008) A single Hox locus in *Drosophila* produces functional microRNAs from opposite DNA strands. Genes Dev 22: 8-13

Sunkar R, Girke T, Jain PK, Zhu JK (2005) Cloning and characterization of microRNAs from rice. Plant Cell 17: 1397-1411

Sunkar R, Kapoor A, Zhu JK (2006) Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in *Arabidopsis* is mediated by downregulation of miR398 and important for oxidative stress tolerance. Plant Cell 18: 2051-2065

Sunkar R (2010) MicroRNAs with macro-effects on plant stress responses. Sem Cell Dev Biol doi:10.1016/j.semcdb.2010.04.001

Talmor-Neiman M, Stav R, Frank W, Voss B, Arazi T (2006) Novel micro-RNAs and intermediates of micro-RNA biogenesis from moss. Plant J 47: 25-37

Unver T, Budak H (2009) Conserved microRNAs and their targets in model grass species *Brachypodium distachyon*. Planta 230: 659-669

Williams L, Grigg SP, Xie MT, Christensen S, Fletcher JC (2005) Regulation of *Arabidopsis* shoot apical meristem and lateral organ formation by microRNA *miR166g* and its *AtHD-ZIP* target genes. Dev 132: 3657-3668

Wu G, Poethig RS (2006) Temporal regulation of shoot development in *Arabidopsis thaliana* by *miR156* and its target *SPL3*. Dev 133: 3539-3547

Xie FL, Frazier TP, Zhang B (2010) Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (*Panicum virgatum*). Planta 232: 417-434

Xie FL, Huang SQ, Guo K, Xiang AL, Zhu YY, Nie L, Yang ZM (2007) Computational identification of novel microRNAs and targets in *Brassica napus*. FEBS Lett 581: 1464-1474

Yin Z, Li C, Han X, Shen F (2008) Identification of conserved microRNAs and their target genes in tomato (*Lycopersicon esculentum*). Gene 414: 60-66

Yin Z, Shen F (2010) Identification and characterization of conserved microRNAs and their target genes in wheat (*Triticum aestivum*). Genet Mol Res 9: 1186-1196

Zhang BH, Pan XP, Cannon CH, Cobb GP, Anderson TA (2006a) Conservation and divergence of plant microRNA genes. Plant J 46: 243-259

Zhang BH, Pan XP, Anderson TA (2006b) Identification of 188 conserved maize microRNAs and their targets. FEBS Lett 580: 3753-3762

Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA (2006c) Evidence that miRNAs are different from other RNAs. Cell Mol Life Sci 63: 246-254

Zhang BH, Pan XP, Cobb GP, Anderson TA (2006d) Plant microRNA: A small regulatory molecule with big impact. Dev Biol 289: 3-16

Zhang BH, Pan XP, Stellwag EJ (2008a) Identification of soybean microRNAs and their targets. Planta 229: 161-182

Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA (2005) Identification and characterization of new plant microRNAs using EST analysis. Cell Res 15: 336-360

Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA (2006e) Computational identification of microRNAs and their targets. Comput Biol Chem 30: 395-407

Zhang BH, Wang QL, Wang KB, Pan XP, Liu F, Guo TL, Cobb GP, Anderson TA (2007) Identification of cotton microRNAs and their targets. Gene 397: 26-37

Zhang J, Zeng R, Chen J, Liu X, Liao Q (2008b) Identification of conserved microRNAs and their targets from *Solanum lycopersicum* Mill. Gene 423: 1-7

Zhang WW, Luo YP, Gong X, Zeng WH, Li SG (2009) Computational identification of 48 potato microRNAs and their targets. Computl Biol Chem 33: 84-93

Zhao B, L.Ge, Liang R, Li W, Ruan K, Lin H, Jin Y (2009) Members of miR-169 family are induced by high salinity and transiently inhibit the NF-YA transcription factor. BMC Mol Biol 10: 29

Zhou XF, Sunkar R, Jin HL, Zhu JK, Zhang WX (2009) Genome-wide identification and analysis of small RNAs originated from natural antisense transcripts in *Oryza sativa*. Genome Res 19: 70-78

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406-3415

Table 2-1. Tobacco miRNAs identified by GSS analysis

miRNA	Query miRNA	Mature Sequence	MN	ML	Arm	PL	A/U	A%	C%	G%	U%	G+C%	A+U%	MFE	MFEI	GSS
nta-miR156a	ath-miR156abcdef	TGACAGAAGAGAGTGAGCAC	0	20	5'	91	0.69	22	23	23	32	46.15	53.85	35.7	0.85	FH457758
nta-miR156b	ath-miR156abcdef	TGACAGAAGAGAGTGAGCAC	0	20	5'	91	0.759	24	21	23	32	43.96	56.04	39.8	1.00	FH457758
nta-miR156c	ath-miR156abcdef	TGACAGAAGAGAGTGAGCAC	0	20	5'	82	0.808	26	20	23	32	42.68	57.32	44.2	1.26	FH457758
nta-miR156d	ath-miR156g	TGACAGAAGAGAGTGAGCAC	1	20	5'	88	0.741	23	23	24	31	46.59	53.41	44.3	1.08	FI020831
nta-miR156e	ath-miR156g	TGACAGAAGAGAGTGAGCAC	1	20	5'	89	0.719	26	18	20	36	38.20	61.80	44	1.29	FH473188
nta-miR156f	ath-miR156g	TGACAGAAGAGAGTGAGCAC	1	20	5'	85	0.778	25	21	22	32	43.53	56.47	48.8	1.32	FH326966
nta-miR156g	ath-miR156g	TGACAGAAGAGAGTGAGCAC	1	20	5'	88	0.962	28	17	25	30	42.05	57.95	44.6	1.21	ET937470
nta-miR156h	ath-miR156g	TGACAGAAGAGAGTGAGCAC	1	20	5'	82	0.808	26	21	22	32	42.68	57.32	38	1.09	ET819266
nta-miR156i	ath-miR156g	TGACAGAAGAGAGTGAGCAC	1	20	5'	91	0.8	26	18	23	33	40.66	59.34	41.9	1.13	FH384749
nta-miR156j	ath-miR156h	TGACAGAAGATAGAGAGCAC	1	20	5'	92	1.111	33	16	22	29	38.04	61.96	42.4	1.21	FI017971
nta-miR156k	ath-miR156h	TGACAGAAGAGAGAGAGCAC	0	20	5'	84	0.846	26	18	25	31	42.86	57.14	38.4	1.07	FI053964
nta-miR156l	ath-miR156h	TGACAGAAGATAGAGAGCAC	1	20	5'	89	0.893	28	19	21	31	40.45	59.55	42	1.17	FH597550
nta-miR156m	ath-miR156h	TGACAGAAGATAGAGAGCAC	1	20	5'	83	0.923	29	20	19	31	39.76	60.24	41.8	1.27	FH496267
nta-miR156n	ath-miR156h	TGACAGAAGATAGAGAGCAC	1	20	5'	93	0.967	31	17	19	32	36.56	63.44	44.1	1.30	FH212536
nta-miR156o	ath-miR156h	TGACAGAAGATAGAGAGCAC	1	20	5'	82	0.821	28	20	18	34	37.80	62.20	42.4	1.37	FH496190
nta-miR158	ath-miR158a	CCCCAAATGTAGACAAATCA	4	20	5'	509	0.753	26	21	18	35	38.70	61.30	127.6	0.65	FH645844
nta-miR159a	ath-miR159b	TTTGGATTGGAGGGAG_TCTT	2	20	3'	306	0.973	35	12	16	36	28.43	71.57	98.4	1.13	FH433954
nta-miR159b	osa-miR159e	CTTGGACTGAAGGGAGCTCCC	3	21	3'	170	0.957	26	20	27	27	47.06	52.94	83.6	1.05	FH010541
nta-miR159c	osa-miR159f	CTTGGACTGAAGGGAGCTCCC	3	21	3'	167	1.114	29	20	25	26	44.31	55.69	73.3	0.99	FH952771
nta-miR160a	ath-miR160abc, ptc-miR160abcd, osa-miR160abcd	TGCCTGGCTCCCCTGTATGCCA	1	22	5'	88	0.652	17	27	30	26	56.82	43.18	30.1	0.60	ET784633
nta-miR160b	ptc-miR160g	TGCCTGGCTCCCTGGATGCCA	0	21	5'	83	1.211	28	27	23	23	49.40	50.60	40.7	0.99	FH452321
nta-miR160c	ptc-miR160g	TGCCTGGCTCCCTGTATGCCA	1	21	5'	88	0.727	18	28	28	25	56.82	43.18	47.8	0.96	FH739060
nta-miR160d	ptc-miR160g	TGCCTGGCTCCCTGTATGCCA	1	21	5'	81	0.842	20	27	30	23	56.79	43.21	45.4	0.99	FH704477
nta-miR162a	ath-miR162ab, ptc-miR162abc, osa-miR162a	TCGATAAACCTCTGCATCCAG	0	21	3'	76	1.063	22	29	28	21	56.58	43.42	41.9	0.97	FH731640
nta-miR162b	ath-miR162ab, ptc-miR162abc, osa-miR162a	TCGATAAACCTCTGCATCCAG	0	21	3'	88	1.556	32	24	24	20	47.73	52.27	36.8	0.88	FH964062
nta-miR164a	ath-miR164ab, ptc-miR164abcde, osa-miR164abf	TGGAGAAGCAGGGCACATGCT	2	21	5'	74	0.727	22	26	23	30	48.65	51.35	38.7	1.08	FI022139
nta-miR164b	ath-miR164ab, ptc-miR164abcde, osa-miR164abf	TGGAGAAGCAGGGCACTTGCT	2	21	5'	75	0.714	20	27	25	28	52.00	48.00	23.1	0.59	FH417064
nta-miR164c	ptc-miR164f	TGGAGAAGCAGGGCACATGCT	0	21	5'	76	0.731	25	20	21	34	40.79	59.21	36.5	1.18	FH438552
nta-miR164d	osa-miR164d	TGGAGAAGCAGGGCACGTGCA	1	21	5'	196	0.721	25	20	20	35	40.31	59.69	62.9	0.80	FH143731
nta-miR166a	ath-miR166abcdefg, ptc-miR166abcdefghijklm, osa-miR166abcdfn	TCGGACCAGGCTTCATTCCCC	0	21	3'	121	0.917	27	25	18	30	42.98	57.02	42.6	0.82	FH965444
nta-miR166b	ath-miR166abcdefg, ptc-miR166abcdefghijklm, osa-miR166abcdfn	TCGGACCAGGCTTCATTCCCC	0	21	3'	89	0.885	26	20	25	29	44.94	55.06	41.9	1.05	FH430060
nta-miR166c	ath-miR166abcdefg, ptc-miR166abcdefghijklm, osa-miR166abcdfn	TCGGACCAGGCTTCATTCCCC	0	21	3'	116	1	29	21	21	29	41.38	58.62	43.7	0.91	FH400103
nta-miR166d	ath-miR166abcdefg, ptc-miR166abcdefghijklm, osa-miR166abcdfn	TCGGACCAGGCTTCATTCCCC	0	21	3'	133	0.593	24	17	19	41	35.34	64.66	48.8	1.04	FH680448
nta-miR166e	ath-miR166abcdefg, ptc-miR166abcdefghijklm, osa-miR166abcdfn	TCGGACCAGGCTTCATTCCCC	0	21	3'	97	0.897	27	24	20	30	43.30	56.70	34.8	0.83	FH597450

nta-miR166f	ath-miR166abcdefg, ptc-miR166abcdefghijklm, osa-miR166abcdfn	TCGGACCAGGCTTCATTCCCC	0	21	3'	174	0.672	24	21	21	35	41.38	58.62	66.3	0.92	FH178749
nta-miR166g	ath-miR166abcdefg, ptc-miR166abcdefghijklm, osa-miR166abcdfn	TCGGACCAGGCTTCATTCCCC	0	21	3'	160	0.83	28	16	23	33	39.38	60.63	61.4	0.97	FH204390
nta-miR166h	ath-miR166abcdefg, ptc-miR166abcdefghijklm, osa-miR166abcdfn	TCGGACCAGGCTTCATTCCCC	0	21	3'	176	0.579	25	15	16	43	31.82	68.18	57.8	1.03	FH117109
nta-miR166i	ath-miR166abcdefg, ptc-miR166abcdefghijklm, osa-miR166abcdfn	TCGGACCAGGCTTCATTCCTC	1	21	3'	329	0.269	14	21	11	53	32.52	67.48	56.56	0.53	FH665134
nta-miR166j	ath-miR166abcdefg, ptc-miR166abcdefghijklm, osa-miR166abcdfn	ATGGACCAGGCTTCATCCTCC	4	21	5'	316	0.906	24	23	25	27	48.73	51.27	87.9	0.57	FH807276
nta-miR167a	ath-miR167ab, ptc-miR167abcd, osa-miR167abc	TGAAGCTGCCAGCATGATCTA	0	21	5'	77	0.808	27	16	23	34	38.96	61.04	36.9	1.23	FI032116
nta-miR167b	ath-miR167ab, ptc-miR167abcd, osa-miR167abc	TGAAGCTGCCAGCATGATCTA	0	21	3'	221	0.907	31	24	12	34	35.29	64.71	42.5	0.54	FI053673
nta-miR167c	ath-miR167ab, ptc-miR167abcd, osa-miR167abc	TGAAGCTGCCAGCATGATCTA	0	21	5'	96	1.32	34	18	22	26	39.58	60.42	38.5	1.01	FI053595
nta-miR167d	ath-miR167ab, ptc-miR167abcd, osa-miR167abc	TGAAGCTGCCAGCATGATCTA	0	21	5'	112	1.367	37	13	23	27	36.61	63.39	30	0.73	FH953139
nta-miR167e	ath-miR167ab, ptc-miR167abcd, osa-miR167abc	TGAAGCTGCCAGCATGATCTA	0	21	5'	71	1.053	28	21	24	27	45.07	54.93	32.8	1.03	FH564066
nta-miR167f	ath-miR167ab, ptc-miR167abcd, osa-miR167abc	TGAAGCTGCCAGCATGATCTA	0	21	5'	100	1.64	41	15	19	25	34.00	66.00	35.2	1.04	FH567995
nta-miR167g	ath-miR167ab, ptc-miR167abcd, osa-miR167abc	TGAAGCTGCCAGCATGATCTG	1	21	5'	205	0.811	29	13	22	36	34.63	65.37	62.6	0.88	FI007537
nta-miR168a	ath-miR168ab	TCGCTTGGTGCAGGTCGGGAA	0	21	5'	311	0.57	22	15	24	39	38.91	61.09	90.1	0.74	FH980352
nta-miR168b	ath-miR168ab	TCGCTTGGTGCAGGTCGGGAC	1	21	5'	113	0.808	19	27	31	23	58.41	41.59	51.2	0.78	FH567195
nta-miR168c	ath-miR168ab	TCGCTTGGTGCAGGTCGGGAC	1	21	5'	146	0.744	22	20	29	29	48.63	51.37	53.8	0.76	FH459764
nta-miR168d	ath-miR168ab	TCGCTTGGTGCAGGTCGGGAC	1	21	5'	135	0.784	21	24	27	27	51.11	48.89	51.2	0.74	FH180059
nta-miR169a	ath-miR169a, ptc-miR169abc, osa-miR169a	CAGCCAAGGATGACTTGCCGA	0	21	5'	89	0.852	26	25	19	30	43.82	56.18	27.1	0.69	FH972338
nta-miR169b	ath-miR169a, ptc-miR169abc, osa-miR169a	CAGCCAAG_ATGACTTGCCCG	3	20	5'	66	1	24	24	27	24	51.52	48.48	34.6	1.02	FH476610
nta-miR169c	ath-miR169a, ptc-miR169abc, osa-miR169a	CAGCCAAGGATGACTTGCCGA	0	21	5'	104	1.097	33	15	22	30	37.50	62.50	44.5	1.14	ET924900
nta-miR169d	ath-miR169a, ptc-miR169abc, osa-miR169a	CAGCCAAGGATGACTTGCCGG	1	21	5'	113	0.806	26	19	24	32	42.48	57.52	51.1	1.06	FH958024
nta-miR169e	ath-miR169a, ptc-miR169abc, osa-miR169a	TAGCCAAGGATGACTTGCCTG	3	21	5'	118	1.03	29	21	22	28	43.22	56.78	48.1	0.94	FH149361
nta-miR169f	ath-miR169a, ptc-miR169abc, osa-miR169a	TAGCCAAGGATGACTTGCCTA	2	21	5'	88	0.84	24	27	20	28	47.73	52.27	34.2	0.81	ET804779
nta-miR169g	ath-miR169a, ptc-miR169abc, osa-miR169a	TACCCAAGGATGACTTGCCTA	3	21	5'	125	0.767	26	20	19	34	39.20	60.80	46.6	0.95	FH615062
nta-miR169h	ath-miR169a, ptc-miR169abc, osa-miR169a	CAGCCAAGAATGACTTGCCGA	1	21	3'	190	0.964	28	18	24	29	42.11	57.89	76.6	0.96	ET768006

nta-miR169i	ath-miR169a, ptc-miR169abc, osa-miR169a	CAGCCAAGGATGAATTGCCGG	2	21	5'	71	0.696	23	21	24	32	45.07	54.93	35	1.09	FH458739
nta-miR169j	ath-miR169a, ptc-miR169abc, osa-miR169a	CGGCCACTGATGACTTGCCGA	3	21	5'	68	1.067	24	29	25	22	54.41	45.59	35	0.95	FH141870
nta-miR169k	ath-miR169a, ptc-miR169abc, osa-miR169a	CAGCCAAGGATGACTTGCCCA	1	21	5'	70	1	24	24	27	24	51.43	48.57	41.4	1.15	FH575132
nta-miR169l	ath-miR169a, ptc-miR169abc, osa-miR169a	CAGCCAAGGATGACTTGCCGA	0	21	5'	68	0.941	24	24	28	25	51.47	48.53	45.1	1.29	FH575132
nta-miR169m	ath-miR169a, ptc-miR169abc, osa-miR169a	CAGCCAAGAATGACTTGCCGA	1	21	5'	68	0.944	25	22	26	26	48.53	51.47	36.2	1.10	FH575132
nta-miR169n	ath-miR169defg, ptc-miR169n	GTAGCCAAGGACGACCTGCCG	4	21	5'	70	1	23	29	26	23	54.29	45.71	37.8	0.99	FH969365
nta-miR169o	ath-miR169defg, ptc-miR169n	TGCAGCCAAG_ATGACTTGCCG	2	21	3'	188	0.845	26	19	24	31	43.09	56.91	75.5	0.93	FI067867
nta-miR169p	ath-miR169defg, ptc-miR169n	GTAGCCAAGGATGACTTGCCT	3	21	5'	107	0.5	21	18	21	41	38.32	61.68	42.9	1.05	FI059630
nta-miR169q	ath-miR169defg, ptc-miR169n	GTAGCCAAGGATGACTTGTCG	3	21	5'	103	0.625	24	17	19	39	36.89	63.11	24.7	0.65	FI059630
nta-miR169r	ath-miR169defg, ptc-miR169n	ATAGCCAAGGATGACTTGCCT	3	21	5'	101	0.632	24	19	20	38	38.61	61.39	38.5	0.99	FH942012
nta-miR169s	ath-miR169defg, ptc-miR169n	GTAGCCAAGGATGACTTGCCT	3	21	5'	97	0.487	20	18	23	40	40.21	59.79	39.7	1.02	FI022404
nta-miR169t	ath-miR169defg, ptc-miR169n	ATAGCCAAGGATGACTTGCCT	3	21	5'	109	0.944	31	17	19	33	35.78	64.22	44.5	1.14	FI045443
nta-miR169u	ath-miR169defg, ptc-miR169n	GTAGCCAAGGATGACTTGCCT	3	21	5'	149	0.797	32	14	15	40	28.86	71.14	73.7	1.71	FH961287
nta-miR169v	ath-miR169defg, ptc-miR169n	GTAGCCAAGGATGACTTGCCT	3	21	5'	106	0.537	21	19	22	39	40.57	59.43	47.6	1.11	FH573350
nta-miR169w	ath-miR169defg, ptc-miR169n	ATAGCCAAGGATGACTTGCCT	3	21	5'	126	1.105	33	16	21	30	36.51	63.49	48.4	1.05	FH573350
nta-miR169x	ath-miR169defg, ptc-miR169n	GTAGCCAAGGATGACTTGTCG	3	21	3'	285	0.87	28	18	21	32	39.65	60.35	105.8	0.94	FH569139
nta-miR169y	ath-miR169defg, ptc-miR169n	GTAGCCAAGGATGACTTGCCT	3	21	5'	104	0.452	18	19	22	40	41.35	58.65	46.6	1.08	FH104969
nta-miR169z	ath-miR169defg, ptc-miR169n	ATAGCCAAGGATGACTTGCCT	3	21	5'	104	0.632	23	22	18	37	40.38	59.62	38.7	0.92	FH228557
nta-miR169aa	ath-miR169defg, ptc-miR169n	T_AGCCAAG_ATGACTTGCCT	3	19	5'	114	0.78	28	20	16	36	35.96	64.04	40.2	0.98	FH083956
nta-miR169ab	ath-miR169defg, ptc-miR169n	ATAGCCAAGGATGACTTGCCT	3	21	5'	107	1.031	31	17	22	30	39.25	60.75	45.6	1.09	FH327284
nta-miR169ac	ath-miR169defg, ptc-miR169n	GTAGCCAAGGATGACATGCCC	4	21	5'	100	0.714	25	16	24	35	40.00	60.00	34.2	0.86	ET930050
nta-miR169ad	ath-miR169defg, ptc-miR169n	ATAGCCAAGGATGACTTGCCC	3	21	5'	106	0.97	30	18	21	31	38.68	61.32	37.3	0.91	FH063630
nta-miR169ae	ath-miR169defg, ptc-miR169n	GTAGCCAAGGATGACTTGCCT	3	21	5'	95	1.125	28	21	25	25	46.32	53.68	39.6	0.90	ET789470
nta-miR169af	ath-miR169defg, ptc-miR169n	GTAGCCAAGGACGACCTGCCG	4	21	5'	70	0.938	21	26	30	23	55.71	44.29	33.2	0.85	FH476610
nta-miR169ag	ath-miR169defg, ptc-miR169n	GTAGCCAAGGACGACCTGCCG	4	21	5'	70	1.063	24	29	24	23	52.86	47.14	36.1	0.98	FH144793
nta-miR169ah	ath-miR169defg, ptc-miR169n	GTAGCCAAGGACGACCTGCCG	4	21	5'	70	1.063	24	29	24	23	52.86	47.14	37.9	1.02	FH144793
nta-miR169ai	ath-miR169defg, ptc-miR169n	GTAGCCAAGGACGACCTGCCG	4	21	5'	69	0.882	22	28	26	25	53.62	46.38	32.4	0.88	FH141870
nta-miR169aj	ath-miR169defg, ptc-miR169n	GCAGCCAAGGATGACTTGCCG	2	21	5'	70	0.889	23	24	27	26	51.43	48.57	42.6	1.18	FH476610

nta-miR169ak	ptc-miR169q	TAGCCAAGGACGACCTGCCAA	3	21	5'	68	1.059	26	26	22	25	48.53	51.47	25.7	0.78	ET767927
nta-miR169al	ptc-miR169q	TAGCCAAAGACGAC_TGCCTA	3	20	5'	101	1.179	33	23	17	28	39.60	60.40	36.9	0.92	FH536979
nta-miR169am	ptc-miR169q	TAGCCAAGAACAACTTGCCGG	3	21	5'	78	1	28	23	21	28	43.59	56.41	32	0.94	ET970604
nta-miR169an	ptc-miR169q	TAGCCAAGGACGACCTGCAGA	4	21	5'	68	1.125	26	26	24	24	50.00	50.00	28.8	0.85	ET998274
nta-miR169ao	ptc-miR169r	TAGCCAATGATGACTTGCCTT	2	21	5'	97	0.541	21	21	21	38	41.24	58.76	42.4	1.06	FH387003
nta-miR169ap	osa-miR169e	TAGCTAAGGATGACTTGCCTG	2	21	5'	69	0.519	20	17	23	39	40.58	59.42	28.1	1.00	ET994654
nta-miR169aq	osa-miR169no	TAGCCAAGGATGACTAGCCTA	2	21	5'	86	1.217	33	21	20	27	40.70	59.30	32.2	0.92	ET972603
nta-miR169ar	ath-miR169a, ptc-miR169abc, osa-miR169a	TAGCCAAGGACGACCTGCCGA	3	21	5'	68	1.063	25	28	24	24	51.47	48.53	33.1	0.95	ET768006
nta-miR169as	ath-miR169a, ptc-miR169abc, osa-miR169a	CAGCCAAGGATGACTTGCCGG	1	21	3'	192	0.81	24	19	26	30	45.31	54.69	87.8	1.01	FH141870.1
nta-miR169at	ath-miR169a, ptc-miR169abc, osa-miR169a	TAGCCAAGGACGACCTGCCGA	3	21	3'	192	1.234	30	26	19	24	45.31	54.69	66.9	0.77	FH141870.1
nta-miR169*	ath-miR169g*	TCCGGCAAGTTGTTCTTGGCT	2	21	3'	76	0.955	28	20	24	29	43.42	56.58	37.4	1.13	ET970604
nta-miR170a	ath-miR170	TGATTGAGCCGTGCCAATATC	1	21	3'	78	0.731	24	18	24	33	42.31	57.69	40.5	1.23	FH969274
nta-miR170b	ath-miR170	TGATTGAGCCGTGCCAATATC	1	21	3'	81	0.88	27	19	23	31	41.98	58.02	25.5	0.75	FI076426
nta-miR170c	ath-miR170	TGATTGAGCCGTGCCAATATC	1	21	3'	77	0.84	27	18	22	32	40.26	59.74	37.2	1.20	FH971960
nta-miR170d	ath-miR170	TGATTGAGCCGTGCCAATATC	1	21	3'	89	0.828	27	15	26	33	40.45	59.55	39.4	1.09	FH999148
nta-miR170e	ath-miR170	TGATTGAGCCGTGCCAATATC	1	21	3'	81	0.724	26	16	22	36	38.27	61.73	37.9	1.22	FH403151
nta-miR170f	ath-miR170	TGATTGAGCCGTGCCAATATC	1	21	3'	76	0.815	29	18	17	36	35.53	64.47	30.8	1.14	FH609240
nta-miR170g	ath-miR170	TGATTGAGCCGCGTCAATATC	1	21	3'	83	1	33	16	19	33	34.94	65.06	34.5	1.19	FH381610
nta-miR170h	ath-miR170	TGATTGAGCCGCGTCAATATC	1	21	3'	87	0.897	30	14	23	33	36.78	63.22	38.1	1.19	FH298029
nta-miR170i	ath-miR170	TGATTGAGCCGTGCCAATATC	1	21	3'	79	0.581	23	15	23	39	37.97	62.03	44.1	1.47	ET948878
nta-miR170j	ath-miR170	AGATTGAGCCGTGCCAATATC	2	21	3'	82	0.778	26	17	24	33	41.46	58.54	34	1.00	FH064515
nta-miR171	osa-miR171g	ATGATGAGCCGAACCAATATC	4	21	3'	72	0.87	28	19	21	32	40.28	59.72	31.5	1.09	FH433016
nta-miR172a	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	AGAATCTTGATGATGCTGCAT	0	21	3'	97	1.259	35	12	25	28	37.11	62.89	42.8	1.19	FH723904
nta-miR172b	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	AGAATCTTGATGATGCTGCAT	0	21	3'	90	1	31	16	22	31	37.78	62.22	42.5	1.25	FI014481
nta-miR172c	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	AGAATCTTGATGATGCTGCAT	0	21	3'	111	1.147	35	18	16	31	34.23	65.77	40.4	1.06	FH399852
nta-miR172d	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	AGAATCTTGATGATGCTGCAT	0	21	3'	96	1.296	36	16	20	28	35.42	64.58	40.9	1.20	FH419018
nta-miR172e	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	AGAATCTTGATGATGCTGCAT	0	21	3'	108	1.321	34	16	24	26	39.81	60.19	39.5	0.92	FH304912
nta-miR172f	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	AGAATCTTGATGATGCTGCAG	1	21	3'	126	0.809	30	14	18	37	32.54	67.46	45.6	1.11	FH723628
nta-miR172g	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	TGAATCTTGATGATGCTGCAT	1	21	3'	68	0.885	34	15	13	38	27.94	72.06	24.4	1.28	FH562896
nta-miR172h	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	GGAATCTTGATGATGCTGCAG	2	21	3'	86	0.667	23	16	26	35	41.86	58.14	39.9	1.11	FH354512
nta-miR172i	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	TGAATCTTGATGATACTACAT	3	21	3'	104	0.943	32	23	12	34	34.62	65.38	33.4	0.93	ET735116
nta-miR172j	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	TGAATCTTGAAGATGCTGCAG	3	21	3'	102	1.032	31	19	20	30	38.24	61.76	40.8	1.05	FH995689

nta-miR172k	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	AGAATCATGATGATGCTGCAT	1	21	3'	111	1.086	34	17	17	32	34.23	65.77	34.4	0.91	FH310965
nta-miR172l	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	TGAATCTTGATGATGCTACAC	3	21	3'	97	0.794	28	25	12	35	37.11	62.89	29.1	0.81	FH354528
nta-miR172m	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	TGAATCTTGATGATGCTACAT	2	21	3'	108	0.757	26	24	16	34	39.81	60.19	33.4	0.78	FH354528
nta-miR172n	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	TGAATCTTGATGGTGCTGCAC	3	21	3'	96	0.771	28	20	16	36	35.42	64.58	44.8	1.32	FH419018
nta-miR172o	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	GAAATCTTGATGATGCTGCAT	2	21	3'	93	1.111	32	15	24	29	38.71	61.29	38.7	1.08	FI036771
nta-miR172p	ath-miR172ab, ptc-miR172abcf, osa-miR172ad	TGAATCTTGATGATGCTACAT	2	21	3'	99	0.875	28	24	15	32	39.39	60.61	34.2	0.88	FI036850
nta-miR172q	ath-miR172cd	TGAATCTTGAAGATGCTGCAT	3	21	3'	124	1.154	36	18	15	31	32.26	67.74	47.7	1.19	ET710327
nta-miR172r	ath-miR172cd	TGAATCTTGATAATGCTGCAG	3	21	3'	111	0.921	32	17	17	34	34.23	65.77	32.9	0.87	FH310965
nta-miR172s	ath-miR172cd	TGAATCTTGATGATGCTCCAC	3	21	3'	86	1.5	35	26	16	23	41.86	58.14	36.5	1.01	FH354512
nta-miR172t	ath-miR172e, ptc-miR172de, osa-miR172b	GGAATCTTGATGATGCTGCAT	0	21	3'	219	0.974	35	14	16	36	29.68	70.32	60.7	0.93	FH413962
nta-miR172u	ath-miR172e, ptc-miR172de, osa-miR172b	AGAATCTTGATGATGCTGCAT	1	21	3'	107	1.31	36	19	19	27	37.38	62.62	41.1	1.03	FH729470
nta-miR172v	ptc-miR172gh	GTG_ATCTTGATGATGCTACAT	4	21	3'	93	0.9	29	24	15	32	38.71	61.29	28.4	0.79	FI036771
nta-miR172w	ath-miR172cd	AGAATCTTGATGATGCTGCAG	0	21	5'	124	0.867	31	15	18	36	32.26	67.74	45.1	1.13	ET710327
nta-miR172x	ath-miR172e, ptc-miR172de, osa-miR172b	AGAATCTTGATGATGCTGCAT	1	21	3'	104	1.061	34	12	23	32	34.62	65.38	47.5	1.32	ET735116.1
nta-miR172y	ath-miR172e, ptc-miR172de, osa-miR172b	AGAATCTTGATGATGCTGCAT	1	21	3'	99	1.143	32	15	24	28	39.39	60.61	42.9	1.10	FI036850
nta-miR172*	ath-miR172b*	GCAGCACCATCAAGATTCAC	1	20	5'	92	1.36	37	16	20	27	35.87	64.13	37.9	1.15	FH419018
nta-miR319a	osa-miR319ab	GTGGACTGAAGGAAGCTCTC	3	20	3'	178	1.149	30	24	19	26	43.26	56.74	70.3	0.91	FH739509
nta-miR319b	osa-miR319ab	CTGGACTGTAGGGTGCTCCA	3	20	3'	243	1.286	26	28	26	20	53.91	46.09	80.5	0.61	FH934552
nta-miR319c	ptc-miR319abcd	TTGGATTGAAGGGAGCTCCA	2	20	3'	166	0.704	23	19	26	33	44.58	55.42	80.2	1.08	FH956766
nta-miR319d	ptc-miR319i	TTGGACTGAAGGGAGCTCCC	1	20	3'	174	0.978	26	22	26	26	47.70	52.30	82.6	1.00	ET896014
nta-miR390a	ath-miR390ab	AAGCTCAGGAGGGATAGCGCC	0	21	5'	86	0.815	26	20	23	31	43.02	56.98	41.9	1.13	FH491982
nta-miR390b	ath-miR390ab	AAGCTCAGGAGGGATAGCACC	1	21	5'	109	0.811	28	18	20	34	38.53	61.47	40.1	0.95	FH554722
nta-miR393a	ath-miR393ab	TCCAAAGGGATCGCATTGATCC	0	22	5'	90	0.611	24	19	17	40	35.56	64.44	36.7	1.15	FH947232
nta-miR393b	ath-miR393ab	TCCAAAGGGATCGCATTGATCC	0	22	5'	91	0.771	30	18	14	38	31.87	68.13	32.8	1.13	FH976110
nta-miR393c	ath-miR393ab	TCCAAAGGGATCGCATTGATCC	0	22	5'	97	0.684	27	19	15	39	34.02	65.98	39.2	1.19	FH956670
nta-miR394a	ath-miR394ab	TTGGCATTCTGTCCACCTCC	0	20	5'	58	0.867	22	28	24	26	51.72	48.28	29.5	0.98	FH156387
nta-miR394b	ath-miR394ab	TTGGCATTCTGTCCACCTCC	0	20	5'	66	0.778	21	27	24	27	51.52	48.48	30.8	0.91	FH076040
nta-miR395a	ath-miR395ade	CTGAAGTGTTTGGGGGAACTC	0	21	3'	63	0.714	24	19	24	33	42.86	57.14	32	1.19	FH986265
nta-miR395b	ath-miR395ade	CTGAAGTGTTTGGGGGAACTC	0	21	3'	96	0.968	31	15	22	32	36.46	63.54	43.2	1.23	FI059310
nta-miR395c	ath-miR395ade	TTGAAGTGTTTGGGGGAACTC	1	21	3'	81	0.517	19	20	26	36	45.68	54.32	37.1	1.00	FH255253
nta-miR395d	ath-miR395ade	CTGAAGTGTTTGGGAGAACTC	0	21	3'	79	0.783	23	23	25	29	48.10	51.90	38.3	1.01	FH137172
nta-miR395e	ath-miR395bcf	CTGAAGTGTTTGGGGGAACTC	1	21	3'	85	0.833	29	18	18	35	35.29	64.71	38.7	1.29	FH345146
nta-miR395f	osa-miR395t	CTGAAGTGTTTGGGGAA_CTC	2	20	3'	79	1	30	22	18	30	39.24	60.76	22.1	0.71	FH572407
nta-miR396a	ath-miR396a	TCCCACAGCTTTCTTGGACTT	3	21	5'	116	1.313	36	22	15	28	36.21	63.79	40	0.95	FI088049
nta-miR396b	ath-miR396b	TTCCACAGCTTTCTTGAACTT	0	21	5'	106	0.789	28	18	18	36	35.85	64.15	42.2	1.11	FH472448
nta-miR396c	ath-miR396b	TTCCACAGCTTTCTTGAACTG	1	21	5'	116	1.162	37	15	16	32	31.03	68.97	37.9	1.05	FI074237
nta-miR396d	ath-miR396b	TTCCACAGCTTTCTTGATCAT	2	21	5'	150	0.632	24	19	19	38	38.00	62.00	45.4	0.80	FH568190
nta-miR397a	ath-miR397a	TCATTGAGTGCAGCGTTGATG	0	21	5'	105	0.432	18	21	19	42	40.00	60.00	39.3	0.94	FH735277
nta-miR397b	osa-miR397b	TTATT_AGTGCAGCGTTGAGG	2	20	3'	224	0.604	25	17	18	41	34.82	65.18	56.8	0.73	FH016551

nta-miR398	ath-miR398a	TGTGTTCTCAGGTCGCCCCTG	2	21	3'	99	0.568	21	20	21	37	41.41	58.59	41.4	1.01	FH375629
nta-miR399a	ath-miR399a	TGCCAAAGGAGATTTGCCCGG	1	21	3'	107	1.071	28	22	23	26	45.79	54.21	39.4	0.80	FH941588
nta-miR399b	ath-miR399a	TGCCAAAGGAGAGTTGCCCTG	1	21	3'	98	0.848	29	19	18	34	37.76	62.24	43.9	1.19	FH364170
nta-miR399c	ath-miR399a	TGCCAAAGGAGAGTTGCCCTG	1	21	3'	87	0.793	26	17	23	33	40.23	59.77	32.2	0.92	ET857336
nta-miR399d	ath-miR399a	TGCCAAAGGAGAGTTGCCCTA	2	21	3'	84	0.69	24	19	23	35	41.67	58.33	41.7	1.19	FH036611
nta-miR399e	ath-miR399a	TGCCAAAGAAGATTTGCCCCA	2	21	3'	80	0.84	26	24	19	31	42.50	57.50	37.7	1.11	FI032447
nta-miR399f	ath-miR399a	TGCCAAAGAAGATTTGCCCCG	1	21	3'	85	0.885	27	22	20	31	42.35	57.65	39.1	1.09	FH533031
nta-miR399g	ath-miR399bc	TGCCAAAGGAGAGCTGCCCTG	1	21	3'	96	0.833	26	22	21	31	42.71	57.29	42.2	1.03	FH960835
nta-miR399h	ath-miR399bc	CGCCAAAGGAGAGCTGCCCTG	2	21	3'	100	0.903	28	23	18	31	41.00	59.00	38.9	0.95	FH945862
nta-miR399i	osa-miR399efg	TGCCAAAGAAGATTAGCCCAG	2	21	3'	84	1.588	32	20	27	20	47.62	52.38	33.7	0.84	FH491959
nta-miR399j	osa-miR399i	CTCCAAAGGAGAGCTGCCCTG	2	21	3'	72	0.818	25	22	22	31	44.44	55.56	28.8	0.90	FH550672
nta-miR399k	ath-miR399bc	CGCCAAAGGAGAGCTGCCCTG	2	21	3'	97	1	29	23	20	29	42.27	57.73	35.8	0.87	FH595714
nta-miR399l	ath-miR399bc	CGCCAAAGGAGAGCTGCCCTG	2	21	3'	91	0.926	27	22	21	30	42.86	57.14	40.9	1.05	FH374208
nta-miR399m	ath-miR399bc	TGCCAAAGGAGAGCTGCCCTG	1	21	3'	96	0.758	26	19	21	34	39.58	60.42	33.6	0.88	ET839331
nta-miR399n	ath-miR399bc	CGCCAAAGGAGAGCTGCCCTG	2	21	3'	90	0.846	24	26	21	29	46.67	53.33	33.3	0.79	ET971344
nta-miR399o	ath-miR399a	TGCCAAAGGAGAGCTGCCCTG	2	21	3'	96	0.758	26	19	21	34	39.58	60.42	33.6	0.88	FH048330
nta-miR399p	ath-miR399a	TGCCAAAGAAGATTTGCCCCG	2	21	3'	85	0.679	22	24	21	33	44.71	55.29	43.5	1.14	ET964451
nta-miR399q	ath-miR399a	TGCCAAAGGAGAGCTGCCCTG	2	21	3'	95	0.833	26	21	21	32	42.11	57.89	43.2	1.08	ET866309
nta-miR400	ath-miR400	TATAAAAGTATTATAAGTCAC	2	21	3'	177	1.357	43	14	11	32	25.42	74.58	25.6	0.57	ET986965
nta-miR413	osa-miR413	CTAGTT_CACTTGTTCTTCAC	2	20	5'	479	1.044	30	17	25	28	41.96	58.04	137.2	0.68	ET991182
nta-miR414a	ath-miR414	TCCTCTTCATCATCATCGTCC	2	21	5'	200	1.565	36	21	21	23	41.00	59.00	50.8	0.62	ET797397
nta-miR414b	ath-miR414	TCATCTTCATCATCATCATCA	1	21	3'	65	1.211	35	17	18	29	35.38	64.62	18	0.78	ET703934
nta-miR414c	ath-miR414	TCGTCGTCATCATCATCGTCA	2	21	3'	316	0.979	29	19	22	30	40.51	59.49	85.7	0.67	ET724796
nta-miR437a	osa-miR437	AAAGTTA_AGAAGTTTGACTT	1	20	3'	184	0.867	35	12	12	41	23.91	76.09	29.7	0.68	ET857652
nta-miR437b	osa-miR437	AAAGTTA_AGAAGTTTGACTC	2	20	3'	175	0.716	33	8.6	12	46	20.57	79.43	36.5	1.01	ET927882
nta-miR476a	ptc-miR476a	TAGAAATCCTTCTTTGCAAAA	2	21	5'	440	1.156	34	21	17	29	37.27	62.73	100.1	0.61	ET954922
nta-miR476b	ptc-miR476bc	TAGTACTTCTTCTTTGAAAAA	2	21	3'	89	0.692	30	16	10	44	25.84	74.16	11.7	0.51	ET759843
nta-miR479	ptc-miR479	AGTGATATTGGTCCGGCTCATC	1	22	5'	78	1.25	32	23	19	26	42.31	57.69	34.9	1.06	FH014960
nta-miR482	ptc-miR482.1	TCTACTCCTCGCCATTCC	2	18	3'	551	1.957	42	18	19	21	37.21	62.79	107.3	0.52	ET784952
nta-miR529	osa-miR529a	CCGTACCCTC_CTCTTCTTC	2	19	5'	162	2.308	37	28	19	16	46.91	53.09	33.7	0.44	ET754007
nta-miR530	ptc-miR530a	TGCATTTGCACCTGCACCTC	1	20	5'	122	0.571	23	19	18	40	36.89	63.11	46.2	1.03	ET832485
nta-miR776	ath-miR776	TCTAAGTCTTCAAATTGATGTT	2	22	5'	57	0.654	30	11	14	46	24.56	75.44	8.9	0.64	ET987053
nta-miR809	osa-miR809abcdefgh	TTTATGTGAGACATGTTAGAAT	3	22	3'	52	1.105	40	9.6	13	37	23.08	76.92	15.3	1.28	FH194316
nta-miR810	osa-miR810b.2	AAGTGATTTAATTAAGCGATA	4	21	3'	61	0.917	36	6.6	18	39	24.59	75.41	9.3	0.62	FH745256
nta-miR815	osa-miR815abc	AAGCGGGATTGAGGAGATGGAC	4	22	3'	259	0.654	20	16	32	31	48.26	51.74	87	0.70	FH900233
nta-miR816a	osa-miR816	TTGACATATTTTACTACTAG	3	20	3'	83	0.529	22	19	18	41	37.35	62.65	13.6	0.44	FI025249
nta-miR816b	osa-miR816	TTGACATATTTTTCTACAAA	3	20	5'	70	0.815	31	19	11	39	30.00	70.00	13.6	0.65	FI012435
nta-miR817	osa-miR817	T_GAACTTGAGGCCCGATTTT	4	20	3'	88	0.613	22	16	27	35	43.18	56.82	34.7	0.91	FH022347
nta-miR819a	osa-miR819abcdefghijk	TGATGTTATAAGGCTTTCTAGC	3	22	5'	379	0.928	34	12	18	36	29.82	70.18	82.8	0.73	FH867946
nta-miR819b	osa-miR819abcdefghijk	TGATGTTATAAGGCTTTCTAGC	3	22	3'	287	1.12	36	17	15	32	32.06	67.94	90.6	0.98	FH866280
nta-miR819c	osa-miR819abcdefghijk	GCTTGTTAGAAGACTTTCTAGC	4	22	3'	48	0.8	25	23	21	31	43.75	56.25	10.6	0.5048	FH945450
nta-miR826a	ath-miR826	TAGCTCTCTTTTTGGATACGTG	4	22	5'	197	0.813	31	16	15	38	30.96	69.04	38.7	0.63	FH729260
nta-miR826b	ath-miR826	TAGTTTTGTTTTGGATACGTT	4	21	3'	238	0.456	22	10	20	48	30.25	69.75	42.7	0.59	FH624853
nta-miR827a	ath-miR827	GTGGCTAACCATCAACAAACT	4	21	3'	238	1.037	36	18	12	34	29.83	70.17	38.9	0.55	FH905715
nta-miR827b	ath-miR827	_TAGATGACCATCAACAAACA	2	20	3'	81	1.409	38	15	20	27	34.57	65.43	30.1	1.08	FH050498
nta-miR827c	ath-miR827	_TAGATGAACATCAACAAACA	3	20	3'	81	1.048	27	20	15	26	34.57	53.09	29.8	1.06	FH050498
nta-miR827d	osa-miR827	TTGGATGACCATCAGCGAAAT	3	21	3'	230	0.728	29	16	15	40	30.87	69.13	42	0.59	FH580824
nta-miR827e	osa-miR827	CAAGATGACCATCAGCAAGAT	4	21	5'	63	1.625	41	19	14	25	33.33	66.67	18	0.86	FH663601
nta-miR828	ath-miR828	TCTTGCTCAAATGAGTATTCCA	1	22	5'	102	0.75	26	21	18	35	38.24	61.76	33.9	0.87	FH429120
nta-miR829a	ath-miR829.2	GCAATTAAAGCTTCAAGGAAA	4	21	5'	59	0.905	32	14	19	36	32.20	67.80	12.9	0.68	FI047045
nta-miR829b	ath-miR829.2	AAATTTAAAGCTTCAAGGTGA	4	21	5'	492	1.675	40	18	18	24	36.38	63.62	104.1	0.58	FH738194
nta-miR830a	ath-miR830	TATCTATTTTGAGAAGAAGAA	3	21	5'	196	0.935	30	19	20	32	38.78	61.22	43	0.57	FH519699
nta-miR830b	ath-miR830	CAACTATTTTGAGAAGTACTT	4	21	5'	47	1.059	38	15	11	36	25.53	74.47	13.5	1.13	FH347570
nta-miR832-5p	ath-miR832-5p	TGCGTAGTTCGGGAATCGAAA	4	21	3'	356	0.84	28	19	19	33	38.48	61.52	83.3	0.61	FH636795

nta-miR833-5p	ath-miR833-5p	TTCTTGTTGTACTCGGTATTGT	4	22	5'	131	0.58	22	9.9	30	38	39.69	60.31	32.8	0.63	FH919355
nta-miR835a-3p	ath-miR835-3p	CGGAGACGATACGCAAGAATC	4	21	3'	521	1.313	28	22	28	21	50.29	49.71	176.7	0.67	FI079781
nta-miR835b-3p	ath-miR835-3p	TAGAGAAGATACGGAAGAAAC	3	21	3'	95	0.882	32	9.5	23	36	32.63	67.37	18.6	0.60	FH203543
nta-miR837a-5p	ath-miR837-5p	TTTAGTTTCTTGTTCGTTTCC	3	21	3'	165	0.918	27	23	20	30	43.03	56.97	42	0.59	FH275012
nta-miR837b-5p	ath-miR837-5p	GTTAGTTTCTTGTTCGTTTGC	4	21	3'	450	0.918	30	16	21	33	37.33	62.67	111.7	0.66	FH906632
nta-miR838a	ath-miR838	TGTTCTTCTACTTCTTGCGTG	4	21	5'	220	0.784	26	15	25	34	40.00	60.00	63.3	0.72	FH398247
nta-miR838b	ath-miR838	TTTTCTTCTACTTCTCTAATA	4	21	3'	152	1.102	36	17	15	32	32.24	67.76	36.6	0.75	FH744663
nta-miR838c	ath-miR838	TTTTCTTCTATTTCTTGCAAT	3	21	3'	250	0.75	26	16	22	35	38.40	61.60	65.2	0.68	FH947863
nta-miR838d	ath-miR838	TTTTCTTCTACTTCTATAACC	4	21	3'	103	1.467	43	14	15	29	28.16	71.84	20	0.69	FH611602
nta-miR844a	ath-miR844	CATTAAGGTTGCTTATAAGCT	4	21	5'	50	0.65	26	8	26	40	34.00	66.00	13.8	0.81	ET785156
nta-miR844b	ath-miR844	TTTAAAGATTGATTATAAGCT	4	21	5'	54	0.652	28	15	15	43	29.63	70.37	7.7	0.48	FI082192
nta-miR844*a	ath-miR844*	TTTTTTATCCATCTTACTAGTT	3	21	3'	118	1.406	38	18	17	27	34.75	65.25	28.5	0.70	FH906408
nta-miR844*b	ath-miR844*	ATATAAGCCATCTCACTACTA	4	21	5'	427	1.013	36	14	14	36	27.87	72.13	75.8	0.64	FH664076
nta-miR845a	ath-miR845a	TGGCTCTGATACCAATTGATG	1	21	5'	635	1.045	33	20	17	31	36.22	63.78	158.8	0.69	FH038181
nta-miR845b	ath-miR845a	GTGCTCTGATACCAATTGATG	2	21	5'	429	0.993	32	21	14	32	35.43	64.57	95	0.63	FH912444
nta-miR845c	ath-miR845a	CAGCTCTGATACCAATTGATA	1	21	3'	331	0.736	24	25	17	33	42.30	57.70	85.3	0.61	FI018518
nta-miR846a	ath-miR846	TAGAATTGAAGTGCTTGTAGT	3	21	5'	212	0.706	28	16	16	40	31.60	68.40	41.3	0.62	FH720254
nta-miR846b	ath-miR846	TTGAATTGAAGTGATTGAAGC	3	21	5'	95	0.667	25	19	18	38	36.84	63.16	22	0.63	ET743069
nta-miR847	ath-miR847	TCATCTCCTCTTCTTCTTGTTC	3	22	3'	70	0.739	24	23	20	33	42.86	57.14	20.9	0.70	FH993619
nta-miR854a	ath-miR854abcd	TTTGAGGATAGGGAGGAGGTG	3	21	3'	122	0.805	27	20	20	34	39.34	60.66	32.7	0.68	FH180611
nta-miR854b	ath-miR854abcd	TTTGAGGATAGGGAGGAGGTG	3	21	5'	130	0.762	25	21	22	32	43.08	56.92	41.4	0.74	FH346535
nta-miR854c	ath-miR854abcd	GTTGAGGATAGGGCGGAGGAG	2	21	5'	121	0.656	17	25	31	26	56.20	43.80	54.8	0.81	FH885315
nta-miR854d	ath-miR854abcd	GTTGAGGATATGGAGGAGGAG	2	21	5'	130	0.675	21	23	25	31	48.46	51.54	41	0.65	FH620540
nta-miR854e	ath-miR854abcd	TTTGAGGATATGGAGGAGGAG	3	21	5'	130	0.622	22	22	22	35	43.85	56.15	37.7	0.66	FH678339
nta-miR854f	ath-miR854abcd	GAGAAGAATAGGGAGGAGGAG	3	21	5'	501	1.061	35	15	18	33	32.93	67.07	117.5	0.71	FH559057
nta-miR854g	ath-miR854abcd	GA_GAGGATAAGGAGGAGGAG	2	20	5'	510	0.735	30	12	16	41	28.24	71.76	114.7	0.80	FH409279
nta-miR859a	ath-miR859	TCTCTCTGTTGTGATGCCAAG	3	21	5'	88	1.167	32	18	23	27	40.91	59.09	19.4	0.54	FH717042
nta-miR859b	ath-miR859	TCGTCTCTGTTGTGAAGAAAAG	4	22	5'	121	1	26	16	31	26	47.11	52.89	42	0.74	FH133509
nta-miR859c	ath-miR859	TCTCTCTGTTGTGATGCCAGT	4	21	3'	413	0.688	27	17	17	39	34.62	65.38	78.5	0.55	FH237129
nta-miR863-5p	ath-miR863-5p	TTATGTCTTGT__ATCTCAAT	2	19	5'	224	0.857	29	14	22	34	36.16	63.84	60.4	0.75	FH668759
nta-miR864-5p	ath-miR864-5p	TCAGGTTTGATTGACCTCAAA	2	21	5'	539	0.824	27	17	24	33	40.45	59.55	181.4	0.83	FH933962
nta-miR865-5p	ath-miR865-5p	ATGAATTTGGAT_TAGTTGAG	2	20	3'	390	0.984	32	18	18	32	36.41	63.59	101.1	0.71	ET720349
nta-miR870	ath-miR870	TAATTTGGTTTTTCTT_GATC	2	20	3'	79	1	33	8.9	25	33	34.18	65.82	15.1	0.56	FH628555
nta-miR1426	osa-miR1426	AGAATCTTGATGAA__TTAAA	3	19	5'	51	1.048	43	5.9	9.8	41	15.69	84.31	7.8	0.98	FH061373
nta-miR1435a	osa-miR1435	TTTCATA_GTCAAACTTTTT	2	19	3'	367	1.231	36	17	18	29	34.33	65.67	90.4	0.72	ET893573
nta-miR1435b	osa-miR1435	TTTCTTAA_TCAAACTTTTG	2	19	5'	45	0.619	29	8.9	16	47	24.44	75.56	8.5	0.77	FH265925
nta-miR1435c	osa-miR1435	CTTCTTAAGTCAAAGTTTTT	2	20	5'	47	0.625	32	6.4	11	51	17.02	82.98	2.2	0.28	FH180299
nta-miR1435d	osa-miR1435	TTTCTTAA_TCAA_CTTTTT	2	18	3'	135	0.96	36	13	14	37	27.41	72.59	27.3	0.74	FH505991
nta-miR1435e	osa-miR1435	TTTCTGAAGTCAA_CTTTTT	2	19	3'	216	1.23	42	11	13	34	23.61	76.39	34.6	0.68	FI031600
nta-miR1436a	osa-miR1436	ACTATATGGGACGGAGGGAGT	2	21	3'	252	0.989	37	10	15	38	25.00	75.00	61	0.97	ET964119
nta-miR1436b	osa-miR1436	ATATTATGAGACGGAGGGAGT	2	21	3'	275	1.084	37	13	15	35	28.00	72.00	74.1	0.96	FH197776
nta-miR1439a	osa-miR1439	TTTTGAAACGGAGGGAGTATT	2	21	3'	261	0.944	32	20	13	34	33.72	66.28	66.5	0.76	FH057665

Note: Nucleotide changes between the predicted miRNAs and the query miRNAs are denoted in red; underscores (_) represent nucleotide deletions in the potential tobacco miRNAs as compared to the query miRNAs. MN: number of mismatches nucleotides; ML: length of the mature miRNA; PL: length of the miRNA precursor sequence; Arm: mature miRNA location within the hairpin structure; GSS: GSS sequence where the miRNA is derive from, only one GSS is listed per miRNA even though some miRNAs are located in more than one GSS sequence.

Table 2-2. Precursor sequences of the 259 identified tobacco miRNAs

miRNA	Precursor
nta-miR156a	TGACAGAAGAGAGTGAGCACAGATGGTGTTTTCTTGCATGATATGTTTATCCCATGCCTGAAGCTATCCGTGCTCTCTGCTCTATCTGCCA
nta-miR156b	TGACAGAAGAGAGTGAGCACATATGGTGTTTTCTTGCATGATATGTTTATCCCATGCCTGAAGCTATGCGTGCTCACAGCTCTATCTGTCA
nta-miR156c	TGACAGAAGAGAGTGAGCACACATGGTATTTTCTTGCATATGATATGCTTGAAGCTATGCGTGCTCACTGCTCTATCTGTCA
nta-miR156d	TGACAGAAGAGAGTGAGCACACATGGTCTTTTCTTGCATGAAATCCGTCTATGCTTGAAGCTATGCGTGCTCGCTGCTCTATCTGTCA
nta-miR156e	TGACAGAAGAGAGTGAGCACACATGGTCTTTTCTTGCATTGATATATATTTATATGCTTGAAGCTATGCGTGCTCACTCTCTATCTGTA
nta-miR156f	TGACAGAAGAGAGTGAGCACACATGATTCTTGCATGGATAATTGCCCATGCTTAGAAGCTATGTGTGCTTTCTCTCTCTCTGTCA
nta-miR156g	TGACAGAAGAGAGTGAGCACAGATGGTGTTTTCTTGCATAGAAAGATGAATATGCTTGAAGCTATGCGTGCTCACTCTCTATCTGTCA
nta-miR156h	TGACAGAAGAGAGTGAGCACACATGGTGTTTTCTTGCATATGATATACTTGAACCTATGCGTGCTCACTGCTCTATCTGTCA
nta-miR156i	TGACAGAAGAGAGTGAGCACACATGGAGTTTTCTTGCATTGATATATATATATATATGCTTGAAGCTATGCGTGCTCGCTGCTCTATCTGTCA
nta-miR156j	TGACAGAAGATAGAGAGCACTAATGATGATATGCTAAGTAAATGTAGGGGCAAAAGCATCTCAATTCATTTGTGCTCTCTATGCTTCTGTCA
nta-miR156k	TGACAGAAGAGAGAGAGCACAATCTGTCATGAGCAAAATGATGATGTTGCTGTTTGTTGGAATGTGCTCTCTCTTCTTCTGCCA
nta-miR156l	TGACAGAAGATAGAGAGCACAGATGATGATAAGATGCTAATTGGAAGCTTTCTGCACCTTAATCCTTTGTGCTCTCTAGTCTTCTGTCA
nta-miR156m	TGACAGAAGATAGAGAGCACAGATAATGAAGTGCATGAAAATTGTCTGCACCTCACTCCTTTGTGCTCTTTATTCTTCTGTCA
nta-miR156n	TGACAGAAGATAGAGAGCACTGATGATGATATGCTAATTAAATTTGTGCAGCAAAAGCATCTCAATTCATTTGTGCTCTCTATGCTTCTGTCA
nta-miR156o	TGACAGAAGATAGAGAGCACAGATAATGAAGTGCTGAAAATTTTCTGCACCTCACTCTTTTGTGCTCTTTATTCTTCTGTCA
nta-miR158	CCCCAAATGTAGACAAATCAACCGTTGGTTCACTTGTCCCTAATACTGAAACTGACATTGGCCATGTACAAGTAGTACTATTACCACCACCAAAAATGCCGATGTTAATGTTATTATTTGTAGTTGTACTAAATAAAACCTCGTCGTGACCGCGACTCATTTGTGGTTCTCCCCAAGGAAATCTTCCTTGCTCATTAGCCCACATTGTTGCTGAATTTGTTCCACTATTGTACTGGTTATTTTGCCATAATTGTGATGAGATCACGGGCTGATGATCGCTACTCATGAGGCTAATGTAATCAAGATTCTGATTTCCCCGGAATAAAGTATCAGAGTAATTCTTACCTTGTCCAACATTATTTACATGTATAATATCTCGTTTAGGTAATTTCTCAGCTTCCACGTTATTGCTCATTAATCCTCTGGCTAACGTGACAGTTTGTCCTGTTTGTGACAATGTTGTCTTGTTTTCAAGTGAGGAAACTGATGGTTTAATTGCTGCTTTGGAA
nta-miR159a	TATACTCCCTCTCGTTCAAAATAATTGATGTTTTAGATTTGGGCACATGGATTAAGGAATTCACTTACTCCTAAAAATTAAGAGGTATTTTGACTAAAATACCCTTAATTAGAATCTGCCTTTTATATCAAGTACACATAAAAGAATGTGTGTCTATGCATTAAAATTGTATATTGAATTAGGAATGTATAAAGGGCAAATTTGGGAAAAATATTTAAATACCTCCTTGATTTTGTGATACATAAATTATTTTGAACCAACTTTTAGAAGCTAAAATATCGATTATTTTGGATTGGAGGGAGTCTT
nta-miR159b	GAGAGCTTTCTTCAGTCCACACATGGGAGGTGATAAGGTTCAATTTGCTGCCGACTCATTCATCCAAATGCTGAGGTTAGATAGTCAAAAGCATCTCAGTGGCTGAATGAATGAAGCGGGAGTACAAGTTGTGTCTTAAGCTTCCTGTGCTTGGACTGAAGGGAGCTCCC
nta-miR159c	GAGAGCTTTCTTAGTCCACAATCAGGGGGCAATATGACTCAATTAGTTGCCGACTCATTCACACAAATGCTAAGGTTTGGATTGAAAAAGCCCTTAGTAACTGAGTGAATGAAGCGGTAGACAAGTTGAATCTTAAGCTTCCTGTGCTTGGACTGAAGGGAGCTCCC
nta-miR160a	TGCCTGGCTCCCCTGTATGCCACACACTTTTCACCAAATCCCGCGAGTTCTTTGATTGGCTGATCAGTGGGGTGGGGGTGCCGAGGAA

nta-miR160b	TGCCTGGCTCCCTGGATGCCAACTAAGAACAGCATTGTCAAAATGTTTGACACTATTCCTAGTTGGCCCCAAGGAGACAGGCA
nta-miR160c	TGCCTGGCTCCCTGTATGCCACACACTTTCACCAATCCGCGAGTTCTTTGATTGGCTGATCAGTGGGTGGCGTGCGAGGAGCCAAGCA
nta-miR160d	TGCCTGGCTCCCTGTATGCCATTTGCAGAGCTCATCGGAATACCGTTGGGTCCCCGTGAATGGCGTATGAGGAGCCAAGCA
nta-miR162a	CTGGAGGCAGCGGTTCATCGATCTGTTCCCTGAACGCAACAGCGCTGGGAATCGGTCGATAAACCTCTGCATCCAG
nta-miR162b	CTGGAGGCAGCGGTTCATCGATCTGTTCCCTGAAAGGCGATAAACAAAATATAGCAACAGGAATCGGTCGATAAACCTCTGCATCCAG
nta-miR164a	TGGAGAAGCAGGGCACATGCTAGATTACTCGGATTTCTATCCTTTGAGAATTTTGCATGTGCCCTCGCCCTCCA
nta-miR164b	TGGAGAAGCAGGGCACTTGCTGGATTACGCTGACCTTCGATCTTCTGAGAATTTTGCATGTGCCCCGATCCTCCA
nta-miR164c	TGGAGAAGCAGGGCACATGCTAAATTACTTGGTAACTTTGATCTTTCAAGAATTTTGCATGTGCTCTTGCTCTCCA
nta-miR164d	TGGAGAAGCAGGGCACGTGCATTACTAACTCATGCACAGGAATTGTCAGTGAAAATTTTTATGGCATTCATCTGACATCATAATTCATAGTACCATTACTGTACTTCCTAGCGTAGTGGAAAATAATGGTTTGGTACTTCACATGTATGTGTTTCCGTTTCCGTTGAGTTAGTTCTTCATGTGCCTGTCTTCCCCA
nta-miR166a	GGGGAATGTTGTCTGGCTCGAAGCCATTATCTCCTCTTGATTTAACCTACTTCTTCTTAACTTCAAAAATAGCAGAACAAACAATCATGAGATAGTGTCGTCGGACCAGGCTTCATTCCCC
nta-miR166b	GGGGAGTGCTGTCTGGTTCGAAACCATTCACCAAATTAGAGTATCAAGAGTTGATGATTTGAATGATTTCGGACCAGGCTTCATTCCCC
nta-miR166c	GGGAAATGTCGTCTGGTTCGAGATCTTTCATTACAAGGCAAAACCCTAAACTAAAGATATTTATATGAGTGTAGTCATCACGTACGTGAATGATTTCGGACCAGGCTTCATTCCCC
nta-miR166d	GGGGAATGTTGTTTGGCTCGAGGATTCTTCAATCTTGATCAAATTCTTTCTATTTTTTATAGATCTGCATATTAATATAGATCTATAATATTGAATCATGATTTTGGTGTCGTCGGACCAGGCTTCATTCCCC
nta-miR166e	GGAGAATGTTGTCTGGTTCGTGACCAATCTGTAAACAGTAATTAGCTAACTGTCAACATTCATCACTTGAATGATCTCGGACCAGGCTTCATTCCCC
nta-miR166f	GGGGAATGTTGTCTGGCTCGAGGTCACTAACTTGATCCAATTTGAACTCTCTTTCCCTTTTTGAGATGTCATCTCCATTAATTAGCAAATAGATCAATAGTTCTTGATTGACATTCTCTATTAAGAGGGATCATGGATCATGGTTTAGTGTCGTCGGACCAGGCTTCATTCCCC
nta-miR166g	GGGGAATGTTGTCTGGCTCGAGGTCACCAACTAGATCCATAAGCTCCACTTTATGCTTATTTGAACATAAAGATTAGTTTTGTAATATAAAGATTCGTTGATTATATGAAAAGGGGTTATGGATCATGAGTTAGTGTTGTCGGACCAGGCTTCATTCCCC
nta-miR166h	GGGGAATGTTGTCTGGCTCGAGGAATCTTTAATCTTGATCAAATTCTTTACACAGATCTATCATTATATATGTTTTATATATTTTTTAGATCTTATATAGATCTATAATTGTTAGTTGTTAGACCTATGAATTTTGATCATATTTTTTTGCGTCGTCGGACCAGGCTTCATTCCCC
nta-miR166i	TTGGAATGAAGACTGATCCAAGATCCTTCTATTTCCTTTTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTTACCCTTTGATTTTTTGGGGTTTTTGATTCTTTTAGTCTAAATCTTAATCAATCTTTTACTTTTTCTCCTCTCTCTCTTTTTTATTTTTTAAATTTGTGTGTGTATTTTCGGAGTGTTTTTGCTTTTTAGCTAAAAGTAGTAATCTAATTCAGTGAGACATTTTGGAGTGATCTCGGACCAGGCTTCATTCCTC
nta-miR166j	ATGGACCAGGCTTCATCCTCCTATACGACATAAGGCGAGGTTCTCTAAATTCGAGGGCAGGGCTCGTAGGTGGTGGAAGTCCTATCTTCTTGGCAGACCAGCAGAGTCTCCTCCCATGACTTGGGACAGGTTCACCCGTATCTTTCTAGATAGGTATATTCCGCCCTCCCAGAGGGAAGAGTTACGGTTTCAGTTTGAGCAGCTCCAACAGGGTTAGATGTCAGTGATCGACTATAAGGCGAGGTTCTCTGAGTTGTCTCACCATGCACTTATGATACTACCTACCGACACATAAAAGAGTGAAGGTTTGTTTCAG
nta-miR167a	TGAAGCTGCCAGCATGATCTAAACTTACCTCTTGTAATTATTAAGTTTGAGGGAGATTAGATCATGTGGTGGCTTCA

nta-miR167b	ATGATAATGCCTCCAGCTTCCTACTACTATAGTATAAATAAATGCTAACAACAAACTCCTC CCTTCACTTTAAGACTCGCTCTCTTTTGCATTTTCTTCTTCTTCTCAATTCAACCACTAGTTA TATATAGCATTTCCTATACATTACTTTTTTGCAATATGAAATGGATACTTCAGATCCAAGTC TGAAAGGGAACAAGTGAAGCTGCCAGCATGATCTA
nta-miR167c	TGAAGCTGCCAGCATGATCTAAACTTGGCCAGTTAATAATTATACATAAATAAGAGAAAT GGATGGCCAAGGTTAGGTCATGCTTTGACAGCCTCA
nta-miR167d	TGAAGCTGCCAGCATGATCTAAACTTGACCAGGTTATTATAAATATATATATATGGAAAGA AAGAAAGAGATGGATGGATGGCCAAGTTAAGGTCATGTTTTGACAGCCTCA
nta-miR167e	TGAAGCTGCCAGCATGATCTAAACTTTTCTAGTTGCTCCGATGAGGAAAGATCAGATCATG TGGCAGCATC
nta-miR167f	TGAAGCTGCCAGCATGATCTAACCTTGGTCAAAATAAAATTGGAAAGGAATAAAGAAAAA TAAATTTTGACTATAGTTAGGTCATGCTCAGACAGCATCA
nta-miR167g	TGAAGCTGCTAGATGACCTGATGGAACCAAACCCATGAAAGCTACCTATATTCTCTTTCAA TAAAAACTTTGCAAAATTCCAATTTCTTTTCCAAATCACTTAGCTAAATTCAAACCCAAAA CACCTCAAGATTATTCTTTGCAAGAAAATATATATATAGCAAGAAATGGACTTTGTTCGTA CCAGATCATGCTGGCAGCTTCA
nta-miR168a	TCGCTTGGTGCAGGTCGGGAACTGACTCACCGGTGACGCTGTCCGCGCCGGAATCGGTGG CGTTAGTTTCGTTGATATTTTGTTTGCATTTATCGATATTTAGAGAGCTCTCTTTAGGACTT ATTTGTTAGTAATTTGTTTTGTTGAAATTAACTTGATGTATGTGTAGATAAGTAGAGAGTGT TTACCCCCGGTTGGTTAAGTAATAGATATGAAGAGAGTTCATTTTTTTTTTTGTTTTTAATAA ATTAGTTATGTATGAAAGTGCGCCTCGTCTATGGCTTAAGTTTGTTCCCGCCTTGCATCAAC TGA
nta-miR168b	TCGCTTGGTGCAGGTCGGGACCTGCCTCGCCGGCGACAATGATGTCAGCTGACGGTGACG GCGGCGTACCGATAAATATACATGCTGTTTGGGTCCCGCCTTGCATCAACTGA
nta-miR168c	TCGCTTGGTGCAGGTCGGGACTAATTCGCCGGTGGCGGCGGCGGCAATCACAACGGCGGT GATTGTTATTAAATGGAGATTAGATGTACGAAGTTATCAACATTTTTTGTTTTGCGGCGAA ATTTGTCCCGCCTTGCATCAACTGA
nta-miR168d	TCGCTTGGTGCAGGTCGGGACTAATTCGCCGGCGGCAATCACGTCGGCGGTGATTATTCTA CGAAGCTTAGATGTACGAAGTTATCAACAATGTTCCTTTGCGGCGAAATTTGGGTCCCGCC TTGCATCAACTGA
nta-miR169a	CAGCCAAGGATGACTTGCCGAAAATTTCCCACCAGCCTTTAATTTGAGGGTTTATATAATT CTTTCAACCGGCAAGCTTTCCTGGGCTA
nta-miR169b	CAGCCAAGATGACTTGCCCGGGATGTTGTAGTCAATGCAACGACTATCGGTAAGTCGTCCT GGCTA
nta-miR169c	CAGCCAAGGATGACTTGCCGACTCTTGCACATTAACAGTAATAAATATATATAGTATGGTT ATTCTTATTGGAAGAGAAAAAGTTGGCAAGTGGTCCTTGGCTA
nta-miR169d	CAGCCAAGGATGACTTGCCGGCAAGATCGATCAATTCTGCTCGATCGAGTACTTTGTAGTT CTCAACTGAAGAGTGGAGATTATCATTTTGATTGGCAAGTCATTTTTGGCTA
nta-miR169e	TAGCCAAGGATGACTTGCCTGCATACTACCCTCCACCTCAAAGGGGTACAGTGCTATTAAT AATTTAGCACAAAACATCTGTTTTGGGGTTGAAATGATAGGCAGTCATCTTGGGCTA
nta-miR169f	TAGCCAAGGATGACTTGCCTACATATCCTCCTCCCTCAAAACGTAGCTCCCTTTTGGGGTT GAAATACTAGGCAGTCATCTTGGGCTA
nta-miR169g	TACCCAAGGATGACTTGCCTAATGTACTATTTTTTCCCCCTAAAAGGGGCATTTGGTACAA GTTAAGAAAGACTCCCCCCCTTTTTATGGAGTTAAGTATATATTAGGAAGTCGTCTTTGGC TTT
nta-miR169h	TCGGCAGGTCGTCCTTGGCTACATTTTACTCTCTTCTGCTCATGTAAGGCTCTGAAAGACAT ATATAATTCGAGGTGAAGTCAGTGATTGAGTTTCACAAATCACATCCAATTGAATTGCTCA TGATTTGTCTAGAGAGTCATTGCAAGAAAAGGTAGGGAGTGGATTGCAGCCAAGAATGAC TTGCCGA
nta-miR169i	CAGCCAAGGATGAATTGCCGGCGTATATCATATTACTTGGCAAATTTTGCCGGCAGTTTGT TCCTTGGCTA
nta-miR169j	CGGCCACTGATGACTTGCCGAGATGCTGCAACCAATGCAACAACTTATCGGCAGGTCGTCC TTGGCTA
nta-miR169k	CAGCCAAGGATGACTTGCCCGAGATGTTGTAGCTAATGCAACGACTTATCGGCAAGGTCG TCCTTGGCTA

nta-miR169l	CAGCCAAGGATGACTTGCCGAGATGTTGTAGCCAATGCAATGACTTATCGGCAGGTCGTCCTTGGCTA
nta-miR169m	CAGCCAAGAATGACTTGCCGAGATGTTGTAGCCAATGCAATGACTTATTGGCAGGTCGTCCTTGGCTA
nta-miR169n	GTAGCCAAGGACGACCTGCCGATAAGTCGTTGCATTGGCTACAACATCTCGGCAAGTCATCCTTGGCTGC
nta-miR169o	CGGCAGGTCGTCCTTGGCTACATTTCACTCTTTTCTGCTCATGTAAGGCTCTGTAAGACGCATATATAGTTCGAGGTGAAGTCAGTAATTGCGTTTCACAAATCATATCCAGTTGAATTGATCTTAATTTGTCCAAAGAGTCATTGCATGAGAGGTAGAGAGTGGATTGCAGCCAAGATGACTTGCCG
nta-miR169p	GTAGCCAAGGATGACTTGCCTGCTCCATTATTTTAAAAGGTGTTTTGATTACACCATTTAATATTGGTAGCCTTTTTTGTTTGGTTTCAGGCAGTCTCCTTGGCTAT
nta-miR169q	GTAGCCAAGGATGACTTGTCGTGCACCATTTTGGAAAGCTCATTTATATATTTATATATGATGGTACAGTCCTTTCTTTCTTGGTTTCAGTCACTATGGCTAA
nta-miR169r	ATAGCCAAGGATGACTTGCCTGCTTCGATCTTGAGGCTGTTGAATTTTATATATAAAATTAGCATTCTTGTTTGATTTCCAGGCAGGTCTCCTTAGCTACT
nta-miR169s	GTAGCCAAGGATGACTTGCCTGCTCCATTTATTTTTAAAAGGGGGTGTTTGATTATACCATTTTTTGCTTTGGATTCAGGCAGTCTCCTTGGTTATC
nta-miR169t	ATAGCCAAGGATGACTTGCCTAAAACCTTTATTTATGTAGGGTTTTAAGCTTCAAATTAAAAGGCTTATGATCCTACAAAATGGTATTGTAGGCATCATCCGAGGCTTT
nta-miR169u	GTAGCCAAGGATGACTTGCCTGCACCATTTTGGAAAGCTCATTAATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATTGTGGTGTCCTTTCTTTCTTGGTTTCAGGCAGGTCACCATGGCTAAC
nta-miR169v	GTAGCCAAGGATGACTTGCCTTCATCTTATGGAAAGGTCTTTGCACCATTAATTGGTGGTGTTCCTTTCCTTATATTGGTTTTTAGGCAGGTCATCTTTAGCTAAC
nta-miR169w	ATAGCCAAGGATGACTTGCCTAGACCATTTTCAAGGGCTTATTTAATTAGCTTAGATAAAAATAAAAAGAGACAGAGAGATATAATCCCTTTTTGCATGGGTTGCAGGCAAGTCATCTTTGGCTAT
nta-miR169x	TGCAGTCTCCTTGGTTATCTTGACATGCTCTTCTCTTTTATGCTAGACTTTTCATTCTTCTATCACTACTAAAAATGTAAATTCCCGAACGGGCATCCCCTCGGAAATTGGCTCGTCGGAAAGGAATTTCTGACGGGAAAGGCCCTTCAGAAATATTTCCGACGATATCCGATTTACTTTTTTTTGGTAGTGTTCAAGTTAATTTACGAAGAAAGAAGAATGAGATAGAAGAGAGGTATAGCATAGTGAAACGAGTCTTATTTGGTAGCCAAGGATGACTTGTCG
nta-miR169y	GTAGCCAAGGATGACTTGCCTGTTCCATTATTTTTAAAAGGCGTTTGACCATTTAATATTGGTGGCCTTTTTTTGCTTGGTTTCAGGCAGTCTCCTTGGCTATC
nta-miR169z	ATAGCCAAGGATGACTTGCCTGCTTCATTCTTACATGCTAATATTGTATTATGTGCCAATTATCAGCATTCTTGTTTGATCACCAGGCAGTCTCCTTGGCTAGT
nta-miR169aa	TAGCCAAGATGACTTGCCTGCGCCATTTTGGAAAGGTCATTAATATATACTACTAATATATATTGATGATACAATCCTTTTCTTTCTTGGTTTCACGCACGTCACTATGGCTAA
nta-miR169ab	ATAGCCAAGGATGACTTGCCTGAACCATTTTCAAGGGCTTATTTAATTAGAGACAGATAGAGATAAGCCATTTTTGCATAGGTTGCAGGAAAGTCATCTTTGGCTAC
nta-miR169ac	GTAGCCAAGGATGACATGCCCAAACTCATATTATGTATATGTTTTAAAAGGCGAGCTATTAGTTCTAGATTATGTGTTTTGGGCTGTTGTCCGCGGTTAC
nta-miR169ad	ATAGCCAAGGATGACTTGCCCGGACCATTTTCAAGGGCTTATTTAATTAGCTAGATAAATAATAAAAGCCTTTTTGCATGAGTTGCAGGCAGTCAACTTTGGCTAT
nta-miR169ae	GTAGCCAAGGATGACTTGCCTTTCATCAATGCATACAAAATGGGAGCAAAAACCCCTTTTGATGGTTGGAATGCGTGGCAAGCATCTTTGGCGAC
nta-miR169af	GTAGCCAAGGACGACCTGCCGATAAGTTGTTGCATTGGTTGCAGCATCTCGGCAAGTCATCAGTGGCCGC
nta-miR169ag	GTAGCCAAGGACGACCTGCCGATAAGTCATTGCATTGGCTACAACATCTCGGCAAGTCATCCTTGGCTGC
nta-miR169ah	GTAGCCAAGGACGACCTGCCGATAAGTCGTTGCATTAGCTACAACATCTCGGCAAGTCATCCTTGGCTGC
nta-miR169ai	GTAGCCAAGGACGACCTGCCGATAAGTCGTTGCATTGGCTGCAACATCTTGACAAGTCTCCTTGGCTGC

nta-miR169aj	GCAGCCAAGGATGACTTGCCGAGATGTTGTAGCTAATGCAACGACTTATCTGCAGGTCGTCCTTGGCTAC
nta-miR169ak	TAGCCAAGGACGACCTGCCAATAAGTCATTGCATTGGCTACAACATCTCGGCAAGTCATTCTTGGCTG
nta-miR169al	TAGCCAAAGACGACTGCCTATAATACCATTTTCAAAAGTTGACTCTCATCGCCAAACTCTTTTGGACAAAATGGTGTAAATAGGCAAGTCATCCTTGGCTA
nta-miR169am	TAGCCAAGAACAACTTGCCGGAATTCGAAAATCATCGCAATTTTGCTTTGAAATTTGCCGGCAAGTCATCCTTGGCTG
nta-miR169an	TAGCCAAGGACGACCTGCAGATAAGTCGTTGCATTAGCTACAACATCTCGGCAAGTCATCCTTGGCTG
nta-miR169ao	TAGCCAATGATGACTTGCCTTCAACCTTTGGAGAGGTATTTACACCATTGATGGTGTCCTTTCCTTATATTGGTTTTCGGCAGGTCATCTTTAGCTA
nta-miR169ap	TAGCTAAGGATGACTTGCCTGCTCCATTATTTTAAAAGGTGTTTTGGTTTCAGGCAGTCTTCTTGGCTA
nta-miR169aq	TAGCCAAGGATGACTAGCCTACTCCATTTGTAAATTAAAAAAAAAAGGATCCTATCCATTTGGTTGCAGGCAGGTCATCCTGGCTA
nta-miR169ar	TAGCCAAGGACGACCTGCCGATAAGTCATTGCATTGGCTACAACATCTCGGCAAGTCATCCTTGGCTG
nta-miR169as	TCGGCAGGTCGTCCTTGGCTACATTTCACTCTCTTCTGCTCATGTAAGGCTTTGTAAGACGCATATATAGTTCGAGGTGAAGTCAGTGATTGCGTTTCACAAATCACATCCAGTTGAATTGCTCATAATTTGTCTAGAGAGTCATTGCATGAAGAGGTAGGGAGTGGATTGCAGCCAAGGATGACTTGCCGG
nta-miR169at	CCGGCAAGTCATCCTTGGCTGCAATCCACTCCCTACCTCTTCATGCAATGACTCTCTAGACAAATTATGAGCAATTCAACTGGATGTGATTTGTGAAACGCAATCACTGACTTCACCTCGAACTATATATGCGTCTTACAAAGCCTTACATGAGCAGAAGAGAGTGAAATGTAGCCAAGGACGACCTGCCGA
nta-miR169*	AGCCAAGGATGACTTGCCGGCAAATTTCAAAGCAAAATTGCGATGATTTTCGAATTCCGGCAAGTTGTTCTTGGCT
nta-miR170a	GATGTTGGTGCGGTTCAATGAGAAAGCAGTTGTCAAACTAGTTTTGACCCTACTTTTTGATTGAGCCGTGCCAATATC
nta-miR170b	GATATGGCCAGGTTCACTCAGACAGATACAGAACATCATTTGAATGTTGGAGTTTTGCTTTGATTGAGCCGTGCCAATATC
nta-miR170c	GATATTGGTGCGGTTCAATGAGAAAGCAGTACTCGACAAATTTTGACTCTACTTTTTGATTGAGCCGTGCCAATATC
nta-miR170d	GGTATTGGCCTGGTTCACTCAGACAACAAGATGAAACTATTTGAATATTGAATGGTGGAGTTTTGGTTTGATTGAGCCGTGCCAATATC
nta-miR170e	GATATTGGTGTGGTTCAATTAGATTACAATGTCTCCACATGATCGTGAAGTTTTGCAATTTGATTGAGCCGTGCCAATATC
nta-miR170f	GATGTTGGAATGGCTCAATCTAATCAAAATTTCCAACATATTTGGTCCTTTAATTTGATTGAGCCGTGCCAATATC
nta-miR170g	GATATTGATGCGACTCAATCTGAAGACGTGATTATATATAAATATAATCAGCCATGTGATTTTGATTGAGCCGCGTCAATATC
nta-miR170h	GATATTGATGCGACTCAATCTGAAGACGTGGTTATATATATAAATAGAGGTTTAGCCATGTGATTTTGATTGAGCCGCGTCAATATC
nta-miR170i	GATATTGGTGCGGTTCAATTAGAAAGCCGAGTTCTTTTTGTTTTAGAACTCTGTTATTTGATTGAGCCGTGCCAATATC
nta-miR170j	GATGTTGGTGAGGTTCAATCCGAAGACGTGTTTACGTTTTGTTTTCGTAAAAAACGATCTCAGATTGAGCCGTGCCAATATC
nta-miR171	GATATTTGTTTGGCTCATCACTTTCTGTTGAATTCATGAGACAGTCCAGAGATGATGAGCCGAACCAATATC
nta-miR172a	GTGTAGCATCATCAAGATTCACATGAAAAGGCAAGGTGAAATGAAATTAATTAAGAGCCATGGCTTTTTGAAGGTGAGAATCTTGATGATGCTGCAT
nta-miR172b	ATGTAGCATCATCAAGATTCACATATTATAAAGGCATGATGATTAACGACCATGGCCTTTTTGGAGGTGAGAATCTTGATGATGCTGCAT

nta-miR172c	CTGCAGCATTATCAAGATTCACATAGAAATCATGTGGAGCAGAAAATTAAATATTTTTCTCAAAATCTGCCCCCAAGTATTTTGAACATGAGAATCTTGATGATGCTGCAT
nta-miR172d	GTGCAGCACCATCAAGATTCACATAGAAATATGGAGCAGAAATTGAAATGAAATCTGCCCTATTAATTTGATATGAGAATCTTGATGATGCTGCAT
nta-miR172e	ATGTAGCATCATCAAGATTCACATAAAAAACGCATGGTGATGAGCTGATGAAATGACATTAATGGCCAAGGCAGCTTTCTGAAGGTGAGAATCTTGATGATGCTGCAT
nta-miR172f	ATGCAGCATCTTTCAAGATTCACATATAGTATATTTTTATGGAGTTCAATTGTGGGAGAAAAAAGTTAAATTTTCTTGCTCCCTATGGCTCCTTAATTTAATATGAGAATCTTGATGATGCTGCAG
nta-miR172g	ATGCAACACTTTCAAGATTTACATATATTTGTCATTAAAATCATATGTGAATCTTGATGATGCTGCAT
nta-miR172h	GTGGAGCATCATCAAGATTCACATAGCCTTTTTAGGGATTTTTGGCCCTATGGCTTATTGATATGGGAATCTTGATGATGCTGCAG
nta-miR172i	ATGCAGCATCATCAAGATTCTCACCTCCAAAAAAGCCATGGTCGTTAATTTTTCTTCATTAATCACCATGCCTTTACAATATGTGAATCTTGATGATACTACAT
nta-miR172j	ATGCAGCATCATCAGATCTCATGTTCAAAAAACTTAGGGGGCAGATTTAGAAAAATATTTCTGCTCCCCATGATTTCTATGTGAATCTTGAAGATGCTGCAG
nta-miR172k	CTGCAGCATTATCAAGATTCACATAGAAATCATGTGGAGCAGAAAATATGTTTCTAAAAATCTGCCACCCCTAAGTTTTTTTTGAATATGAGAATCATGATGATGCTGCAT
nta-miR172l	ATGCAGCATCATCAAGATTCTCACCTTCAAAAAGCCATGGCTCTTAATTAATTTCATTTCACCTTGCCTTTTCATGTGAATCTTGATGATGCTACAC
nta-miR172m	ATGCAGCATCATCAAGATTCTCACCTTCAGAAAGCTGCCTTGGCCATTAATGTCATTTCATCAGCTCATCACCATGCGTTTTTTATGTGAATCTTGATGATGCTACAT
nta-miR172n	ATGCAGCATCATCAAGATTCTCATATCAAATTAATAGGGCAGATTTCATTTCAATTTCTGCTCCATATTTCTATGTGAATCTTGATGGTGCTGCAC
nta-miR172o	ATGTAGCATCATCAAGATCACATGCAAACGCATGGTGATGAAATGATATTAATGGCATGGCAGCTTTTGAGTGAAATCTTGATGATGCTGCAT
nta-miR172p	ATGCAGCATCATCAAGATTCTCACCTTCAAAAGCTGCCATGGCCATTAATATCATTTCATCACCATGCGTTTTGCATGTGAATCTTGATGATGCTACAT
nta-miR172q	CTGCAGCATCATCAAGATTCTCATATTAAATTAAGGAGCCATAGGGAGCAAGAAAATTTAATTTTTTTCTCCCACAATTGAACTCCATAAAATATACTATATGTGAATCTTGAAGATGCTGCAT
nta-miR172r	ATGCAGCATCATCATGATTCTCATATTCAAAAAAAACTTAGGGGTGGCAGATTTTTAGAAACATATTTTCTGCTCCACATGATTTCTATGTGAATCTTGATAATGCTGCAG
nta-miR172s	CTGCAGCATCATCAAGATTCCCATATCAATAAGCCATAGGGCCAAAAATCCCTAAAAAGGCTATGTGAATCTTGATGATGCTCCAC
nta-miR172t	ATGCATCATCATCAAGATTCACAAACTTGAAAGGAGAAAGGATCTACATACGGGTTAGGTCGAACCCAGTATATTTGCTCAAATAGTGTATTTATCTTAAGAAATTTATTAAATATGTACATATATTAAATTTTGAAATCATTTATTACATTTTGAGGTCGTTGTTCTAAAATCCAAGATTCTCCTAAATTTCATATGGGAATCTTGATGATGCTGCAT
nta-miR172u	CTGCAGCATTATCAAGATTCACATAGAAATCATGTGGGGCAGAAAACATAGGTTCTAAAAATCTAACCCCAAGTTCTTTGAACATGAGAATCTTGATGATGCTGCAT
nta-miR172v	ATGCAGCATCATCAAGATTTCACTCAAAAGCTGCCATGCCATTAATATCATTTCATCACCATGCGTTTGCATGTGATCTTGATGATGCTACAT
nta-miR172w	GACGTCGTAGTAGTTCTAAGAGTATAATTTAATTCCTCGGTATCCCTCGTTCTTTTAAATTAAAAAAAGAGGGTGTTAACTTGAGGTATTTTATATGATATACACTTAGAACTTCTACGACGTA
nta-miR172x	ATGTAGTATCATCAAGATTCACATATTGTAAAGGCATGGTGATTAATGAAGAAAAATTAACGACCATGGCTTTTTTGGAGGTGAGAATCTTGATGATGCTGCAT
nta-miR172y	ATGTAGCATCATCAAGATTCACATGCAAAACGCATGGTGATGAAATGATATTAATGGCCATGGCAGCTTTTGAAGGTGAGAATCTTGATGATGCTGCAT
nta-miR172*	GCAGCACCATCAAGATTCACATAGAAATATGGAGCAGAAATTGAAATGAAATCTGCCCTATTAATTTGATATGAGAATCTTGATGATGCTGC

nta-miR319a	GGGAGCTCCCTTCAGTCCAAGCACAGATAAAAAGGGATGAATTTACCTCCCGCATCATTCA CACATTCACAGTAGCGTATAGAAATACATATATGATTTTCTACTGTGTTTGAATGAATGAG TCGGCAGCTATCCAAATCCCCTTCATCCACCTATGAGTGGACTGAAGGAAGCTCTC
nta-miR319b	TAACTGCACTGACATCAGTTGCTTAGCCTTACATCGCCCTTCCAAGGAATCTCAGGAAGGC ACCGCGTAATAGTTGGTGTCAGAGCCTAGGCTCGACATCGGACGAGGGAGCGCATTGCCA TTACCACCATTTCTGACCATGGTGAATCATGGGGACCGCACTGCGACCCTTGAACAGACGG CCCCTGAACTAAGAACCAGCCTAGGGAAAAGGTTGAACGACCTGGACTGTAGGGTGCTCC A
nta-miR319c	TGGAGCTCCTTGAAGTCCAACAGAGGATCTAACAGGTAAGATTGAGCTGCTGACCTATGG ATTCCTCAGCCCTATCTATTATTTGATTGGATAGGTTTGTGGGTTGCATATGTCAGGAGCTT CATTGCCCTAAGTTGGATCCCTTTTTGGATTGAAGGGAGCTCTA
nta-miR319d	AGGAGCTCTCTCCAGTCCAGTCCGAGGCAGATCGAAAGGCTATAAACACAGCTGATGACT CGTTGATTCTTAAGCACGTCAACTAGGGTTGAGAAATTGACGTATTTTGGGATTCAACGAT GCATGAGCTATGTTTAGCTATCGCTGTCACGTCTTGGACTGAAGGGAGCTCCC
nta-miR390a	AAGCTCAGGAGGGATAGCGCCATGGATGAATATTAGTAACACTATTGCTTGTTTTAGCCAT CTGTAGCGCTATCCATCCTGAGTTT
nta-miR390b	AAGCTCAGGAGGGATAGCACCAGAGAGAGATTATCTTAATGAGCTTAATTAACACTATTT AGGCTTTCTTTTTTGACTATCGATCTGTAGCGCTATCCATCCTGAGTTT
nta-miR393a	TCCAAAGGGATCGCATTGATCCCATTTCATTTGTGTTTCTATTTCTTTCAGAATTAAAATAC TGAGTTTGGATCATGCTATCCCTTTGGA
nta-miR393b	TCCAAAGGGATCGCATTGATCCCATTTCACTTTGTATTTCTATTAATTCAAGAAAATTAATA CTGAATTTGGATCATGTTATCCCTTTGGA
nta-miR393c	TCCAAAGGGATCGCATTGATCCCATTTCATTTGTGTTTCTATATTTCTTTCCAGAAATAATT AAAATACTGAGTTTGGATCATGCTATCCCTTTGGA
nta-miR394a	TTGGCATTCTGTCCACCTCCATCCATATCTTTCAATGAGGAGGGGGCCAGAGTGCCAA
nta-miR394b	TTGGCATTCTGTCCACCTCCATTGCTAGCTCCATATCTTTCAATGAGGAGGGGGCCAGAGT GCCAA
nta-miR395a	GAGTTCTCCTGAATCACTTCATTGGGATTTCAATAATGCCTACTGAAGTGTTTGGGGGAAC TC
nta-miR395b	GAGTTCTCCTGAATCACTTCATTGGGATGTAAAAAAATGCCCTTAAATTGAAAGGGTATTA ATTATTAATGCCTACTGAAGTGTTTGGGGGAACTC
nta-miR395c	GAGTTCCCTTGACCGCTTCATGAGGGCTTACTTCTCCATAAATTGTTGTTTGAAGCTCTGTT GAAGTGTTTGGGGGAACTC
nta-miR395d	GAGGTCCCCAAACCGCTTCAAGAGGGCTTAATACCTTAATTTGTGGTTGTATGCCCTCCTG AAGTGTTTGGGAGAACTC
nta-miR395e	GAGTTCTCTCAAATCACTTTACTGAGTTATATATAATTCCCTTTATTAAGGGAAAATCTACT CACTGAAGTGTTTGGGGGAACTC
nta-miR395f	GGTCTTCCCCAATCCACCCAAGGTAAGATTGATAAATATTACTAATTATTAATGCCTACTG AAGTGTTTGGGGAACTCC
nta-miR396a	TCCCACAGCTTTCTTGGACTTTCAGAGCTTTAAATCTCATCTCTCAAATTAAAACTATCCAC AATTGCCTAAAGAGAGATCAAAATTCAGCTAGAAGAAGTTCAAGAAAGCTGTGA
nta-miR396b	TTCCACAGCTTTCTTGAACTTCTTTCCCTCTGAATTTTCATCTCTAGCAATTGAGTTATTGAT ATGAGATAAGATTAAGCTCTGAAAGTTCAAGAAAGCTGTGGGA
nta-miR396c	TTCCACAGCTTTCTTGAACTGCATATATTTGAAAACTCACACCATATATAACCAATTATTAT TATTGGAGGTGAAAATAAAATAAAAAATTGTTGCGGTTCAATAAGGCTGTGGGA
nta-miR396d	TTCCACAGCTTTCTTGATCATCAACTCTTCACTGCAGATAAATTGTTCTGTAACTTTTCTAG GCCGTATGATTTGCAGTGATCTTCTCCCCTTTGGGATGGCGATGTCAAGAATATGTTTTTCT TGATGATCAAGATTATTAACTGGGAA
nta-miR397a	TCATTGAGTGCAGCGTTGATGTCAATTTCTGTGAAGTCTTTAATTTGGCATTTGGTCTTTTT GCCACTTTTTCCAATTCGTTGTCATCTGCGCTGCACTCAATCA
nta-miR397b	ACTTGGTGCTGCATTAGTCCCTTTTGCTTTCATTTTTTGGAAATCAGTTCAATTGGAAGAAT TAAGTACTTTCAGTATATTTTGCAGTAATACTTCGTATGACCTATATGCATTTGCTGTTATT ACTCCGTCTTCTTTTGCGTCAATACACAAACTTCCCGTTAAAATATCAGGAAGACTACTGT ATTTTATTTTTCTGAGGGTTTATTAGTGCAGCGTTGAGG

nta-miR398	CAGGGGCAATCTGAGATCACATGTTGGTACTTTTCTTTTTTCTCTGTAAATGACAAGTGTGG TTTCAATATCTCAATATGTGTTCTCAGGTCGCCCCTG
nta-miR399a	CAGGGCAAAAACTCCATTGGCAGGCAGCTAACTAAAGGCATGCAAATTGCAATTATAGTT TTACGTGTTCTTCTAAGTGTGCCCTCTGCCAAAGGAGATTTGCCCGG
nta-miR399b	TAGGGCTACTTTCTATTGGCATGCACTTAATTATGTATATATAGTGACTCCATCACATATAT ATATTTCAGCTGACATGCCAAAGGAGAGTTGCCCTG
nta-miR399c	TAGGGTTACTCTTTATTGGCATGCAATAATTTAGTGACTCTGTTTGATCATAAATATCAGCT GACGTGCCAAAGGAGAGTTGCCCTG
nta-miR399d	TAGGGTAGCTCTCCGTTTGGCAGATATTTAGTTATCAACACGTATTTACGTACATTTGTTGT CTGCCAAAGGAGAGTTGCCCTA
nta-miR399e	CTGGGCTAATCTTCTTTGGCAAACAGTCACACTTCACTATTATGTGCAAGTTTGGTCATTGC CAAAGAAGATTTGCCCCA
nta-miR399f	CTGGGCTAATCTTCTTTGGCAAATAGCCATAATTCACTACCATTTGGTGCAAATTTGGTGA CATTGCCAAAGAAGATTTGCCCCG
nta-miR399g	TAGGGCTACTCTCTATTGGCATGCCGTTATATGTGTAGTGACTCCATCACATATATATATAT ATTTCAGCTGACATGCCAAAGGAGAGCTGCCCTG
nta-miR399h	TAGGGCTACTCTCTATTGGCATGCACTTAATTATGTATAGTGATTCCACTTAATCATCACAT ATATACTTCAGTTGACACGCCAAAGGAGAGCTGCCCTG
nta-miR399i	CGGGGCAAATCTCTTTGGCAATGAGAACCAAACTTGCACCAATGGTAGTGAAGTAGGGCA GTTTGCCAAAGAAGATTAGCCCAG
nta-miR399j	TAGGGCTACTCTCTATTGGCATGCAGTTAATTATATGTATTTCAGCTGACACTCCAAAGGA GAGCTGCCCTG
nta-miR399k	TAGGGCTACTCTCTATTGGCATGCACTTATATATGTATAGTGACTCAACTTGATATCATAAC ATATTCAGCTGACACGCCAAAGGAGAGCTGCCCTG
nta-miR399l	TAGGGCTACTCTCTATTGGCATGCAGTTATATATGTAGTGACTCCACTTCATAATATAATTT AGCTGACACGCCAAAGGAGAGCTGCCCTG
nta-miR399m	TAGGGCTACTCTGTATTGGCATGCAGTTATTTATTGTGACTCTACTTAATCATCATATATAT ATTTGAGCTGACATGCCAAAGGAGAGCTGCCCTG
nta-miR399n	TAGGGCTACTCTCTATTGGCATGCCGTTATATGTATTGTGACTCCCTCACATATAAATTTCA GCTGACACGCCAAAGGAGAGCTGCCCTG
nta-miR399o	TAGGGCTACTCTGTATTGGCATGCAGTTATTTATTGTGACTCTACTTAATCATCATATATAT ATTTGAGCTGACATGCCAAAGGAGAGCTGCCCTG
nta-miR399p	CTGGGCTAATCTTCTTTGGCAAATAGCCCTACTTCACTATCATTTGGTGCAAGTTTGGTGTC ATTGCCAAAGAAGATTTGCCCCG
nta-miR399q	TAGGGCTACCCTCTATTGGCATGCAGTTAATATATGTGTTGTGACTCCATCATATATATATA TTTCAGCTGACATGCCAAAGGAGAGCTGCCCTG
nta-miR400	GGGACTCAATAAGTTTATATATTTTCGTCGAAAAGGATTTAATTAAACCCACTTTACCATA CATAGCTCTGGTACTGACAAACATTAGTTGAAACCCAAGTAAAATTGACAATTCTTAAAAA AAATTGAAATAAAATTTACATATTTTAGAAACTATATAAAAGTATTATAAGTCAC
nta-miR413	CTAGTTCACTTGTTCTTCACATCAACAACTGCTACTACAGTGTGGAGGTATTTTATTCAAAG AGCAGGAATAGCAATGGAGGGATTGTCTTTGCGCCAAGCCATTACCAAATGCTGGACTGC TAAGGTGTTGCCTAGAATTAAGCCAATAATGCAGGCTTTACCTTTGTGTATTGTGTGGTAA TTATGGAAGAGAAGGAATATCCTTAAATATGGAGAGGCAGTCTCAGTAAGCAGAGTTATC TACCAAGTTTCAACAACATTGCAATCTTTGGTCAAGGTGAGAAAACTAGGTCTTCATGTAC CCCACAGATGGCATGACTTGTTGTTCATGCTGGAAAACTATACTCCACGACTGAAGTTTGA GAAAGTCACCTGGGAATTTCCATTGGAAGGGTGGATCAAAGTTAACACAGATGGAACATC TAGGGGAAATTCGGGGAGAAGTGCTATTGGTTTTTGTGTGAGGGATGAAGCTGG
nta-miR414a	TCCTCTTCATCATCATCGTCCGAATAGCATCAGCAACATACAACAGAAACCCTAACTCTGA AACTGAAACTGAAGAAGAAGAGCGTTTCATGGAAGCCAGGTACAGTGGATAACGAGTTCC TCAACAAGAAAAGCTCTAAGATTTGCTGTATATTTCACAAGGAAAAGCCTTTTGATGAGGA TGACAGTGATGACGATGA
nta-miR414b	TTAGGATAATGAACGGAGGTGATAGTCATGACATGAATGAATCATCATCTTCATCATCATC ATCA

nta-miR414c	GGATCGTGTGAACGAGACGAATCCTAAGCGGAATACCACTCTTTGCTTCATTGTGCCCGAG AAAGATATCATAAAAGGGAGTACTATTCCTTCTAAGATTGAGCTGTTTTGCAAAAGGACTT TCTCAGATGATAAATACGTAGGGGAGCTGCTCGTTTCTCTAGCTCCTCGTTATAAAGGAGA ATGTACTTTTCCAGTGCAAAGAAATGATTCAAATGAGAACAAAAGTTTTGGAACTTTGACG TTTTCGCATGCTTTAGGAGATAAATTCATAGTTACGAACTCCTCGTAGTCGTCGTCGTCATC ATCATCGTCA
nta-miR437a	GAGTCAAACGATTTGTTCTTTAAAAATGACTTTCTCATATACCTTCTTTAAATATTTTGAAT TGTTAGTTATTGTGACACATAGTACCCTTTATGTAGTTTCCAATTATATAAATTTTATTTAA AAAAAAAAAACTTAAAGATTCTATGTCCGAATTTATGGTCAAAGTTAAGAAGTTTGACTT
nta-miR437b	GAGTCACATGGATTATTCTTTGACTACGATTGTTTTAGGTACCTTTTAAATATTTTGAATTG TTCATTATTGTGATTTATAATATTTTTATATAGTTTTTAAATATATAAATTTTATTTAAAAAA CTTAAAGATTTTATATCCTAATTCATGGTCAAAGTTAAGAAGTTTGACTC
nta-miR476a	TAGAAATCCTTCTTTGCAAAAATGCTCCCTTTCTTCTTCATAAGTATTTCCTTCAAAAATAA ATAAAATAAAAAAAATTTAGTTCATTGCTTATTCATGATCTACCGAACTACGCGAGTTTGA TTCTCACCGGATGTGAGATACGTAGGCAACCCTCATCGGGTCCAACCCCACCTTTTGCTAA AATAGCCAAAAACAAATAAATAATAATAAAAAGAAAAAAACATGTCAAAATTTTAATTCT GTCATAAAAAAGTCGGGTGACGCTATTTTGTCAAGGCATAGCCGAATATTCCCGAAAGGG ACGCCGGAAGGCCGACTTTGAATAAACAACCACTTTTTGGGTCATGTTTAAGATTTGGTCC AGTTGACCCCCACAGCCTTAAAAATCTTCGTCCCCGAGGCGTTGAAAGGTCGTGTTTGCAA TATTGAGTTTTCTA
nta-miR476b	CCTTTCAAGAAGTCCTGTAATAATTCATTTTCATTCGAATGCTTCTTTTTTTAATAAGTTAT AGCTAATAGTACTTCTTCTTTGAAAAA
nta-miR479	AGTGATATTGGTCCGGCTCATCATCTCTAGACTGTCTCATGAATTCAACAGAAAGTGATGA GCCAAACCAATATCACG
nta-miR482	GGAAGAACAGGAGGTAGCTCGAGTAATTACAACTTTGTACTGGCAGTCTGGCACACATTA AAAGACGCTCTAGCATACAGTCCCCCGACAAAGCAAGAAGAGGAATGGCAGCCACAAAA GAATTAAAGACAGTAGGCTGATTATTAAAATAAAAAATTCCACAGCTATAGATAATTCAG ATAACCACTAACCAGATTTCGCAATTTCATCACTTCATCATTCAACTTAGAGATCTCTTCCT GCAAAAAAAAAAAAACATAAAGAATATAGTTGTTTTGTAATTAAAGAACGCAACTTAAAAA CTTTAATCCAAAATTAAACATGAATTCCACAAAGAAAATAATCAGAGAAAAAGGACGGTG AAAAAAATTGGAAAATGAAAGAGGAAGAAGAACCTGATGATCACGGACTAAAGTTTGTC GATCCCAAAAGAAGAAGGTGGCGAAAACAAAAACGATGAAGAGCAAAGCCAGACGCTTT GGCCGTTTCAGATAGTAAGCCGGATTAATGTACCTGAAAGTAGTGGGAAGACGATCTACT CCTCGCCATTCC
nta-miR529	CCGTACCCTCCTCTTCTTCTCCCTTACCCTTACAAAACCTTAGACCTAACCCCTAAAAATGA TAGAGAGACGGACGCCCTCCTCTGAAGAGAACTCAACCGCAGCCTAAAAAAAGAAAAAC GGAAGTAAGAACAAAAGAGGGAGGAGCGAACGGACATATGG
nta-miR530	TGCATTTGCACCTGCACCTCTACTATTCTCTCTTTCTTTCTTTCTTTCTTGTTTCTCTATGGTT CTTTATTATCTTAGGAAGGAAATGAAATGAAAGGAAAGGAGGTGCAGGTTCAATTGCA
nta-miR776	TCTAAGTCTTCAAATTGATGTTTTCAAGTCTGTTAATTTTTCAAGTTGAAGTTTAAA
nta-miR809	ATTTCAACATATTTCACATAAGATAGTAAATTTATGTGAGACATGTTAGAAT
nta-miR810	TATCGGTATGGACATGATACTATATTAATTTATTATAGTGAAGTGATTTAATTAAGCGATA
nta-miR815	ATTCGTGTTTTTAGCGTTGTTTGATGTGATTTGAGGGCTCGACTAAGTTCGTATGGTATTTT AGGACTTGTTGGTATGTTTGGTTGAGGTCCCGGGGGCCTCGGGTGAATTTCAGATGCTTAA CGGATCGAAATTTAGACTTGTGTAATTGCTGAAGCTTTCAAGCTTCTGTTGTAATCGCACCT GCGGTAGGGTGGCCGTAGGTGCGGACACGCACATGCTAGAGGCCAAGTGCAGAAGCGGG ATTGAGGAGATGGAC
nta-miR816a	CTGTTGTATTATGTCTCAGATGAGCTGGCTAGTTTCTTGACTGACAACCTTATGCCCTTGAT TTTGACATATTTTACTACTAG
nta-miR816b	TTGACATATTTTTCTACAAAATCTGCTCTTAAATTCAAAAGGCCAGACTTTTTGGTTCAAAT CTTGTCAA
nta-miR817	GGACTGGGCTTTTAAGTTAGGTAGTGGATTCAAATTAAAAGGTTCGTCCTTTCGTTTGGAT CCGGACTTGAACTTGAGGCCCGATTTT

nta-miR819a	TGATGTTATAAGGCTTTCTAGCATCGAGGATTGACATTCCACAACCATGGCTCCAAAAAAC AACTCCTGATATTTGAAAAGAAATTGTGACCACATGTAAGATATAATCAACTAAGGCTATA ATAGAGGATCTAAATGGTGACTATTTTGCTCTATCGGTTGATGAATCTTTAGACATGTCATT GCAAAGAGCAAGCTGTTATTGTTTTGAAATATGATGATAAAAAAGATTTGTGATGGAAAT GTTTGTTCTTGTTCATGTTCAAGATATTAGTGATTTATTGATAAATAAACCTATTGTTGATA TATTTTTTAATTGTTTTTAAGTTTATCTTATGTATGTGGACAATCTTATGATGAGACAAGCA AAGTCATATTA
nta-miR819b	TCAAGAAAAGTCTTACCACATCAATCAAAATATTTGGACGAAGCCAATATTCATGTTTAGC TTGACCATATTGCATGTCAAATGCAACATGAATCGACTACTGTTGCATTAGATTTTCACAA TTTCTTTATAAACGTTGTATAGCAGTCATAATCTAGCAAAAACAAACTGTGAAGATCTATT ATAACAACAACAATCAATTCATGTTGCATATGAGAGGCAATCTGATTAAGCATGAATATA GGCTTCACTTTAATGCTTCAATTGATGTTATAAGGCTTTCTAGC
nta-miR819c	ATTATCAAGCACTTCGCGAGCTACTGGCTTGTTAGAAGACTTTCTAGC
nta-miR826a	TAGCTCTCTTTTTGGATACGTGTTTGTTTGTTTGTTTAATACAACACTAAAAAGGGAAATGA TTACAAGATTCTCTATGAATTCGTATCCCTTTTGTAAAATGTAATCTCCTCATTCTTTTGTGT ATACTTCTATAAATGGGCCCTTAAAACTTTCAGGTCATATATAAACTTAACACGCCTTTAA AAATGAGGAGA
nta-miR826b	CATATGCAATTCAAAGATATTGTTGTGAAAGCTGCTAATGTGGTCTTTGTGAATATATATA GAGTGTTTGAGTTTTTTTTTTGTTCTATTTATTTATCATTTTTCGTTCTTTCCTTCTTTTTCTG TTCGGAATTTGCGCAAATTGATTTATGTCATTTTCTGTTTGTCAGTAGTTTAGGAAGAACTT GGTATTTGGTTGAGGCACTGATAGTAGTAATAGTTTTGTTTTGGATACGTT
nta-miR827a	AGTTGTGACAGTTTAGCTATCGTTTTGTATTGATTTCAAAAATATTAGTTACCCAAAAATA AATTGAACGATGGATTAAGATATTCACCTTCAATAATCTCTACTTCTTTCACTAGTTTTTTC ACAGCCGATTGAAACGAATTTCTTATTATAATGATCATTACTAGGCAAACCTAACATATTT GATGCAAACCTAATATATTTGATCCAAGACAAAGTGGCTAACCATCAACAAACT
nta-miR827b	TGTTTGTTGATGTTCATCTAACCATAGGATAATGGGTGCAAAAACTATGCAGAATGATGAA TAGATGACCATCAACAAACA
nta-miR827c	TGTTTGTTGATGGTCATCTATTCATCATTCTGCATAGTTTTTGCACCCATTATCCTATGGTTA GATGAACATCAACAAACA
nta-miR827d	ATTCTTTGTGTGGTCGTGGAGTAGACAAAAATTGTCGAATCACGTAAATCTTGTCTTGTCTT GTGAAATTTATTGCTTCTCTCCTAAATATCATTTTCATTCTATTATCTTGACACGCTTCCTCA ACACAAGTTTATATTTTAGTTATGCTTTGGTAAGATAACTTTATGTGACTCAAAAAATAGT GACTTTTATGTTATAAAACTTTATTGGATGACCATCAGCGAAAT
nta-miR827e	CAAGATGACCATCAGCAAGATTAATATCGAATCCTAATAATTAATCCAAATGGTGGTCAA CTA
nta-miR828	TCTTGCTCAAATGAGTATTCCACAATGTCACTCTCCAGAAATGTTTGAAAGTCTCTTTTCAG TCTCTAGTTGACATAGCTTGAGATACTCATTTGGGCATGA
nta-miR829a	GCAATTAAAGCTTCAAGGAAATGAGGAACTTATTTTTGTTTTCTTTAAGAGCTAACTGC
nta-miR829b	AAATTTAAAGCTTCAAGGTGAGAATCATTCTTCCATATCTGTTCTGATTAACCTGAAATCC AAAATCGATGATATACATAAGTCATAATACATCATACGGGGCTAGTCATGCCCGAGAACT AGCGAGCGGAGTGCAAATGCTCAAAACGACTAGTCGGGTCATTACACTAACCTTGGTAGA ATAGTGACGAGGTCGAGTCATAGCCCTACTAGAAATAATAAGAAACAAGATGAAAGATTT AACAATATATTAATAATGACAATGGGAATGAACCAAGTAAGGGTATGCCACAATTTAGCT GCACAAAATAAAGGCAAATAATACATCACGAAAGAAAATAAGGATAGTATGTCACGGAA CAAAACAACCAAAAACACAAGTGAAGAAATAAAAGGTCTCAACAAGAGTCACTATCGAG GTACCGCCTTTTAGCCTCAAATCATAAATAAATTCACAGTCTTTCTTTATATCACCGCGGGA GCTTTACATTT
nta-miR830a	TATCTATTTTGAGAAGAAGAAAAGAGTTTCAGGCAGACGTTTCCATCAGTTTCATCCACGG ATGCTCGATTTCTACTATATTTTCCTCTGTAGATTTTTGGAGTCGGAATAGGTACAGTATAC ACTAGAGTCCCACTTCGGCATGTTAGAACAAGGAACATTCAGCTTGAAGTTTGTTCTCCTA GTCAAAAAGAAC
nta-miR830b	CAACTATTTTGAGAAGTACTTTTTTACAAAAGTACTCTTCAAAAGTA

nta-miR832-5p	TTTTGATCTCGTGAAGTACGTATGAGAGGCCAAATTACCACAAAACCAAACCACTTTTCTA GGAACCTGAACTTATCAGCAAACATGACGGTTAGTTTAAAACAAATAACAGACAGGTAAT CTCACATTGGGGTGTGATGCACCTATACAGTCAGGGGAGTTTCTACATGCTTGCAACAAAG ACATGTGTCATTCTGGCTTCTCCTTAGGCTTTTTCTTTTATTTATCACCATTTTCTTTCTTTGT TTTTATTTTATTTTATTATTTGCAGTTAAAAATGTGACCGGATCCGATGAGGATTGCCTACG TATCACGATGCCGCGTGAATCAGATCTTGCGTAGTTCGGGAATCGAAA
nta-miR833-5p	TTCTTGTTGTACTCGGTATTGTTGAGCCGTGGGCTATGTGTGGTTATGATATTGGAATTTTT GTTATGGAAATGTTGTGGTATATAGGCACATGTGGTATAATTATCTATGCGTGTACGGAGA GACAAGGC
nta-miR835a-3p	GATTCCTGATCGTATTTGTATGGGATGCATGCCGTGGGGCGCGACGGATAAGGACAATACT GTCTATCCAGTGCGAGAGCTGCCTGCAGGCGATACGACAAACAGTACCGGCGGGAGGTTG GGCCCATCGTCCTTCTCGTCACCTTCTCGCAAGCTGACAGGTTTGGCGTCAACTTCTACCGA TAACAAAAACAACAACAATAATGAGGCAGGTGGTGGTGAAGGTCGCGCTATTAGTGTGTC AGACGGCGGCAAGTCTTCAAAGGAAGCGCTTGAAGAGAAGGCGAAAAACAAATTCAAGC ACACCAAATGGAAGGCGCCAGTAGTCAAGGCAGTTTCGACATTACTGTTGAATAAAATTG ATTATTTCATCTCACTGATGGATGAGAAGGAAGAGCAGGCGATCGAAGTCGCGTGCGCGG TGAAGTCCGTAGAGAGTCTTATCGCCACCTCCTTGCAGACGGTGCAGAGTAGTGTTGCGGC CATTATCTCACAAAGTACGGAGACGATACGCAAGAATC
nta-miR835b-3p	GCTTTTGTATGATTTCTTTGAGAAACTAAGGCATTTATTGAGGTTTATTGGCATATTTCAGG TATCAAGTGATTTAGAGAAGATACGGAAGAAAC
nta-miR837a-5p	GTGTGAAGCAACATAGAAACGAAGGTGGTCTTCATATTCATGAAATCAACAGGCCTATTG TTCCCTGGCATCTGAAATATTAGTCCTGGACATAGAAACGCAAATAACGCCGTCGCTATGA TCGGAGCTGCCCAATCCTTCATCTTTAGTTTCTTGTTCGTTTCC
nta-miR837b-5p	GGCAAGAAGAAAAAGAAAACTGAAAAGAGTGCAGCAGAAATTGCATTGCTAGTAGAGAA TGTTATGGCTGAGCTTGAAGTTGTTGCTGAAGAGGATGCGGAACTGAATAGACAGTCTAA ACCTGCTATAAATAAACTAAAAAAGCTTCCTCTTCTCATAGACGTCCTATCAAAGTGAGTG AGTTTGTCGTTTTATTTTTTGTGGTAGTAACTAAATTTGTCTTTTTGGAGGCTAATATTTCTA GACTCACACTTTTAAATTTTCAATTTGCAGGAAGCAACTTCAGCAAGAATTTTTAGATCAT GGAGTATTGACTCTTTTGAGGACTTGGCTTGAGCCCCTTCCAGATGGTAGCCTGCCCAATA TAAACATTCGTGCTGCAATCCTAAAAATCCTTACTGATGTATGAACGTGGTTTCTAGATTC AAGTTAGTTTCTTGTTCGTTTGCC
nta-miR838a	TGTTCTTCTACTTCTTGCGTGAAGCATTCGAAGGCTGGACTATTTGCTTGCTTGCACTGAAT GCATTACTTTCTCTTGCCTTTGGAATTGTCGAAGAAGGGATACAGAAGTGTTGGTTCGACA ATTTGATTCTGCTTTTATCCATACTTGGACTGGTAAAGTTTAAGTTGGTGTGGATGTATTGG GAAACAAAAAAATTACTGAAGAAGAAGGAGGAGCG
nta-miR838b	AGATTGGAGGAGAATGAAGAAGAAAGTCTTCCAACAAAGACAGCTCTTACGAAGTTTTTA GAGGGATAATTTATAAAATGCAACACACGTATCTATAACATGTATAATAAGCTTGTACAAT TTCTCCTCCTTTTTCTTCTACTTCTCTAATA
nta-miR838c	GTTGTCTGAAAGTAGAAGAAGAGTACTAGGAAGGAAAGCGGTGCTGTCGTGGCACGATTA TCTCTAAGAGGCGTGCGGGTTTGAGGAAACTGGGTTTGATATTATTATATTATATATTATA TAGAAGAGGCAGTGAGGTCGACCATGGCATAAATGTAATTTCGCTCGCCATAAAGTCTCT ATCCATTTCACATCTTCGACTTTCAGATTCTTTTTATGTCTCTTGATCTTTTCTTCTATTTCTT GCAAT
nta-miR838d	AGTTAAGATAAAAGAGAATGAGAAAAGTAACTATGAGTTCAAACAAAAAGAACTGTAAT CCAGATGCATAGGAACATTTTTATTTTCTTCTACTTCTATAACC
nta-miR844a	CATTAAGGTTGCTTATAAGCTAAGGGAAGCTTTGGTTTGAGTTTTTAATG
nta-miR844b	TTTAAAGATTGATTATAAGCTAGTTTCCACGAATTCCTTATCGTAGTTTTTGAC
nta-miR844*a	AACTAGAAAGGAGAAAAGAAATTTAAGTTGGTGTTAAAGACACAAACCAATTGGTGTCCA ACTGTCCTCCATGTCAGGATGCAAACACAAGCTAAAATTTTTATCCATCTTACTAGTT

nta-miR844*b	ATATAAGCCATCTCACTACTACTTGTGTAAAAAAAATCTTTTATAGTGCATCTTAGTTTCTC ATGTGTTCGTTTGTTTATTTGTAAGCAAAAAAAGTTTCTTTTTGTTGGTCTGAATATGAATA TGCATTTATCAAGAATTGTGGTTATTACTTATGGCAATGATAGATCGCAAAGAATATTAAA ATTAACTTTCTAAATAATAGTGTACATGCTTTTAACAATAAAACAATTTTATATGGCAACG CGAAGAAGCATTAGACGGCAAATTTGTAGAAACGATAACCTTTTAACATAAGTTGAACAC TATGCATACCTGACATGCCAAACCTAGAATGCTTTTTTCCGAAGAAATTCATTCATTATGA GATTCTCTTTGAAATACTTTATATAGTTATACACAATTAAAACAATGAGAAGATTTATAT
nta-miR845a	TGGCTCTGATACCAATTGATGTTTTAACTTTAATACCACACAAGAGGGGGGGGGTGATTTG TGTGGTGTCCAATTTTTTGCTTAAATTGATAATGGAAAGACCTGGTTCTTCTAAGTGTTCCT TAACCTACTATTGCCGAAATAGTAAATACGGAAAGTAAAAAACATAAGTATTTTACGTGA AAAATACCTGGCTCAAAAGGTGAAAAAACCACGACCTACTTCTTAGTAGGATTTTCCCAA ACTCTCCACTAGATCATTGAGCCAAAAACTGTATTTACAGAAACACTTTTGTAAACCTAGG ATTAACTCTAACCCCATTGTGGTAACCAGCCTCTAACTGTTGCGACAACTTCAAGTTAACT CTAACTTGAATACTCTGAGTACCTATTACAATTGCCTCTAGATAAAGCTAAAAGGTACAAT ATGAAAACACCTACTACAAATGAACTAGAATAAAAGACAGACACTTGGAACTGGTTCTTC TATCTGGTTCATGTAGCTTCAGGTTTGCACACTTGAATCACACATGAATTGCTTACAAAAT GCCTTGCTATTTTTGCTCTCAATTCACGTTTAACTTCTGCTTTTGTGTTTCCCTGTAGAATGA GAACATCCCGTAATATATAGAGTTA
nta-miR845b	GTGCTCTGATACCAATTGATGGATTATAATAAGCAGCGGAAGCAATCCCAGTATCAAAAT CACATGTAATCTAGACAACTAGTATATACGGGGTTTCACGGGTTACCTCTTGAAACGTGAA CATATCTCTTCTATCACGTGAGCCTTCAGTTCCATAACTTCTCAATGGAATCTCCACTCCCC ATATTAAGATTCAGATACTAATCAAGTAAATTAAATTACTGACGATTTAATTTATTGATTA TTTTCTTTAGACTTTCGCTTAACTTATTTCATGTGTCGGATACAAAATCCACCGGCCGGGTT TACACATGAAAACTTATAAGCTTTCATAAAGGTGTATCATCAATCTCTAAACCGAGACATG GATTCCATCAACTAGCTATTACTTCACCAATGTACATTATTATTATCCAATTTACCAGGCAT
nta-miR845c	TACCAGTTGGAGTAGAGCTGCTTGAAATTCTCGAGGTCGATTCAGACGACTCTGGCCCATC GAACAGTCACCGTCAAATGGAGATTCAATTCCCACTTGCTCTGTTTCCTTGCCCTTCTGTTT AATGAACCCAATGGCCTCCTTGATCTCAAACAACTCCTTATCAGTTTGTGTATTTTGATTAG TCAAGCTAGTCATACCCTCCATTAATTTCTGCAACTGTTCTTGAATTTCTGAAATTTTACCA TCCATTCCGTCAATAGGGTTGTCTTGATTTTTATCCCGACCCATGTCGATGCCAAGAATCGA GCAGCTCTGATACCAATTGATA
nta-miR846a	TAGAATTGAAGTGCTTGTAGTCTCTGGAAAACTCTTATAGTCATTCTCTTATTTGTTTCCTA TATAATTTGAGTGCTGTTAAGATATTGCATATTTTGAAACAGCACTCTGGAAAAAATGTAA TTATATGAGCATAAATACCTATTTGATGCTAATTTTCTGTTCTAACAAGTCATTATTTCCTC CTATGGATGCAGCATTCGTCGCTTTTG
nta-miR846b	TTGAATTGAAGTGATTGAAGCAGATTCATTCCATCTGAGCACACCATTCAATCTTGGGAGT GACTTTTCTCTTCTGCAATTAATCTTGTCGATTT
nta-miR847	TACAAAAAGAGAGAGAGAGAGAGAGATGGTGCACCTTTTTTCCTGTCCTCATCTCCTCTTC TTCTTGTTC
nta-miR854a	CATTTTCTCCAAGGCTTCAACTAATGGCACATTTATAGAGAGACTCTTCATCATATCAATG AACTTCTTGAATTAATTCTCGCCATTTTTCTTGGCAAGCCTTTGAGGATAGGGAGGAGGTG
nta-miR854b	TTTGAGGATAGGGAGGAGGTGGCTTAGGCAATGGTGCCTTAGTTTTTTGAACTACCAGTTC CGGTATATCAATCACATAATGCATATACAGGTTCACATCCTCTTGAGTCTCTTCCACCGTGT CATCAAT
nta-miR854c	GTTGAGGATAGGGCGGAGGAGGCCGAGGTAAGGGTGCTTTGGCTTTGGGCACCACCAGCT CGGGCATATCAATCACGTGTTCCCTAGATGGTTTCACGGCCTCTTGTGTCTCCTCTTCAAC
nta-miR854d	GTTGAGGATATGGAGGAGGAGGCCTTGGCATTGGTGCCTTAGCCTTTGGCACTACCGGTTC CAGTATGTCAACTATGTACTCCCGAGAGAGGTTCACTTCTTCTTGAGTCTCCTACACATTGA CTTCAAT
nta-miR854e	TTTGAGGATATGGAGGAGGAGTCTTAGGCAATTGTGCCTTATCCTTTTGCACTACCAGTTC TGGTATGTTAATAATGTGCTCCCTAGACGGGTTCACTTCCTCTTGAGTCTCGTCCACATTGT CATCAAA

nta-miR854f	GAGAAGAATAGGGAGGAGGAGAAAAAAGAGGAGAAATTCTGGCAGGTACCTATAGTGGC TACTTCTATAAACAGGTACAAAAGTTAAACACTAAGCCTTAAAAAAGGCCATTCGGTGAA AATACTTGCATGTTCCCTAGAACTAGAAGATGGGACTATGGTTTTCCATCACATCCACAAA TTGGTAGTTGTATAATTGGATTAGTAGGTACCTTCTTTGTAAACTTGAAATTACAAGCAGT CAAGCAAAACTTAGAAATTTAAGTTTTAGATTTTCTACTAACACTGTGAAAATATAATATT TTACACGTTTTACTGTAGGGTTGTCTCTGTCACATCCACAAAAAAGTTAAATAATTTCTTGT AATATTAATCATGTTAGTTTTCGTTAGTGTACGTGGTGATTTCAGATTTGGGATGCATTGTG ACGACAGAATGGCAGAACCTTTCACTATAAAACAAACTAAATACCTAACATTTTTTTTTCT TTCAAATTCTTCTT
nta-miR854g	GAGAGGATAAGGAGGAGGAGGAGAAATTGAATAGAGGAGAGAAACGTCTATTTGAAAGC TTATGGGTAGTGGTTGCTATATTTTTTGGGAATATAATATTGTATAGCCGCTCTTAAAAATA TAGCCGAAATATATATATTTTTTGTATACTTTATGTATATTTATATATACATTCTTTATGATA TATATAAATTTTATATACTTTTTCGGCTATCGAATATAAATAGTTTCTGGCGCGGGTTACAA GTAAAAAAAGTTCATATTTTTTCCGGCTATTTGGCAAAATGTATTTCTTGGCTAAGCAGAT GCAATTAGGTGTAAGAAAATGTACATTTATAGATTTGTACCAATTTTCCTTTGTTTTATAGA ATAAACATACATTATTTTCCGCTAGTTTTGTTATCTCTATTGTACTTTCTAGATTGTGCCAA CGCCTTACTATTTTTAAAACTATCCCTTTCTATTAGTTTGTATATAAATATATTTTTTTCCTG TTCTTGCAGTTCTTTT
nta-miR859a	TCTCTCTGTTGTGATGCCAAGTGTTACGCTCTAAAGCAGATTTTCCATTGGCAAAATACAA GAAATGTTGGGCGGACAACAAAGTAGA
nta-miR859b	TCGTCTCTGTTGTGAAGAAAAGGGAAGGGATCGGATTCGCCTTCTTCAAGGAGATAGAGC TAAAGGAGTCGGCGGCTCAGTGAAGAAGAAAGACTCTTTTTCTTTTTTCCAATGGGGAGGG
nta-miR859c	ATTGGATATTGACGAGAGATGTAATTTCCTAGTTATAAAACGAGAATACCACATTTTTTGG CACATGAATTCCAACTATAGGCAACTCTTTTTTAGTTTTAGATCTCTGAATTCTCGGTTAAA GCTGTTTCTTTCGCTTCCCAAATGTTTGCTTGGTTGCTGTTTCCTTTCATGGTTTGTTTTTAG TTCCTTTTTCATTAATAATGTTGTTCCGAGACATGAAACTAAACAAGGTTCATATCTAAGTA TATTGATCGCTAATTTAATTGCTAATTATTCACAGTTTAGAAAGCACTCCGTCTTAGTTTTC CGGCCCTGCCTTCACTTGGGCAGTTTAAATTACTGATAGTTCAAACAGGATAGTAGTATTA CATGTTGTCTCAAGAAAAAAGTCTCTCTGTTGTGATGCCAGT
nta-miR863-5p	TTATGTCTTGTATCTCAATTCCGTGGACAGTAGTATCCTATCTCTTCATTTCTACTCAGAAT GATTGTTTCTGGTGTAACTTCAGGTATAAGAGAGAGAGAAGTTGATTTTCATAAGGCTATT TCAGATAAGCCACTTGTATGCAGTTTGGCTAGTGAAAAGCTGAAAAGGATTAGAGATGAT GAGATGGTTGTAACTGCCCTTTTCAAGAGAATAAGATAGTT
nta-miR864-5p	TCAGGTTTGATTGACCTCAAATTTTGCACACAAGTCATAAATGACACAACAGACCAATGCC AACTCCCAAAACCAAAATTTGAGCCCCATAATATAACGACCCAGCCGGTCATTTTAAGTGT ATTAGTCCCCATCCCATATTTACTGTTTTCCCCGTATTTGTTTCTACTTATTTGACTTACCGG GAGGTCTTGTTTTTGGTGTCAGAGTGATTTGGGATACTTAGTCCCTAAAATATAAGCTAAG TCTTAGGATTTTAACTATAGTCAGAAGTGTATAAAGATGACTCTGGAATGGAGTTCCATCG GTTCCGTTAGCTACGTTGGGTGATTTTGGACTTAAGAGCATATCCAGAATGTGAATTTGAG GTTTGTAGCTAATTTAGGCTTAAAATGGCGAAAGTCGAATTTTTGGGAAGTTTGACCGGGG GGGGGGGGTAACTTTTTGATATCGGGGTCGGATTCCTATTATGGAAGTTGGACTAGGTTCG TAATGTTGAATATGACTTGTGTGCAAATTCTGAGGTCAATCGGACGTGG
nta-miR865-5p	TTTAGCTCATCTACTAAAAATGATTGTACTATCCTGTCACGCATTGACATTGCAAAAAGAT AGCAGATAACACAACAAATGTTTTGGTGGATTATCATTTGTTATCTTACATTACATTATAAT CAGCTGTACCAAGATATTCTAGATAGTGACCAGTGCAGTTGCAACTAAATTAGCCTCCTCT CCAACCATTACCATGATTAATGATTTATGGTTCTCCAGTAACTACATGCAACTTGTAATAT AACATTGAGGAGTGTGGTAGGATAGAGGGGATCCTTTAACCTTTAACTAGAGATTTCGGTC TCGAGCCCTAGGAATGAAGAAATTCTTGGCAGAGAGATTTTCTCTTTAGTGACCCCTAAGC AACATGAATTTGGATTAGTTGAG
nta-miR870	GATCAAGAAGATGAAGGGATTAGCAGCGATCGTAAACATGGATATAATAGAGAGGTTTTT AATTTGGTTTTTCTTGATC
nta-miR1426	AGAATCTTGATGAATTAAAATTAGATTTTATAATTTTATTCTCTAGAAAAA

nta-miR1435a	AAGAAGTAGAAGATGAAGAAGCAATTAATAATAGGAAAAATATTGGAGTTAGAAGCTAA GAAATATTTTTACCCAAGCACTAGAGTTCTAAAAGAACAAACAAACCAACAAGCTTGTTCT TATTTCCTCCCCTCTTATATATGAACTCTCATACAGAAGCACTAAGCTTAACCAAACAAAT GCAGAGTTTGTCAAACGTGCCTAACACAGAGGGGAGAGAGAAAGGGTAGCAGAGAAGTT AAGAGCTCTTGAGTTTTGCACTCTAAACTCTGTGGCTTTTGTAATACGAGTTCAAAGCAAG CAAATAACATTGTTATTTTTTATTTTTTATGTGGCACCATTTTGCTCTTTCATAGTCAAACTT TTT
nta-miR1435b	TTTCTTAATCAAACTTTTGTGTGTGATCATTTGTTTGAAAAGAAT
nta-miR1435c	CTTCTTAAGTCAAAGTTTTTTTTTGTTGAAATTAATATTTAATTAAG
nta-miR1435d	TGAAAGTTGAAAAAAGAAATTGCAGAAGTTATTCCTTAAAAAGAAGTCACATGAACTTTC TCATAAGAGTAGATTTTTAAGTTGATGAAATTCAATCTTTTGCAGATGACTTTCCCTTTTCT TAATCAACTTTTT
nta-miR1435e	AAGAAGGACACCGGAAATATCAAAAGAAAAATAATTTCTCTTTAAATGAATTGATGTGAA TCTAAATTTATTCAGGTATATTTATAAATAAAAATAAAGGGTAATAGTGAATTTTCATCTT ACGACTATGAAATTGCACTTTCTTTTATGACTTACAAAAAATGAAGAAAAATAAAATTATT ACCTGAATTACATGGTTTCTGAAGTCAACTTTTT
nta-miR1436a	ACTCTTTCCCTCCCATATTAGTTGTCAAAATTACTAAAATTACTTGGCCCAAAATATTTGTT AATTTAGAAAGATAAGATAGGATTAATTAATTTTTTTCTCACTTTGACCTTAATATTAATTA TTTTTGAAAGTAAATAATATTATAGAATGCACATTCATTGAGAGAAATATGATAATTAGTG AAAGCAATTTTATTTATGATTTCTAAGTAGTATAAAGGAATATGATACTATATGGGACGGA GGGAGT
nta-miR1436b	ACTTCATCCTTCTCATAATATGTGTCATGGTTGCTAAAAATAATTGTCTCAAATTATTTATC ATTTTAGAAATTCAAGACAAAATTAATTATATTTTTCCTTTTTAACCCTTAATAATCAATTG CATTCACGTTTAGTAAAAATATGGATAACTTCATTAATAGAGATGACACATAAATTGAGAA AACATTTAATAAGGGTAGAGTAGTCAAAACTATTCTTTCTTAAGGGCGTGCACAAGAGAA TCATGACATATATTATGAGACGGAGGGAGT
nta-miR1439a	GATATTCCCTCTGTTTCAAAAAGATTCACTCTTTCCTTCTTAGTCTGTTCCAAAAGGATTGG CACCTTTCAATATTTCCTTAATAGGAAGCTTTTATAGCCACACAAATGCTATGGTAGGTTTC ACACCACAAGTTTCAAAAGTTTTACATCCACACAAATGTTATGGCTTATTTAAGACCACAT ATCTCAAAAAACTCTTCTTTATTATAGTTTGTGCCAAATCAAACTAAGACAAACTTTTTGA AACGGAGGGAGTATT

Table 2-3. Major characteristics of the predicted tobacco pre-miRNAs

Characteristic	Minimal	Maximum	Median	Average	Standard Deviation
Precursor Length (nt)	45	635	97	141	108
A%	14	43	28	28	5.4
C%	5.9	29	19	19	4.7
G%	9.8	32	21	20	4.5
U%	16	53	32	32	5.8
G + C %	15.69	58.41	39.58	39.43	7.52
A + U %	41.59	84.31	60.42	60.57	7.52
A/U	0.269	2.308	0.875	0.912	0.261
MFE (-kcal/mol)	2.2	181.4	39.8	45.93	26.19
MFEI	0.275	1.714	0.936	0.91	0.23

Table 2-4. Sense/Antisense miRNA pairs identified in tobacco

Pair	GSS	miRNA	Mature sequence	Length (nt)
1	ET768006.1	nta-miR169l	CAGCCAAGGATGACTTGCCGA	21
	ET768006.1	nta-miR169ar	TAGCCAAGGACGACCTGCCGA	21
2	FH141870.1	nta-miR169as	CAGCCAAGGATGACTTGCCGG	21
	FH141870.1	nta-miR169at	TAGCCAAGGACGACCTGCCGA	21
3	ET710327.1	nta-miR172w	AGAATCTTGATGATGCTGCAG	21
	ET710327.1	nta-miR172q	TGAATCTTGAAGATGCTGCAT	21
4	ET735116.1	nta-miR172x	AGAATCTTGATGATGCTGCAT	21
	ET735116.1	nta-miR172i	TGAATCTTGATGATACTACAT	21
5	FH310965.1	nta-miR172k	AGAATCATGATGATGCTGCAT	21
	FH310965.1	nta-miR172r	TGAATCTTGATAATGCTGCAG	21
6	FH354512.1	nta-miR172h	GGAATCTTGATGATGCTGCAG	21
	FH354512.1	nta-miR172s	TGAATCTTGATGATGCTCCAC	21
7	FH354528.1	nta-miR172e	AGAATCTTGATGATGCTGCAT	21
	FH354528.1	nta-miR172m	TGAATCTTGATGATGCTACAT	21
8	FH354528.1	nta-miR172a	AGAATCTTGATGATGCTGCAT	21
	FH354528.1	Nta-miR172l	TGAATCTTGATGATGCTACAT	21
9	FH419018.1	nta-miR172n	TGAATCTTGATGGTGCTGCAC	21
	FH419018.1	nta-miR172d	AGAATCTTGATGATGCTGCAT	21
10	FI036771.1	nta-miR172o	GAAATCTTGATGATGCTGCAT	21
	FI036771.1	nta-miR172v	GTGATCTTGATGATGCTACAT	21
11	FI036850.1	nta-miR172p	TGAATCTTGATGATGCTACAT	21
	FI036850.1	nta-miR172y	AGAATCTTGATGATGCTGCAT	21
12	FH050498.1	nta-miR827c	TAGATGAACATCAACAAACA	20
	FH050498.1	nta-miR827b	TAGATGACCATCAACAAACA	20

*Letters denoted in red indicate nucleotide changes between sense and antisense miRNAs

Table 2-5. Potential target genes of the identified tobacco miRNAs

miRNA	Targeted protein	Target function	Targeted genes or EST homologs of genes in other plant species
156	superman	Development	BE559491
	DNA-binding protein 1	Transcription Factor	DW004197
	kinase, putative		EB425044
	LIGULELESS1 protein, putative	Development	AM807766
	H-Protein precursor	Metabolism	EB436218
	hydrolase, alpha/beta fold family protein	Metabolism	FG634482
	ORF; able to induce HR-like lesions	Cell Death	BP135646
	squamosa promoter binding-like protein	Transcription Factor	EB444606
	3-oxo-5-alpha-steroid 4-dehydrogenase, putative	Metabolism	EB450875
	external rotenone-insensitive NADPH dehydrogenase	Metabolism	FG138705
	prolyl 4-hydroxylase	Development	FS417650
	putative red chlorophyll catabolite reductase	Metabolism	DV158165
	RecName: Full=B2 protein.	Stress Response	FG164488
	mitochondrial small heat shock protein	Stress Response	EB424610
	beta-amylase 1	Metabolism	EB434926
	ATP:citrate lyase	Development	FG167990
	extra sporogenous cells	Development	AM799421
	predicted protein		DV160561, EB425160, FG163073
	hypothetical protein		EB425345
	unnamed protein product		FG187335, FG150861, FG156435, DV158476,
	PREDICTED: hypothetical protein		DW000397, EB442636, EB677559, AM787035, FG640925, EB426154, EB426255, FG164766, FS393265, FG636537, FG168385
	conserved hypothetical protein		EB427769
	unknown		EH619226
158	PREDICTED: hypothetical protein		B427309
	unnamed protein product		FG182728
	unknown		DV998841
159	glutamine synthetase GS58	Metabolism	CV017169
	chromoplast-specific lycopene beta-cyclase	Development	FG148334
	oxidoreductase	Metabolism	FG167430
	PREDICTED: hypothetical protein		EB443826, FG146888, FG193261, FS394461, DW003222
	unnamed protein product		AM790652
160	glyceraldehyde-3-phosphate dehydrogenase	Metabolism	DV158995
	cutin deficient 2	Stress Response	AM783156
	PREDICTED: hypothetical protein		AM793353
162	dicer-1, putative	Metabolism	BP528819
	hypothetical protein SNOG_04965		EB678746
164	translation initiation factor eIF(iso)4E	Metabolism	DV160670

	VAP27	Stress Response	EB427431
	inorganic phosphate transporter	Stress Response	EH623241
	formyltetrahydrofolate deformylase	Metabolism	AM806229
	CONSTANS interacting protein 6	Transcription Factor	FG192737
	LEA1-like protein	Stress Response	EB440322
	glutamine synthetase GS56	Metabolism	EB426177
	MPF1	Transcription Factor	EB436448
	NADH dehydrogenase	Stress Response	EB441823
	methionine synthase	Metabolism	FG144455
	nam-like protein 18	Transcription Factor	EB443280, EH617135
	conserved hypothetical protein		FS432977
	PS60	Development	FG150485
	ERD1 (EARLY RESPONSIVE TO DEHYDRATION 1); ATP binding / ATPase/	Stress Response	DW002464
	Integrase core domain containing protein	Transcription Factor	EB443754
	glycogenin, putative	Metabolism	AM804369
	salicylic acid-induced protein 19	Stress Response	EH622111
	cysteine proteinase aleuran type	Metabolism	DV157790
	putative dienelactone hydrolase family protein	Metabolism	EB427875
	Biotin carboxyl carrier protein subunit of of Het-ACCase (BCCP2)	Metabolism	DV160864
	WRKY transcription factor, putative	Transcription Factor	DV162181
	FUS-interacting serine-arginine-rich protein 1, putative	Transcription Factor	EB447994
	PREDICTED: hypothetical protein		EB426905, FG159402, DV158728
	PREDICTED: similar to RNA binding motif protein 5		FG188951
	PREDICTED: coproporphyrinogen oxidase		FG180150
	unnamed protein product		FG202823, DV160965, DW000461, FG179392
	PREDICTED: hypothetical protein isoform 1		FG186666
	predicted protein		DV159557, EB427000
	putative auxin growth promotor protein	Development	EB450015
	hypothetical protein		FS416083
	unknown		EB440943
166	S-adenosyl-methionine-sterol-C-methyltransferase	Metabolism	FG198645
	bZIP transcriptional activator RSG	Transcription Factor	FG197643
	putative cytosolic cysteine synthase 7	Metabolism	FG164511
	PHAVOLUTA-like HD-ZIPIII protein	Transcription Factor	DV162377
	PREDICTED: hypothetical protein isoform 3		FG143036
167	NAD-malate dehydrogenase	Metabolism	FG133966

	putative hydroxymethylglutaryl-CoA lyase	Metabolism	DW002132
	PREDICTED: hypothetical protein		EB681652, EB448185
	predicted protein		FG133622, FG136127
	unnamed protein product		FG171938
168	actin related protein 2/3 complex, subunit 1	Development	AM814390
	AGO1-1	Stress Response	FG137295
169	S-phase kinase-associated protein 1A	Metabolism	AM807876
	phenylalanine ammonia-lyase 4	Development	FG144215
	squalene epoxidase	Metabolism	DV160820
	Gag/pol polyprotein, 3'-partial, putative	Transcription Factor	AM804568
	Pc12g10670	Stress Response	AM832554
	DNA-damage inducible protein ddi1, putative	Metabolism	DV158761
	hypothetical protein PPL_03932		AM784231
	cytochrome C oxidase, putative	Metabolism	DV159009
	CCAAT-binding transcription factor subunit B	Transcription Factor	AM832441
	RecName: Full=Pleiotropic drug resistance protein 1	Stress Response	BP528280
	arabinogalactan protein	Development	EB683943
	angustifolia		FG135788
	hypothetical protein Atu1592		AM809201
	nucellin, putative	Development	EH665829
	ribosomal protein L3B	Development	DV158371
	SBT1	Development	FG161907
	kinase, putative	Metabolism	AM804777
	YA6		FG160256
	pyruvate kinase	Metabolism	BQ842885
	thiosulfate:cyanide sulfurtransferase-like protein	Metabolism	DV999177
	o-linked n-acetylglucosamine transferase, ogt, putative	Metabolism	FG142825
	wound-inducible carboxypeptidase	Stress Response	AM846244
	RecName: Full=Photosystem II 22 kDa protein	Metabolism	EB435486
	putative polyprotein		AM806099
	ribulose-1,5-bisphosphate carboxylase small subunit	Metabolism	CV019286
	PREDICTED: similar to ubiquitin-related		FS381448
	predicted protein		DW000411, BP130805, AM786192, FS407719, FG135631, DV999481, DW003220, EB427563, EH623345, BP130420, EH615494, FS416248
	unnamed protein product		DV157545, FG160606, FG170096, FG161007, FS384301, EB447894, EB677908
	hypothetical protein		AM832748, BP526804, FG641643
	PREDICTED: hypothetical protein		FG141018, FG156434, EH666174, EH623371, DV160352, AM838211, AM845406, AM835846, EB679313, DV161265, DW000969

	conserved hypothetical protein		FG173399, EB443248, DV999787
	unknown		CV020194
170	ATP-dependent transporter, putative	Metabolism	BP533608
	hairy meristem	Development	FG136779
	Magnesium and cobalt efflux protein corC, putative	Cellular Transport	AM784659
	protein binding protein, putative	Development	FS416186
	predicted protein		FG141163
	PREDICTED: hypothetical protein		DW001281
	hypothetical protein		FG158889
	unknown		FG174873
171	Glycine dehydrogenase	Metabolism	FG135194
	CYP72A57	Metabolism	AM804287
	Putative mandelonitrile lyase, related	Metabolism	AM826614
	ABC transporter family protein	Cellular Transport	AM820005
	unnamed protein product		DW002970, FG203042, FS423912, EB427599, EB445881
	conserved hypothetical protein		FS409961
	PREDICTED: hypothetical protein		FG169695
	unknown		FG639442
172	Protein yippee-like.	Transcription Factor	FG644675
	anther ethylene-upregulated protein ER1	Transcription Factor	FG165844
	poly A binding protein, cytoplasmic 1 a	Transcription Factor	FG181563
	putative Rho GDP dissociation inhibitor	Development	DV157624
	Full=Pirin-like protein	Metabolism	AM781883
	Retrotransposon gag protein		AM797596
	putative lipoic acid synthase	Stress Response	FG201394
	apetala 2-like protein	Transcription Factor	FG196047, FG167355
	hypothetical protein SORBIDRAFT_04g026360		AM791290
	A-type cyclin	Development	FG621503
	UDP-glucuronate decarboxylase 1	Metabolism	DV160013
	allergen-like protein BRSn20	Stress Response	DW004910
	D-glycerate 3-kinase, chloroplast precursor, putative	Metabolism	EG650182
	ubiquitin activating enzyme 2	Metabolism	FG174792
	cyclin A-like protein		FG167945
	pectin acetylesterase, putative	Metabolism	EB427406
	cellulose synthase	Metabolism	AM813957
	PHAP2A protein	Transcription Factor	AM823467, FG175128
	PERK1-like protein kinase		FG168645
	Potassium transporter, putative	Cellular Transport	FG145341
	cytochrome b5 isoform Cb5-D	Metabolism	DV162449
	isopentenyl diphosphate isomerase 2	Metabolism	EB426446
	skip-2, putative		FG158904
	Auxin-responsive protein IAA1, putative	Transcription Factor	BP526688
	hypoxanthine-guanine phosphoribosyltransferase	Metabolism	AM816727
	histone acetyltransferase	Transcription Factor	EB428908

	serine carboxypeptidase, putative	Metabolism	AM845590
	nucleic acid binding protein, putative	Transcription Factor	FG640746
	hypothetical protein OsI_36426		FG639259
	ring finger protein, putative	Transcription Factor	EB449998
	WRKY DNA-binding protein	Transcription Factor	FG184696
	copper-containing amine oxidase	Metabolism	FG190820
	unnamed protein product		DW001320, EB438511, FG634880, FS408685, DV159237, EB682809, BP128422, FG147855
	conserved hypothetical protein		BP525578, EB679502
	predicted protein		FG149486, DV161945, DV158983, EB443695
	hypothetical protein		AM799732
	unknown protein		DV159112
	PREDICTED: hypothetical protein		AM831286, AM831939, EB426071, EB448001, AM846453, FG187709
	unknown		AM790527
319	PHAP2B protein	Transcription Factor	FG171579
	4-alpha-glucanotransferase	Metabolism	FG149364
	inositol or phosphatidylinositol kinase, putative	Metabolism	FG173520
	PREDICTED: similar to RING finger protein 122		FG181569
	SlTCP3		AM824285
	gamma-glutamylhydrolase 2	Metabolism	FG175238
	GAMyb-like1	Development	FG173519
	G protein beta subunit 2	Signal Transduction	FG637658
	unnamed protein product		FG635743, BP526144
	predicted protein		FS428299
	PREDICTED: hypothetical protein		FG161938
	conserved hypothetical protein		BP529940
	hypothetical protein		BP135022
390	elongation factor 1-alpha, putative	Transcription Factor	FG139910
	ATP synthase epsilon chain, mitochondrial, putative	Metabolism	EB683637
	predicted protein		AM804787
	PREDICTED: hypothetical protein		FS399504
	hypothetical protein		AM827396
393	protein disulfide-isomerase precursor	Metabolism	EB677770
394	Glucan endo-1,3-beta-glucosidase precursor, putative	Metabolism	FG148537
	F-box family protein	Metabolism	FG167650
	bZIP	Transcription Factor	FG143354
	PREDICTED: hypothetical protein		EB451712
	hypothetical protein		AM784871
395	ADP,ATP carrier protein, mitochondrial	Metabolism	FG136770
	ADP,ATP carrier protein precursor-like	Metabolism	FG158746
	serine carboxypeptidase, putative	Metabolism	FG135687
	beta-mannosidase enzyme		FS428578
	putative pyruvate dehydrogenase E3 subunit	Metabolism	EB437215
	PREDICTED: similar to mago		AM801308

	nashi, putative		
	AP2 domain-containing transcription factor	Transcription Factor	FG133359
	PREDICTED: hypothetical protein isoform 1		EB451940
	Polyprotein, putative		AM794604
	unnamed protein product		DV162284, FG635107, FG189414
	conserved hypothetical protein		FS384226
	predicted protein		FG157504, AM827911
	PREDICTED: hypothetical protein		AM841736, AM835374, DV159064, EB436913
	hypothetical protein		FG144930
396	calcineurin B-like interacting protein kinase	Stress Response	FG181278
	hypothetical protein RCOM_1491960		AM826953
	PREDICTED: similar to high mobility group box		FG188949
	SDL-1 protein	Metabolism	EH665019
	31 kDa ribonucleoprotein, chloroplastic	Transcription Factor	DV158935
	catalytic, putative		EB450427
	ribosomal protein L24-like protein	Transcription Factor	EB678262
	oxidoreductase, short chain dehydrogenase/reductase family	Metabolism	AM820385
	UPA17		FG137771
	CD63 antigen	Metabolism	FG181974
	Two pore calcium channel protein 1A	Cellular Transport	FS405053
	Ftsh-like protease	Stress Response	AM802161
	CXIP2 (CAX-INTERACTING PROTEIN 2); electron carrier/ protein	Metabolism	FG200330
	2-oxoglutarate-dependent dioxygenase	Metabolism	BP533160
	PREDICTED: hypothetical protein isoform 1		EB425511
	cysteine proteinase precursor	Metabolism	EB677689
	60S ribosomal protein L6, YL16-like	Transcription Factor	BP129814, EB677373
	mitochondrial processing peptidase	Cellular Transport	EB449176
	cysteine protease	Metabolism	AM835825
	fkbp-type peptidyl-prolyl cis-trans isomerase 2, chloroplast	Metabolism	EB433134
	unnamed protein product		EB429949
	predicted protein		FG192398, AM841671, FG156875, EB442055
	conserved hypothetical protein		EB424646
	PREDICTED: hypothetical protein		FG160221, EB427614, FG164325, FG167468, BP526271
397	laccase 1a	Metabolism	FG146277
	laccase, putative	Metabolism	FG177428, EB427021, AM801395
	diphenol oxidase	Metabolism	FG193650
	putative LysR-family regulatory protein	Transcription Factor	FS422307
	unnamed protein product		FG143000

	conserved hypothetical protein		BP133995
	PREDICTED: hypothetical protein		FG163250, FG169921
398	Blue copper protein precursor, putative	Metabolism	EH621792
399	ATP-dependent clp protease ATP-binding subunit clpx, putative	Metabolism	FG143637
	Eukaryotic translation elongation factor 2, like	Transcription Factor	FG188383
	RAB GTPase activator	Metabolism	FG636335
	mannitol dehydrogenase, putative	Metabolism	FG172582
	carbohydrate transmembrane transporter/ phosphate transmembrane	Cellular Transport	AM828792
	CPK related kinase 5	Metabolism	FG134143
	STY-L protein		FG138167
	NADPH:protochlorophyllide oxidoreductase	Metabolism	CV015965
	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase	Metabolism	FG625883
	mitochondrial ADP/ATP carrier protein	Metabolism	AM816871
	Root phototropism protein, putative	Developoment	BP128739
	pentatricopeptide repeat-containing protein, putative	Transcription Factor	BP135252
	unnamed protein product		FG154138, FG148716, BP128452, FG202483
	PREDICTED: hypothetical protein		FG634076, FG622030
	unknown		EB426039
400	PREDICTED: hypothetical protein		DV162153, FG140899
413	regulator of gene silencing	Transcription Factor	EH620795
	ATP-dependent RNA helicase, putative	Transcription Factor	FS385696
	unnamed protein product		DV161811
414	triacylglycerol lipase, putative	Metabolism	DV160414
	cinnamoyl-CoA reductase	Metabolism	EB426648
	ribosomal protein L7.	Transcription Factor	EH614684
	Floral homeotic protein PMADS 1, Green	Development	AM833905
	fumarylacetoacetate hydrolase, putative	Metabolism	FG639908
	CONSTANS interacting protein 5	Development	FG157668
	acyl-CoA oxidase ACX3	Metabolism	FG144716
	nucleosome assembly protein 1-like protein 2	Transcription Factor	FG154005
	30S ribosomal protein S17, putative	Transcription Factor	FS423826
	DNA binding protein, putative	Transcription Factor	FS377435
	RNase H family protein	Transcription Factor	FS408702
	DC1 domain containing protein	Transcription Factor	BP129153
	protein binding protein, putative		BP133434, FG153380
	KELP; DNA binding / transcription coactivator/ transcription	Transcription Factor	DW002370
	transporter, putative	Cellular Transport	DW003824
	ripening regulated protein	Metabolism	EB433985

DDTFR8		
sugar transporter, putative	Cellular Transport	EB449943
UPA21	Metabolism	AM832770
dead box ATP-dependent RNA helicase, putative	Transcription Factor	FG136678
putative flower-specific thionin	Stress Response	CN949710
main allergen 15 kDa oleosin	Stress Response	EB678103
hypothetical protein		AM781834
NAEGRDRAFT_57082		
PREDICTED: similar to eukaryotic translation elongation factor 1		AM821431
transcriptional activator FHA1	Transcription Factor	FG176447
calcium-binding protein	Signal Transduction	FG163611
PH4	Transcription Factor	FG199044
Pollen Ole e 1 allergen and extensin	Stress Response	FG137148
Dof28	Transcription Factor	FS395366
MYB transcription factor MYB127	Transcription Factor	FS374207
elongation factor-like protein	Transcription Factor	CV016769
nonclathrin coat protein zeta1-COP	Development	DV159654
nematode resistance-like protein	Stress Response	DV160751
nucleosome assembly protein 1-like protein 3	Transcription Factor	DV161202
PREDICTED: hypothetical protein isoform 1		DV999881
aldehyde dehydrogenase (NAD+)	Metabolism	DW001613
CONSTANS	Transcription Factor	EB680464
PREDICTED: hypothetical protein isoform 3		EH619199
heat shock protein Hsp90	Stress Response	AM787091
ATP binding protein, putative	Metabolism	FG637956
protein kinase, cAMP-dependent, regulatory, type II, beta	Signal Transduction	FG181413
hypothetical protein RCOM_1621800		FG177040
Ser/Thr protein kinase	Signal Transduction	FG162172
molecular chaperone Hsp90-2	Stress Response	FG133458
cig3	Transcription Factor	FG146766
glycine-rich protein precursor	Stress Response	CV016898
inorganic pyrophosphatase	Metabolism	FG200826
4-coumarate--CoA ligase 2, 4CL 2	Metabolism	FG169453
SET nuclear oncogene	Transcription Factor	FG181846
Vacuolar protein sorting-associated protein 41	Cellular Transport	EB102899
RNA-binding protein AKIP1-like	Transcription Factor	FG173868
unnamed protein product		DW002668, EB440419, FG138537, EB438538, FG643799, FG140979, FS407090, EB677597, FG141543, FG201769, FG147249, FG176854, BP530861, DW002343, FG164701, FG167238, FG199462, EB439826,
predicted protein		EB449631, AM805824, FG179262, DV158249, DV159465, DV160994, EB430739, EB438536, EB681526, EH619957, FG146023, FG170595,

			FG143925, FG152631
	protein with unknown function		DV159661
	hypothetical protein		FS416775, DW001325, AM798691, EB448580, EH622581
	conserved hypothetical protein		FG156101, DV999118, EB447933, FG175776, FG149677, DV999970, EB448670, EB677292, FG159434, FG153169, AM782636
	PREDICTED: hypothetical protein		DW000876, EB429242, EB432691, EB435290, EB440991, EB449722, AM814196, FS410675, DV157938, FG645047, FG136400, FG174533, BP133280, EB428385, EB446197, DV999560, EB441417, EB448318, AM805890, AM821122, AM812430, AM800287, FG640452, FG645008, FG149077, FG135686, FG155165, FG180580, FG164263, EB425880, FG160241, DV158861, EB426533, EB439887, EB679456, EB683324, FG187695
	unknown		EB681640, CV016507, EB447326, FS382161, FG632268
437	allergen-like protein BRSn20	Transcription Factor	CV019877
	hypothetical protein SORBIDRAFT_05g001830		DW003339
	cytochrome b5 isoform Cb5-B	Cellular Transport	DW003635
	CYP84A14v1		EB679929
	putative proline-rich protein	Development	AM823941
	phytoene desaturase	Metabolism	FS389399
	Vesicle-associated membrane protein, putative	Cellular Transport	EB439066
	hypothetical protein PTRG_11991		FG152095
	DNA-binding protein NtWRKY3	Transcription Factor	FG642501
	2-isopropylmalate synthase A	Metabolism	FG168228
	protein kinase 1b	Signal Transduction	FG155075
	ubiquitin-protein ligase, putative	Metabolism	DV162630
	unnamed protein product		FG144125, BP530921
	predicted protein		EB428507, DV159589, EH624236
	conserved hypothetical protein		EB643455
	PREDICTED: hypothetical protein		EB679209, FG191867
476	Phospho-2-dehydro-3-deoxyheptonate aldolase 1	Metabolism	DW000920
	cellulose synthase-like protein CslG	Metabolism	AM847566
	jasmonate ZIM-domain protein 1	Stress Response	FG173937
	hypothetical protein SORBIDRAFT_04g001030		EB452161
	glycosyltransferase	Metabolism	EB447137
	unnamed protein product		EH616948, AM834377
	PREDICTED: hypothetical protein		EB681590
479	Nodulation signaling pathway 2 protein, putative	Signal Transduction	EB427599
	GRAS family transcription factor	Transcription Factor	FG147444
482	aquaporin, MIP family, PIP subfamily	Cellular Transport	FG635594

	endomembrane protein emp70, putative	Cellular Transport	FS419227
	unnamed protein product		FS384026
	protein with unknown function		FG156406
	predicted protein		FG165848
	conserved hypothetical protein		CV018394
529	putative 60S ribosomal protein L7-like protein	Transcription Factor	DV159219
	Putative methyltransferase family protein	Metabolism	EH665954
	PREDICTED: hypothetical protein		FG163174
530	pectin methylesterase	Metabolism	FG182954
	hypothetical protein		AM807829
	DDB_G0275057		
	hypothetical protein		EB442388
776	glycine rich protein-interacting protein		BP136121
809	phosphate translocator	Cellular Transport	FG135239
810	lipid transfer protein 1 precursor	Stress Response	EB429408
	DNA-binding protein 3	Transcription Factor	AM789177
	unnamed protein product		AM795783
	predicted protein		DV160588, AM812167, DV158136
	conserved hypothetical protein		EB678653
	PREDICTED: hypothetical protein		FG636175
	unknown		EB450416
815	C-4 sterol methyl oxidase 2	Metabolism	DV160940
	predicted protein		FG145248
816	ubiquitin fusion-degradation protein-like	Metabolism	FG182701
	PREDICTED: hypothetical protein isoform 2		FG179733
	tRNA-dihydrouridine synthase A, putative	Transcription Factor	FG160290
	glutamate decarboxylase isozyme 1	Metabolism	DV160147
	putative helicase	Transcription Factor	AM806994
	endo-1,4-beta-glucanase	Development	FG138648
	hypothetical protein		EB679690
	POPTRDRAFT_573245		
	Glycine dehydrogenase	Metabolism	FG140474
	NTGP1		EB447632
	L-ascorbate oxidase, Ascorbase, ASO	Metabolism	FG171954
	NAD dependent epimerase/dehydratase, putative	Metabolism	DV999890
	pectin methylesterase PME2.1	Metabolism	EB442369
	salt responsive protein 1	Stress Response	FG163164
	collagenase 3	Development	FG179961
	peroxidase 12	Metabolism	FG152034
	unnamed protein product		FG623348, EB682228, CV016412, FG638887, DV160197
	predicted protein		FG170624, BP135813, EH623231, FG643891
	protein with unknown function		EH615307
	PREDICTED: hypothetical protein		FS434007, DW004554, FG634765, EB425377, FG198283, AM815091,

			AM820847
	unknown		EB446733, FS431689
817	putative isopropylmalate synthase	Metabolism	DW000268
	polypeptide with a gag-like domain	Transcription Factor	AM826142
	cystatin-like protein	Metabolism	FG639909
819	high affinity potassium transporter 2	Cellular Transport	AM837068
	RAD23-like	Stress Response	FG132867
826	elongation factor-1 alpha	Transcription Factor	DV158473
	putative Myb-like DNA-binding protein	Transcription Factor	FG179243
	hypoxia-responsive family protein	Stress Response	EB446560
	predicted protein		FG164981, AM823266, AM781500
	conserved hypothetical protein		FG158541, DV160091
	PREDICTED: hypothetical protein		FG174593
	unknown		DV157767
827	annexin	Stress Response	EB683148
	f-box family protein	Metabolism	FG150387
	DC1 domain containing protein	Transcription Factor	EB425372
	polypyrimidine tract binding protein, putative	Transcription Factor	FG168039
	serine/threonine protein phosphatase 2a regulatory subunit A,	Transcription Factor	FG178213
	EBP1	Development	FG146138
	glucosamine-fructose-6-phosphate aminotransferase, putative	Metabolism	FG138603
	TGACG-sequence-specific DNA-binding protein TGA-1A	Transcription Factor	FG184743
	hypothetical protein OsJ_21756		AM847043
	multifunctional protein		EH621465
	ADP-glucose pyrophosphorylase	Metabolism	EB429110
	protein kinase 2 beta chain	Stress Response	AM794768
	unnamed protein product		DV160523, AM799973, BP128647, FS423744
	predicted protein		FS408231, FG165734, FG195590
	conserved hypothetical protein		FG158163, BP128381, BP129746
	PREDICTED: hypothetical protein		FG153218, FG642348, BP534506, FG192418
	unknown		DW003860, BQ842961
828	enolase	Metabolism	FG168029
	beta-galactosidase 1 precursor	Metabolism	FG134436
	cytochrome P450	Metabolism	EB447290
	hypothetical protein		DV159314
	PREDICTED: hypothetical protein		FG621802
829	invertase inhibitor	Transcription Factor	EB425155
	nitrate transporter, putative	Cellular Transport	DV998825
	glycosyl hydrolase family-like protein	Metabolism	FG135254
	At1g67070-like protein		FG156537
	PREDICTED: hypothetical protein		DV999755, EB426493, EB424814
830	auxin-repressed protein	Development	DV160676
	allantoinase, putative	Metabolism	DW000559
	hypothetical protein DDB_G0283773		AM813001
	catalytic, putative	Metabolism	EB440747

	o-methyltransferase, putative	Metabolism	EB435759
	PHB1	Transcription Factor	EB682996
	sec10, putative	Transcription Factor	FG189777
	angustifolia		FG147339
	PREDICTED: hypothetical protein isoform 2		EB678087
	PREDICTED: hypothetical protein		EB432224, FS424081, BP532960, AM804199, EH621534
832	predicted protein		EB429240
833	aminophospholipid ATPase	Metabolism	AM825704
	unnamed protein product		BP532821
835	cation proton exchanger	Cellular Transport	FS420775
	PREDICTED: hypothetical protein		FS402179
	unknown		EB438970
837	CT099		EB426950
	beta-1,3-glucanase	Metabolism	FS416173
	dihydropyrimidinase, putative	Metabolism	BP130299
	FAM114A2		FG180355
	AG-motif binding protein-1	Transcription Factor	FG133085
	50S ribosomal protein L1p, putative	Transcription Factor	FG182955
	arginine decarboxylase	Metabolism	FG172113
	RWD domain-containing protein	Stress Response	EB449524
	transport inhibitor response 1	Transcription Factor	FG162945
	unnamed protein product		FG132871
	predicted protein		AM832558, EB427362, FG158766
	conserved hypothetical protein		DV162614
	PREDICTED: hypothetical protein		AM818218, DV161965
838	WRKY transcription factor NtEIG-D48	Transcription Factor	FG172760
	oligopeptide transporter OPT family	Cellular Transport	FG192769
	CHR18; ATP binding / DNA binding / helicase/ nucleic acid binding	Transcription Factor	EB440200
	hydroxy-methyl-glutaryl-coenzyme A reductase	Metabolism	FG182643
	alcohol dehydrogenase, putative	Metabolism	EB681185
	chloroplast-localized protein	Development	CV016364
	Carbonic anhydrase, chloroplastic	Metabolism	DV998890
	osmotic stress-induced zinc-finger protein	Stress Response	EB424711
	one-helix protein		EB432134
	lactoylglutathione lyase, putative	Metabolism	DW004313
	Protein AFR, putative	Signal Transduction	EB451948
	nematode resistance protein	Stress Response	AM799879
	secretory carrier membrane protein	Cellular Transport	FG639137
	xyloglucan endotransglucosylase-hydrolase XTH6	Stress Response	FG150239
	harpin inducing protein 1-like 18		CN498791
	multidrug resistance pump, putative	Cellular Transport	DW000606
	BIPINNATA		DV160465
	NADP-malic enzyme	Metabolism	EH664808
	4-nitrophenylphosphatase, putative	Metabolism	FS421237

	bHLH transcriptional regulator	Transcription Factor	BP131678
	Negatively light-regulated protein, putative	Stress Response	FS391155
	unnamed protein product		FG185268, FG166081, FS404395, FG635734, FS395514, EB440356, AM824105, FG163653, DW000621, EH624111, AM822057, FG635453, DV999564
	protein with unknown function		EB432557
	predicted protein		FG163234, FG132948, EB438429, FG643654, EB441957, BP137101, AM802885, BP534464, FG151336,
	unknown protein		FG158789, EB434433,
	conserved hypothetical protein		EB426258, FG135373
	hypothetical protein		FG176169
	PREDICTED: hypothetical protein isoform 1		FG164972
	PREDICTED: hypothetical protein isoform 3		EB439424
	PREDICTED: hypothetical protein		EB679749, FG163824, BP529720, EB430724, FG191641, AM840811, BP526798, FG644937, FG190392, FG152784, EB443042, DW000855, EB433675, EH666113, FG155517
	unknown		DV999897, FS374771, AM806524, AM842564, FG631550, DV158711, FG156831, BP534648, EB426335
844	resistance protein PSH-RGH6		FG138058
	Patellin-6, putative	Cellular Transport	FG140263
	EIL4	Transcription Factor	EB682741
	kelch repeat-containing F-box family protein, putative	Metabolism	FG183596
	ctp synthase, putative	Metabolism	DV159592
	unnamed protein product		BP535352, CV018224, FS422608
	PREDICTED: hypothetical protein		EB440279, AM828935, DV162641
	unknown		FG169066
845	nitrate transporter, putative	Cellular Transport	DW002467
	putative polyprotein	Development	DV161173
	hypothetical retrotransposon	Transcription Factor	AM845093
	polypeptide with a gag-like domain	Transcription Factor	FG193955
	RNA helicase PRH75	Development	DV157674
	glycerol-3-phosphate dehydrogenase 2 (mitochondrial)	Metabolism	FG181494
	retroelement pol polyprotein-like	Stress Response	BP137391
	glucosyltransferase	Metabolism	BP128870
	voltage-dependent anion channel	Metabolism	DV157589
	SAMT	Development	DW000692
	predicted protein		FG173359, FG137749, FG199806
	unnamed protein product		EB679175
	hypothetical protein		EH621931, EH623313
	PREDICTED: hypothetical protein		EH619422, FG182880, EB428698
846	xyloglucan endotransglucosylase-hydrolase XTH6	Development	BP527930
	littorine mutase/monooxygenase	Metabolism	FG175688
	programmed cell death, putative	Stress Response	EB677896

	polypeptide with a gag-like domain	Transcription Factor	AM817506
	50S ribosomal protein L17, chloroplastic	Development	EB680673
	coatomer protein complex subunit beta 2	Development	FS419351
	unnamed protein product		FG148315
	conserved hypothetical protein		EB441981, EB683567, EB444815
	hypothetical protein		FG142623
	PREDICTED: hypothetical protein		DW003320, EB447370
847	vacuolar citrate/H+ symporter	Metabolism	EB437180
	catalase 1	Stress Response	EB441492
	Mitochondrial import receptor subunit TOM20	Mitochondrial Transport	EG649732
	PREDICTED: similar to MYB transcription factor MYB139	Transcription Factor	FS394167
	amidase-like protein	Metabolism	EB680222
	protein translocase, putative	Cellular Transport	EB682123
	ovate-like protein	Transcription Factor	BP528963
	Ran binding protein-1	Metabolism	EB446361
	polygalacturonase 7	Development	FG630713
	26S protease regulatory subunit 6A homolog	Metabolism, Stress Response	EB427349
	Histone deacetylase HDT1, Histone	Transcriptional Regulator	EB430949
	phosphofructokinase, putative	Metabolism	FG150832
	PREDICTED: similar to UDP-glucuronic acid decarboxylase 1 isoform 2	Metabolism	FG139639
	putative ethylene response protein	Stress Response	BP131158
	40S ribosomal protein S2, putative	Development	DV159165
	putative VAMP protein SEC22	Metabolism	EB424642
	germin-like protein	Development	EH615784
	Pyruvate kinase isozyme G, chloroplastic	Metabolism	FG637058
	putative 26S proteasome subunit RPN2a	Development	FS427798
	CYP82E8	Metabolism	FS383006
	srpk, putative	Metabolism	BP530488
	thaliana 60S ribosomal protein L7 (At2g44120)	Development	DV158858
	low temperature and salt responsive protein	Stress Response	EH616830
	WRKY transcription factor, putative	Transcription Factor	FG177384
	unnamed protein product		DW000611, FG134467, FS405008, EB440647, FG177137, AM790804, AM837910, EB446770, FS430244, FS425845
	predicted protein		EB438918, EH622543, DW001040, EB426863, BP532363, EB427230, EB681084, FG638482, FG643509, BP130137,
	conserved hypothetical protein		AM794630, FG135961, DV160259, EB447642
	hypothetical protein		BP525754, FG170779, BP133732,
	PREDICTED: hypothetical protein		DW002271, EB427123, FG623759,

			FG198752, FS436048, DV161994, FG635780, FS411331, BP134202, BP135223, DW000843, DW001098, DW005098, EB424606, EH664643, EB438659, FG159484, FS381763, CV017091, EB440127
	unknown		EB424989, DV162599, EB425278,
854	shaggy-like kinase 91	Stress Response	FG136395
	zinc finger CONSTANS-related	Transcription Factor	AM848332
	tRNA, putative	Development	DV998926
	RNA-binding gricine-rich protein-1c	Transcription Factor	BP531231
	beta-amylase 1	Metabolism	EB426982
	cytokinin-regulated kinase 1	Development	FG141828
	sugar transporter, putative	Development	EB443251
	Cyclic nucleotide-gated ion channel, putative	Metabolism	FG142690
	Cop11 protein	Signal Transduction	EB682558
	phospholipase A1	Metabolism	FG165328
	6-phosphogluconate dehydrogenase family protein	Metabolism	FG139030
	zinc finger protein, putative	Transcription Factor	FS433248
	ribonucleotide reductase	Metabolism	DV161199
	ribose-phosphate pyrophosphokinase 4 / phosphoribosyl diphosphate	Metabolism	EB438220
	Putative isopenicillin N epimerase, identical	Metabolism	EB449488
	B-type cyclin	Development	FG139249
	nucleic acid binding protein, putative	Transcription Factor	FG137115
	hypothetical protein PANDA_017642		FG181840
	calcineurin B-like interacting protein kinase	Signal Transduction	DW003067
	jasmonic acid-amino acid-conjugating enzyme	Stress Response	EB442375
	unnamed protein product		DV161617
	hypothetical protein SDM1_56t00017		FS410927
	predicted protein		AM783870, EH620448, FG198249, FG175937, DV999925
	conserved hypothetical protein		FS383851, EB443904
	PREDICTED: hypothetical protein		EB441984, DW005157, DV161209, AM832908, FG196469, FG173720
	unknown		FG198005, DW001649
859	poly(p)/ATP NAD kinase, putative	Metabolism	FG161614
	Putative retrotransposon protein, identical	Stress Response	BP527716
	ZTL	Development	EB441385
	glutathione reductase	Stress Response	FG200359
	predicted protein		EB426054, EB430922, FG156161
	conserved hypothetical protein		FG194403
	hypothetical protein		FG144230
	PREDICTED: hypothetical protein		FS393941, FG168343
863	putative 40S ribosomal protein S9	Development	CN949785

	acyl carrier protein	Metabolism	DW003069
	Peroxidase 2 precursor, putative	Stress Response	EB683937
	P-rich protein EIG-I30	Stress Response	FG639172
	I-box binding factor	Transcription Factor	EB425765
	sulfate adenylyltransferase	Metabolism	EB426185
	beta-galactosidase 2 precursor	Metabolism	FG198990
	iaa-amino acid hydrolase 4	Development	FS398654
	unnamed protein product		EH623156, DW001096
	conserved hypothetical protein		FG202895
	unknown protein		AM811338
	unknown		DW000523
864	multidrug resistance protein ABC transporter family	Cellular Transport	AM836701
	ADH-like UDP-glucose dehydrogenase	Metabolism	EH663613
	F5M15.26	Development	AM818991
	predicted protein		DW002770
	hypothetical protein		FG645049
	PREDICTED: hypothetical protein		FS435982
	unknown		FS427121
865	Nt-iaa2.3 deduced protein	Transcription Factor	EB428225
	beta-ketoacyl-coa synthase family protein	Metabolism	EG650201
	transposon protein, putative, CACTA, En/Spm sub-class	Transposon Mobility	AM844890
	alpha-l-fucosidase, putative	Metabolism	FG174195
	Glyceraldehyde-3-phosphate dehydrogenase A	Metabolism	DV999146
	hypothetical protein NitaMp042		AM845578
	nucleoporin, putative	Nuclear Transport	FG135286
	protein binding protein, putative	Development	FG187073
	pathogenesis-related protein 10	Stress Response	CV017755
	subtilisin-like protease preproenzyme	Development	FS431723
	Thioredoxin domain-containing protein, putative	Stress Response	CV017023
	DNA binding protein, putative	Transcription Factor	DW000301
	porphobilinogen synthase, putative	Metabolism	DW001718
	putative anthocyanidine rhamnosyl-transferase	Stress Response	DW001930
	vacuolar ATP synthase subunit ac39, putative	Metabolism	DW002727
	Protein CCC1, putative	Stress Response	EB427889
	Probable cinnamyl alcohol dehydrogenase 1, CAD	Development	EB683825
	ascorbate peroxidase	Stress Response	EH620021
	inositol polyphosphate kinase	Metabolism	AM802720
	MAX2 (MORE AXILLARY BRANCHES 2); ubiquitin-protein ligase	Development	AM839437
	CDP-diacylglycerol synthase	Metabolism	AM842538
	seven-transmembrane-domain protein 1	Cell Signaling	EB425001
	gibberellin 2-oxidase 2	Development	EB425640
	UDP-glucose:protein transglucosylase-like	Metabolism	EB447810

	Putative gag-pol polyprotein, identical	Stress Response	EB682710
	Putative retrotransposon protein, identical	Stress Response	AM843276
	PREDICTED: R2R3 MYB60 transcription factor	Transcription Factor	AM845039
	unnamed protein product		EB424939, FG175716, FG173829, DV999558, FS382447,
	predicted protein		AM842295, AM801483, EB427718
	conserved hypothetical protein		FG638527, FG153708
	hypothetical protein		FG160812, BP135290
	PREDICTED: hypothetical protein isoform 2		EB435865
	PREDICTED: hypothetical protein		EB439284, FG176712, BP133145, AM794766, DV159744, FS401278
	unknown		FS424251
869	progesterone 5beta reductase-B	Metabolism	FG157317
	transcription activator TGA1a - tobacco.	Transcription Factor	EB448325
	Putative polyprotein, identical	Development	AM817969
	unnamed protein product		AM824686, FG644947
	predicted protein		AM788322, FG160951
	PREDICTED: hypothetical protein		FG142606, FG157545, DW000085
	unknown		CV021515, EB677961
870	stress-associated protein 3	Stress Response	DV160063
	Omega-3 fatty acid desaturase, endoplasmic reticulum	Stress Response	DV160430
	40S ribosomal protein S7-like protein	Metabolism	DW005192
	actin 1	Development	EH620716
	S-locus-specific glycoprotein S6 precursor, putative	Development	EH622223
	starch synthase VI precursor	Metabolism	FG163218
	ATP binding protein, putative	Metabolism	FG202531
	UPF0497 membrane protein	Development	DV160871
	Aspartic proteinase nepenthesin-2 precursor, putative	Stress Response	EB678964
	dual-specific kinase DSK1	Metabolism	EB680855
	putative p-coumarate 3-hydroxylase	Development	EH619191
	alpha-actinin	Development	AM791971
	TGB12K interacting protein 2	Stress Response	FG640444
	putative FtsH protease	Metabolism	FG192053
	putative		FS409589
	PREDICTED: hypothetical protein		EB428574, EH664831
1426	Rar1	Stress Response	DV157766
	PB1 domain-containing protein	Metabolism	EB428587
	calcium dependent protein kinase	Metabolism	EH616787
	serine/threonine protein kinase SAPK8-like protein	Metabolism	FG166619
	obtusifoliol-14-demethylase	Metabolism	FG134650
	ATP binding protein, putative	Metabolism	FG154136
	rac-GTP binding protein, putative	Development	FG165528
	Avr9/Cf-9 rapidly elicited protein 146	Stress Response	FG160993
	endopeptidase inhibitor	Metabolism	BP132686

	Pirin-like protein	Transcription Factor	BP136323
	Ycf1	Stress Response	BP528054
	Phenylalanine ammonia-lyase	Metabolism	EB425087
	D-3-phosphoglycerate dehydrogenase	Metabolism	EH615067
	kinesin	Development	AM838477
	MADS-box protein 15	Transcription Factor	FG645580
	predicted protein		EB446422, AM800915, FS378682, EB428773
	conserved hypothetical protein		EC391384, DV158708
	hypothetical protein		EB442951, FG142069
	PREDICTED: hypothetical protein isoform 2		EB678617
	PREDICTED: hypothetical protein		FG143854, FG146281, DV999617, EB445042, EB451689, FG633874
	unknown		DW000438
1435	IAA-amino acid hydrolase ILR1 precursor, putative	Development	EB426487
	integrase	Metabolism	AM787912
	auxin response factor 2	Transcription Factor	FG136307
	beta-glucosidase, putative	Metabolism	FS374856
	putative centromere protein	Development	BP528905
	Pathogenesis-related protein R major form	Stress Response	EH621953
	NBS-LRR resistance protein-like protein	Stress Response	AM783753
	Zn finger protein	Transcription Factor	AM829152
	CBL-interacting serine/threonine-protein kinase, putative	Metabolism	FG145611
	breast cancer metastasis-suppressor 1-like protein	Transcription Factor	FG188053
	GRAS1	Stress Response	FG157320
	Potyvirus VPg interacting protein	Stress Response	BP532595
	WD-40 repeat protein	Transcription Factor	AM815877
	high affinity nitrate transporter protein	Stress Response	FG158222
	SAM (and some other nucleotide) binding motif	Metabolism	FG139117
	Ammonium transporter 1 member 1	Stress Response	EB429670
	cell division control 20	Metabolism	FG186676
	metal transporter	Stress Response	DV159888
	yth domain-containing protein, putative	Transcription Factor	EB427225
	Kunitz trypsin inhibitor	Development	EB684161
	ornithine decarboxylase, putative	Metabolism	EH615970
	glutaredoxin family protein	Stress Response	AM818893
	neryl diphosphate synthase 1	Metabolism	AM822999
	acyl-CoA thioesterase, putative	Metabolism	FG642272
	wd40 protein, putative	Transcription Factor	FG621790
	hypothetical protein SORBIDRAFT_07g024970		BP129626
	chalcone synthase	Stress Response	EB426939
	aldehyde dehydrogenase, putative	Metabolism	EB427935
	cytochrome P450	Metabolism	EB431984
	DNA-binding protein 4	Transcription Factor	EB681468

	cytochrome c oxidase family protein-like	Metabolism	AM793195
	chaperonin-60 beta subunit	Stress Response	FG151784
	short-chain dehydrogenase, putative	Metabolism	FG183181
	Protein kinase APK1B, chloroplast precursor, putative	Metabolism	DV159586
	ccaat-binding transcription factor subunit A, putative	Transcription Factor	EB679967
	Beta transducin-like protein, putative	Metabolism	FG170858
	white-brown-complex ABC transporter family	Metabolism	BP528984
	auxin influx carrier component	Stress Response	EB441827
	ATPase	Metabolism	AM803909
	H(+)-transporting ATPase	Metabolism	FG138662
	IAA9 protein	Transcription Factor	FG159405
	DNA (cytosine-5) methyltransferase	Metabolism	FG148122
	putative protein		FG157222
	unnamed protein product		AM786029, AM789338, EH614411, BP130551, EH663834, BP132025, EH623586, EB428475, EB447640, EB684091, AM824835
	predicted protein		EH620704, AM844832, EB432651, FG138825, DV162192, EB682612, FS428120, EH622211
	conserved hypothetical protein		FG201735, BP534720
	hypothetical protein		EB431943
	PREDICTED: hypothetical protein isoform 2		FG170582, FS402685, AM842075, BP130886
	PREDICTED: hypothetical protein		BP136088, DV157569, FG641194, FG152500, FG185995, EB447447, EB451463, DV999056, FG201610, BP534234, EB424886, EB427453, FG152304, FS380632, FS405740, EB436875, EB440010, EB680033, FS385751
	unknown		FG163476, DW002651, FG168856
1436	inorganic pyrophosphatase	Metabolism	FG158642
	predicted protein		FS388877, FS393461, BP532541
	PREDICTED: hypothetical protein		BP530665, FG173225, EB448316, FG186192
1439	Integrase core domain containing protein	Metabolism	AM840276
	acetyl-CoA synthetase	Metabolism	FG203109
	unnamed protein product		DV159539
	predicted protein		EB424621
	PREDICTED: hypothetical protein		EB450501
1442	Aquaporin PIP2.1, putative	Development	CK720599
	N-carbamoylputrescine amidase	Metabolism	BP534113
	Bax inhibitor 1	Stress Response	EB426602
	UPF0737 protein mc410	Metabolism	EB679472
	hypothetical protein smi_1934		AM798599
	xyloglucan endotransglucosylase/hydrolase 12	Metabolism	FG138402

	CBL-interacting serine/threonine-protein kinase, putative	Metabolism	FG177455
	ring domain containing protein	Transcription Factor	DW001993
	AP2 domain-containing transcription factor	Transcription Factor	EH665541
	amino acid transporter, putative	Metabolism	BP130193
	cell death associated protein	Stress Response	BP526050
	Superoxide dismutase	Stress Response	DW004580
	cytochrome P450	Metabolism	FG192274
	G protein beta-subunit-like protein	Metabolism	FG151480
	luminal binding protein	Stress Response	FG133604
	Chaperonin CPN60-1, mitochondrial	Stress Response	FG179274
	acetolactate synthase SuRB	Metabolism	FG136262
	unnamed protein product		AM822896, FG139648
	unknown protein		FG152328
	predicted protein		FG137870, FG133711,
	PREDICTED: hypothetical protein, partial		BP128734
	PREDICTED: hypothetical protein		FG168127, FG172317, FG642258, AM780077, EB424744, FG640136, EB451442, EB439328, EB435588, FG135805, FG184669, DV999761, EB677963
	unknown		FS434170
1445	constitutive plastid-lipid associated protein	Stress Response	DV159099
	9-cis-epoxy-carotenoid dioxygenase 2	Stress Response, Development	FG155526
	BTB-POZ and MATH domain protein	Transcription Factor	EB442834
	light harvesting chlorophyll a/b-binding protein	Metabolism	FG136315
	predicted protein		FG156235
	unnamed protein product		FS433193
	conserved hypothetical protein		DV998941
	PREDICTED: hypothetical protein		EB439245, FS415498
	PREDICTED: hypothetical protein isoform 2		FG165097
	unknown		EH618297
1446	Triosephosphate isomerase, chloroplastic, TIM	Metabolism	DV999593
	peroxidase	Stress Response	FG158108
	elicitor-inducible protein EIG-J7	Stress Response	DV158275
	vicilin	Development	DV158572
	Heat shock protein 82	Stress Response	EB430029
	poly(A)-binding protein	Metabolism	EB680143
	short-chain dehydrogenase/reductase (SDR) family protein	Metabolism	FG162137
	unnamed protein product		EB451268
	unknown protein		DW004734
	conserved hypothetical protein		EB440229
	hypothetical protein		FG185887
	PREDICTED: hypothetical protein		DV157751
	unknown		EH665665, DW000421

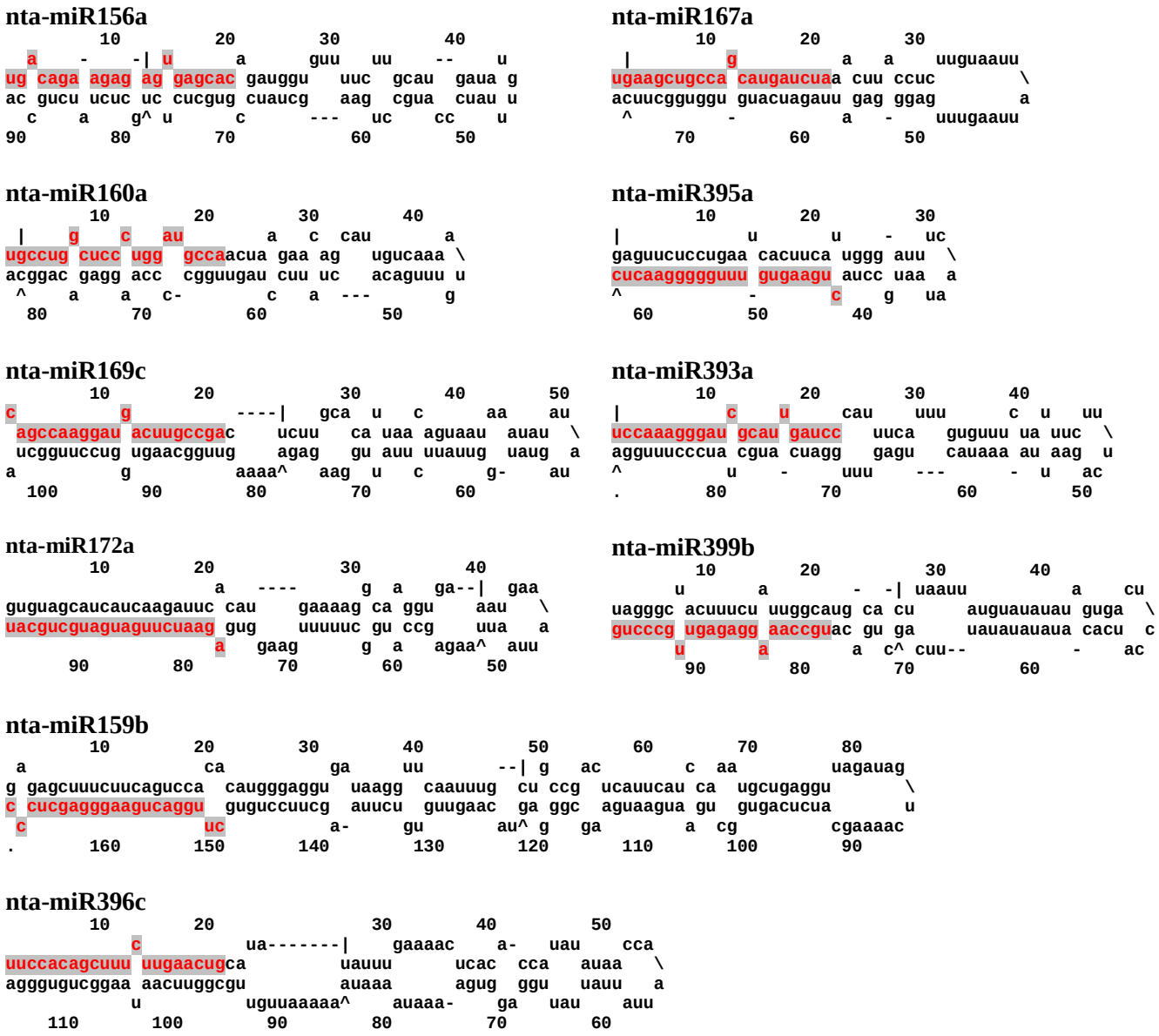


Figure 2-1. Predicted secondary hairpin structure for 10 of the tobacco miRNAs identified in this study. Mature miRNA sequences are shaded and pre-miRNAs may be longer than what is presented in this figure.

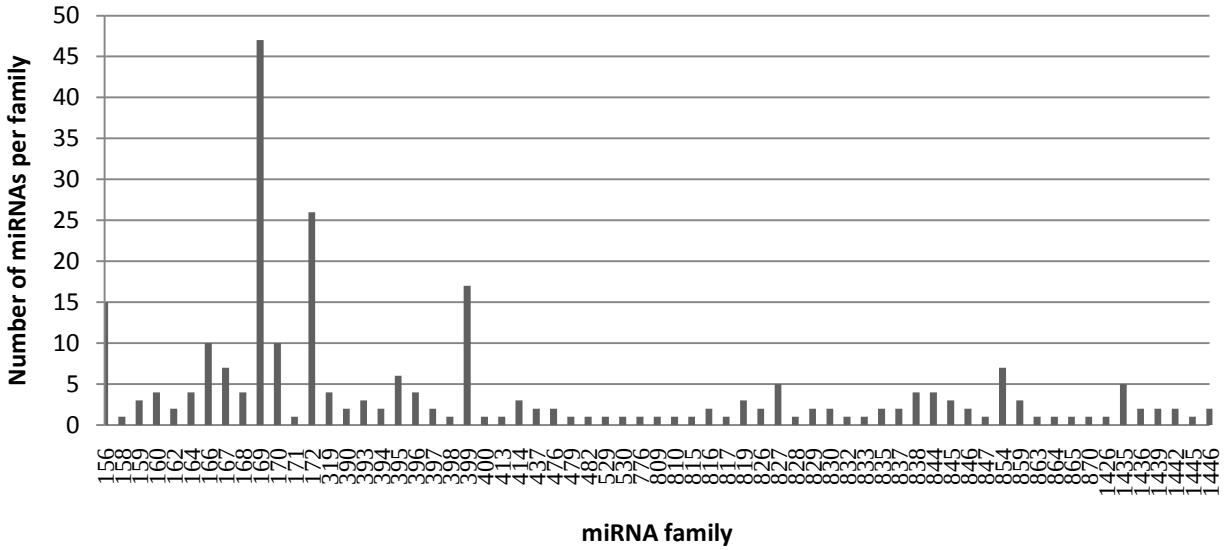


Figure 2-2. Size of tobacco miRNA families.

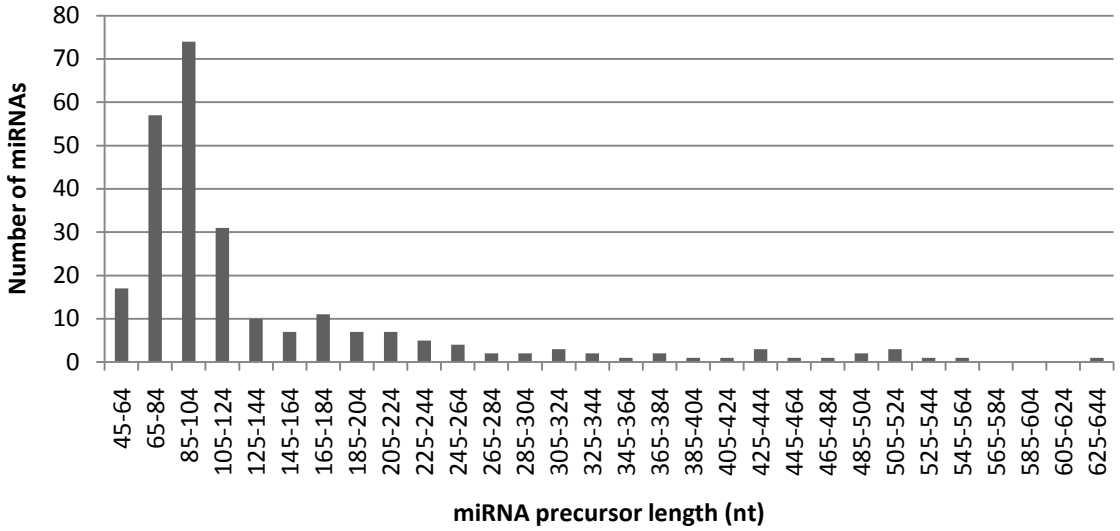
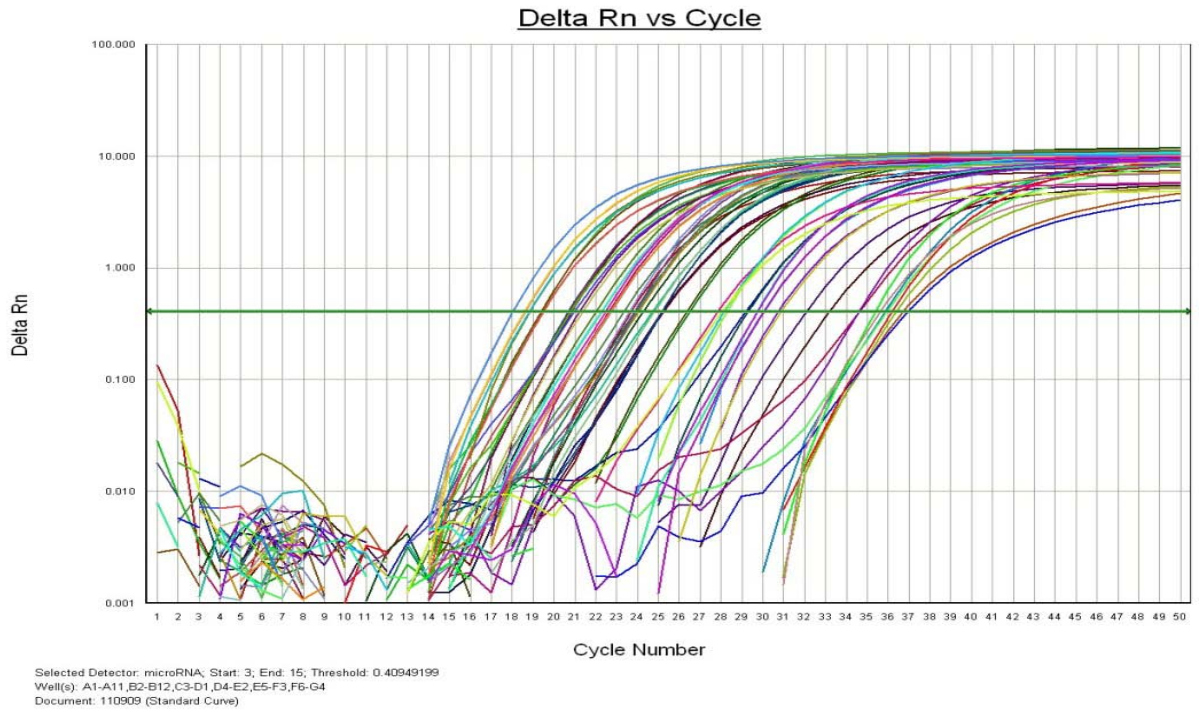


Figure 2-3. Size distribution of pre-miRNAs in tobacco.

a



b

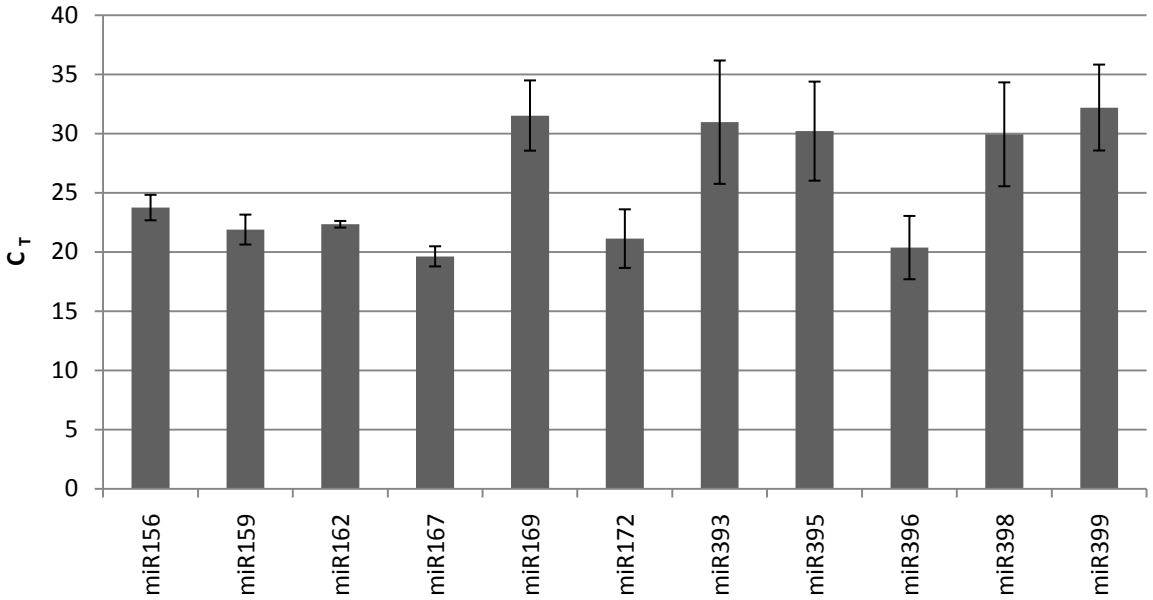


Figure 2-4. qRT-PCR confirmation of 11 predicted tobacco miRNAs. **a** Amplification plot depicting tobacco miRNA expression in young tobacco leaves. **b** C_T analysis of qRT-PCR data showing miRNA expression levels in young tobacco leaves.

a nta-miR172q and nta-miR172w

nta-miR172q

```
        10         20         30              40           50
c          a          u     uaaa----   a    c-  ua---|    c        u
 ugcagcauc  ucaagauuc cauau        uua ggag  ca     gggag aagaaaau \
 acgucguag aguucuaag guaua         aau ccuc  gu     cccuc uuuuuuua u
u          a          u     ucauauaa   a    aa  uaaca^    -        a
  120        110        100         90         80         70         60
```

```
CTGCAGCATCATCAAGATTCTCATATTAAATTAAGGAGCCATAGGGAGCAAGAAAATTTAATTTTTTTCTCCCACAATTGAACTCCATAAAATATACTATATGTGAATCTTGAAGATGCTGCAT
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GACGTCGTAGTAGTTCTAAGAGTATAATTTAATTCCTCGGTATCCCTCGTTCTTTTAAATTAAAAAAAGAGGGTGTTAACTTGAGGTATTTTATATGATATACACTTAGAACTTCTACGACGTA
```

nta-miR172w

```
        10         20             30          40           50
g          u          a      ----  uu   u    -   au---|    guuc     a
 acgucguag  aguucuaag guaua     au  aau ccuc ggu     cccuc    uuuua \
 ugcagcauc ucaagauuc cauau     ua  uua ggag uca     gggag    aaaau a
a          u          a      agua  uu   u    u   auugu^    aaa-     u
  120      110        100         90         80         70         60
```

b nta-miR169l and nta-miR169ar

nta-miR169l

```
        10         20         30
c|                    gau  ug a  c
 agccaaggaugacuugccga    gu  u gc a
 ucgguuccugcuggacggcu    ca  a cg a
a^                     auu  gu a  u
       60         50         40
```

```
CAGCCAAGGATGACTTGCCGAGATGTTGTAGCCAATGCAATGACTTATCGGCAGGTCGTCCTTGGCTA
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTCGGTTCCTACTGAACGGCTCTACAACATCGGTTACGTTACTGAATAGCCGTCCAGCAGGAACCGAT
```

nta-miR169ar

```
        10         20              30
          c   c       ua------|    uu
uagccaagga  gac  ugccga         aguca  g
gucgguuccu cug acggcu           ucggu  c
          a   a       cuacaaca^    ua
       60         50         40
```

Figure 2-5. Examples of two sense and antisense miRNA pairs identified in tobacco. **a** nta-miR172q and nta-miR172w are a sense/antisense pair found at GSS sequence ET710327.1. **b** nta-miR169l and nta-miR169ar are a sense/antisense pair found at GSS sequence ET768006.1.

a GSS: FH457758

b

Atacacaaataaaaaagtctagagagagatagagggagtgatggggtggcagatagagcagagagcacggatagcttcaggcatgggataaacat
atcatgcaagaaaacaccatct**gtgctcactctcttctgtca**cccatcacctccctttgctagtttttggcaaaggggattaggataaaaccttacaccct
atagttctgagtaccataccaataattccatatgttcataaacacagatgcacgcatacataaaacagaaagaaatatatatagagagagtggtgggggt
gacagatagagctgtgagcacgcatagcttcaggcatgggataaacatatcatgcaagaaaacaccatat**gtgctcactctcttctgtca**ccttcctttt
gttagaataatttttgacagagggtgttctgatgtcccttcacccccctaggtccgctcctgagtgctaataataccatatgttcataaacacagatgcgcatat
acataaaaacagaaagaaatccacagagagagagtggtgggggggtgacagatagagcagtgagcacgcatagcttcaagcatatcatatgcaagaa
aataccatgt**gtgctcactctcttctgtca**cctccctttatttacccctttatgtcgttaaccctagattttaattgggatgtaattcactccagccgacagaga
aataagaaaccctagatctgaacattagcaactactgtttcagaattatgattttgatttatgatttgattgtgtgaggatgatgg

c

nta-miR156a

```
          10          20          30          40
   a      -     -| u       a       guu    uu     --     u
ug caga  agag  ag  gagcac  gauggu    uuc   gcau   gaua g
ac gucu  ucuc  uc  cucgug  cuaucg    aag   cgua   cuau u
  c     a     g^  u       c       ---    uc     cc     u
90          80          70            60          50
```

nta-miR156b

```
          10          20          30          40
         -     a-|        a       guu    uu     --     u
ugacaga  agag   gugagcac  uauggu    uuc   gcau   gaua g
acugucu  ucuc   cacucgug  guaucg    aag   cgua   cuau u
         a     ga^        c       ---    uc     cc     u
90          80          70            60          50
```

nta-miR156c

```
          10          20          30         40
         -     -|         a       auu    uu      u
ugacaga  agag  agugagcac  cauggu    uuc   gcaua \
acugucu  ucuc  ucacucgug  guaucg    aag   cguau g
         a     g^         c       ---    uu      a
 80          70          60            50
```

Figure 2-6. Identification of the miR156a-miR156b-miR156c cluster found in tobacco GSS sequence FH457758. **a** A depiction of the organization of the miRNA cluster. **b** The tobacco GSS sequence and the location of the 3 pre-miRNAs shaded in grey and the location of the 3

mature miRNAs underlined in red. **c** The predicted secondary structures of nta-miR156a, nta-miR156b, and nta-miR156.

nta-miR156a	(3'-5')		CACGAGUGAGAGAAGACAGU
			||||| |||||||||||
SBP (EB444606)	(5'-3')	712	UUGCUCUCUCUCUUCUGUCA 730
nta-miR164a	(3'-5')		UCGUACACG-GGACGAAGAGGU
			||| | ||| ||||||||||
Translation initiation factor eIF(iso)4E (DV160670)	(5'-3')	674	AGCUUAUGCGGCUGCUUCUCCA 695
nta-miR166a	(3'-5')		CCCCUUACUUCGGACCAGGCU
			|||| ||||||||||||:
PHAVOLUTA-like HD-ZIPIII protein (DV162377)	(5'-3')	231	UGGGA-UGAAGCCUGGUCCGG 250
nta-miR169a	(3'-5')		AGCCGUUCAGUAGGAACCGAC
			||||||: ||||||||||||
CCAAT-binding transcription factor subunit B (AM832441)	(5'-3')	86	UCGGCAGUUCAUCCUUGGCUA 106
nta-miR172a	(3'-5')		UACGUCGUAGUAGUUCUAAGA
			|||||||||||||:||||||
Apetala 2-like protein (FG196047)	(5'-3')	388	CUGCAGCAUCAUCAGGAUUCU 408
nta-miR396a	(3'-5')		UUCAGGUUCUUUCGACACCCU
			|| :|||||||:|| | ||
SDL-1 protein (EH665019)	(5'-3')	855	CAGCGUAAGAAAGUUGCGUGA 873

Figure 2-7. Six selected tobacco miRNAs and one of their predicted target genes. Complementary base pairing is indicated by solid lines whereas G:U wobble base pairing is indicated by dotted lines. Note: these miRNAs may target other genes besides the ones shown.

CHAPTER 3: Titanium dioxide nanoparticles affect the growth and microRNA expression of tobacco (*Nicotiana tabacum*)

Abstract:

Titanium Dioxide (TiO_2) is one of the most widely used pigments in the world. Due to its heavy use in industry and daily life, such as food additives, cosmetics, pharmaceuticals, and paints, many residues are released into the environment and currently TiO_2 nanoparticles are considered an emerging environmental contaminant. Although several studies have shown the effect of TiO_2 nanoparticles on a wide range of organisms including bacteria, algae, plankton, fish, mice, and rats, little research has been performed on land plants. In this study, we investigated the effect of TiO_2 nanoparticles on the growth, development, and gene expression of tobacco, an important economic and agricultural crop in the southeastern United States as well as around the world. We found that TiO_2 nanoparticles significantly inhibited the germination rates, root lengths, and biomasses of tobacco seedlings after 3 weeks of exposure to 0.1%, 1%, 2.5%, and 5% TiO_2 nanoparticles and that overall growth and development of the tobacco seedlings significantly decreased as TiO_2 nanoparticle concentrations increased. Overall, tobacco roots were the most sensitive to TiO_2 nanoparticle exposure. Nano-TiO_2 also significantly influenced the expression profiles of microRNAs (miRNAs), a recently discovered class of small endogenous non-coding RNAs (~20-22 nt) that are considered important gene regulators and have been shown to play an important role in plant development as well as plant tolerance to abiotic stresses such as drought, salinity, cold, and heavy metal. Low concentrations (0.1% and 1%) of TiO_2 nanoparticles dramatically induced miRNA expression in tobacco seedlings with miR395 and miR399 exhibiting the greatest fold changes of 285 fold and 143 fold, respectively. The results of this study show that TiO_2 nanoparticles have a negative impact on tobacco growth

and development and that miRNAs may play an important role in tobacco resistance to heavy metals by regulating gene expression.

Introduction:

Titanium Dioxide (TiO_2) is a common whitening agent used in the production of foods, sunscreens, cosmetics, plastics, medicines, papers, pharmaceuticals, and paints (Trouiller et al. 2009). Estimates have shown that approximately 4 million metric tons of bulk TiO_2 are produced annually around the globe with the United States leading production with 1.3 million metric tons (Lubick 2009). TiO_2 production also accounts for 70% of the total amount of manufactured pigments worldwide (Baan et al. 2006). TiO_2 is becoming increasingly used in industry due to its white pigment, ability to block ultraviolet light, and photocatalytic properties that are commonly used to destroy pollutants in soil, air, and water/wastewater treatment (Higarashi and Jardim 2002; Kubo et al. 2005).

Due to its heavy use in industry and daily life, many TiO_2 residues are released into the environment and currently, TiO_2 nanoparticles are being considered as an emerging environmental contaminant. Studies analyzing the health effects of TiO_2 nanoparticle exposure, have led it to be classified by the International Agency for Research on Cancer as a group 2B carcinogen, meaning that it is potentially harmful to humans (Baan et al. 2006). Accordingly, mounting evidence has shown that TiO_2 nanoparticles create lipid peroxidation products and reactive oxygen species that induce epithelial cell injury and oxidative DNA damage in cell lines, mice, and rats (Falck et al. 2009; Gurr et al. 2005; Kang et al. 2008; Minhua et al. 1998; Trouiller et al. 2009; Vileno et al. 20007; Wang et al. 2008; Wang et al. 2007). Since TiO_2 nanoparticles are a potential hazard to human health, studies in rat models have shown that lung cancer can be induced after exposure to nano-TiO_2 (Borm et al. 2004; Heinrich et al. 1995; Pott

and Roller 2005). It should be noted, however, that no studies have shown a correlation between TiO_2 nanoparticle exposure and the development of lung cancer in industry workers (Boffetta et al. 2004; Fryzek et al. 2003).

As the manufacturing and industrial use of TiO_2 increases, the amount of TiO_2 nanoparticles released into the environment is also likely to increase. This is due in part to the high use of TiO_2, and the ability of TiO_2 nanoparticles to travel freely through wastewater treatment plants and enter neighboring soils and streams (Kiser et al. 2009; Lubick 2009). Therefore, TiO_2 is also being investigated as a potential environmental contaminant and the effect of TiO_2 in water (Adams et al. 2006; Hsu and Chang 2000; Kaegi et al. 2008; Moore 2006) and on soils (Higarashi and Jardim 2002; Mattigod et al. 2005; Nagaveni et al. 2004; Quan et al. 2005), bacteria (Kubo et al. 2005; Wang et al. 2009; Yao et al. 2007), algae (Hartmann et al. 2009; Hund-Rinke and Simon 2006; Wang et al. 2008), fungi (Hur et al. 2005; Lu et al. 2006; Maneerat and Hayata 2006; Sichel et al. 2007), plankton (Hund-Rinke and Simon 2006; Lovern and Klaper 2006; Lovern et al. 2007), and fish (Federici et al. 2007; Reeves et al. 2008; Zhang et al. 2007) have been analyzed. A recent study has also observed that TiO_2 nanoparticles cause genetic toxicity to *Salmonella typhimurium* (Pan et al. 2010). It is surprising, given the increasing concern of TiO_2 nanoparticles to the environment, that only two studies have analyzed the effect of TiO_2 nanoparticles on plants (Asli and Neumann 2009; Zheng et al. 2005) and no study has been performed on understanding the effect of nanoparticles on regulatory RNAs.

microRNAs (miRNAs) are a recently discovered class of small regulatory endogenous RNAs (~20-22 nt) that do not code for proteins (Bartel 2004; Zhang et al. 2006). miRNAs play an important role in plant growth and development as they have been shown to aid in leaf development (Bao et al. 2004; Bowman 2004; Emery et al. 2003; Juarez et al. 2004; Kim 2005;

Mallory et al. 2004b; Palatnik et al. 2003; Zhong and Ye 2004), stem and root growth (Guo et al. 2005; Laufs et al. 2004; Mallory et al. 2004a), organ maturation (Achard et al. 2004; Aukerman and Sakai 2003; Chen et al. 2004; Mlotshwa et al. 2006), and the plant's ability to withstand abiotic stresses (Allen et al. 2005; Jones-Rhoades and Bartel 2004; Lu et al. 2005; Sunkar and Zhu 2004; Zhang et al. 2005). miRNAs target messenger RNAs (mRNAs) and regulate gene expression by marking these mRNAs for degradation or repressing protein translation (Bartel 2004; Vaucheret 2006). Recent studies have shown that miRNAs mediate the expression of more than 30% of genes coding for protein (Lewis et al. 2005; Xie et al. 2005) and this number is expected rise as more miRNAs and their target sequences are identified.

miRNAs have been shown to play an important role in many plant biological and metabolic processes. Specifically, certain miRNAs have been identified in plant responses to abiotic stresses. For example, miR156, miR169, miR395, miR399, and miR399* expression levels are altered during the response of *Arabidopsis* to phosphate starvation (Fujii et al. 2005; Hsieh et al. 2009). miR398ab and miR408 have been shown to be up-regulated in drought-exposed *Medicago truncatula* plants (Trindade et al. 2010) and miR398 has been implicated in plant response to copper and oxidative stress (Abdel-Ghany and Pilon 2008; Sunkar et al. 2006). miRNA expression has also been shown to be altered under both cold (Zhang et al. 2009; Zhou et al. 2008a) and salt (Ding et al. 2009; Jia et al. 2009) stress.

In this study, we used tobacco as a model species to analyze the effect of TiO_2 nanoparticles on the growth, development, and miRNA expression levels of an agricultural plant. Tobacco is an important economic and agricultural crop in the southern United States as well as around the world. Although much research has been dedicated to this crop, no experiments have shown the effects of TiO_2 nanoparticles on tobacco growth and development. We found that as

the concentration of nano-TiO_2 increased, the germination rates, root lengths, biomasses, and leaf counts of 3-week old tobacco seedlings decreased in a concentration dependent manner. We also found that the expression levels of miRNAs were either up- or down-regulated in response to increasing TiO_2 nanoparticle concentrations. Ultimately, we were able to conclude that TiO_2 nanoparticles negatively affect tobacco growth and that miRNAs may play an important role in providing tobacco plants with resistance to heavy metal toxicity.

Materials and Methods:

Tobacco Treatment

Tobacco (*Nicotiana tabacum*) seeds were sterilized in 70% ethanol for 2 minutes followed by 10% bleach for 15 minutes. The seeds were then rinsed with sterile distilled water until no bleach odor remained. Sterilized seeds were sowed on Petri dishes containing Murashige and Skoog (MS) medium containing Gamborg's B5 Vitamins (0.44% salts, 1% sucrose, 0.8% agar, pH 5.8) and 0%, 0.1%, 1%, 2.5%, or 5% Titanium dioxide (TiO_2) nanoparticles. TiO_2 nanoparticles were purchased from Sigma. The part size is less than 25 nm with more than 99.5% purity. A total number of 25 seeds were sowed per plate and there were a total number of 5 plates for the control and each treatment with different TiO_2 nanoparticle concentrations. The plates were inverted at a 75° angle to promote downward root growth and allowed to grow for 3 weeks on a 16 hour day/8 hour night cycle.

Measuring Germination Rate, Leaf Count, Root Length, and Biomass

After 3 weeks of growth, the germination rate of all 5 plates for the control and each concentration of nano-TiO_2 was recorded. We followed the procedure as described in (Lin and Xing 2007) in which we recorded a seed as germinated if the radical or cotyledon was observed protruding from the seed coat. The number of leaves per seedling per plate for the control and

different concentrations of TiO_2 was also recorded. Seedlings were removed from the media and the root lengths of the control, 0.1% TiO_2, and 1% TiO_2 plates were measured (in cm) using a standard ruler. The root lengths of 2.5% and 5% seedlings were not measured because no roots were formed. In the case where more than one root was formed, all roots were measured but only the largest root measurement was used in calculating average root length. Once root lengths were recorded, seedlings were then placed on a weigh boat and the weight of biomass measured using a Discovery 214C balance (OHaus, Pine Brook, NJ, USA). After the biomass of each plate was recorded, the seedling tissue was immediately placed in liquid nitrogen and stored at -80°C until RNA extraction.

Total RNA Extraction

Total RNA was isolated from the 3-week old seedlings using the mirVana miRNA Isolation Kit (Ambion, Austin, TX, USA) according to the manufacturer's protocol. Total RNA was then quantified and assessed for quality using a Nanodrop ND-1000 (Nanodrop Technologies, Wilmington, DE, USA). RNA samples were stored at -80°C until further analysis.

Analyzing microRNA expression changes using RT-PCR and qRT-PCR

Applied Biosystems TaqMan microRNA Assays were employed to detect and quantify tobacco miRNAs using stem loop real-time PCR according to the manufacturer's instructions. There were two steps in the TaqMan miRNA Assays: a) reverse transcription of the mature miRNA to a longer single-stranded cDNA sequence using a miRNA-specific stem-looped primer, and b) quantitative real-time PCR. Briefly, a single-stranded miRNA cDNA was generated from 1 µg of the total RNA from the control, 0.1%, and 1% TiO_2 nanoparticle tissue samples by reverse transcription using the Applied Biosystems TaqMan microRNA Reverse Transcription Kit and miRNA-specific stem-looped RT primers provided in the kit. Each

reaction was repeated in triplicate. In total, we investigated changes in the expression levels of 11 different miRNAs (miR156a, miR159a, miR162a, miR167a, miR169a, miR172a, miR393a, miR395a, miR396a, miR398a, and miR399b) and 2 stress related genes (Alcohol Dehydrogenase (ADH) and Alcohol Peroxidase (APX)). Two housekeeping genes (Elongation Factor 1 (EF1) and Tubulin) were used as reference genes to normalized expression values. All reactions were performed in triplicate and the results were analyzed using the $\Delta\Delta C_T$ method.

Results:

Nano-TiO_2 Inhibited Tobacco Seed Germination and Plant Growth

After 3 days of planting, seeds started to germinate. At low concentrations (0.1% and 1.0%), nano-TiO_2 did not affect tobacco seed germination (Figure 3-1A). However, nano-TiO_2 significantly affected tobacco seed germination at higher concentrations ($p=0.016$). When 2.5% TiO_2 nanoparticles were added into the MS medium, the germination rate of seeds decreased from 98-100% to 93.60±3.58%. Also, the germination of tobacco seeds was completely inhibited by exposure to 5.0% nano-TiO_2. Although tobacco seeds germinated with a higher percentage after exposure to 2.5% nano-TiO_2, it was hard for these seeds to produce roots and fully developed cotyledons. Therefore, the germinated seedlings quickly died within about one week.

After three weeks, there was a significant obvious decrease in plant growth and development after exposure to elevated TiO_2 nanoparticle concentrations as evidenced by decreases in root lengths ($p<0.001$), leaf numbers ($p<0.001$), as well as plant biomasses ($p<0.001$) (Figures 3-1 and 3-2). Roots were the most sensitive to TiO_2 nanoparticle exposure (Figure 3-1B). Under normal conditions, the root lengths reached to 1.43±0.15 cm after 3 weeks of culture; however, after exposure to 0.1% and 1% nano-TiO_2, the root lengths were significantly reduced to 0.74±0.05 and 0.21±0.04 cm ($p<0.001$), respectively. 2.5% nano-TiO_2

completely inhibited the appearance and elongation of roots. An interesting phenomenon is that TiO_2 nanoparticle exposure influenced differential root patterns. Tobacco is a dicot; generally speaking, one seed only produces one major root. However, after exposure to 1.0% TiO_2, a majority of tobacco seedlings produced two roots although both were very short with an average of 0.21±0.04 cm after 3 weeks of growth. One of the potential reasons is that nano-TiO_2 exposure can damage the major root apical meristem and therefore allow the seedlings to produce an additional root from the damaged location.

Exposure to nano-TiO_2 also influenced tobacco leaf development as demonstrated by a reduction in the total number of leaves and the leaf size. After exposure to nano-TiO_2, tobacco leaves grew much smaller compared to the untreated control (Figure 3-2). The averages for total leaf number were 4.07±0.19, 3.92±0.06, 4.19±0.24, and 1.72±0.63 for the control, 0.1% nano-TiO_2, 1% nano- TiO_2, and 2.5% nano- TiO_2 groups, respectively (Figure 3-1C). Although TiO_2 nanoparticle exposure significantly reduced the total number of seedling leaves at higher concentrations (2.5%; $p<0.001$), no significant differences were observed between the untreated control and low nano-TiO_2 treatments ($p=0.536$ for 0.1% and $p=0.626$ for 1%).

TiO_2 nanoparticle exposure significantly reduced plant growth and biomass even at very low concentrations (Figures 3-1D and 3-2; $p<0.001$). For the untreated control, the average biomass was 146±49 mg. Nano-TiO_2 significantly inhibited tobacco plant growth and plant biomass was reduced to 81±21 mg after exposure to 0.1% nano- TiO_2 ($p<0.001$). However, there was no significant difference in plant biomass between the 0.1% and 1 % TiO_2 nanoparticle treatments ($p=0.431$). When 2.5% nano-TiO_2 was added into the medium, almost all seedlings died within one week of germination and minimal plant biomass was harvested (Figure 3-1D).

TiO_2 nanoparticles affect miRNA expression in tobacco

TiO_2 nanoparticle exposure significantly affected miRNA expression (Figure 3-3). Among the 11 miRNAs investigated in this study, the miRNA expression levels were up-regulated at both concentrations (0.1% and 1.0%) after nano-TiO_2 exposure for all miRNAs except miR-156 at high concentration (1.0%). According to their response to nano-TiO_2 exposure, these 11 miRNA could be classified into 3 groups: 1) miRNAs whose expression levels were increased as TiO_2 nanoparticle concentrations increased; the majority of tested miRNAs belonged to this group, which included miR169, miR393, miR395, miR399, miR172, and miR396. 2) miRNAs whose expression levels decreased as TiO_2 nanoparticle concentrations increased; there are two miRNAs (miR159 and miR156) that belonged to this group. 3) miRNAs whose expression levels were not significantly changed as TiO_2 nanoparticle concentrations increased; there are three miRNAs (miR162, miR167, and miR398) that belonged to this group.

Among the 11 tested miRNAs, miR395 and miR399 were the two most sensitive to nano-TiO_2 exposure exhibiting the greatest fold changes to 285 fold and 143 fold, respectively. This suggests that these two miRNAs play an important function in tobacco seedlings facing nanoparticle or metal stress. It was also observed that these two miRNAs were up-regulated under different environmental abiotic stresses, such as drought and salinity stress in other plant species (Jones-Rhoades and Bartel 2004; Zhao et al. 2007). Six miRNAs (miR169, miR159, miR393, miR398, miR396 and miR172) are moderately sensitive to nano-TiO_2 exposure, exhibiting at least a 10 fold change after 3 weeks of growth on the medium containing 0.1% or 1.0% of TiO_2. However, miR156 and miR162 were less sensitive to nano-TiO_2 exposure as they exhibited only less than a 2 fold change after 3 weeks of treatment. According to previous studies, miR156 regulates plant leaf development (Schwab et al. 2005; Zhang et al. 2006) and no study has reported that miR156 plays a function in stress response. In this study, we observed

that miR156 expression is up-regulated by 0.1% nano-TiO_2 but inhibited by 1.0% nano-TiO_2. One potential reason for this change is that nano-TiO_2 inhibited tobacco seedling development and therefore the decrease in expression was due to a decrease in overall plant growth (Figures 1and 2).

TiO_2 nanoparticles affect the expression of stress-related genes in tobacco

In this study, we also observed that TiO_2 nanoparticle treatment affected the expression levels of stress-related genes. Both APX and ADH were significantly up-regulated by both 0.1% and 1% nano-TiO_2 concentrations after 3 weeks of exposure. ADH and APX showed relatively the same fold increase in expression in 0.1% TiO_2 nanoparticle-exposed seedlings with ADH increasing 3.9 fold and APX increasing 3.6 fold. At a higher concentration (1.0%), TiO_2 nanoparticles increased the expression of ADH to 5.73 fold while the expression of APX was kept around a 3.5 fold change. This suggests that nanoparticle treatment also induces abnormal expression of stress-related genes and plants may have similar mechanisms to handle both nanoparticle toxicity as well as other abiotic stresses.

Discussion:

Nanoparticles are used in a variety of applications (paints, papers, manufactured products, etc.) and are interesting to study because their large surface area to volume ratio often times leads to an increase in reactivity (Gao and Zhang 2001). As nanotechnologies become increasingly used, nanoparticles will continue to be an environmental concern and the phytotoxicity of these nanoparticles on terrestrial plants must be elucidated. In this study, we analyzed the effect of TiO_2 nanoparticles on the growth and development of tobacco seedlings. We found that as TiO_2 nanoparticle concentrations increased, the germination rates, leaf counts,

root lengths, and overall biomasses of the tobacco seedlings decreased. We also found that TiO_2 nanoparticles influenced miRNA expression levels in tobacco.

The results of our study correspond to the results of other studies in which germination rates of ryegrass and corn were inhibited by Zn and ZnO nanoparticles (Lin and Xing 2007) and black cumin and wheat germination rates were inhibited by Zn^{2+} (El-Ghamery et al. 2003). Our results are slightly different from those found by Zheng and colleagues (2005) in which they found that low concentrations of nano-sized TiO_2 enhanced the germination rates of naturally aged spinach seeds (Zheng et al. 2005). While TiO_2 nanoparticles did not enhance the germination rate of the tobacco seeds, we did not observe a linear decrease in germination rate either and germination rates rather decreased as TiO_2 nanoparticle concentrations reached 2.5% and above. One potential reason is that Zheng and colleagues only used TiO_2 to treat spinach seeds while we added TiO_2 into the medium for longer exposure time. The seed coat has been shown to play a vital role in protecting the embryo from dangerous environmental factors (Wierzbicka and Obidzinksa 1998). Lin and Xing (2007) offered the explanation that even though seeds are exposed to nanoparticles, the nanoparticles cannot invade the seed coat and effect germination (Lin and Xing 2007). Therefore it is only after the seed begins to germinate that it comes into contact with the nanoparticles. In this study, we noticed a corresponding decrease in overall tobacco growth with an increase in TiO_2 nanoparticle concentration, suggesting that the nanoparticles have a negative impact on tobacco growth after seed germination has begun.

In this study, we also observed a decrease in average root length of tobacco seedlings as TiO_2 nanoparticle concentrations increased. Seedlings grown in 0.1% nano-TiO_2 exhibited a decrease in average root growth by 48% as compared to the control. An even greater inhibition

of root growth was observed in 1% nano-TiO_2 tobacco seedlings as the average root growth was reduced by 85%. At 2.5% nano-TiO_2 no root growth was observed. Therefore, we were able to show that TiO_2 nanoparticles have a negative impact on the development of tobacco roots.

The phytotoxicity of other metal oxide nanoparticles on the root growth of a variety of plant species has also been evaluated. Results have shown that nano-Al_2O_3 negatively impacts root development in corn (Lin and Xing 2007) and that ZnO nanoparticles damage epidermal and cortical cells in the roots of ryegrass (Lin and Xing 2008) leading to an overall decrease in plant growth. The mechanism by which TiO_2 nanoparticles affect root growth in tobacco still needs to be investigated; however, a recent study by Asli and Neumann (2009) showed that TiO_2 nanoparticles affected the water transport ability of maize roots. The authors were able to show that TiO_2 nanoparticles formed aggregates along the root cell wall that blocked the absorption of water, ultimately leading to reduced shoot and leaf development (Asli and Neumann 2009). Therefore, TiO_2 nanoparticles might inhibit tobacco root development by interfering with the pore on the root cell wall and reducing water uptake.

It should be noted that tobacco seedlings grown on 1% nano- TiO_2 produced more than one root. If TiO_2 nanoparticles do adhere to the root cell walls of tobacco seedlings, the growth of more than one root would be necessary in order to increase water uptake and promote plant development.

Overall biomass of tobacco seedlings decreased as TiO_2 nanoparticle concentration increased. A 47% decrease in biomass, as compared to the control, was observed for 0.1% nano-TiO_2 seedlings. Interestingly, seedlings grown on 1% nano-TiO_2 showed only a 37% reduction in overall biomass. This result is surprising since a decrease in root development and an overall decline in plant health was noticed in these plants. The increase in biomass could be attributed to

an intake of TiO_2 nanoparticles. A recent study has shown that carbon-coated iron nanoparticles can be transported through the xylem in pumpkin plants (Corredor et al. 2009) and alternate studies have demonstrated that copper (Lee et al. 2008) along with Zn and ZnO (Lin and Xing 2008) nanoparticles can also be taken in by plant roots. It is possible that, over the 3 week course of this experiment, TiO_2 nanoparticles entered into the tobacco seedlings through pores in the root cell wall. This would cause the slight increase in biomass that was observed for tobacco seedlings exposed to 1% nano-TiO_2.

As the TiO_2 nanoparticle concentration increased, a visible reduction in plant health occurred. Tobacco seedlings exposed to increasing levels of nano-TiO_2 exhibited yellowing, wilting, reduced leaf sizes and leaf counts, reduced root growth, and a decrease in shoots. In the case of tobacco seeds grown on 2.5% nano-TiO_2, no roots were formed. The results from our study do not agree with those by Zheng et al (2005) in which they concluded that nano-TiO_2 improved the growth of spinach plants (Zheng et al. 2005). Our results, however, do agree with others and appear to be symptoms of heavy metal toxicity (El-Ghamery et al. 2003; Lin and Xing 2008) and indicate cell death (Abraham 1997). Since TiO_2 nanoparticles have been shown to induce the formation of reactive oxygen species leading to oxidative DNA damage and lipid peroxidation (Falck et al. 2009; Federici et al. 2007; Trouiller et al. 2009), a future study is needed to determine the mechanisms of TiO_2 nanoparticle toxicity on tobacco. Alternate studies, as suggested by Murashov (2006), are also needed to determine if disassociation of the titanium ion is a cause of phytotoxicity rather than the TiO_2 nanoparticles themselves (Murashov 2006).

miRNAs are a newly discovered class of small regulatory RNAs (Bartel 2004) that have been shown to play a role in plant response to many environmental stresses such as drought (Zhao et al. 2007), salinity (Jia et al. 2009), and cold (Zhang et al. 2009). miRNAs have also

been identified in plant response to heavy metals. For instance, Sunkar, Kapoor, and Zhu (2006) have shown that miR398 targets two related superoxide dismustase genes in *Arabidopsis* and by down-regulating this miRNA, the plant has a higher tolerance to copper and zinc metals (Sunkar et al. 2006). Other investigations have analyzed the effect of cadmium, aluminum, and mercury on miRNA expression levels in *Medicago truncatula* and rice and have found that miRNA expression levels are altered after exposure to heavy metals (Huang et al. 2009; Zhou et al. 2008b).

In this study, we analyzed changes in expression levels of 11 miRNAs in response to TiO_2 nanoparticles. We found that all 11 miRNAs were up-regulated and, in comparison to the control plants, miR399b and miR395a in tobacco seedlings exposed to 1% nano-TiO_2 exhibited the greatest fold changes of 143 and 285, respectively. The expression levels of two stress-related genes, ADH and APX, were also up-regulated after exposure to TiO_2 nanoparticles. Since miRNAs do play a significant role in negatively regulating gene expression, we can assume that the up-regulation of these miRNAs is affecting the expression levels of many important developmental genes. This would explain, in part, the observed decrease in overall plant growth.

Conclusion:

In conclusion, TiO_2 nanoparticles had a negative impact on the growth and development of 3 week old tobacco seedlings. An overall reduction in germination rates, leaf counts, root lengths, and biomasses were observed in seedlings exposed to 0.1%, 1%, 2.5%, and 5% nano-TiO_2. This result could be due to the interference of TiO_2 clusters with the root cell wall, ultimately inhibiting water uptake, or the induction of reactive oxygen species leading to oxidative DNA damage and lipid peroxidation of the cell membrane in tobacco cells. More

research needs to be performed in order to elucidate the mechanisms of TiO_2 nanoparticle toxicity and to determine if phytotoxicity is caused by dissociation of the titanium ion. As the manufacturing of TiO_2 continues to increase, TiO_2 nanoparticles will continue to be released into the environment. Studies have shown that TiO_2 nanoparticles can reach depths of between 40 cm to 370 cm in soil suspensions (Fang et al. 2009). Therefore, the phytotoxicity of TiO_2 on important crop plants, such as tobacco, must be investigated.

microRNA expression levels in tobacco were also altered after exposure to nano-TiO_2. In particular, miR399 and miR395 exhibited the greatest fold increases of 143 and 285, respectively. Our study analyzed changes in expression levels of 11 conserved miRNAs. In order to identify novel tobacco miRNAs that may play a role in plant tolerance to heavy metal stress, other experiments such as high through-put deep sequencing or direct cloning will be needed. We found that TiO_2 nanoparticles affected miRNA expression levels in tobacco and since they function as important gene regulators, miRNAs may play a vital role in tobacco tolerance to heavy metal stress

References

Abdel-Ghany SE, Pilon M (2008) MicroRNA-mediated Systemic Down-regulation of Copper Protein Expression Response to Low Copper Availability in *Arabidopsis*. J Biol Chem 283: 15932-15945

Abraham S (1997) Studies on cytological changes induced by muriate of potatsh in *Allium cepa*. Cytologia 62: 291-294

Achard P, Herr A, Baulcombe DC, Harberd NP (2004) Modulation of floral development by a gibberellin-regulated microRNA. Development 131: 3357-3365

Adams LK, Lyon DY, Alvarez PJJ (2006) Comparative eco-toxicity of nanoscale TiO_2, SiO_2, and ZnO water suspensions. Water Res 40: 3527-3532

Allen E, Xie ZX, Gustafson AM, Carrington JC (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell 121: 207-221

Asli S, Neumann PM (2009) Colloidal suspensions of clay or titanium dioxide nanoparticles can inhibit leaf growth and transpiration via physical effects on root water transport. Plant, Cell Environ 32: 577-584

Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes. Plant Cell 15: 2730-2741

Baan R, Straif K, Grosse Y, Secretan B, Ghissassi FE, Cogliano V (2006) Carcinogenicity of carbon black, titanium dioxide, and talc. Lancet Oncology 7: 295-296

Bao N, Lye KW, Barton MK (2004) MicroRNA binding sites in *Arabidopsis* class IIIHD-ZIP mRNAs are required for methylation of the template chromosome. Developmental Cell 7: 653-662

Bartel DP (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 116: 281-297

Boffetta P, Soutar A, Cherrie JW, Granath F, Andersen A, Anttila A, Blettner M, Gaborieau V, Klug SJ, Langard S, Luce D, Merletti F, Miller B, Mirabelli D, Pukkala E, Adami H-O, Weiderpass E (2004) Mortality among workers employed in the titanium dioxide production industry in Europe. Cancer Causes Control 15: 697-706

Borm PJA, Schins RPF, Albrecht C (2004) Inhaled particles and lung cancer, part B: Paradigms and risk assessment. International J Cancer 110: 3-14

Bowman JL (2004) Class III HD-Zip gene regulation, the golden fleece of *ARGONAUTE* activity? Bioessays 26: 938-942

Chen CZ, Li L, Lodish HF, Bartel DP (2004) MicroRNAs modulate hematopoietic lineage differentiation. Science 303: 83-86

Corredor E, Testillano PS, Coronado M-J, Gonzalez-Melendi P, Fernandez-Pacheco R, Marquina C, Ibarra MR, de la Fuente JM, Rubiales D, Perez-de-Luque A, Risueno M-C (2009) Nanoparticle penetration and transport in living pumpkin plants: in situ subcellular identification BMC Plant Biol 9

Ding D, Zhang LF, Wang H, Liu ZJ, Zhang ZX, Zheng YL (2009) Differential expression of miRNAs in response to salt stress in maize roots. Annals Botany 103: 29-38

El-Ghamery AA, El-Kholy MA, El-Yousser MA (2003) Evaluation of cytological effects of Zn^{2+} in relation to germination and root growth of *Nigella sativa* L. and *Triticum aestivum* L. . Mutation Res 537: 29-41

Emery JF, Floyd SK, Alvarez J, Eshed Y, Hawker NP, Izhaki A, Baum SF, Bowman JL (2003) Radial patterning of *Arabidopsis* shoots by class IIIHD-ZIP and KANADI genes. Curr Biol 13: 1768-1774

Falck G, Lindberg H, Suhonen S, Vippola M, Vanhala E, Catalan J, Savolainen K, Norppa H (2009) Genotoxic effects of nanosized and fine TiO_2. Hum Experimental Toxicol 28: 339-352

Fang J, Shan X-q, Wen B, Lin J-m, Owens G (2009) Stability of titania nanoparticles in soil suspensions and transport in saturated homogenous soil columns. Environ Pollut 157: 1101-1109

Federici G, Shaw BJ, Handy RD (2007) Toxicity of titanium dioxide nanoparticles to rainbow trout (*Oncorhyncus mykiss*): Gill injury, oxidative stress, and other physiological effects. Aquatic Toxicol 84: 415-430

Fryzek JP, Chadda B, Marano D, White K, Schweitzer S, McLaughlin JK, Blot WJ (2003) A Cohort Mortality Study among Titanium Dioxide Manufacturing Workers in the United States. J Occupational Environ Medicine 45: 400-409

Fujii H, Chiou TJ, Lin SI, Aung K, Zhu JK (2005) A miRNA involved in phosphate-starvation response in *Arabidopsis*. Curr Biol 15: 2038-2043

Gao L, Zhang Q (2001) Effects of amorphous contents and particle size on the photocatalytic properties of TiO_2 nanoparticles. Scripta Mater 44: 1195-1198

Guo HS, Xie Q, Fei JF, Chua NH (2005) MicroRNA directs mRNA cleavage of the transcription factor *NAC1* to downregulate auxin signals for *Arabidopsis* lateral root development. Plant Cell 17: 1376-1386

Gurr J-R, Wang ASS, Chen C-H, Jan K-Y (2005) Ultrafine titanium dioxide particles in the absence of photoactivation can induce oxidative damage to human bronchial epithelial cells. Toxicol 213: 66-73

Hartmann NB, Von der Kammer F, Hofmann T, Baalousha M, Ottofuelling S, Baun A (2009) Algal testing of titanium dioxide nanoparticles- Testing considerations, inhibitory effects and modification of cadmium bioavailability. Toxicol 269: 190-197

Heinrich U, Fuhst R, Rittinghausen S, Creutzenberg O, Bellmann B, Koch W, Levsen K (1995) Chronic Inhalation Exposure of Wistar Rats and two Different Strains of Mice to Diesel Engine Exhaust, Carbon Black, and Titanium Dioxide. Inhalation Toxicology 7: 533-556

Higarashi MM, Jardim WF (2002) Remediation of pesticide contaminated soil using TiO_2 mediated by solar light. Catalysis Today 76: 201-207

Hsieh L-C, Lin S-I, Shih AC-C, Chen J-W, Lin W-Y, Tseng C-Y, Li W-H, Chiou T-J (2009) Unconvering Small RNA-Mediated Responses to Phosphate Deficiency in Arabidopsis by Deep Sequencing. Plant Physiology 151: 2120-2132

Hsu J-P, Chang Y-T (2000) An experimental study of the stability of TiO_2 partciles in organic water mixtures. Coll Surfaces A: Physiochem Engineering Aspects 161: 423-437

Huang SQ, Peng J, Qiu CX, Yang ZM (2009) Heavy metal-regulated new microRNAs from rice. J Inorg Biochem 103: 282-287

Hund-Rinke K, Simon M (2006) Ecotoxic Effect of Photocatalytic Active Nanoparticles (TiO_2) on Algae and Daphnids. Environ Science Poll Res 13: 225-232

Hur J-S, Oh S-O, Lim K-M, Jung JS, Kim J-W, Koh YJ (2005) Novel effects of TiO_2 photocatalytic ozonation on control of postharvest fungal spoilage of kiwifruit. Postharvest Biol Tech 35: 109-113

Jia XY, Wang W, Ren L, Chen Q, Mendu V, Willcut B, Dinkins R, Tang X, Tang G (2009) Differential and dynamic regulation of *miR398* in response to ABA and salt stress in *Populus tremula* and *Arabidopsis thaliana*. Plant Mol Biol 71: 51-59

Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell 14: 787-799

Juarez MT, Kui JS, Thomas J, Heller BA, Timmermans MCP (2004) microRNA-mediated repression of *rolled leaf1* specifies maize leaf polarity. Nature 428: 84-88

Kaegi R, Ulrich A, Sinnet B, Vonbank R, Wichser A, Zuleeg S, Simmler H, Brunner S, Vonmont H, Burkhardt M, Boller M (2008) Synthetic TiO_2 nanoparticle emission from exterior facades into the aquatic environment. Environ Poll 156: 233-239

Kang SJ, Kim BM, Lee YJ, Chung HW (2008) Titanium dioxide nanoparticles trigger p53-mediated damage response in peripheral blood lymphocytes. Environ Mol Mutagenesis 49: 399-405

Kim VN (2005) MicroRNA biogenesis: coordinated cropping and dicing. Nature Rev Mol Cell Biol 6: 376-385

Kiser MA, Westerhoff P, Benn T, Wang Y, Perez-Rivera J, Hristovski K (2009) Titanium Nanomaterial Removal and Release from Wastewater Treatment Plants. Environ Sci Tech 43: 6757-6763

Kubo M, Onodera R, Shibasaki-Kitakawa N, Tsumoto K, Yonemoto T (2005) Kinetics of ultrasonic disinfection of *Escherichia coli* in the presence of titanium dioxide particles Biotech progress 21: 897-901

Laufs P, Peaucelle A, Morin H, Traas J (2004) MicroRNA regulation of the *CUC* genes is required for boundary size control in *Arabidopsis* meristems. Development 131: 4311-4322

Lee W-M, An Y-J, Yoon H, Kweon H-S (2008) Toxicity and Bioavailability of Copper Nanoparticles to the Terrestrial Plants Mung Bean (*Phaseolus radiatus*) and Wheat (*Triticum aestivum*): Plant Agar Test for Water-Insoluable Nanoparticles. Environ Toxicol Chem 27: 1915-1921

Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120: 15-20

Lin D, Xing B (2007) Phytotoxicity of nanoparticles: Inhibition of seed germination and root growth. Environ Poll 150: 243-240

Lin D, Xing B (2008) Root Uptake and Phytotoxicity of ZnO Nanoparticles. Environ Sci Tech 42: 5580-5585

Lovern SB, Klaper R (2006) Daphnia magna mortality when exposed to titanium dioxide and fullerene (C_{60}) nanoparticles Environ Toxicol Chem 25: 1132-1137

Lovern SB, Strickler JR, Klaper R (2007) Behavioral and physiological changes in Daphnia magna when exposed to nanoparticle suspensions (titanium dioxide, nano-C_{60}, and $C_{60}HxC_{70}Hx$). Environ Sci Tech 41: 4465-4470

Lu J-w, Li F-b, Guo T, Lin L-w, Hou M-f, Liu T-x (2006) TiO_2 photocatalytic antifungal technique for crops diseases control. J Environ Sci 18: 397-401

Lu SF, Sun YH, Shi R, Clark C, Li LG, Chiang VL (2005) Novel and mechanical stress-responsive microRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. Plant Cell 17: 2186-2203

Lubick N (2009) Hunting for engineered nanomaterials in the environment. Environ Sci Tech: 6446-6447

Mallory AC, Dugas DV, Bartel DP, Bartel B (2004a) MicroRNA regulation of *NAC*-domain targets is required for proper formation and separation of adjacent embryonic, vegetative, and floral organs. Curr Biol 14: 1035-1046

Mallory AC, Reinhart BJ, Jones-Rhoades MW, Tang GL, Zamore PD, Barton MK, Bartel DP (2004b) MicroRNA control of *PHABULOSA* in leaf development: importance of pairing to the microRNA 5 ' region. EMBO J 23: 3356-3364

Maneerat C, Hayata Y (2006) Antifungal activity of TiO_2 photocatalysis against *Penicillium expansum* in vitro and in fruit tests. International J Food Microbiol 107: 99-103

Mattigod SV, Fryxell GE, Alford K, Gilmore T, Parker K, Serne J, Engelhard M (2005) Functionalized TiO_2 Nanoparticles for Use for in Situ Anion Immobilization Environ Sci Tech 39: 7306-7310

Minhua X, Jianmin M, Jianhua G, Zuhong L (1998) Photocatalytic TiO_2 nanoparticles damage to cellular membranes and genetic supramolecules. Supramol Sci 5: 511-513

Mlotshwa S, Yang ZY, Kim YJ, Chen XM (2006) Floral patterning defects induced by Arabidopsis *APETALA2* and *microRNA172* expression in *Nicotiana benthamiana*. Plant Mol Biol 61: 781-793

Moore MN (2006) Do nanoparticles present ecotoxicological risks for the health of the aquatic environment? Environ International 32: 967-976

Murashov V (2006) Comments on "Particle surface characteristics may play an important role in phytotoxicity of alumina nanoparticles" by Yang, L., Watts, D.J., Toxicology Letters, 2005, 158, 122-132. Toxicol Lett 164: 185-187

Nagaveni K, Sivalingam G, Hegde MS, Madras G (2004) Photocatalytic degradation of organic compounds over combustion-synthesized nano-TiO_2. EvironSci Tech 38: 1600-1604

Palatnik JF, Allen E, Wu XL, Schommer C, Schwab R, Carrington JC, Weigel D (2003) Control of leaf morphogenesis by microRNAs. Nature 425: 257-263

Pan XP, Redding JE, Wiley PA, Wen L, McConnell JS, Zhang BH (2010) Mutagenicity evaluation of metal oxide nanoparticles by the bacterial reverse mutation assay Chemosphere 79: 113-116

Pott F, Roller M (2005) Carcinogenicity study with nineteen granular dusts in rats. Eur J Oncology 10: 249-282

Quan X, Zhao X, Chen S, Zhao H, Chen J, Zhao Y (2005) Enhancement of *p,p'*-DDT photodegradation on soil surfaces using TiO_2 induced by UV-light. Chemosphere 60: 266-273

Reeves JF, Davies SJ, Dodd NJF, Jha AN (2008) Hydroxyl radicals (.OH) are associated with titanium dioxide (TiO_2) nanoparticle-induced cytotoxicity and oxidative DNA damage in fish cells. Mut Res 640: 113-122

Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D (2005) Specific effects of microRNAs on the plant transcriptome. Developmental Cell 8: 517-527

Sichel C, de Cara M, Tello J, Bianco J, Fernandez-Ibanez P (2007) Solar photocatalytic disinfection of agricultural pathogenic fungi: *Fusarium* species. Applied Catalysis B: Environ 74: 152-160

Sunkar R, Kapoor A, Zhu J-K (2006) Posttranscriptional Induction of Two Cu/Zn Superoxide Dismutase Genes in *Arabidopsis* Is Mediated by Downregulation of *miR398* and Important for Oxidative Stress Tolerance Plant Cell 18: 2051-2065

Sunkar R, Zhu JK (2004) Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. Plant Cell 16: 2001-2019

Trindade I, Capitao C, Dalmay T, Fevereiro MP, dos Santos DM (2010) *miR398* and *miR408* are up-regulated in response to water deficit in *Medicago truncatula* Planta 231:705-716

Trouiller B, Reliene R, Westbrook A, Solaimani P, Schiestl RH (2009) Titanium Dioxide Nanoparticles Induce DNA Damage and Genetic Instability *In vivo* in Mice. Cancer Res 69: 8784-8789

Vaucheret H (2006) Post-transcriptional small RNA pathways in plants: mechanisms and regulations. Genes Development 20: 759-771

Vileno B, Lekka M, Sienkiewicz A, Jeney S, Stoessel G, Lekki J, Forro L, Stachura Z (20007) Stiffness Alterations of Single Cells Induced by UV in the Presence of NanoTiO_2. Environ Sci Tech 41: 5149-5153

Wang J, Zhang X, Chen Y, Sommerfeld M, Hu Q (2008) Toxicity assessment of manufactured nanomaterials using the unicellular green alga *Chlamydomonas reinhardtii*. Chemosphere 73: 1121-1128

Wang JJ, Sanderson BJ, Wang H (2007) Cyto- and genotoxicity of ultrafine TiO_2 particles in cultured human lymphoblastoid cells. Mut Res 628: 99-106

Wang S, Chang L-Y, Wang Y-J, Wang Q, Yang C-H, Mei R-H (2009) Nanoparticles affect the survival of bacteria on leaf surfaces. FEMS Microbiol Ecology 68: 182-191

Wierzbicka M, Obidzinksa J (1998) The effect of lead on seed imbibition and germination in different plant species. Plant Sci 137: 155-171

Xie XH, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES, Kellis M (2005) Systematic discovery of regulatory motifs in human promoters and 3 ' UTRs by comparison of several mammals. Nature 434: 338-345

Yao KS, Wang DY, Ho WY, Yan JJ, Tzeng KC (2007) Photocatalytic bactericidal effect of TiO_2 thin film on plant pathogens. Surface Coatings Tech 201: 6886-6888

Zhang BH, Pan XP, Cobb GP, Anderson TA (2006) Plant microRNA: A small regulatory molecule with big impact. Dev Biol 289: 3-16

Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA (2005) Identification and characterization of new plant microRNAs using EST analysis. Cell Res 15: 336-360

Zhang J, Xu Y, Huan Q, Chong K (2009) Deep sequencing of Brachypodium small RNAs at the global genome level identifies microRNAs involved in cold stress response. BMC Genomics 10: 449

Zhang X, Sun H, Zhang Z, Niu Q, Chen Y, Crittenden JC (2007) Enhanced bioaccumulation of cadmium in carp in the presence of titanium dioxide nanoparticles Chemosphere 67: 160-166

Zhao BT, Liang RQ, Ge LF, Li W, Xiao HS, Lin HX, Ruan KC, Jin YX (2007) Identification of drought-induced microRNAs in rice. Biochem Biophys Res Comm 354: 585-590

Zheng L, Hong F, Lu S, Liu C (2005) Effect of Nano-TiO2 on Strength of Naturally Aged Seeds and Growth of Spinach. Biol Trace Elem Res 104: 83-91

Zhong RQ, Ye ZH (2004) Amphivasal vascular bundle 1, a gain-of-function mutation of the *IFL1/REV* gene, is associated with alterations in the polarity of leaves, stems and carpels. Plant Cell Phys 45: 369-385

Zhou XF, Wang GD, Sutoh K, Zhu JK, Zhang WX (2008a) Identification of cold-inducible microRNAs in plants by transcriptome analysis. Biochim Biophys Acta-Gene Reg Mech 1779: 780-788

Zhou ZS, Huang SQ, Yang ZM (2008b) Bioinformatic identification and expression analysis of new microRNAs from *Medicago truncatula*. Biochem Biophyl Res Comm 374: 538-542

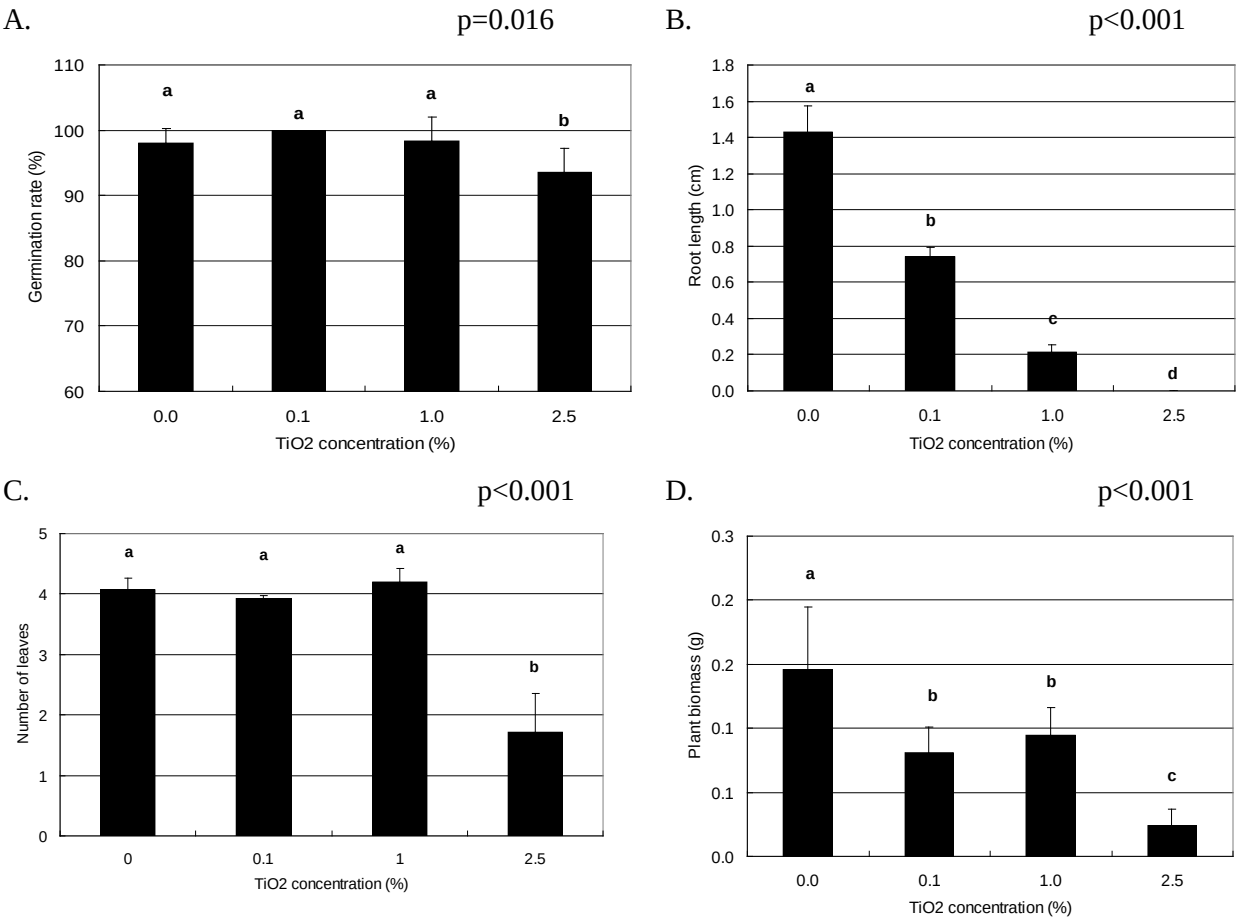


Figure 3-1. Effect of nanoparticle TiO_2 on tobacco seed germination, growth, and development. A) Germination rate; B) Root length; C) Total number of leaves; and D) Plant biomass.

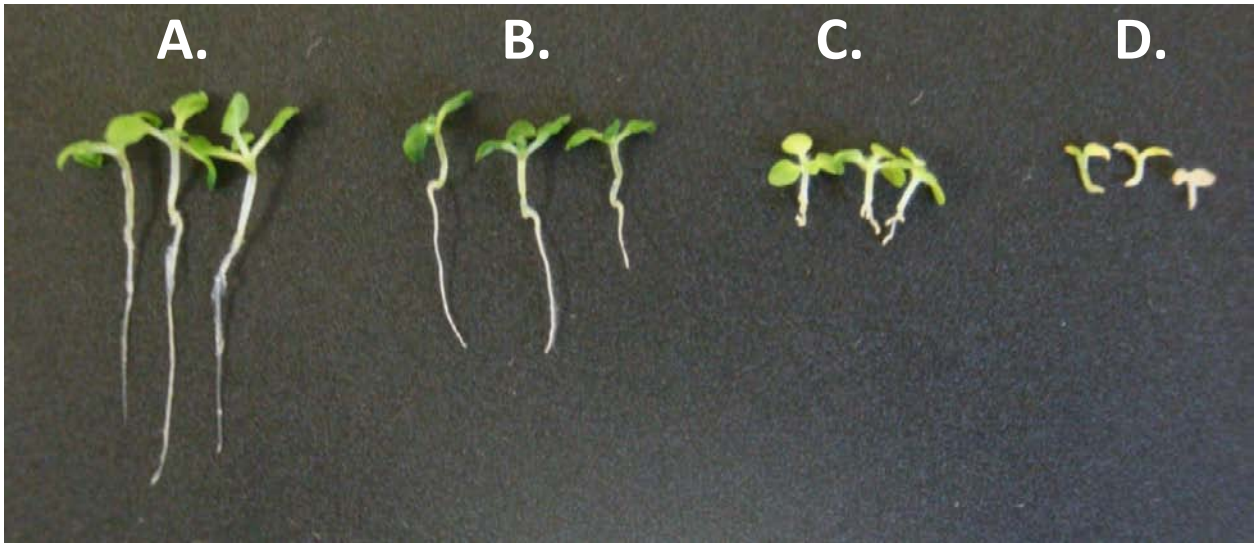
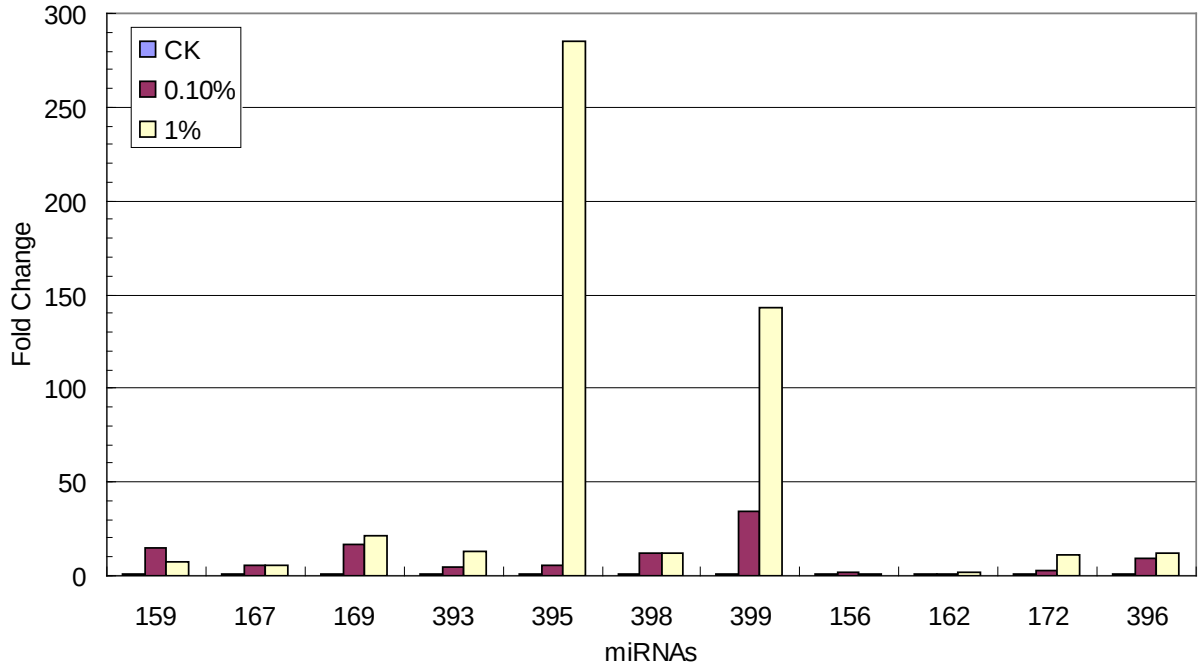


Figure 3-2. Effect of nano-TiO_2 on tobacco growth and development after 3 weeks of treatment. A) MS media (control). B) MS +0.1% nano- TiO2. C) MS+1% nano- TiO2. D) MS+2.5% nano-TiO2. E) MS+5% nano- TiO2.

A.



B.

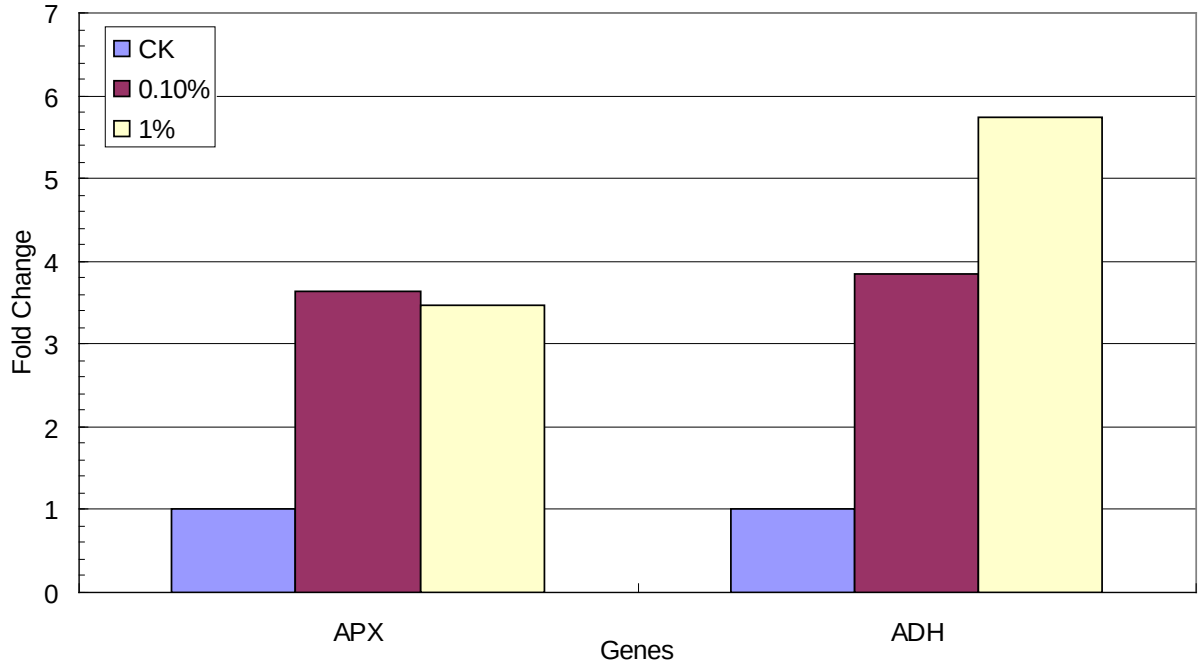


Figure 3-3. Effect of nanoparticle TiO2 on the expression of microRNAs (A) and two stress-related genes (B).

CHAPTER4: Salt and Drought Stresses Induce the Aberrant Expression of microRNA Genes in Tobacco

Abstract:

Drought and salinity stresses are two major environmental abiotic stresses that significantly affect plant growth and development. Although many studies have shown that abiotic stresses induce the aberrant expression of protein-coding genes, the regulatory mechanisms regulating these genes remain unclear. MicroRNAs (miRNAs) are a newly discovered extensive class of endogenous small RNAs that post-transcriptionally regulate the expression of many protein-coding genes including those that control plant growth, development, and responses to environmental stresses. In this study, we employed tobacco as a model system to investigate the effect of drought and salinity stresses on this important gene regulator. Our results show that both salinity and drought stresses significantly altered miRNA expression profiles in a dose-dependent manner in young tobacco seedlings. Salinity stress changed the miRNA expression levels from a 6.86 fold down-regulation to a 616.57 fold up-regulation. Alternatively, miRNAs were down-regulated by 2.68 fold and up-regulated 2810 fold under drought conditions. miR395 was the most sensitive to both drought and salinity conditions and was up-regulated by 616 and 2810 folds by 1.00% PEG and 1.00% NaCl, respectively. miRNAs are the most sensitive to salinity treatment at the lowest concentration (0.10%) while the up-regulated miRNAs are the most sensitive at the highest concentration (1.00%). However, miRNAs are more sensitive to lower concentrations of PEG rather than higher levels of drought treatment. Salinity and drought stresses also changed the expression profiles of ADH and APX. The results of this study suggest that miRNAs may play an important role in plant response to environmental abiotic stresses. Further investigation of miRNA-mediated gene regulation may help elucidate the molecular mechanism of plant tolerance to abiotic stresses and has the

potential to create a miRNA-based biotechnology for improving plant tolerance to drought and salinity stresses.

Introduction:

Unlike animals, plants are stationary organisms that have evolved mechanisms to cope with a wide range of environmental and climate changes (Sunkar 2010). Over the past century, global warming has led to a rise in seawater levels (Meehl et al. 2005) and a slow but gradual increase in the surface temperature of the Earth (Abu-Asab et al. 2004; Peng et al. 2004). This has caused previously wet regions to become more arid and the deposition of salt into low-lying grass and farm lands (Houghton 2005). Aside from global warming, rain-fed plants can often times experience dry conditions and can also be naturally exposed to high concentrations of salt in soil (Sunkar 2010). Drought and salt stresses are two of the more severe and wide-ranging environmental stresses that significantly affect crop growth and productivity. Although much research has been dedicated to elucidating gene expression during plant exposure to dry and brackish conditions, the mechanisms underlying the regulation of gene expression remain largely unknown.

miRNAs are short sequences (~21nt) of endogenous non-coding RNA that negatively regulate gene expression at the post-transcriptional level (Bartel 2004). miRNAs have been shown to play an important role in a variety of plant biological and metabolic processes including organ maturation (Aukerman and Sakai 2003; Guo et al. 2005; Juarez et al. 2004; McHale and Koning 2004), hormone signaling (Liu et al. 2009b; Mallory et al. 2005), developmental timing (Achard et al. 2004; Poethig et al. 2004), response to pathogens (Hewezi et al. 2008; Navarro et al. 2006; Sullivan and Ganem 2005), and response to environmental

abiotic stresses such as drought (Zhao et al. 2007), salinity (Zhao et al. 2009), heavy metals (Huang et al. 2009), and cold (Zhou et al. 2008).

Recent studies have shown that drought and salinity stresses are able to induce the differential expression of thousands of protein-coding genes (Aprile et al. 2009; Matsui et al. 2008). However, the regulatory mechanisms underlying gene expression in response to drought and salt stresses are poorly understood. miRNAs, an important class of gene regulators, have been implicated to play an important role in plant tolerance to abiotic stresses (Jones-Rhoades and Bartel 2004; Sunkar and Zhu 2004). The expression of miR393, for example, has been shown to be influenced by abiotic stress conditions (Sunkar and Zhu 2004; Zhao et al. 2007) and miR393 itself has been shown to target stress-related genes in *Arabidopsis* and rice (Gao et al. 2010b). miR169, miR395, and miR398 expression have also been shown to be induced under other environmental stress conditions such as high salt (Zhao et al. 2009), sulfate starvation (Jones-Rhoades and Bartel 2004), and heavy metal toxicity (Sunkar et al. 2006), respectively.

Tobacco is an important agricultural and economic crop in the United States, as well as around the world. It is cultivated by more than 100 countries around the globe. One new study has investigated the use of tobacco as a potential biofuel crop (Andrianov et al. 2010). In our previous study, we identified 259 tobacco miRNAs using a well-established *in silico* approach. In this experiment, we analyzed the effect of sodium chloride (NaCl) and polyethyleneglycol (PEG) on the expression levels of 9 different miRNAs as well as 2 stress related genes in tobacco. NaCl was used to simulate salt stress whereas PEG was used to simulate drought stress. These 9 miRNAs were selected based on previous studies and all of them are related to plant development and stress response. The results of this study show that NaCl and PEG have an effect on miRNA expression and on the expression of stress-related genes in tobacco. One

miRNA, miR395, was significantly up-regulated after exposure to high NaCl and high PEG conditions. Given the results of this study, we believe miRNAs may play an important role in tobacco tolerance to salt and drought stresses.

Materials and methods:

Seed sterilization, media preparation, and tobacco treatment

Tobacco (*Nicotiana tabacum*) seeds were sterilized by soaking in 70% ethanol for 2 minutes followed by 10% bleach for 15 minutes. The seeds were then rinsed with distilled water (approximately 4 times) until no bleach odor remained. Basic medium used in these experiments contained the following: Murashige and Skoog (MS) salts supplemented with Gamborg's B5 Vitamins, 1% sucrose, and 0.8% agar, and was adjusted to a pH of 5.8 after the addition of sodium chloride (NaCl) (99.85% purity, Acros Organics, Geel, Belgium) and polyethyleneglycol (PEG) 6000 (Hampton Research, Aliso Viejo, CA). Sterilized tobacco seeds were sown on Petri dishes containing MS medium and 0%, 0.1%, 0.25%, 0.5%, and 1% NaCl to simulate varying degrees of salt stress. Sterilized tobacco seeds were also sown on Petri dishes containing MS medium with 0%, 1%, 2.5%, 5%, and 7.5% PEG to simulate varying degrees of drought stress. Approximately 25 tobacco seeds were placed on each plate and there were a total number of 5 plates for each concentration of treatment (NaCl and PEG), including control plates. The plates were placed under a 16 hour day/8 hour night cycle for 3 weeks until seedling removal.

Total RNA Extraction

Three week old tobacco seedlings were removed from the plates and immediately frozen in liquid nitrogen. The seedling tissue was placed at -80°C until RNA extraction. Total RNA was isolated from the 3 week old seedlings using the mirVana miRNA Isolation Kit (Ambion, Austin, TX) according to the manufacturer's protocol. Total RNA was then quantified and assessed for

quality using a Nanodrop ND-1000 (Nanodrop Technologies, Wilmington, DE). RNA samples were stored at -80°C until further analysis.

Analyzing microRNA expression changes using RT-PCR and qRT-PCR

Applied Biosystems TaqMan microRNA Assays were employed to detect and quantify tobacco miRNAs using stem loop real-time PCR according to the manufacturer's instructions. There were two steps in the TaqMan miRNA Assays: a) reverse transcription of the mature miRNA to a longer single-stranded cDNA sequence using a miRNA-specific stem-looped primer and, b) quantitative real-time PCR. Briefly, a single-stranded miRNA cDNA was generated from 1 µg of the total RNA from 3 biological replicates collected from three individual NaCl plates (0%, 0.1%, 0.25%, 0.5%, and 1%) as well as from 3 biological replicates collected from three individual PEG plates (0%, 1%, 2.5%, 5%, and 7.5%). This was completed by reverse transcription using the Applied Biosystems TaqMan microRNA Reverse Transcription Kit and miRNA-specific stem-looped RT primers provided in the kit. In total, we investigated changes in the expression levels of 9 different miRNAs (miR159a, miR167a, miR169a, miR172a, miR393a, miR395a, miR396a, miR398a, and miR399b) and two stress related genes (Alcohol Dehydrogenase (ADH) and Alcohol Peroxidase (APX)). Elongation Factor 1 α (EF1α) was used as a reference gene to normalize expression values. All biological samples (3) were repeated in duplicate and the results were analyzed using the $\Delta\Delta C_T$ method.

Results:

NaCl and PEG effected tobacco growth

Tobacco seeds were planted on Petri dishes containing MS medium supplemented with 0%, 0.1%, 0.25%, 0.5%, and 1% NaCl. Seeds were allowed to grow for 3 weeks in each medium and were then removed and stored at -80°C until total RNA extraction. At the time of tissue

removal, we observed a gradual decrease in tobacco growth as the concentration of salt in the media increased (data not shown). We also noted that the roots of the tobacco plants exhibited a different growth pattern as the salt concentration increased. The roots grew in a more twisted and crooked pattern under higher salt concentrations as compared to the straight, uniform roots of the control plants, which were not exposed to NaCl. A decrease in root health and rigidity was observed for the plants exposed to salt as these were thinner and more prone to break when removed from the media than the control plants. The number and size of tobacco leaves per plant also gradually decreased as the salt concentration increased.

Tobacco seeds were also placed on Petri dishes containing MS medium supplemented with 0%, 1%, 2.5%, 5%, and 7.5% PEG. After 3 weeks, the plants showed consistent growth across the control and PEG concentration plates. Tobacco plants grown in higher concentrations of PEG exhibited longer root lengths compared to those grown on the control plates. Also, the majority of plants grown in 7.5% PEG grew more than one root. In contrast to the plants grown in NaCl, tobacco plants grown in PEG had the same number and size of leaves across all concentrations, including the control plates.

Salinity stress alters miRNA expression levels in tobacco

Salinity treatment significantly altered miRNA gene expression in a dosage-dependent manner (Figure 4-1a). At the tested concentration range, three miRNAs (miR159, miR167, and miR169) were down regulated by salinity stress and the expression inhibition was decreased as the salinity concentration increased. In contrast, salinity stress induced the over expression of two miRNAs (miR 172 and miR 396) and the fold changes of these miRNAs increased as the salinity concentration increased. At the lowest concentration (0.1%), salinity stress inhibited the expression of miR395 and miR399; however, both miRNAs were over expressed in a dose-

response manner under the treatment of 0.25-1.0% salinity. It was observed that miR398 was not as sensitive as other miRNAs, however, it was down-regulated by salinity treatment at all tested concentrations. The largest fold change for miR398 was observed at 0.50% salt treatment with a 2.5 fold change.

Among the 9 tested miRNAs, all miRNAs, except miR398, exhibited more than a 3 fold change under certain NaCl treatments. Of these miRNAs, miR395 was the most sensitive to salinity stress and was down-regulated 3.36 fold at 0.1% salinity treatment while it was up-regulated by 616.37 fold at 1.00% NaCl treatment. At low concentrations, miR159 is the most sensitive to salinity stress with a down-regulation of 6.86 fold.

One more interesting phenomenon we observed was that the down-regulated miRNAs were the most sensitive to salinity treatment at the lowest concentration (0.10%) while the up-related miRNAs were the most sensitive at the highest concentration (1.00%). As we know, miRNAs negatively regulate gene expression which targets specific biological function. Our data may suggest that these miRNAs have synergistic activities during salinity stress.

Drought stress alters miRNA expression levels in tobacco

We also analyzed the effect of PEG, a chemical that simulates drought conditions when added to plant growth media, on miRNA expression levels in tobacco (Figure 4-2a). Surprisingly, all 9 miRNAs were significantly up-regulated after exposure to 1% PEG. Similar to miRNA expression changes after exposure to 1% NaCl, miR395 was the most up-regulated with a change in expression of greater than 2800 fold at 1% PEG. miR169 showed the second greatest change in expression level of 122 fold at 1% PEG.

miRNA expression changes showed 1 of 2 trends as the PEG concentration in the media increased: 1) a decrease in expression followed by a sudden increase at 7.5% PEG or 2) a

decrease in expression over the course of all PEG concentrations. It should be noted, however, that even with a decrease in expression, all of the miRNAs were still up-regulated in comparison to the controls. miRNAs that fell into the first category included miR167, miR393, miR399, miR172, and miR396. For example, miR393 was up-regulated approximately 30 fold after exposure to 1% PEG. The expression level of miR393 decreased to a 1.9 fold up-regulation at 2.5% PEG and decreased again to a 1.8 fold up-regulation at 5% PEG. When exposed to 7.5% PEG, however, miR393 increased in expression to a 3.5 fold up-regulation. miRNAs that fell into the second category included miR159, miR169, and miR395. miR395 was up-regulated 2810.5 fold after exposure to 1% PEG. The expression of miR395 decreased to a 230 fold up-regulation at 2.5% PEG and continued to decrease to a 45 fold up-regulation at 5% PEG. Of the 4 drought treatments, miR395 was the least expressed at 7.5% PEG with an up-regulation of 28.5 fold. Interestingly, miR398 alternated between varying degrees of gene expression. After exposure to 1% and 5% PEG, miR398 was more up-regulated (39 fold and 23 fold, respectively) as compared to when exposed to 2.5% and 7.5% PEG (3 fold and 13 fold, respectively).

NaCl and PEG affect the expression of stress-related genes in tobacco

We also investigated changes in the expression levels of two stress-inducible genes, ADH and APX, in tobacco after exposure to NaCl and PEG (Figure 4-1b and Figure 4-2b). We found that both of these genes were up-regulated under all 4 salt concentrations and that the expression of the genes was consistently the same (approximately a 1 fold increase) as the salt concentrations increased. We also found that both of these genes were up-regulated after exposure to 7.5% PEG. Interestingly, APX was down-regulated after exposure to 1%, 2.5%, and 5% PEG and was only up-regulated after exposure to 7.5% PEG, the highest concentration tested. ADH and APX exhibited different expression patterns under salinity and drought

treatments, which suggest that tobacco may have different mechanisms to handle drought and salinity stresses.

Discussion:

Recent studies have shown that the expression of miRNAs, an important class of gene regulators, is altered after abiotic stress treatment (Ding et al. 2009; Huang et al. 2009; Jia et al. 2009; Lv et al. 2010). However, most of these studies have been performed in model organisms such as *Arabidopsis* and rice. In this report, we investigated changes in miRNA expression levels after exposure to salt and drought stress in tobacco, an important agricultural and economic crop.

Using RT-PCR and qRT-PCR, we analyzed changes in 9 different miRNAs in 3 week old tobacco seedlings after exposure to salt and drought conditions. We found that miR395 was significantly up-regulated after exposure to 1% NaCl. We also found miR395 to be up-regulated after exposure to 7.5% PEG. Interestingly, miR395 has only been shown to function in plant response to sulfate deprivation by targeting sulfur transporter genes (Jones-Rhoades and Bartel 2004; Kawashima et al. 2009; Lu and Huang 2008). The results of our study suggest an alternative role for miR395 in response to high salinity and drought stresses.

miR399, a miRNA involved in regulating phosphate homeostasis in *Arabidopsis* (Chiou et al. 2006; Pant et al. 2008), was up-regulated after exposure to both NaCl and PEG. In contrast to our results, Fujii et al (2005) found that there was no significant up or down-regulation of miR399 after exposure to salt and drought stresses (Fujii et al. 2005). However, our results show that miR399 was up-regulated 6-fold and 13-fold after exposure to 1% NaCl and 7.5% PEG, respectively. The development of stem-loop RT-PCR and TaqMan qRT-PCR analysis has provided a reliable and sensitive method to determine miRNA expression in plants (Chen et al. 2005). Therefore, small changes in the expression level of miR399 can be detected using this

method. Since miR399 is only induced under stress conditions (Sunkar and Zhu 2007), we believe that miR399 may have other unconventional roles and play a part in tobacco tolerance to salt and drought conditions.

Two miRNAs, miR396 and miR172, were up-regulated in tobacco seedlings after exposure to 1% NaCl and 1% PEG. miR396 has been shown to function in leaf development (Liu et al. 2009a) and expression of miR396 has been shown to be induced under high salt, cold, and drought stresses (Liu et al. 2008). Interestingly, over-expression of miR396 in tobacco leads to an increased tolerance to drought stress (Feng-Xi and Di-Qiu 2009). miR396 expression is also up-regulated in rice after exposure to high salinity and transgenic over-expression of this miRNA in rice led to plants with reduced salt tolerance (Gao et al. 2010a). Therefore, miR396 may not only play an important role in tobacco development but may also function in tobacco tolerance to environmental stress. miR172 has been shown to play a role in the phase change between vegetative and reproductive growth and contributes to floral organ identity (Lauter et al. 2005; Mlotshwa et al. 2006). A recent study has suggested a role for miR172 in plant resistance to cold stress (Liu et al. 2008). The results of our study suggest a novel function for miR172 in regulating tobacco tolerance to salt and drought conditions.

miR169, a miRNA known to be induced under high salinity (Zhao et al. 2009), was found to be down-regulated in tobacco after exposure to increasing concentrations of NaCl. It is possible that miR169 is only induced under extreme conditions greater than 1% salt. Surprisingly however, this miRNA was significantly up-regulated in tobacco seedlings after exposure to 1% PEG. The expression levels of this miRNA, however, significantly decreased as the concentration of PEG increased. These results are consistent with those of others that show

miR169a and miR169c expression is down-regulated after exposure to extremely dry conditions (Li et al. 2008).

Another miRNA, miR159, was shown to be highly induced after exposure to 1% PEG and then decreasingly less up-regulated in tobacco under higher PEG concentrations. miR159 has been shown to target MYB101 and MYB33 transcripts, two factors that positively regulate the accumulate of the plant hormone ABA (Reyes and Chua 2007). Consistent with our findings, this miRNA has also been shown to be up-regulated under drought stress (Reyes and Chua 2007) and has been implicated to provide plant tolerance to environmental stress by functioning through hormone and abiotic stress signaling networks (Phillips et al. 2007).

miR393 and miR398 are two additional miRNAs that have been shown to be differentially expressed under abiotic stress conditions. For example, miR393 expression levels are altered under high salinity and cold (Liu et al. 2008) as well as under drought conditions (Zhao et al. 2007). We found that miR393 was up-regulated after exposure to 1% NaCl and after exposure to all concentrations of PEG. miR393 is speculated to cease plant growth and development during times of environmental stress by targeting TIR1, a positive regulator of plant growth (Shukla et al. 2008). miR398 has been found to be up-regulated in response to copper metal toxicity (Yamasaki et al. 2007). miR398 targets superoxide dismustases, genes that scavenge free radicals, and has been shown to be down-regulated during times of oxidative stress (Shukla et al. 2008; Sunkar et al. 2006). miR398 was down-regulated in tobacco seedlings exposed to all concentrations of NaCl, suggesting that the salt might have induced stress by creating an oxidative environment inside the tobacco cells. Interestingly, miR398 was up-regulated in tobacco seedlings after exposure to all concentrations of PEG. This result is

consistent with the results of Trindade et al. in which they found miR398 to be differentially expressed in water deficit *Medicago truncatula* plants (Trindade et al. 2010).

APX and ADH are two genes whose expression levels have been shown to be up-regulated under environmental stress conditions. In this study, we analyzed the effect of NaCl and PEG on APX and ADH expression. We found that both genes were up-regulated after exposure to 1% NaCl and 7.5% PEG. These findings indicate that these environments caused changes in the levels of gene expression as well as changes in miRNA expression.

Conclusion:

Global warming and nutrient depletion of soils due to over-farming has led to a world-wide reduction in growth and productivity of several important crops such as soybean, maize, and wheat. Tobacco, an important agricultural and economic crop, has recently been investigated as a potential biofuel crop. In this paper, we analyzed the expression levels of 9 different miRNAs in tobacco seedlings exposed to 0.1- 1% NaCl as well as 1- 7.5% PEG. We used NaCl to simulate abiotic salt stress and PEG to simulate abiotic drought stress. We found that individual miRNA expression profiles varied between the two different stresses, indicating that salt and drought stresses induce differential miRNA expression through different mechanisms, such as oxidative stress or inhibition of plant growth. We also found that salt and drought conditions induced the expression of APX and ADH, two stress-related plant genes, in tobacco. Therefore, we believe that miRNAs may play a key role in developing tobacco plants with a greater tolerance to salt and drought stress.

References

Abu-Asab MS, Peterson PM, Shetler SG, Orli SS (2004) Earlier plant flowering in spring as a response to global warming in the Washington, DC, area. Biodivers Conserv 10: 597-612

Achard P, Herr A, Baulcombe DC, Harberd NP (2004) Modulation of floral development by a gibberellin-regulated microRNA. Development 131: 3357-3365

Andrianov V, Borisjuk N, Pogrebnyak N, Brinker A, Dixon J, Spitsin S, Flynn J, Matyszczuk P, Andryszak K, Laurelli M, Golovkin M, Koprowski H (2010) Tobacco as a production platform for biofuel: overexpression of *Arabidopsis* DGAT and LEC2 genes increases accumulation and shifts the composition of lipids in green biomass. Plant Biotech J 8: 277-287

Aprile A, Mastrangelo AM, De Leonardis AM, Galiba G, Roncaglia E, Ferrari F, De Bellis L, Turchi L, Giuliana G, Cattivelli L (2009) Trancriptional profiling in response to terminal drought stress reveals differential responses along the wheat genome. BMC Gen 10: 279

Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes. Plant Cell 15: 2730-2741

Bartel DP (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 116: 281-297

Chen CF, Ridzon DA, Broomer AJ, Zhou ZH, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, Lao KQ, Livak KJ, Guegler KJ (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. Nuc Acids Res 33: e179

Chiou TJ, Aung K, Lin SI, Wu CC, Chiang SF, Su CL (2006) Regulation of phosphate homeostasis by microRNA in *Arabidopsis*. Plant Cell 18: 412-421

Ding D, Zhang LF, Wang H, Liu ZJ, Zhang ZX, Zheng YL (2009) Differential expression of miRNAs in response to salt stress in maize roots. Annals Botany 103: 29-38

Feng-Xi Y, Di-Qiu Y (2009) Overexpression of *Arabidopsis mir396* Enhances Drought Tolerance in Transgenic Tobacco Plants. ACTA BOTANICA YUNNANICA 31: 421-426

Fujii H, Chiou TJ, Lin SI, Aung K, Zhu JK (2005) A miRNA involved in phosphate-starvation response in *Arabidopsis*. Curr Biol 15: 2038-2043

Gao P, Bai X, Yang L, Lv D, Li Y, Cai H, Ji W, Guo D, Zhu Y (2010a) Over-expression of *osa-MIR396c* decreases salt and alkali stress tolerance. Planta 231: 991-1001

Gao P, Bai X, Yang L, Lv D, Pan X, Li Y, Cai H, Ji W, Chen Q, Zhu Y (2010b) *osa-MIR393*: a salinity- and alkaline stress-related microRNA gene. Mol Biol Rep: 10.1007/s11033-010-0100-8

Guo HS, Xie Q, Fei JF, Chua NH (2005) MicroRNA directs mRNA cleavage of the transcription factor *NAC1* to downregulate auxin signals for *Arabidopsis* lateral root development. Plant Cell 17: 1376-1386

Hewezi T, Howe P, Maier TR, Baum TJ (2008) *Arabidopsis* small RNAs and their targets during cyst nematode parasitism. Mol Plant-Microbe Interact 21: 1622-1634

Houghton J (2005) Global Warming. Rep Prog Phys 68: 1343-1403

Huang SQ, Peng J, Qiu CX, Yang ZM (2009) Heavy metal-regulated new microRNAs from rice. J Inorg Biochem 103: 282-287

Jia XY, Wang W, Ren L, Chen Q, Mendu V, Willcut B, Dinkins R, Tang X, Tang G (2009) Differential and dynamic regulation of *miR398* in response to ABA and salt stress in *Populus tremula* and *Arabidopsis thaliana*. Plant Mol Biol 71: 51-59

Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Molecular Cell 14: 787-799

Juarez MT, Kui JS, Thomas J, Heller BA, Timmermans MCP (2004) microRNA-mediated repression of *rolled leaf1* specifies maize leaf polarity. Nature 428: 84-88

Kawashima CG, Yoshimoto N, Maruyama-Nakashita A, Tsuchiya YN, Saito K, Takahashi H, Dalmay T (2009) Sulphur starvation induces the expression of *microRNA-395* and one of its target genes but in different cell types. Plant J 57: 313-321

Lauter N, Kampani A, Carlson S, Goebel M, Moose SP (2005) microRNA172 down-regulates *glossy15* to promote vegetative phase change in maize. Proc Nat Acad Sci 102: 9412-9417

Li W-X, Oono Y, Zhu J, He X-J, Wu J-M, Iida K, Lu X-Y, Cui X, Jin H, Zhu J-K (2008) The *Arabidopsis* NFYA5 transcription factor is regulated transcriptionally and posttranscriptionally to promote drought resistance. Plant Cell 20: 2238-2251

Liu D, Song Y, Chen Z, Yu D (2009a) Ectopic expression of *miR396* suppresses *GRF* target gene expression and alters leaf growth in *Arabidopsis*. Phys Plant 136: 223-236

Liu H-H, Tian X, Li Y-J, Wu C-A, Zheng C-C (2008) Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. RNA 14: 836-843

Liu Q, Zhang YC, Wang CY, Luo YC, Huang QJ, Chen SY, Zhou H, Qu LH, Chen YQ (2009b) Expression analysis of phytohormone-regulated microRNAs in rice, implying their regulation roles in plant hormone signaling. Febs Lett 583: 723-728

Lu XY, Huang XL (2008) Plant miRNAs and abiotic stress responses. Biochem Biophys Res Comm 368: 458-462

Lv D-K, Bai X, Li Y, Ding X-D, Ge Y, Cai H, Ji W, Wu N, Zhu Y-M (2010) Profiling of cold-stress-responsive miRNAs in rice by microarrays. Gene 459: 39-47

Mallory AC, Bartel DP, Bartel B (2005) MicroRNA-directed regulation of *Arabidopsis* AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. Plant Cell 17: 1360-1375

Matsui A, Ishida J, Morosawa T, Mochizuki Y, Kaminuma E, Endo T, Okamoto M, Nambara E, Nakajima M, Kawashima M, Satou M, Kim J, Kobayashi N, Toyoda T, Shinozaki K, Seki M (2008) *Arabidopsis* transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. Plant Cell Phys 49: 1135-1149

McHale NA, Koning RE (2004) MicroRNA-directed cleavage of *Nicotiana sylvestris PHAVOLUTA* mRNA regulates the vascular cambium and structure of apical Meristems. Plant Cell 16: 1730-1740

Meehl GA, Washington WM, Collins WD, Arblaster JM, Hu A, Buja LE, Strand WG, Teng H (2005) How much more global warming and sea level rise? Science 307: 1769-1772

Mlotshwa S, Yang ZY, Kim YJ, Chen XM (2006) Floral patterning defects induced by *Arabidopsis APETALA2* and *microRNA172* expression in *Nicotiana benthamiana*. Plant Mol Biol 61: 781-793

Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JDG (2006) A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. Science 312: 436-439

Pant BD, Buhtz A, Kehr J, Scheible WR (2008) MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. Plant J 53: 731-738

Peng S, Huang J, Sheehy JE, Laza RC, Visperas RM, Zhong X, Centeno GS, Khush GS, Cassman KG (2004) Rice yields decline with higher night temperature from global warming. Proc Nat Acad Sci 101: 9971-9975

Phillips JR, Dalmay T, Bartels D (2007) The role of small RNAs in abiotic stress. FEBS Lett 581: 3592-3597

Poethig S, Hunter C, Park MY, Peragine A, Wu G, Yoshikawa M (2004) Regulation of developmental timing in plants by miRNAs. Dev Biol 271: 551-551

Reyes JL, Chua NH (2007) ABA induction of *miR159* controls transcript levels of two *MYB* factors during *Arabidopsis* seed germination. Plant J 49: 592-606

Shukla LI, Chinnusamy V, Sunkar R (2008) The role of microRNAs and other endogenous small RNAs in plant stress responses. Biochim Biophy Acta-Gene Reg Mech 1779: 743-748

Sullivan CS, Ganem D (2005) MicroRNAs and viral infection. Mol Cell 20: 3-7

Sunkar R (2010) MicroRNAs with macro-effects on plant stress responses. Semi Cell Dev Biol In Press: 10.1016/j.semcdb.2010.1004.1001

Sunkar R, Kapoor A, Zhu J-K (2006) Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in *Arabidopsis* is mediated by downregulation of *miR398* and important for oxidative stress tolerance. Plant Cell 18: 2051-2065

Sunkar R, Zhu JK (2004) Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. Plant Cell 16: 2001-2019

Sunkar R, Zhu JK (2007) Micro RNAs and short-interfering RNAs in plants. J Integ Plant Biol 49: 817-826

Trindade I, Capitao C, Dalmay T, Fevereiro MP, dos Santos DM (2010) *miR398* and *miR408* are up-regulated in response to water deficit in *Medicago truncatula* Planta 231:705-716

Yamasaki H, Abdel-Ghany SE, Cohu CM, Kobayashi Y, Shikanai T, Pilon M (2007) Regulation of copper homeostasis by micro-RNA in *Arabidopsis*. J Biol Chem 282: 16369-16378

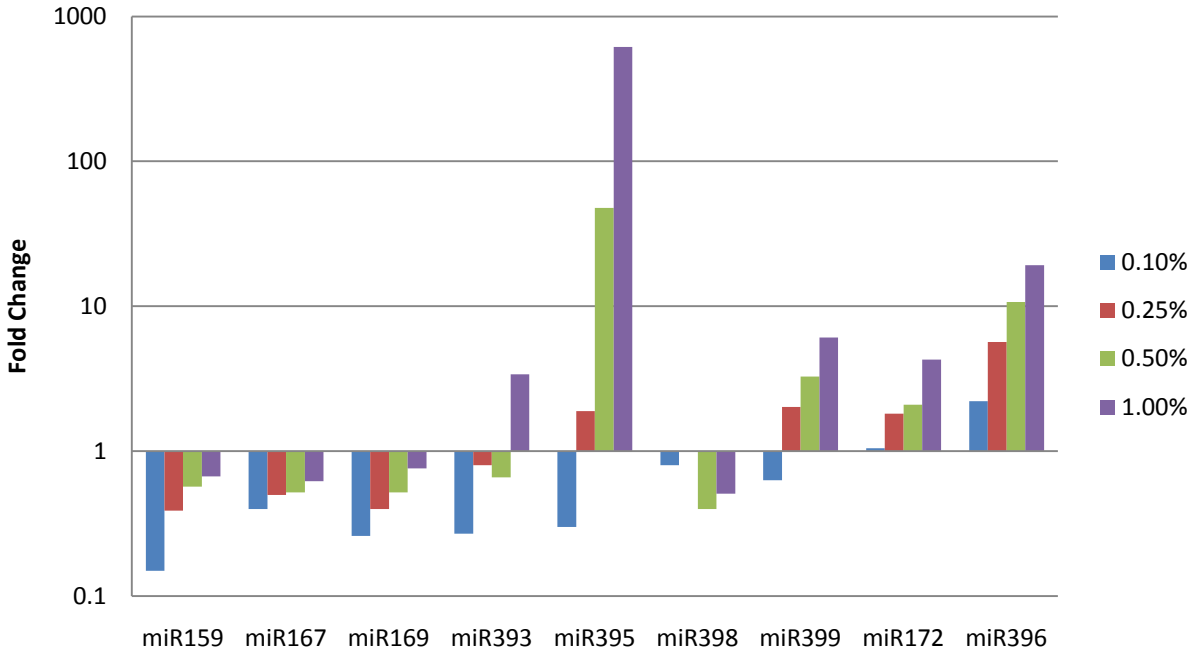
Zhang BH, Anderson TA (2006) MicroRNAs: New players in plant responses to agrochemicals and environmental stress. 232nd Amer Chem Soc (ACS) Nat Meet, San Francisco, CA, USA, Sep 10-14, 2006

Zhao B, L.Ge, Liang R, Li W, Ruan K, Lin H, Jin Y (2009) Members of miR-169 family are induced by high salinity and transiently inhibit the NF-YA transcription factor. BMC Mol Biol 10:29

Zhao BT, Liang RQ, Ge LF, Li W, Xiao HS, Lin HX, Ruan KC, Jin YX (2007) Identification of drought-induced microRNAs in rice. Biochem Biophys Res Comm 354: 585-590

Zhou XF, Wang GD, Sutoh K, Zhu JK, Zhang WX (2008) Identification of cold-inducible microRNAs in plants by transcriptome analysis. Biochim Biophys Acta-Gene Reg Mech 1779: 780-788

a.



b.

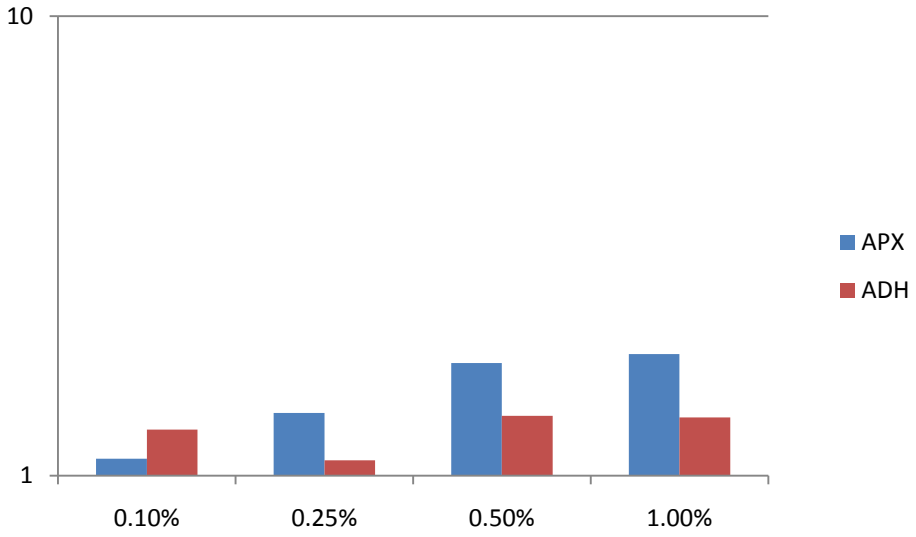
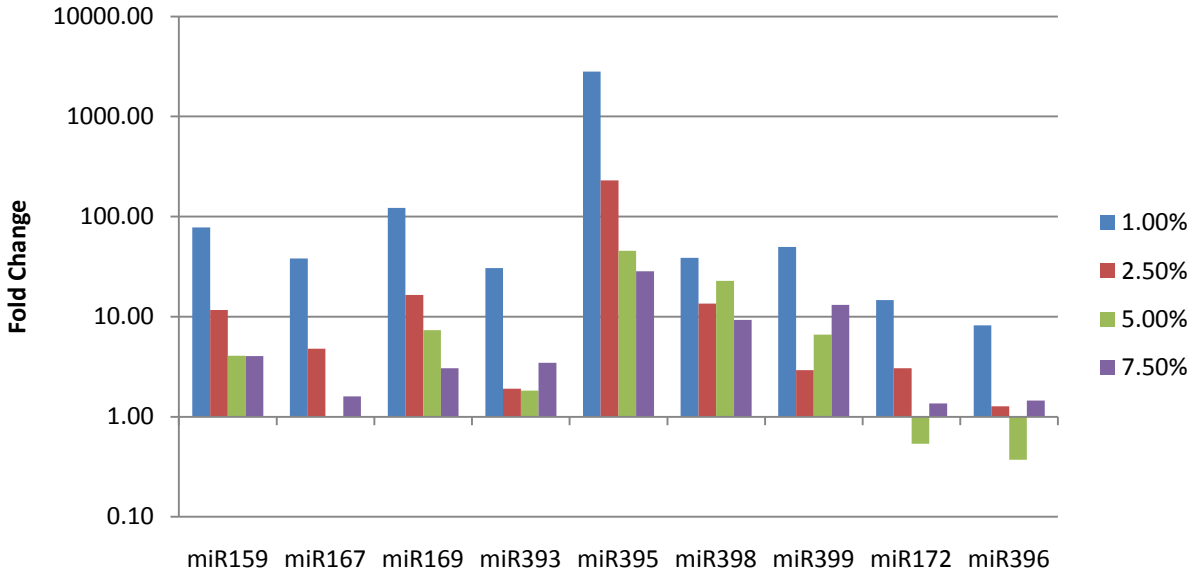


Figure 4-1. NaCl induces gene expression changes in young tobacco seedlings. **a** Changes in 9 miRNA expression levels under 0.1%, 0.25%, 0.5%, and 1% NaCl. **b** Changes in 2 stress-related genes under 0.1%, 0.25%, 0.5%, and 1% NaCl

a.



b.

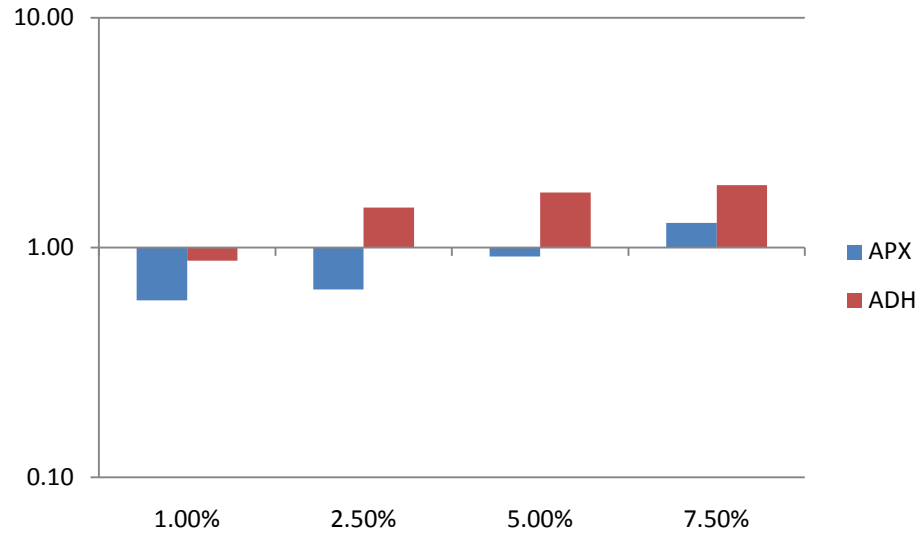


Figure 4-2. PEG induces gene expression changes in young tobacco seedlings. **a** Changes in 9 miRNA expression levels under 1%, 2.5%, 5%, and 7.5% PEG. **b** Changes in 2 stress-related genes under 1%, 2.5%, 5%, and 7.5% PEG