The miR159-*GAMYB* pathway: silencing and function of *GAMYB* homologues amongst diverse plant species.

Zihui Zheng

October 2018

A thesis submitted for the degree of Doctor of Philosophy of

The Australian National University

Statement of Authorship

The research carried out in this thesis was conducted at The Australian National University between March 2014 and May 2018. I declare that I am the sole author of this thesis, and I have fully acknowledged and referenced the ideas and work of others, whether published or unpublished, in my thesis. My thesis does not contain work extracted from a thesis, dissertation or research paper previously presented for another degree or diploma at this or any other university, except for parts of sections 3.2.1, which were from my Masters' thesis and included for completeness of the chapter; and Figure 3.7C, whose image was generated by lab colleague Gigi Wong. Parts of my introduction (Figures 1.1 and corresponding text) have been updated and modified from my Masters' thesis.

Zihui Zheng

zihui.zheng@anu.edu.au

17th October 2018

Word count: ~32,807

Acknowledgements

First and foremost, I want to thank my supervisor Tony Millar for his outstanding supervision. He is a very nice, patient and friendly supervisor. I have followed him from my Master to PhD for six years, and his status was my Master convenor to PhD supervisor. It's my pleasure to work with him and very informative to my whole life. In science, Tony has guided and taught me academic knowledge and critical thinking at the same time. Tony helps me a lot not only in my academic career, but also encouraging me in my difficult time. As PhD is a tough journey, I have felt depressed and had troubles sometimes. Tony has guided me to have a positive attitude and passion on science and life. He has brought out my potentials and led me from a science stranger becoming to professional participator in plant science.

I would also like to thank my co-supervisors Mingbo Wang and Rob Allen, who have always provided academic advice and technical helps. Furthermore, I very appreciate Junyan (Mary) Li, who is my Master co-supervisor and PhD advisor, but also my good friend. She was my first guider in the Millar lab and taught me lots of biological techniques. Besides, I have always shared my happiness and sadness with her and we have enjoyed a good time in my PhD journey. Millar Lab is more likely to be a big family. Therefore, I really thank helps from my lab members, Rob Allen, Mary Li, Renee Li, Marlene Reichel, Ira Deveson, Gigi Wong, Maria Alonso-Peral, Allan Lohe and Meachery Jalajakumari. Here, I am really appreciating Marlene Reichel. She was former PhD student and has most well organisation and critical thinking in science. We had collaboration in the study of RNA secondary structure in miRNA regulation. Thanks a lot for her helps always. Moreover, special thanks go to Meachery Jalajakumari, who I had

collaboration with tobacco experiments. Even though lab members have come and gone, our common experience in the Millar lab would be shared in our memory.

My PhD study has worked on different plant species, Arabidopsis, tobacco and rice. Besides our lab, I have learned biotechnical skills from other labs in our school. I would like to thank Whitney Lab for their helps with my tobacco experiments. Dr. Spencer Whitney has given me professional advice for my tobacco project. Rosemary Birch taught me tobacco techniques, including tobacco transformation. She is one of the best technician I have met and plays very patient to me. Also, I would like to thank helps and advice from other lab members in Whitney Lab, Elena Martin Avila and Brendon Conlan. Referring to my rice part, I really want to thank von Caemmerer Lab, Hugo Alonso provided me theoretical advice and Riya Rebecca Kuruvilla taught me rice techniques. Special thanks go to Riya, who taught me rice transformation and helped with my troubleshooting. They are all excellent scientists. Additionally, I would like to thank Weihua Chen and Deyun Qiu, who are excellent postdoctoral scientists in Masle Lab. We have often exchanged views on the scientific questions and troubleshooting, and I have got many helps from them. Particularly, Deyun helped with my troubleshooting of rice transformation and phenotyping. Generally, I would like to thank all staff in Research School of Biology, ANU College of Science, including plant services, IT and technical officer.

Furthermore, I'm also thankful to Enhui Shen, who was a Chinese visitor in CSIRO in 2017. He helped and taught me to analyse RNA sequencing data. Moreover, special thanks go to the Sumie Ishiguro Lab, Department of Biological Mechanisms and Functions, Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya,

Japan. They provided the Gateway Destination pGWB602Ω vector. And thanks to the Liwang Qi Lab, Laboratory of Cell Biology, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, People's Republic of China. They provided pinetree *LaMYB33* DNA sample.

Lastly, I would like to thank my family for their love and encouragement. I have always loved them. Also, I am really appreciating my good friends in Australia, Mary Li, Yiming Li, Weihua Chen, Deyun Qiu, Yu Zhou and Wiei Yih Hee. Thanks for your company in my PhD journey.

I have been in Australia for six years, including two years of Master and four years of PhD. Even though I had both happiness and depression, this period is still a good time in my life.

Publications

Chapter 3 has been published here:

Zheng, Z., Reichel, M., Deveson, I., Wong, G., Li, J., & Millar, A. A. (2017). Target RNA secondary structure is a major determinant of miR159 efficacy. *Plant physiology*,**174**(3): 1764-1778.

Work of Reichel, M., and Deveson, I. that contributed to the paper has been removed from the Chapter 3.

Abstract

MicroRNAs (miRNAs) are a class of small RNAs that regulate gene expression in eukaryotes. In plants, many miRNA families mediate silencing of target genes, which are involved in plant development and plant defence. For my thesis, I have been investigating the miR159-*GAMYB* pathway, which appears conserved from basal vascular plants to angiosperms. GAMYB transcription factors have been demonstrated to have conserved roles of programmed cell death (PCD) in both the seed aleurone and the anther tapetum in a number of different plant species. However, what the functional role of *GAMYB* is in vegetative tissues remains unknown. In Arabidopsis, miR159 is critical for proper growth, as its inhibition has a strong negative impact on vegetative growth, due to deregulated *GAMYB* expression. However, *gamyb* loss-of-functional mutants display a wild-type phenotype, as their expression is silenced to phenotypically inconsequential levels by miR159 in vegetative tissues. This raises two questions: (1) how is *GAMYB* so strongly silenced; (2) why is *GAMYB* strongly and widely transcribed in vegetative tissues for it to be then completely repressed by miR159? These two questions were the focus of my thesis.

Firstly, how the Arabidopsis *MYB33* and *MYB65* are so strongly silenced in the model plant Arabidopsis was investigated. Both genes were predicted to contain a distinctive RNA secondary structure abutting the miR159 binding site, composed of two stem-loop (SL) structures; whereas such SL structures were not predicted to form in other *GAMYB-like* genes that are targeted less efficiently by miR159 for expression regulation. Functional analysis found that the RNA structure in *MYB33* correlated with strong silencing efficacy; introducing mutations to disrupt either SL attenuated miR159 efficacy, while introducing mutations to form an artificial stem-loop structure adjacent to a miRNA-binding site restored strong miR159-mediated silencing. Although how these

predicted structures promote miR159-mediated silencing are not determined, we speculate that the stem-loop structures in the vicinity of the miR159 binding site promotes accessibility of the binding site, where if adjacent sequences form strong stem structures, they are less likely to base-pair with binding site nucleotides, maintaining high accessibility of the binding site. Interestingly, the RNA SL structures are predicted to reside in *GAMYB-like* homologues of numerous angiosperm and gymnosperm plant species, arguing that these structures have been integral in the miR159-*GAMYB* regulatory relationship over a long period of time. In addition, these structures are not present in the Arabidopsis *GAMYB-like* homologues that are not transcribed in vegetative tissues, suggesting that selection for such structures only occurs for homologues transcribed in vegetative tissues as to prevent their expression and demarcating them as sensitive targets of miR159.

Secondly, to investigate the functional role of the miR159-*GAMYB* pathway, target *MIMIC159* (*MIM159*) transgenes, which can inhibit endogenous miR159 activity, were expressed in a number of Arabidopsis ecotypes, as well as in tobacco and rice. Inhibiting miR159 in all three plant species resulted in similar phenotypic outcomes, predominantly stunted growth and irregular leaf shape. This implies that the function and expression of the miR159-*GAMYB* pathway is strongly conserved in distant plant species. This raises several questions: why is *GAMYB* widely transcribed if its expression is strongly silenced by miR159 throughout the plant to result in little to no phenotypic impact; and why has this been strongly conserved across multiple plant species. When miR159 activity is inhibited in *MIM159* tobacco leaves, pathways related to plant defence response are most up-regulated. This included *PATHOGENESIS-RELATED PROTEIN* (*PR*) mRNA levels that were 100-1000s fold up-regulated compared to wild type, and correlated with deregulated *GAMYB* expression. This finding suggests that the miR159-*GAMYB* pathway

is involved in the plant defence response to biotic stress. However, *PR* expression is not up-regulated in Arabidopsis or rice when miR159 is inhibited, suggesting that despite the conserved nature of the miR159-*GAMYB* pathway, there are species-specific differences in its function.

Abbreviations

5’-RACE: 5’ Rapid Amplification of Complementary DNA Ends
AGO: ARGONAUTE
amiRNA: artificial miRNA
AP2: *APETALA2*
CC-NBS-LRR: coiled-coil/nucleotide binding site/leucine rich repeat
CDS: coding DNA sequences
CP1: CYSTEIN PROTEINASE 1
DMS: dimethyl sulphate
ETI: effector-triggered immunity
GO: Gene Ontology
HR: hypersensitive response
IPS1: INDUCED BY PHOSPHATE STARVATION1
MIM: artificial miRNA target *MIMIC*
MIR: miRNA gene
miRISC: miRNA induced silencing complex
miRNA: microRNA
NBS-LRR: NUCLEOTIDE-BINDING SITE AND LEUCINE-REPEAT
Nt: *Nicotiana tabacum*
nt: nucleotide
PAMP: pathogen-associated molecular pattern
PAR-1c: *PHOTOASSIMILATE-RESPONSIVE PROTEIN 1c*
PCD: programmed cell death
PP2A: PROTEIN PHOSPHATASE 2A
PR: PATHOGENESIS-RELATED gene
pri-miRNA: primary miRNA
pre-miRNA: precursor miRNA
PRR: pattern recognition receptor
PTI: PAMP-triggered immunity
qRT-PCR: quantitative real-time PCR
RBP: RNA binding protein
RNAi: RNA interference
R protein: resistance protein

SAR: systemic acquired resistance

SEM: standard error of the mean

siRNA: small interfering RNA

SL: stem-loop

SP: *SPONGE*

SPL: *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE*

sRNA: small RNA

STTM: short tandem target mimics

ta-siRNA: trans-acting small interfering RNA

TCP: *CINCINNATA-like TEOSINTE BRANCHED1*, *CYCLOIDEA*, and *PCF*

TIR-NBS-LRR: toll-interleukin-1 receptor homology/nucleotide binding/leucine rich repeat

TMV: *Tobacco mosaic virus*

Table of contents

Chapter 1

Introduction

1.1 Plant miRNAs

The discovery of RNA interference (RNAi) mediated by 21-24 nucleotide (nt) small RNAs (sRNAs) has been a landmark discovery regarding our understanding of gene expression regulation and the regulatory potential of the genome. Commonly known as gene silencing in plants, these sRNAs guide effector complexes to target genes via sequence complementarity. The two best studied classes of sRNAs are microRNAs (miRNAs) and small interfering RNAs (siRNAs). Generally, siRNAs operate via a transcriptional silencing mechanism, where they are required for genome stability by targeting transposon and repetitive elements to ensure they remain inactive. By contrast, miRNAs generally operate via a post-transcriptional mechanism, and target protein-coding genes to regulate their expression (reviewed in Bologna and Voinnet, 2014; Borges and Martienssen, 2015; Chen, 2012; Ito, 2012).

With the development of sRNA high-throughput sequencing technologies, a multitude of miRNAs have now been discovered in diverse plant species (reviewed in Budak and Akpinar, 2015; Ma et al., 2015). Many of these miRNAs are highly conserved, and target mRNAs that encode regulatory proteins that play fundamental roles in a variety of biological processes, including development, morphogenesis, hormone signalling, stress response and plant defence response (reviewed in Wu, 2013; Luo et al., 2013; Rubio-Somoza & Weigel, 2011; Sunkar et al., 2012; Weiberg et al., 2014). Thus, miRNA are often considered as master regulators of gene expression in plants.

1.2 miRNA biogenesis in plants

In plants, canonical miRNA biogenesis starts with transcription of a *MICRORNA* (*MIR*) gene by RNA polymerase II (Pol II) into an transcript, which forms an imperfect stem-loop structural RNA, namely a primary miRNA (pri-miRNA) (Xie et al., 2005). The pri-

miRNA is then processed to form a precursor miRNA (pre-miRNA) by a complex formed by the dynamaic interaction of DICER-LIKE1 (DCL1), the RNA-binding protein DAWDLE (DDL), the double-stranded RNA binding protein HYPONASTIC LEAVES1 (HYL1), the zinc finger protein SERRATE (SE) and the G-patch domain protein TOUGH (TGH) in nuclear Dicing bodies (D-body). Subsequently, the pre-miRNA is processed to generate a miRNA:miRNA* duplex, which is then methylated by the methyltransferase HUA ENHANCER1 (HEN1) and exported into the cytoplasm, possibly by the Exportin 5 homolog HASTY (HST). The mature miRNA strand is incorporated into an ARGONAUTE (AGO) protein to form the miRNA induced silencing complex (miRISC) to carry out the silencing of complementary mRNAs. In Arabidopsis, the AGO protein family contains ten members. The selection of AGO binding is dependent on the thermodynamic stability of the 5 end of mature miRNA strand (Figure 1.1, reviewed in Budak and Akpinar, 2015; Chen, 2009; Rogers and Chen, 2013; Wu, 2013).

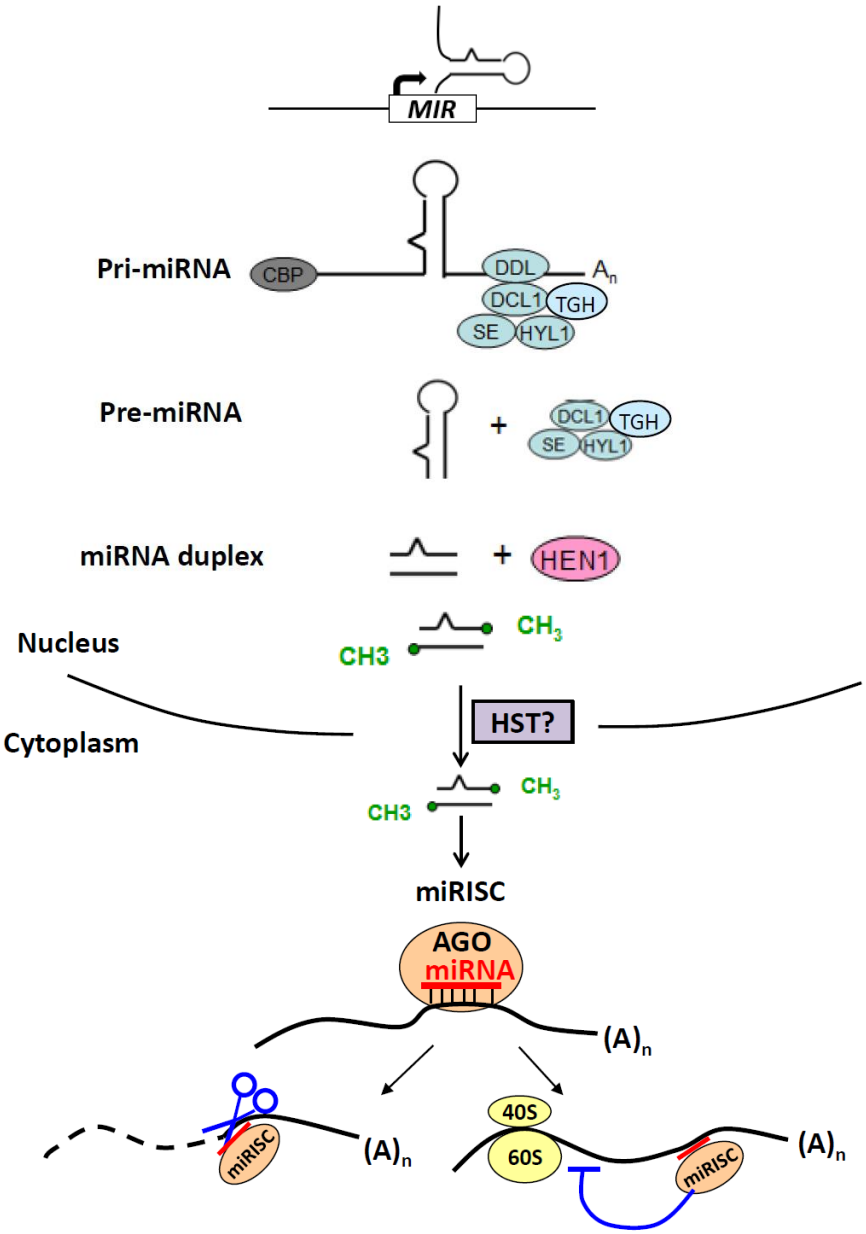


Figure 1.1. Canonical miRNA biogenesis and action in plants. The *MIRNA* gene (*MIR*) is transcribed into a pri-mRNA by RNA polymerase II. The pre-miRNA and miRNA duplex are successionally formed via processing by a complex of DICER-LIKE1 (DCL1), DAWDLE (DDL), HYPONASTIC LEAVES1 (HYL1), SERRATE (SE) and TOUGH (TGH) in the nuclear Dicing bodies (D-body). After methylation by the methyltransferase HUA ENHANCER1 (HEN1), the miRNA duplex is exported into the cytoplasm possibly by the HASTY (HST) and the mature miRNA strand is incorporated into an AGO protein. miRNA associated with an AGO protein is the core component of the miRISC and binds the complementary mRNA target. miRNA-mediated gene silencing occurs via two mechanisms: target mRNA degradation and/or translational repression.

1.3 Mechanisms of miRNA action

In plants, miRNA-mediated silencing operates via two mechanisms: target mRNA degradation and/or translational repression (Figure 1.1). After miRISC formation, the miRNA acts as a sequence-specific guide to direct the miRISC to target mRNAs. It has been widely accepted that highly complementary miRNA-target pairs with perfect central matches facilitate transcript cleavage at the position pairing to 10-11 nucleotides of miRNA (reviewed in Chen, 2009; Huntzinger and Izaurralde, 2011). This cleavage mechanism is mainly dependent on the P-element-induced whimpy testis (PIWI) domain of AGO, with differential RNaseH-like fold and endonuclease activity (Huntzinger and Izaurralde, 2011; Mallory and Vaucheret, 2010). The 3'-end product of these cleaved target mRNAs can be detected by a 5' Rapid Amplification of Complementary DNA Ends (RACE) miRNA cleavage assay, an approach of identifying authentic miRNA targets (Llave et al., 2002). With the development of high-throughput sequencing techniques, degradome sequencing is established to detect and quantitate cleaved miRNA targets via mapping the 5' end of degraded RNA fragments to a global transcriptome (Addo-Quaye et al., 2008; Ma et al., 2015).

In addition to target mRNA degradation, evidence suggests that miRNA-mediated translational repression of target mRNA is present in plants. The miR172 family was originally identified to repress expression of *APETALA2* (*AP2*) and *AP2*-like target genes via a translational repression mechanism. This result was determined by RNA gel blot and immunoblot analyses, which showed minimal changes in *AP2* transcript abundance, but strong repression of AP2 protein levels in the *MIR172* over-expressing plants (Aukerman and Sakai, 2003). Further analysis of the *miRNA action deficient* (*mad*) mutant reveals that suppression of miRNA-mediated silencing of target genes at protein but not mRNA levels is present in many miRNA families, which suggests that miRNA-

mediated translational repression is widespread in plants (Rogers and Chen, 2013; Yang et al., 2012). Furthermore, some studies have described miRNA-mediated translational repression of target genes is dependent on AGO1 and AGO10 (Beauclair *et al*., 2010; Lanet et al., 2009). Although the mechanism behind miRNA-mediated translational repression is not clear in plants, the process of translational repression is able to be reversible and de-repressed in animals (Bhattacharyya et al., 2006; Flynt and Lai, 2008). Therefore, this provides a possibility that plant miRNA-mediated translational repression may be also reversible and more suitable to regulate stress-responsive genes.

In conclusion, miRNA-mediated regulation is likely a combination of mRNA degradation and translational repression in plants. However, the balance between two mechanisms in miRNA regulation and what factors render target mRNAs processed into different mechanisms remain unknown.

1.4 Determinants of miRNA-target recognition

1.4.1 Plant miRNA-mRNA target interaction requires high complementarity

In plants, all known miRNA-target gene pairs have high complementarity (Figure 1.2A). Within the miRNA-target pairing, the positions and numbers of mismatches tolerated for efficient miRNA targeting have been determined empirically: only one mismatch in the so-called "seed region" (nt 2-12 of the 5' end of miRNA), with no mismatch at the cleavage site (nt 10-11); and three additional mismatches (no more than two continuous mismatches) tolerated in the region complementary to the 3' end of miRNA (Schwab *et al*., 2005). These empirical parameters were further confirmed in *Nicotiana benthamiana* quantitative transient assays, which were used to evaluate miRNA silencing efficacies against different targets with different mismatch positions (Liu et al., 2014). Furthermore,

molecular studies have demonstrated that over-expression of several miRNA targets with high complementarity does not cause developmental defects, suggesting that they are efficiently silenced by miRNA (Mallory *et al*., 2005; Palatnik *et al*., 2003; Wang *et al*., 2005; Wu and Poethig, 2006). Therefore, it is widely accepted that the high complementarity in a miRNA-target pair is sufficient to facilitate strong miRNA-mediated silencing of targets in plants. This principle is the basis of many different miRNA-related software, including the bioinformatic programs for prediction of miRNA targets (Dai and Zhao, 2011; Dai et al., 2018; Song et al., 2010; Sun et al., 2014) and the design of artificial miRNAs (amiRNAs; Li et al., 2013a; Schwab et al., 2006).

1.4.2 Are there factors beyond complementarity required for plant miRNA recognition?

A number of studies have found that amiRNAs with high complementarity analogous to endogenous miRNA-target pairs, display variable efficacies while some are not effective as the others (Deveson *et al*., 2013; Li et al., 2013a; Schwab *et al*., 2006). In addition, miR159 with two artificially central mismatches still silences *MYB33* expression (Li et al., 2014b). Therefore, we suspect that high complementarity is critical but not an absolute determinant for a strong miRNA silencing outcome, where factors beyond complementarity regulate miRNA silencing efficacy in plants.

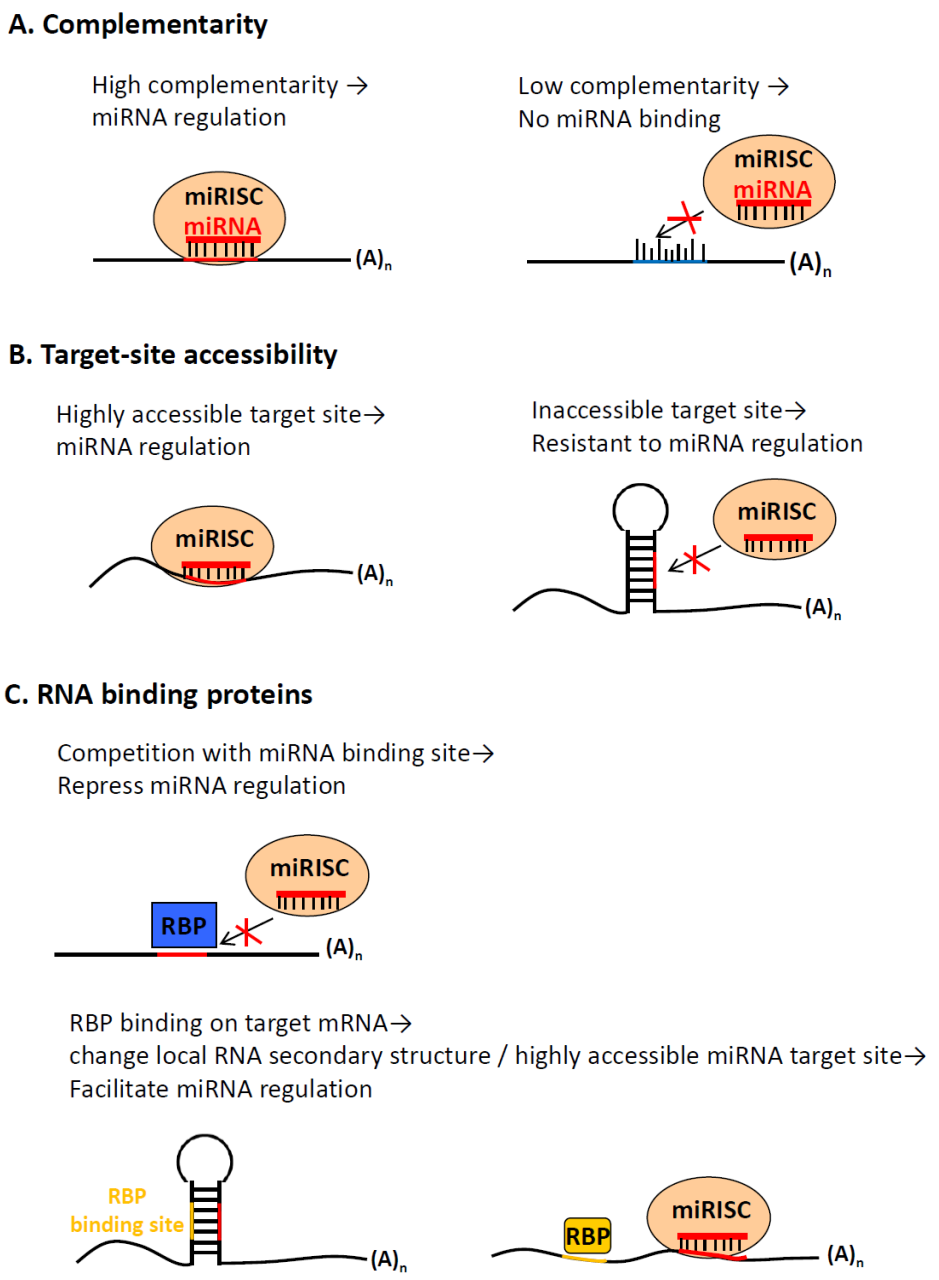


Figure 1.2. Determinants of miRNA-target recognition and interaction. (A) High complementarity between miRNA-target pairs is required for miRNA-mediated regulation; whereas low complementary mRNA is unable to be recognized by miRNA. **(B)** If the target site in an open RNA structure, it is highly accessible to miRNA binding, whereas low target-site accessibility, due to being located in a strong RNA structure, prevents miRNA binding. **(C)** RNA binding proteins (RBPs) may sterically inhibit miRNAs interaction with a binding site, attenuating miRNA-mediated silencing; or the RBP-mRNA interaction may change local RNA secondary structure to generate a highly accessible miRNA target site, facilitating miRNA binding.

1.4.2.1 Target-site accessibility

Such factors beyond complementarity have mainly been studied in animal systems. For example, target-site accessibility can influence miRNA-target recognition and has been incorporated in bioinformatic prediction programs of miRNA targets in animals (Figure 1.2B, Kertesz et al., 2007). It has been experimentally validated by introduction of mutations in flanking regions that decrease target-site accessibility, leading to dramatically reduced miRNA-mediated regulation, with effects as important as sequence complementarity of the binding site (Kertesz et al., 2007). Additionally, tests of stepwise-reduced accessibility in a siRNA target site have been designed via folding the target sequences in different RNA secondary structures, including a totally open structure, partial stem-loop structures and a long stem, which gradually increase stability of RNA secondary structure. A linear correlation was found between increased RNA structure stability and decreased siRNA-mediated cleavage efficiency (Ameres et al., 2007). Furthermore, microarray analysis of 11 mammalian miRNA families reveals highly enriched A and U content in regions flanking their functional miRNA target sites rather than non-functional target sites; and in regions flanking to the conserved miRNA target sites rather than non-conserved target sites. Therefore, AU-rich regions flanking the miRNA binding site could promote high target-site accessibility (Grimson et al., 2007).In addition, analysis of RNA secondary structures of predicted miRNA targets across the human transcriptome found that miRNAs targeting candidates with high target-site accessibility are more likely to be bound by miRISC (Wan et al., 2014). Based on the above findings, they suggest that high accessibility of the miRNA target site is required for efficient miRNA-mediated expression regulation in animals.

Interestingly, target-site accessibility has been incorporated into the plant bioinformatic target prediction program, psRNATarget (Dai and Zhao, 2011), despite there being no

direct evidence showing the significance of this parameter in plant miRNA-target recognition. However, two indirect lines of evidence suggest target-site accessibility does impact plant miRNA-target gene recognition. Firstly, the bioinformatic analysis of synonymous codon usage around predicted miRNA target sites has been performed on four different plant genomes (Gu et al., 2012). This analysis revealed AU-rich synonymous codon usage in the approximate 30-nt regions flanking the miRNA binding sites compared to the whole genome: a feature speculated to increased target-site accessibility of miRNA binding (Gu et al., 2012). Secondly, another RNA secondary structure mapping approach of the Arabidopsis transcriptome indicates that RNA structures in the 21-nt miRNA binding site are not as strong as structures in the 50-nt upstream and downstream flanking regions (Li et al., 2012b). These two previous reports suggest less structural RNAs are present in the miRNA target sites. Although some potential evidences have been given to RNA secondary structure and miRNA target-site accessibility in plants, there still need further work to investigate whether RNA structure can determine plant miRNA regulation.

1.4.2.2 RNA-binding proteins

In animals, RNA binding proteins (RBPs) that bind within the vicinity of the miRNA binding site have been found to modulate miRNA-mediated regulation by two mechanisms (Figure 1.2C). One is an antagonised interplay between the RBP and the miRISC complex. For example, the human dead end 1 (Dnd1) protein can modulate miRNA regulation in wide range of targets in human cells and zebrafish primordial germ cells via binding the 3' UTR of mRNA and preventing association with the miRNA. Such an interaction suppresses miRNA-mediated gene silencing (Kedde *et al*., 2007). On the other hand, RBPs can play collaborative roles, promoting miRNA regulation. A human protein, Pumilio-1 (PUM1) interacts with the 3' UTR of the *p27* mRNA resulting in a

change of the local RNA secondary structure which promotes accessibility of the miRNA target site to miR-221-RISC complex binding (Kedde et al., 2010). In this case, the two factors of RBPs and target-site accessibility/RNA secondary structure work together to mutually control miRNA-mediated regulation. To date, no such mechanisms of RBPs or RNA secondary structure determining miRNA-mediated regulation have been discovered in plants.

Therefore, this thesis in part provides further evidence that sequence complementarity alone is not sufficient to determine strong miRNA regulation and that additional factors impact on miRNA-mediated regulation in plants. Understanding these factors may assist in refinement of the miRNA target prediction and modification of miRNA decoy systems that inhibit endogenous miRNA activity in plants.

1.5 Approaches of investigating miRNA functions in plants

As previously mentioned, miRNAs have been shown to play important roles in many different plant development and growth processes. However, studying the functional role of miRNAs using loss-of-function mutations can be challenging. This is due to the fact that most miRNA families conserved across multiple plant species are composed of multiple functionally redundant members. Hence, for many of these families, performing a traditional T-DNA loss-of-function approach to simultaneously knocking out an entire family is not practical. Some previous studies have used gain-of-function approach to overexpress a particular *MIRNA* gene in transgenic plants and observing the resulting mutant phenotypes (reviewed in Garcia, 2008; Jones-Rhoades et al., 2006). However, as these experiments express miRNAs at levels much higher than that of their endogenous levels, and in tissues where the miRNA is not normally expressed, these mutant phenotypes may represent additional effects rather than authentic endogenous role of the

miRNA (Garcia, 2008). An alternative strategy of investigating miRNA function is to express a miRNA-resistant version of a target gene via the introduction of synonymous mutations in the miRNA binding site, preventing target recognition by miRISC (Garcia, 2008; Jones-Rhoades et al., 2006). Considering that miRNAs often target a gene family with multiple members, determining which of family member(s) is subjected to functionally significant miRNA regulation is difficult. Furthermore, over-expression of the target gene under a constitutive promoter or native promoter could result in transgenic artefacts (Li and Millar, 2013), which cannot accurately describe *in vivo* miRNA target gene function. Therefore, a clear understanding of miRNA function in plants still remains challenging, and establishing an efficient loss-of-function system to investigate plant miRNA function is highly desirable.

In Arabidopsis, a long non-protein-coding RNA gene, *INDUCED BY PHOSPHATE STARVATION1* (*IPS1*), has been found to contain a highly complementary motif to miR399, but with a bulge of three nucleotides at the miRNA cleavage position within the target site. This is thought to inhibit miR399-mediated cleavage of the *IPS1* transcript, which then sequesters the miR399-RISC complex, resulting in de-regulation of *PHO2*, the primary target gene of miR399 (Franco-Zorrilla et al., 2007). Based on this mechanism, target mimicry has been engineered, where the *IPS1* RNA is used as a backbone and the sequence of the miR399 binding site can be altered to target a specific miRNA, thereby acting as a miRNA decoy to inhibit endogenous miRNA activity *in planta* (Figure 1.3B, Franco-Zorrilla et al., 2007). These modified *IPS1* genes were called miRNA target *MIMIC*s (*MIM*s), and were applied to 73 miRNA families or subfamilies in Arabidopsis under the control of the constitutive 35S promoter (Todesco et al., 2010). In this paper, *MIM*s targeting 15 conserved miRNA families resulted in phenotypic defects that were consistent with phenotypes resulting from over-expressing miRNA-

resistant versions of the corresponding miRNA target genes. Consistently, molecular analyses confirmed miRNA levels decreased in these *MIM* plants. Therefore, the *MIM* decoy system in plants serves as an efficient transgenic tool for inhibiting miRNA activity and de-regulating expression of all miRNA targets simultaneously. However, there is still a limitation. Referring to several miRNA families, the *MIM* plants displayed weaker phenotypic defects compared to expression of miRNA-resistant version of target or miRNA loss-of-functional mutant (Todesco et al., 2010).

In addition to *MIM*s, alternative approaches have been established to inhibit endogenous miRNA activity in plants based on a similar mechanism of expressing a miRNA decoy. One strategy is short tandem target mimics (STTMs), which contain two *MIM* sites linked by a short spacer (Figure 1.3C). Further experiments have validated the 48-nt spacer is most optimal for the miR165/166 family, because this length of spacer can fold into a strong stem of RNA secondary structure to stabilize the *STTM* transcript and render two *MIM* sites highly accessible to miR165/166 (Yan et al., 2012). STTM transgenic systems have showed high efficacies of inhibiting miR165/166, miR156/157 and miR160 activity, resulting in severe phenotypic defects in the STTM plants and the decreased levels of miRNAs (Yan et al., 2012). However, STTM159 and STTM396 have showed relatively low efficacies of miR159 and miR396 targeting (Liang et al., 2014; Reichel et al., 2015).

Another strategy is so-called miRNA *SPONGE*s (*SP*s), which are artificial transgenes containing 15 miRNA binding sites with mismatches at the cleavage site (Figure 1.3C, Reichel et al., 2015). Phenotyping results of *SP* transgenic Arabidopsis have exhibited variable efficacies towards different miRNA families: *SP*s strongly inhibit miR159 and miR165/166 activity, low efficacies towards miR164 and miR390 families and even no impact on miR156 and miR172 activity (Reichel et al., 2015).

Although applications of these three miRNA decoys in Arabidopsis have displayed varied efficacies towards different miRNA families, in many cases they have successfully inhibited miRNA activity, overcoming the problem of functional redundancy (Reichel et al., 2015). Therefore, these tools are being widely used to understand the biological and functional significance of miRNA-mediated regulation in plants.

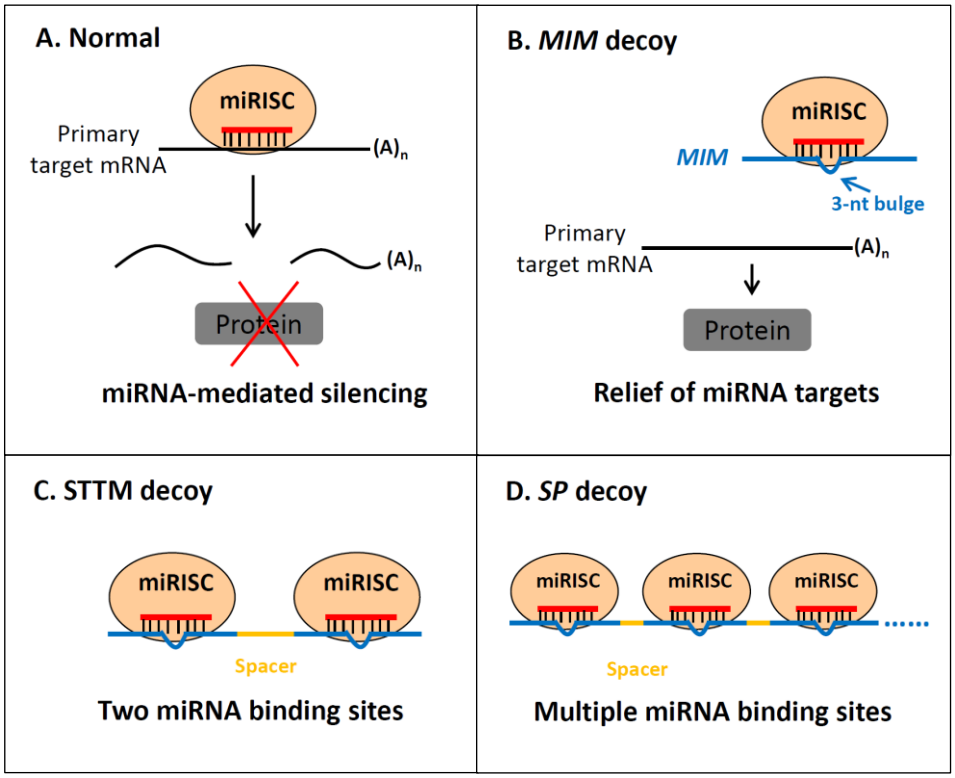


Figure 1.3. miRNA decoys inhibit endogenous miRNA activity in plants. (A) In a wild-type cell, the miRNA mediates silencing of target gene expression. **(B)** In a transgenic cell expressing an artificial miRNA target *MIMIC* (*MIM*), the *MIM* sequesters the miRISC complex as it is unable to cleave the *MIM* decoy due to a three-nt bulge at the miRNA cleavage site. As a result, primary target is relieved from miRNA-mediated regulation (Todesco et al., 2010). **(C)** Short tandem target mimic (STTM) contains two *MIM* sites linked by a short spacer (Yan et al., 2012). **(D)** miRNA *SPONGE* (*SP*) has 15 miRNA binding sites linked by 4-nt spacers (Reichel et al., 2015).

1.6 miRNAs play important roles in a variety of biological processes

1.6.1 Plant development

As mentioned above, genetic strategies have been applied to plants to elucidate the functional significance of miRNA-mediated regulation in plant growth and development (reviewed in Chen, 2009; Luo et al., 2013; Rubio-Somoza & Weigel, 2011). For example, loss-of-functional mutants of genes encoding DCL1, HYL1, AGO1 and HEN1 proteins, which are involved in miRNA biogenesis and processing, display various defects in vegetative tissues, including fused cotyledons, no meristem and abnormal leaf shape (Chen et al., 2002; Schauer al., 2002; Vaucheret et al., 2004). In addition, expression of *MIM*s targeting 15 different miRNA families in Arabidopsis results in strong impacts on rosette development (Todesco et al., 2010). Therefore, many different miRNA families appear to play important roles in vegetative development.

There exists a complex network of miRNA regulation in leaf development. miR319-*CINCINNATA-like TEOSINTE BRANCHED1, CYCLOIDEA,* and *PCF* (*TCP*) and miR164-*CUP-SHAPED COTYLEDON* (*CUC*) regulatory pathways work in collaboration to control the boundary region in the shoot meristem, necessary for proper leaf morphogenesis. Expression of a miR319-resistant version of *TCP3* results in fused cotyledons and defects in shoot formation, which resembles phenotypic defects displayed in the *cuc1.cuc2* loss-of-functional mutant (Koyama et al., 2007). Further molecular studies demonstrate that the TCP3 transcription factor activates *MIR164A* transcription by directly binding to its promoter, resulting in further repression of the *CUC* genes that are regulated by miR164 (Koyama et al., 2010). In addition, miR396, which can be transcriptionally activated by TCP4, modulates cell proliferation in leaves (Rodriguez et al., 2010).

Furthermore, miRNAs cooperatively modulate leaf polarity in vegetative growth. Arabidopsis miR165/miR166 preferentially accumulate in the abaxial side of leaf primordia and restrict expression of their target genes, *PHAVOLUTA* (*PHV*), *PHABULOSA* (*PHB*) and *REVOLUTA* (*REV*), in the adaxial side (Emery et al., 2003; Mallory et al., 2004b; McConnell et al., 2001). The *phb.phv.rev* triple mutant displayed abaxialised cotyledon and failure in formation of apical meristem (Emery et al., 2003). By contrast, accumulation of *TAS3* precursor of *trans*-acting small interfering RNA (ta-siRNA) in the adaxial side of leaf primordia limits expression of *AUXIN-RESPONSE FACTOR 3* (*ARF3*) in the abaxial side (Schwab et al., 2009). Therefore, miR165/miR166 regulation of *PHV*/*PHB*/*REV* and ta-siRNA regulation of *ARF3/ARF4* play opposite roles in modulating leaf polarity in Arabidopsis (reviewed in Chen, 2009; Rubio-Somoza and Weigel, 2011). Therefore, different miRNA-target pathways work together to enable vegetative growth.

Moreover, many miRNA families are implicated in controlling flowering. Of these, miR156 and miR172 are two major miRNA families that determine flowering time (reviewed in Teotia and Tang, 2015). During plant development from the juvenile to the adult vegetative phase, and then to the flowering and reproductive development, miR156 levels gradually decline while miR172 levels increase (Teotia and Tang, 2015). Supporting this, miR156 over-expression delays flowering (Wu and Poethig, 2006; Zhang et al., 2015), whereas miR172 over-expression accelerates flowering (Jung et al., 2007). Additionally, evidence supporting a role of the miR159-*GAMYB* pathway in controlling flowering via *LEAFY* expression has been presented (Achard et al., 2004; Li et al., 2013c). Many other miRNA families have been demonstrated to control flowering, including

miR319, miR390 and miR393 (reviewed in Teotia and Tang, 2015). Together, miRNAs are implicated in controlling flowering in complex signalling pathways.

1.6.2 Abiotic stress response

As master regulators of gene expression, many plant miRNA families have been found to be involved in a plant's adaptive response to stress. Firstly, abiotic stresses alter the expression levels of many different miRNA families, conserved in diverse plant species (Wang et al., 2017; Zhang, 2015). For example, the levels of miR159, miR167, miR393 and miR408 are up-regulated under drought stress in distant plant species, in at least three different plant species (Figure 1.4A, Zhang, 2015). In addition, stress has opposite impacts on miRNA levels in different tissues. For example, miR159 levels increase in wheat leaves whereas miR159 levels decrease in the roots of wheat plant exposed to the drought stress (Figure 1.4B, Akdogan et al., 2016). Likewise, additional five miRNA families return opposing expression changes between wheat leaves and roots under the drought stress (Figure 1.4B, Akdogan et al., 2016).

However, these lines of evidence that miRNA levels change in response to stress, are insufficient to investigate the roles of miRNA in abiotic stress response. Therefore, further transgenic evidence is provided by the miRNA over-expression in transgenic plants, which changes tolerance to abiotic stresses (reviewed in Zhang, 2015). For example, *MIR169A* over-expression in Arabidopsis renders the resulting Arabidopsis plant hypersensitive to drought stress (Li et al., 2008). However, *MIR169C* over-expression in tomato enhances drought tolerance of these plants (Zhang et al., 2011b).

In addition, as post-transcriptional regulators of gene expression, miRNAs are involved in stress response via targeting stress-responsive genes or transcription factors that

regulate downstream stress responsive pathways. For example, miR398 regulates target gene *COPPER/ZINC SUPEROXIDE DISMUTASE* (*CSD*), which encodes copper superoxide dismutase (SOD) enzyme in response to oxidative stress and heat stress (Guan et al., 2013); and the miR164 target gene *N*AC (*OMTN*), which encodes NAC transcription factor involved in rice drought stress response (Fang et al., 2014).

Small RNA sequencing technologies have been used to identify the altered miRNA expression in the whole transcriptome and miRNA over-expression transgenic strategies to investigate the role of a miRNA member in stress response. However, further studies need to investigate function of the whole miRNA family in abiotic stress response via loss-of-function genetic strategies, such as miRNA decoys that inhibit multiple miRNA members simultaneously.

A.

Drought-responsive	Arabidopsis	Rice	Barley	Wheat	Maize
miR159	U	U	U	U	U
miR167	U	U			U
miR393	U	U	D	U	
miR408	U	U	U		

B.

Wheat Drought-responsive	Leaf	Root
miR159	U	D
miR169	D	U
miR319	U	D
miR399	U	D
miR528	U	D
miR5568	U	D

Figure 1.4. Altered expression of miRNA families responding to abiotic stresses. (A) Altered miRNA expression responding to drought stress is conserved in multiple plant species (Zhang, 2015). **(B)** Drought stress has opposite impacts on wheat miRNA levels in different tissues (Akdogan et al., 2016). Red block with "U" represents up-regulated miRNA levels; green block with "D" represents down-regulated miRNA levels; blank represents unknown.

1.6.3 Plant defence responses against pathogens

In plants, there are complex immunity systems that defend against invading pathogens. Plant innate immune receptors contain two major classes: pattern recognition receptors (PRRs) and resistance (R) proteins. PRRs can detect pathogen-associated molecular patterns (PAMPs) and active PAMP-triggered immunity (PTI). However, successful pathogen effectors can suppress PTI. To counter this, R proteins recognize specific pathogen effectors and trigger resistance responses, defined as effector-triggered immunity (ETI) (reviewed in Spoel and Dong 2012; Weiberg et al., 2014).

Deep sequencing of pathogen infected plants reveals altered expression of many miRNA families and their targets in Arabidopsis and rice (Li et al., 2014c; Zhang et al., 2011a), which suggests that miRNAs involve in plant resistance to pathogens. In the PTI immune process, Arabidopsis miR160a and miR393 accumulation can be induced by the PAMP peptide flg22, contributing to the PTI defence (Li et al., 2010; Padmanabhan et al., 2009).

Regarding ETI resistance, plant genomes contain large families of resistance (*R*) genes encoding disease resistance (R) proteins with a nucleotide binding site and leucine-rich repeat motifs (NBS-LRR), which can recognise pathogen effectors and trigger plant defence responses, such as hypersensitive response (HR) cell death (Coll et al., 2011; Hofius et al., 2009; Weiberg et al., 2014). In plants, many miRNA families have been demonstrated to be involved in defence responses against pathogens via mediating regulation of target genes encoding NBS-LRR. In medicago, three 22-nt miRNA families (miR1507, miR2109 and miR2118) are predicted to target diverse classes of *NBS-LRR* genes. The miRNA-mediated cleavage of these *NBS-LRR* transcripts triggers biogenesis of secondary siRNAs, phasiRNAs, which are predicted to be conserved in other legumes species (Zhai et al., 2011). In addition, miR482, a homologue member of the miR2118

family, is present in other plant species. In tomato and cotton, miR482 is validated to cleave *NBS-LRR* mRNAs and generate secondary phasiRNAs. Furthermore, miR482-mediated silencing of *NBS-LRR* is suppressed after pathogen infection (Shivaprasad et al., 2012; Zhu et al., 2013). Shivaprasad et al (2012) proposed a hypothesis that after the first pathogen attack, miRNA-mediated silencing of *NBS-LRR* is inhibited and up-regulated NBS-LRR proteins enhance plant defence response, which protects plants from further pathogen attack.

In *Solanaceae*, two another tobacco miRNA families, miR6019 (22-nt) and miR6020 (21-nt), can mediate transcript cleavage at the site of *N-terminal domain similar to the Toll and Interleukin-1 Receptors* (*TIR*) of the *TIR-NBS-LRR*s; and 22-nt miR6019 triggers biogenesis of secondary siRNAs (Li et al., 2012a). Additionally, transient co-expression of miR6019/6020 and their target *TIR-NBS-LRR* in *Nicotiana benthamiana* attenuate plant resistance against *Tobacco mosaic virus* (*TMV*), demonstrating that these two miRNA families play functional roles in plant defence response (Li et al., 2012a). Beyond tobacco, multiple other 22-nt miRNA families from tomato and potato are predicted to target *R* genes and trigger secondary siRNA biogenesis (Li et al., 2012a), suggesting the miRNA-*NBS-LRR* pathways involved in plant defence response could be conserved in *Solanaceae*.

As so many miRNA families target *R* genes and trigger secondary siRNA biogenesis in the similar mechanisms, what functional roles of these miRNA-*R* pathways? There is a speculation that miRNA-mediated silencing of *R* genes has coevolved with the number of *R* gene loci copies to constrain gene dosage. To counter evolving pathogen effectors, plants require adjusting miRNA regulation of the *R* genes, achieving the optimised

balance between resistant and susceptible contributions in plant-pathogen interactions (Li et al., 2012a).

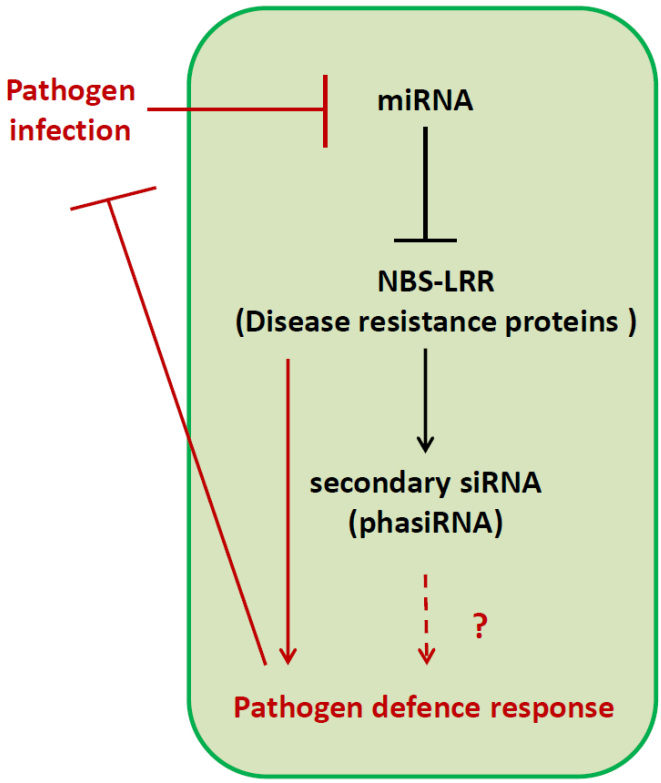


Figure 1.5. miRNA-*NBS-LRR* regulation plays an important role in plant defence against pathogen. In plant effector-triggered immunity (ETI), some miRNA families have been demonstrated to target *NBS-LRR* genes, which encode disease resistance proteins that recognise pathogen effectors and promote plant defence response. Meanwhile, these miRNA-mediated cleavages of the *NBS-LRR* transcripts produce secondary siRNAs, like phasiRNAs, whose function is not clear. There is a hypothesis that, after the first pathogen attack, miRNA-mediated silencing of *NBS-LRR* is inhibited and up-regulated NBS-LRR proteins enhance plant defence responses, which protect plants from further pathogen attack (Shivaprasad et al., 2012; Weiberg et al., 2014).

1.7 The miR159-*MYB* pathway in Arabidopsis

The Arabidopsis miR159 family has been extensively studied as a model for plant miRNA-mediated gene regulation (Palatnik et al., 2007; Allen et al., 2007; Allen et al., 2010; Li et al., 2014b; Li et al., 2016). The miR159 family has three different members, miR159a, miR159b and miR159c, which only differ from one another by one or two different nucleotides in their mature sequences (Table 1.1). Of the three isoforms, miR159a and miR159b are the most abundant, appearing to be expressed in all Arabidopsis tissues (Allen et al., 2007; Li et al., 2016). By contrast, miR159c is possibly non-functional, as this family member is only expressed at very low levels (Allen et al., 2010). Additionally, miR159a and miR159b display functional redundancy; the single *mir159a* or *mir159b* loss-of-function mutant appears wild-type, whereas the *mir159ab* double mutant displays strong pleiotropic developmental defects (Figure 1.6, Allen et al., 2007).

In Arabidopsis, miR159 is bioinformatically predicted to potentially regulate more than twenty targets, including eight conserved *MYB* targets with a highly complementary miR159 binding site (Table 1.1, Allen et al., 2010; Palatnik et al., 2007). However, genetic analysis reveals that miR159-mediated regulation of only two of predicted target genes, *MYB33* and *MYB65*, account for the developmental defects of *mir159ab*, as all defects are suppressed in a *mir159ab.myb33.myb65* quadruple mutant (Figure 1.6, Allen et al., 2007). This defined the functional specificity of Arabidopsis miR159 being restricted to *MYB33* and *MYB65*, but also raised the question of the functional significance of miR159-mediated regulation of the additional predicted targets, including those that harbour a highly conserved miR159 binding site.

MiR159 mediates the post-transcriptional silencing of *MYB33* via two mechanisms. Firstly, the 5'-RACE miRNA cleavage assay has detected miR159-specific cleavage products of *MYB33* (Allen et al., 2010). Secondly, miR159 mediates translational repression of *MYB33*, which becomes apparent in *MYB33* over-expression plants, where there are very high levels of *MYB33* mRNA, but little MYB33 protein (Li et al., 2014b).

In Arabidopsis, miR159 strongly silences *MYB33*/*MYB65* expression in vegetative tissues, a requirement for normal vegetative growth. There are multiple lines of evidence for this. Firstly, the vegetative phenotypes of the wild-type and *myb33.myb65* plants appear indistinguishable, arguing that *MYB33* and *MYB65* have no functional impact in normal vegetative growth (Figure 1.6, Allen et al., 2007). Secondly, micro-array analysis found almost identical transcriptomes in the shoot apex regions of the *myb33.myb65* and wild-type plants (Alonso-Peral et al., 2010). Thirdly, a *MYB33*:GUS transgene fails to express in vegetative tissues, which is in stark contrast to the miR159-resistant version, *mMYB33*:GUS which is strongly expressed in vegetative tissues (Millar and Gubler, 2005). Therefore, the miR159-*MYB33/MYB65* pathway in vegetative tissues seems to be futile, where the strong *MYB33* and *MYB65* transcription occurs only to be completely silenced resulting in no clear physiological outcome. Given this, the functional roles of *MYB33* and *MYB65* in Arabidopsis vegetative tissues still remain unclear.

Table 1.1. miR159 targets predicted by bioinformatics. Sequence alignment of miR159 isoforms and the miR159 binding sites in bioinformatically predicted targets. Nucleotides underlined represent the cleavage site, lower case blue letters represent mis-matches to the miR159a and miR159b sequences, and G:U pairing is shown in uppercase blue letters. Conserved *MYB* targets are shown in the yellow block (Allen et al., 2010).

miR159a	AUCUCGAGGGAAGUUAGGUUU-5′
miR159b	UUCUCGAGGGAAGUUAGGUUU-5′
miR159c	UCCUCGAGGGAAGUUAGGUUU-5′
MYB33	UGGAGCUCCCUUCAuUCCAAu-3′
MYB65	UGGAGCUCCCUUCAuUCCAAu-3′
MYB81	UcGAGUUCCCUUCAuUCCAAu-3′
MYB97	AuGAGCUCUCUUCAAaCCAAA-3′
MYB101	UAGAGCUUCCUUCAAaCCAAA-3′
MYB104	UGGAGCUCCCUUCAuUCCAAG-3′
MYB120	AGcAGCUCCCUUCAAaCCAAA-3′
DUO1	UGGAGCUCCaUUCGAUCCAAA-3′
TCP2	AGGGGgaCCCUUCAGUCCAAu-3′
ACS8	UcGAGUUUCUUUCAAUCCAAA-3′
OPT1	UAGAGCUUUCUUCAuUCCAAc-3′
Zinc/Cu SODM	UGGAcCUCaCUUCAAUCCAuu-3′
MRG1	UAGAGCcCCCUUCAAaCCAAA-3′
MRG-similar	UAGAGCcUCCUUCAAgCCAAA-3′
PHD(ATX3)	UAGAGCUCUCUUaAGUCUAAA-3′
Anion/ExProt	AAGAGCUCCgUUCAGUCCAcG-3′
NPH3 Prot	AAaAGCUUCCUaCGAUCCAAG-3′
NAS2	UAGAGCUUUCUUguAUCCAAu-3′
	U
SPL	AuGAGCUCUCUUCAAUCCAAA-3′
	G
PPDK	AAGAGUUUCCUUCAAUCCAAA-3′

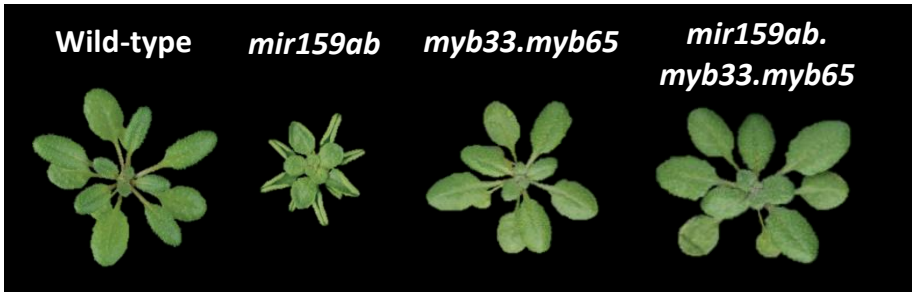


Figure 1.6. The *myb33* and *myb65* knockout alleles rescue the *mir159ab* mutant phenotype. Compare to the wild-type Arabidopsis, the *mir159ab* loss-of-functional mutant displays up-curled leaves on rosette. The introduction of *myb33* and *myb65* knockout alleles into the wild-type plant has no distinguishable phenotypic difference; *myb33* and *myb65* into the *mir159ab* background rescues the *mir159ab* mutant phenotype (Allen et al., 2007).

1.8 Molecular characterisations of miR159-*GAMYB* in diverse plant species

The miR159 family in plants is ancient and highly conserved, present from basal vascular plants to angiosperms (Table 1.2, Axtell et al., 2007; Li et al., 2011). Based on the deep sequencing data of sRNAs from a variety of plant species, it is apparent that miR159 is highly abundant and widely expressed (Fahlgren et al., 2007; Jeong et al., 2011; Rajagopalan et al., 2006; Szittya et al., 2008; Yao et al., 2007). In most examined species, the miR159 family is composed of multiple members (Table 1.3); this includes Arabidopsis (Allen et al., 2007; Allen et al., 2010), rice (Jones-Rhoades and Bartel, 2004), soybean (Subramanian et al., 2008), and maize (Zhang et al., 2009).

Among different miRNA families, miR159 is closely related to the miR319 family. Based on the high sequence similarity and conserved miRNA processing pattern, these two miRNA families are speculated to originate from a common ancestor (Li et al., 2011). During evolution, miR159 and miR319 have become divergent in their sequences and expression patterns, which results in their functional specialisation (Palatnik et al., 2007). This has been shown in Arabidopsis, where miR159 and miR319 share 17 of 21 identical nucleotides. Moreover, miR159 is abundant and widely expressed throughout the plant, whereas miR319 is expressed at much lower levels and in specific tissues (Axtell and Bartel, 2005). Although miR319 can regulate *MYB33*/*MYB65*, as miR319 abundance is much lower than miR159, miR319 does not significantly contribute to the regulation of *MYB33*/*MYB65* in Arabidopsis (Palatnik et al., 2007).

Based on bioinformatic prediction and experimental validation, miR159 mediates regulation of a family of *GAMYB* or *GAMYB-like* genes in diverse plant species, including Arabidopsis (Millar and Gubler, 2005; Schwab et al., 2005), rice (Tsuji *et al*., 2006), strawberry (Csukasi et al., 2012) and pinetree (Li et al., 2013b). All of these *GAMYB*

genes from different plant species contain a conserved miR159 binding site, which is highly complementary to their corresponding miR159 sequence and located in the conserved region between BOX1 and BOX2 of the *GAMYB* genes (Millar and Gubler, 2005; Tsuji *et al*., 2006). Therefore, it appears that the miR159-*GAMYB* pathway is conserved in diverse plant species at the molecular level.

Table 1.2. miR159 contains conserved sequences from basal land plants to angiosperms. Sequence alignments of mature miR159 from distant plant species, conserved nucleotides are shown in black and distinct nucleotides are shown in red.

Arabidopsis	AUCUCGAGGGAAGUUAGGUUU-5′
Soybean	AUCUCGAGGGAAGUUAGGUUU-5′
Poplar	AUCUCGAGGGAAGUUAGGUUU-5′
Rice	GUCUCGAGGGAAGUUAGGUUU-5′
Maize	GUCUCGAGGGAAGUUAGGUUU-5′
Pine	ACCUCGAGGGAAGUUAGGUUC-5′
Lycopods	CCCUCGAGGGAAGUUAGGUUC-5′

Table 1.3. The miR159 family is composed of multiple members. The miR159 family has multiple members in plant species, which has been reported in Arabidopsis (Allen et al., 2007; Allen et al., 2010), rice (Jones-Rhoades and Bartel, 2004), maize (Zhang et al., 2009) and soybean (Subramanian et al., 2008). Some miR159 isoforms have the same mature miR159 sequences. Nucleotides in red letters represent differences in sequence composition.

	miR159 sequence	
Arabidopsis	a	UUUGGAUUGAAGGGAGCUCUA
	b	UUUGGAUUGAAGGGAGCUCUU
	c	UUUGGAUUGAAGGGAGCUCCU
Rice	a/b	UUUGGAUUGAAGGGAGCUCUG
	c	AUUGGAUUGAAGGGAGCUCCA
	d	AUUGGAUUGAAGGGAGCUCCG
	e	AUUGGAUUGAAGGGAGCUCCU
	f	CUUGGAUUGAAGGGAGCUCUA
Maize	a/b/f/j/k	UUUGGAUUGAAGGGAGCUCUG
	b/d	CUUGGAUUGAAGGGAGCUCCU
	g	UUUGGAGUGAAGGGAGUUCUG
	h/i	UUUGGAGUGAAGGGAGCUCUG
Soybean	a/e	UUUGGAUUGAAGGGAGCUCUA
	b/f	AUUGGAGUGAAGGGAGCUCCA
	c	AUUGGAGUGAAGGGAGCUCCG

1.9 *GAMYB* has conserved function in plant development

GAMYB and *GAMYB-like* genes, as major miR159 targets, encode conserved R2R3 MYB transcription factors, which potentially regulate a variety of downstream events. To date, numerous studies have investigated functional roles of *GAMYB* homologues in different plant development processes. GAMYB was initially identified as a positive transcriptional regulator of GA-dependent α-amylase and hydrolytic enzymes expression in barley seed aleurone cells (Gubler et al., 1995; Gubler et al., 1999). Furthermore, GAMYB has been shown to positively regulate GA-mediated programmed cell death (PCD) in the seed aleurone of both barley and Arabidopsis (Alonso-Peral et al., 2010; Guo and Ho, 2008), indicating GAMYB function is necessary for seed germination.

Additionally, *GAMYB* expression is required for anther fertility. In anthers of both the Arabidopsis *myb33.myb65* and rice *gamyb* loss-of-functional mutants, the tapetum fails to undergo PCD, but instead abnormally expands to display tapetal hypertrophy, resulting in pollen abortion and male sterility (Aya et al., 2009; Millar and Gubler, 2005). Moreover, GAMYB also plays an important role in fruit development of strawberry and tomato (da Silva et al., 2017; Vallarino et a., 2015).

As *GAMYB* genes encode transcription factors, they likely participate in the regulation of a variety of downstream genes. Several transcriptome analyses have revealed potential candidates of GAMYB expression regulation in plants. In Arabidopsis, microarray analysis of the *mir159ab* shoot apex region revealed 166 genes differentially expressed compared to wild type. Of these, many up-regulated genes in the *mir159ab* shoot apex tissues were aleurone-related genes, which have been demonstrated to have down-regulated expression in *myb33.myb65.myb101* seeds (Alonso-Peral et al., 2010). This correlation with *GAMYB* expression makes them strong candidates for genes under

GAMYB expression regulation. Of these, *CYSTEIN PROTEINASE 1* (*CP1*), which is demonstrated to promote PCD, is most-upregulated by *GAMYB* expression and serves as a good marker gene for Arabidopsis MYB protein activity (Alonso-Peral et al., 2010; Li et al., 2016). In rice, transcriptomic analysis of *STTM159* transgenic rice grains identified numerous changes in the gene regulatory networks. Of these, PCD related pathways are up-regulated and cell cycle related pathways are down-regulated in the STTM159 rice plants. This finding leads to the speculation that deregulated *GAMYB* expression results in activation of PCD in rice grains (Zhao et al., 2017). Therefore, GAMYB downstream events need to be further identified in vegetative tissues, which is necessary to determine the functional roles of the miR159-*GAMYB* pathway.

1.10 Scope and aims of thesis

As miRNAs have been identified as master regulators of gene expression in a wide range of plant processes, understanding their function is obviously critical to advance our currently limited understanding of gene expression regulation in plants. This thesis has focused on the miR159-*GAMYB* pathway, a conserved regulatory module in plants, investigating its function in distant plant species, with a specific focus on the role of this regulatory module in plant defence responses. Further, this representative model is used to explore the importance of target RNA secondary structure determining miR159 silencing efficacy in Arabidopsis, gaining new insights of mechanisms of plant miRNA-target recognition.

In plants, high complementarity in miRNA-target has been widely accepted as the predominant factor to determine strong miRNA silencing outcomes. Chapter 3 aims to identify whether this dogma is biologically relevant *in planta* and to further investigate factors beyond complementarity determining miRNA silencing efficacy via the miR159-

GAMYB transgenic system in Arabidopsis. Curiously, the *MYB33* and *MYB65* genes are bioinformatically predicted to have a strong stem-loop RNA secondary structure abutting the miR159 binding site, which appears to be conserved across plant species. Thus, I established a series of *MYB33* constructs with altered the conserved RNA secondary structure and compared miR159 silencing efficacies against these mutant *MYB33* constructs. Transformant analysis found that disruption of the conserved RNA structures attenuated miR159 efficacy, while introducing an artificial long stem structure restored strong miR159-mediated silencing. This is the first *in vivo* study that demonstrates, in addition to complementarity, target RNA secondary structure is a major determinant of miRNA silencing efficacy in plants.

Secondly, Chapter 4 aims to utilise a *MIM159* decoy system to inhibit miR159 function to investigate the role of the miR159-*GAMYB* pathway in Arabidopsis, tobacco and rice; representative model species for the dicot and monocot plants. Phenotyping of *MIM159* transformants reveals inhibiting miR159 activity in all three species produced similar phenotypic outcomes, including stunted growth and abnormal leaf development. RNA sequencing of *MIM159* tobacco leaves found strong up-regulation of *PATHOGENESIS-RELATED PROTEIN* (*PR*) expression, contributing to enhanced plant defence. However, *PR* expression is not up-regulated in Arabidopsis or rice when miR159 function is inhibited. These conflicting lines of evidence provide a new insight that the ancient miR159-*GAMYB* pathway could be functionally divergent in different species; and *GAMYB* expression has a possible function that modulates plant defence response.

Chapter 2

Materials and methods

2.1 Plant materials and growth conditions

2.1.1 Arabidopsis

Arabidopsis thaliana ecotype Columbia-0 (Col-0) was used in most Arabidopsis experiments. The *myb33* mutant was in ecotype Col-6 with a *glaborous 1* background mutation leading to a no trichome phenotype (Millar and Gubler, 2005). The *mir159ab* double T-DNA mutant was in the Col-0 background (Allen *et al*., 2007). *Arabidopsis thaliana* ecotypes C24 and Landsberg *erecta* (Ler-1) were used in *MIM159* transformation.

After harvesting of seeds, sterilisation was performed by exposure to chlorine gas for 4-8 hours in a desiccator jar (chlorine gas was generated by mixing 100 mL of commercial bleach with 3 mL of 36% concentrated hydrochloric acid). Sterilised seeds were sown on soil (Debco Plugger soil mixed with Osmocote Extra Mini fertilizer at 3.5 g/L) or on agar plates containing Murashige and Skoog Basal Medium (MS). This was followed by stratification for 48 hr at 4^{o}C in the dark. Plants were grown under long-day conditions (16 hr light / 8 hr dark, at 150 μmol/m^2/sec at 22^{o}C). Only the *35S-MYB* primary transformants were grown under short-day conditions (10 hr light / 14 hr dark, 150 μmol/m^2/sec at 22^{o}C), which can elongate the vegetative growth and provide more details of the rosette phenotypes.

2.1.2 Tobacco

Nicotiana tabacum TN90 was used in tobacco experiments, including stable transformation of tobacco. Tobacco seeds were sterilised by exposure to chlorine gas for 2-3 hours in a desiccator jar (chlorine gas was generated by mixing 100 mL of commercial bleach with 3 mL of 36% concentrated hydrochloric acid). Sterilised seeds were sown on

agar plates containing MS and 3% Sucrose, followed stratification for 48 hrs at 4ºC in the dark. Around three-week-old tobacco plants were transplanting on soil (Debco Plugger soil mixed with Osmocote Extra Mini fertilizer at 7 g/L). Tobacco plants were grown under long-day conditions (16 hr light / 8 hr dark, at 500 μmol/m^2/sec at 25ºC).

2.1.3 Rice

Japonica rice (*Oryza sativa* L.) cultivars Kitaake was used in rice experiments, including the rice transformation. After glumes were removed, rice seeds were sterilised by washing once with 70% ethanol and twice with 33-50% commercial bleach, followed by three washes with sterilised water. Sterilised and dry seeds were placed in Gellan Gum culture containing MS medium and 3% Sucrose. Two-week-old rice plants were transplanted to Rice Mix soil (including 80% peat, 10% perlite, 10% vermiculite, dolomite lime, macro and micro nutrients; mixed with Osmocote Extra Mini fertilizer at 7 g/L). Rice plants were grown under 12 hr light / 12 hr dark, at 400 μmol/m^2/sec at 28ºC).

2.2 Gateway cloning

2.2.1 Generation of *35S-MYB* and *35S-mMYB* constructs

MYB gene sequences were amplified with gene-specific primers (Appendix Primer Table 1, that had a NetPrimer program score higher than 80) that contained attB1 and attB2 recombination sequences to allow integration into a Gateway donor vector. PCR amplification was performed using high fidelity KOD Hot Start DNA Polymerase (Novagen) with the following cycling conditions: 1 cycle of 95ºC/2 min, 35 cycles of 95ºC/20 sec, 55ºC/10 sec, 70ºC for 20 sec/kb extension time according to amplicon size, and 1 cycle of 70ºC for 10 min. PCR products were analysed by agarose gel electrophoresis and products of the desired sizes were excised from the gel and purified using the Wizard® SV Gel and PCR Clean-Up System (Promega). The BP reactions were

performed with Gateway BP ClonaseTM II enzyme mix (Invitrogen), to integrate the purified PCR products into the pDONOR/ZEO vector (Invitrogen) to generate pENTRY-*MYB* vectors. The BP enzyme reaction mixture was transformed into *E. coli* Alpha-Select Gold Efficiency competent cells (Bioline) by heat shock. After recovery at 37°C for one hour, the bacteria were grown overnight at 37°C on low salt Luria Broth (LB) plates containing Zeocin (Invitrogen) at a concentration of 50 μg/mL. Positive clones were subcultured in liquid low-salt Luria broth (LSLB) overnight at 37°C, and plasmids were extracted from the culture using an AxyPrepTM Plasmid Miniprep Kit (Axygen). Plasmids were then screened by restriction enzyme digestion and were then subject to DNA sequencing to confirm the correct *MYB* constructs, using ABI PRISM BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems) under recommended cycling conditions. The sequencing products were precipitated and analysed at John Curtin School of Medical Research, Australian National University, Canberra.

The pENTRY-*mMYB* vectors, which served as miR159-resistant versions, were obtained by performing site-directed mutagenesis of the miR159 binding site on the sequence confirmed pENTRY-*MYB* vectors (Appendix Primer Table 2, Liu and Naismith, 2008). The mutated miR159 binding site in the pENTRY-*mMYB* vectors was confirmed by restriction enzyme digestion and DNA sequencing.

All entry clones were integrated into the pDESTINATION vector pGWB602Ω (Nakamura *et al*., 2010) to generate the corresponding binary vectors expressing 35S-*MYB* via LR reactions using Gateway LR ClonaseTM II enzyme mix (Invitrogen). The LR enzyme reaction mixture was transformed into *E. coli* via electroporation. Transformed bacteria were grown overnight at 37°C on LB plates containing Spectinomycin at a concentration of 50 μg/mL, and positive clones were subcultured in

liquid LB overnight. After plasmid extraction, the expression vectors of *35S-MYB* were screened and confirmed by restriction enzyme digestion.

2.2.2 Generation of *MYB81-33* and *mMYB81-33* constructs

The *MYB81-33* construct was generated by mutating the miR159 binding site of *MYB81* to make it identical to the miR159 binding site of *MYB33*. The pENTRY-*MYB81-33* and pENTRY-*mMYB81-33* (miR159-resistant version) vectors were obtained by performing site-directed mutagenesis of the miR159 binding site on the sequencing confirmed pENTRY-*MYB81* vectors (Appendix Primer Table 4). The mutations in the pENTRY-*MYB81-33* and *mMYB81-33* vectors were confirmed by restriction enzyme digestion and DNA sequencing. The entry clones were recombined into the destination vector pGWB602Ω to generate the corresponding binary vectors expressing 35S-*MYB81-33* and *35S-mMYB81-33* via Gateway LR reactions (Invitrogen). Expression vectors were screened and confirmed by restriction enzyme digestion.

2.2.3 Generation of *MYB33-mSL* constructs

The *MYB33-mSL1*, *MYB33-mSL2* and *MYB33-mSL2-2* constructs were obtained by performing site-directed mutagenesis of the SL1 or SL2 region of a *MYB33* genomic fragment in the pDONR/Zeo vector (Appendix Primer Table 5). This 4356 bp *MYB33* genomic fragment includes the 1991 bp of genomic sequences upstream of the *MYB33* start codon, the whole *MYB33* coding region and 585 bp of sequences downstream of the *MYB33* stop codon (Li et al., 2014b). Following this, the *mMYB33-mSL2-2*, which served as a miR159-resistant version, was achieved by site-directed mutagenesis of the miR159 binding site on the sequencing confirmed pENTRY-*MYB33-mSL2-2* vector. The mutations in the pENTRY-*MYB33-mSL* vectors were confirmed by restriction enzyme digestion and DNA sequencing. The entry clones were recombined into the destination

vector pMDC123 (Curtis and Grossniklaus, 2003) via Gateway LR reactions (Invitrogen). Expression vectors were screened and confirmed by restriction enzyme digestion.

2.2.4 Generation of *MYB81-SL* constructs

The *MYB81-SL* construct, including the *MYB81* genomic sequence (1361 bp) plus 15 nt of the 5' UTR and 9 nt of the 3' UTR, was synthesized by IDT (USA) and sub-cloned into pDONR/Zeo vector. *MYB81-SL-6nt* was generated by performing site-directed mutagenesis of deleting the 6-nt between two SLs in the pENTRY-*MYB81-SL* vector. *mMYB81-SL* and *mMYB81-SL-6nt*, which served as miR159-resistant versions, were achieved by site-directed mutagenesis of the miR159 binding site on the sequencing confirmed pENTRY-*MYB81-SL* and pENTRY-*MYB81-SL-6nt* vectors, respectively (Appendix Primer Table 6). The mutations in the pENTRY-*(m)MYB81-SL/SL-6nt* vectors were confirmed by restriction enzyme digestion and DNA sequencing. The entry clones were recombined into the destination vector pGWB602Ω via Gateway LR reactions (Invitrogen). Expression vectors were screened and confirmed by restriction enzyme digestion.

2.2.5 Generation of the *MIM159* decoy constructs

The artificial target mimics *MIM159* for Arabidopsis and tobacco transformation (Todesco et al., 2010) were obtained from the European Arabidopsis Stock Centre (NASC) and sub-cloned into pDONR/Zeo via Gateway BP reaction. Sequencing confirmed entry clones of *MIM159* were then integrated into the Gateway destination vector pMDC32 (Curtis and Grossniklaus, 2003) for Arabidopsis transformation; or integrated into Gateway destination vector pGWB602Ω for tobacco transformation, via Gateway LR reactions (Invitrogen). Expression vectors were screened and confirmed by restriction enzyme digestion.

To generate rice *MIM159* construct, firstly, a 661 bp rice *IPS1* gene (AY568759) was synthesized by IDT (USA) and sub-cloned into pDONR/Zeo via Gateway BP reaction. Next, site-directed mutagenesis was performed on the pENTRY vector to generate the rice miR159 binding site with a three-nt bulge at the cleavage site (Appendix Primer Table 8), which confirmed by the following restriction enzyme digestion and DNA sequencing. The entry clones were recombined into the modified destination vector pMDC32-Ubi (in which a Ubiquitin promoter had replaced a 35S promoter) via Gateway LR reactions (Invitrogen). The expression vector was screened and confirmed by restriction enzyme digestion.

2.2.6 Generation of *mNtGAMYB2*, *mOsGAMYB* and *mLaMYB33* constructs

Tobacco *GAMYB2* (*NtGAMYB2*) and rice *GAMYB* (*OsGAMYB*) gene sequences were amplified with gene-specific primers, which contained attB1 and attB2 recombination sequences, on the tobacco and rice cDNA templates, respectively (Appendix Primer Table 7 and 8). PCR amplification was performed using high fidelity KOD Hot Start DNA Polymerase (Novagen) with the following cycling conditions: 1 cycle of 95^{o}C/2 min, 35 cycles of 95^{o}C/20 sec, 55^{o}C/10 sec, 70^{o}C for 20 sec/kb extension time according to amplicon size, and 1 cycle of 70^{o}C for 10 min. PCR products were analysed by agarose gel electrophoresis and products of the desired sizes were excised from the gel and purified using the Wizard® SV Gel and PCR Clean-Up System (Promega). The BP reactions were performed with Gateway BP ClonaseTM II enzyme mix (Invitrogen), to integrate the purified PCR products into the pDONOR/ZEO vector (Invitrogen) to generate pENTRY-*NtGAMYB2* and pENTRY-*OsGAMYB* vectors. The *LaMYB33* construct was obtained from the Liwang Qi Lab (Laboratory of Cell Biology, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, People's Republic of China)

and sub-cloned into the pDONR/Zeo vector. Plasmids were then screened by restriction enzyme digestion and were then subject to DNA sequencing to confirm the *NtGAMYB2 and OsGAMYB* sequences.

The pENTRY-*mGAMYB* vectors, which served as miR159-resistant versions, were obtained by performing site-directed mutagenesis of the miR159 binding site on the verified pENTRY-*GAMYB* vectors (Appendix Primer Table 7 and 8). The mutated binding site in the pENTRY-*mGAMYB* vectors was confirmed by restriction enzyme digestion and DNA sequencing.

All entry clones were integrated into the destination vector pGWB602Ω to generate the corresponding binary vectors expressing 35S-*mGAMYB* via Gateway LR reactions (Invitrogen). Expression vectors were screened and confirmed by restriction enzyme digestion.

2.3 Site-directed mutagenesis

Site-directed mutagenesis (Liu and Naismith, 2008) was performed on the verified pENTRY vector to introduce the desired mutations at the expected sites. The primers for mutagenesis contained non-overlapping sequences at the 3' end and complementary sequences, including the mutated nucleotides, at the 5' end. The non-overlapping sequences were larger than the complementary sequences and had a 5-10°C higher melting temperature to minimize primer-dimers. PCR reactions were performed using KOD Hot Start DNA Polymerase (Novagen) with 50 ng of pENTRY plasmid template, 100 ng forward primer and 100 ng reverse primer, at the setting of 1 cycle of 98ºC for 1 min, 25 cycles of 98ºC/10 sec, 55ºC/20 sec and 72ºC for 20 sec/kb extension time according to the amplicon size, and 1 cycle of 72ºC/10 min. The PCR products were

analysed by agarose gel electrophoresis and digested with 2 µL *DpnI* enzyme at 37ºC for 4 hours. After digestion, the PCR products were purified using Wizard® SV Gel and PCR Clean-Up System (Promega) and transformed in *E. coli* Alpha-Select Gold Efficiency competent cells (Bioline) by heat shock. Plasmids from positive clones of pENTRY vectors after antibiotic selection (Zeocin 50 µg/mL) were extracted and confirmed by enzyme digestion and DNA sequencing.

2.4 Transformation of Arabidopsis

Gateway expression vectors were transformed into the *Agrobacterium tumefaciens* strain GV3101 by electroporation (Hellens et al., 2000), and incubated on LB plates containing Rifamycin (50 µg/mL), Gentamicin (25 µg/mL) and the appropriate antibiotic for plasmid selection at 28°C for 48 hours. Single colonies were obtained and individually subcultured in 15 mL of liquid LB containing the same antibiotics. After growth at 28°C overnight, the plasmid was extracted from the culture using the AxyPrep™ Plasmid Miniprep Kit (Axygen). Constructs were confirmed by restriction enzyme analysis of extracted plasmids. Then, 1 mL of the liquid culture was inoculated into a 250 mL liquid LB culture with Gentamicin (25 µg/mL) and the appropriate antibiotic for plasmid selection, and incubated overnight at 28°C with constant shaking at 220 rpm. *Agrobacterium* was harvested by centrifugation for 15 min at 5,000 r.p.m, and resuspended in 500 mL of infiltration medium containing 5% sucrose and 0.03% surfactant Silwet L-77 surfactant (Clough and Bent, 1998). Floral dipping was performed via submergence of the floral parts of the plants in the *Agrobacterium* solution for 30 sec. Plants were then covered with plastic and incubated in the dark for 24 hours at room temperature before being placed back into the growth chamber. Seeds were subsequently harvested and sterilised as described above. Transformants were selected by growing

seeds on agar plates containing MS medium and antibiotics for selection. After 6-7 days growth, transformants were identified and transplanted onto soil.

2.5 Transformation of tobacco

The *MIM159*-pGWB602Ω and *mNtGAMYB2*-pGWB602Ω binary expression vectors were individually transformed into the *Agrobacterium tumefaciens* strain GV3101 by electroporation (Hellens et al., 2000), and incubated on LB plates containing Rifamycin (50 μg/mL), Gentamicin (25 μg/mL) and Spectinomycin (50 μg/mL) at 28°C for 48 hours. A single colony was subcultured in 4 mL of liquid LB containing the same antibiotics. After growth at 28°C overnight, plasmid was extracted from the culture using the AxyPrepTM Plasmid Miniprep Kit (Axygen). The correct tobacco *MIM159* constructs were confirmed by restriction enzyme analysis of extracted plasmids. Then 50 μL of the liquid culture was inoculated into a 25 mL liquid LB culture with Gentamicin (25 μg/mL) and Spectinomycin (50 μg/mL), and incubated overnight at 28°C with constant shaking at 220 rpm. *Agrobacterium* was harvested by centrifugation for 10 min at 5,000 r.p.m, washed once with MS liquid medium, and resuspended in 30 mL MS liquid medium to the concentration of 0.1 OD_{600} value. Transformation of *Nicotiana tabacum* leaf tissue was performed via immersing young leaf strips into *Agrobacterium* culture for 10 mins and transferring leaf strips onto co-cultivation agar plates, containing MS medium, 30 g/L sucrose, 1 mg/L 6-benzylaminopurine (BA) and 0.1 mg/L 1-Naphthaleneacetic acid (NAA). After incubation for 2 days in dark, leaf strips were transferred onto regeneration and selection medium plates that contained MS medium, 30 g/L sucrose, 1 mg/L BA, 0.1 mg/L NAA, 300 mg/L Timentin and 25 mg/L BASTA. Plates were incubated in the dark for two weeks and were then transferred to light and incubated at 25^{o}C. (500 μmol/m^{2}/sec at 25^{o}C). Adventitious buds were initiated from callus 3-4 weeks post-inoculation. Adventitious buds were selected and transferred onto fresh regeneration and selection

medium as necessary. When plantletshad developed a distinct shoot structure, shoots were excised and transferred onto root agar plates that contained MS media supplemented with 30 g/L sucrose, 300 mg/L Timentin and 25 mg/L BASTA. Once roots were established, the putative T0 tobacco transformants were transferred to soil and further cultivated under 16 hr light / 8 hr dark, at 500 μmol/m^2/sec at 25ºC. Seeds were subsequently harvested and sterilised as described above. T1 transformants were selected by growing seeds in agar tissue culture pots containing MS medium, 30 g/L sucrose and 15 g/L BASTA for selection. After 3-4 weeks of growth, transformants were identified and transplanted onto soil.

2.6 Transformation of rice

The *MIM159*-pMDC32-Ubi binary vector was transformed into the *Agrobacterium* strain AGL1 by electroporation (Hellens et al., 2000), and incubated on LB plates containing Rifamycin (25 g/L) and Kanamycin (50 g/L) at 28°C for 3 days. A single colony was sub-cultured in 50 mL of liquid LB containing the same antibiotics. After growth at 28°C overnight, plasmid DNA was extracted from the culture using the AxyPrepTM Plasmid Miniprep Kit (Axygen). The correct rice *MIM159* constructs were confirmed by restriction enzyme analysis of extracted plasmids. Next, 100 μL of the liquid culture was inoculated into a 50 mL liquid LB culture with Rifamycin (25 g/L) and Kanamycin (50 g/L), and incubated overnight at 28°C with constant shaking at 220 rpm. *Agrobacterium* at a cell density of 0.2-0.3 OD_{600} was harvested by centrifugation for 25 min at 5,000 r.p.m and resuspended in 40 mL 2N6-AS liquid medium, containing 30 g/L sucrose, 10 g/L glucose, 3.98 g/L CHU[N6] Basal Mixture, 100 mg/L myo-inositol, 300 mg/L casamino acids, N6 vitamins (×1000) and 20 mg/L acetosyringone.

To generate rice callus, sterilised rice seeds (*Oryza sativa* L. Kitaake) were placed on the

N6D Gellan Gum plates, containing 30 g/L sucrose, 3.98 g/L CHU[N6] Basal Mixture, 100 mg/L myo-inositol, 300 mg/L casamino acids, 2.878 g/L L proline, N6 vitamins (×1000) and 0.2 mg/mL 2,4-D. Rice calli were cultured for four weeks under light at 28ºC. The rice calli were collected (around 5mL of loosely packed volume) and gently mixed by inversion with 40 mL *Agrobacterium* culture for 5 mins. After removing excess bacteria with sterilised paper towels, the calli was co-cultured with *Agrobacterium* on filter paper on solid 2N6-AS Gellan Gum medium for three days at 28ºC in the dark. Afterwards, calli were collected and washed five times with sterilised water. The calli were resuspended in sterilised water including 150 g/L Timentin and dried on the sterilised paper towel. Next, the calli were transferred to N6D medium containing 35 g/L hygromycin and cultivated for 2 weeks in the dark at 28ºC. Rice calli were transferred into fresh N6D medium containing 35 g/L hygromycin as necessary. Normally after four weeks of selection, the resistant calli had small nodular embryos on their surface. The surviving healthy calli were transferred to regeneration Gellan Gum medium plates, containing 30 g/L sucrose, 30 g/L sorbitol, 4.33 g/L MS, 2.0 g/L casamino acids, B5 vitamins (×1000), 20 µg/mL NAA, 2 mg/mL Kinentin, 150 g/mL Timentin and 35 g/mL hygromycin. The calli were transferred to fresh media every 10-14 days. Once observable shoot structures had formed, plantlets were transferred to MS medium containing 30 g/L sucrose, 4.33 g/L MS, B5 vitamins (×1000) and 35 mg/mL hygromycin in tissue culture pots. Only one plantlet derived from an individual piece of resistant callus was taken through to this stage of the selection process. After extensive rooting, healthy seedlings were transferred to Rice Mix Soil and grown under 12 hr light / 12 hr dark 400 µmol/m^2/sec at 28ºC. Seeds were subsequently harvested and sterilised as described above. T1 seeds were growing in Gellan Gum tissue culture pots containing MS medium, 30 g/L sucrose and B5 vitamins. After 2 weeks of growth, seedlings were transplanted onto Rice Mix soil. T1 transformants were selected by genotyping of hygromycin-

resistant gene.

2.7 RNA extraction

TRIzol® (Invitrogen) was used to extract total RNA from plant tissues with the following modifications to the manufacturer's protocol: (1) approximately 500 mg of plant material was used with 1 mL of Trizol for each extraction; (2) homogenisation of tissues was carried out using a mortar and pestle; (3) the chloroform extraction step was repeated twice, and; (4) precipitation of RNA was carried out overnight at -20°C to maximize the recovery of small RNAs. The concentration of each sample was measured using a Nanodrop spectrophotometer. The quality of each RNA sample was then visually examined by denaturing 1 μg each sample with RNA loading buffer at 65°C for 5 mins, followed by 1% agarose gel electrophoresis.

2.8 DNase treatment of RNA samples

30 μg of total RNA was treated with RQ1 RNase-Free DNase (Promega) and RNaseOut Recombinant RNase Inhibitor (Invitrogen) in a 100 μL reaction volume following the manufacturer's protocol. Treated RNA was then purified using a RNeasy Plant Mini Kit (QIAGEN). Post purification, the concentration of each sample was again assessed using a Nanodrop spectrophotometer. The quality of purified RNA was then examined by denaturing 1 μg sample with RNA loading buffer at 65°C for 5 mins, followed by 1% agarose gel electrophoresis.

2.9 cDNA synthesis

cDNA synthesis was carried out using SuperScript® III Reverse Transcriptase (Invitrogen) and an oligo dT primer according to manufacturer's protocol. For each

sample, 2 μg of total RNA was used. The 10 μL reaction was then diluted 50 times in nuclease-free water and used for subsequent qRT-PCR.

2.10 Quantitative real-time PCR (qRT-PCR)

For qRT-PCR, 9.6 μL of each diluted cDNA sample was added to 10 μL of SensiFAST SYBR (Bioline) mix with 0.4 μL of qRT-PCR primers at 10 μmol. The qRT-PCR reactions were carried out on a Rotor-Gene Q real time PCR machine (Qiagen) in triplicate, under the following cycling conditions: 1 cycle of 95°C/5 min, 45 cycles of 95°C/15 sec, 60°C/15 sec, 72°C/20 sec. Fluorescence was acquired at the 72°C step. A 55°C to 99°C melting cycle was then carried out. The *CYCLOPHILIN* (AT2G29960) was used to normalise Arabidopsis mRNA levels for each gene; the *PROTEIN PHOSPHATASE 2A* (*PP2A*, Liu et al., 2012) was used to normalise tobacco mRNA levels for each gene; and *ACTIN* (Os03g50890, Caldana et al., 2007) was used to normalise rice mRNA levels for each gene, using the comparative quantitation program in the Rotor-Gene Q software package provided by Qiagen. The average values of three technical replicates were used for analysis.

2.11 qRT-PCR assays for mature miRNAs

TaqMan sRNA assays (ABI) were used to measure mature miRNA levels according to the manufacturer's instructions, except for the following modifications: (1) For each RNA sample, the retro-transcription was multiplexed with looped-RT primers for both the miRNA and the control Arabidopsis sRNA *snoR101*, tobacco sRNA *snoR101* and rice sRNA *snoR14*, respectively. (2) The cDNA synthesised (15 μL) was diluted with 86.4 μL nuclease-free water, and 9 μL of the diluted cDNA was used in each PCR reaction. Each cDNA was assayed in three technical replicates using a Rotor-Gene Q real time PCR machine (QIAGEN) under the cycling conditions described above. The abundance of

miR159 in each assessed species was normalised to corresponding sRNAs using the comparative quantitation program provided by the Rotor-Gene Q software (QIAGEN).

2.12 RNA sequencing of wild type and *MIM159* tobacco leaves

RNA was isolated from newly emerged leaves of eight-week-old wild-type and *MIM159* tobacco plants. Each sample consisted of tissues sampled from three different plants; for the *MIM159* analyses, tissues were sampled from multiple independent lines. After RNA samples treated with DNase (Promega) and purified using RNeasy Plant Mini Kit (QIAGEN) as described above, RNA samples were used to construct cDNA libraries on which Illumina deep sequencing was performed (Novogene). Novogene mRNA sequencing offered Illumina HiSeq platforms with paired-end 150 bp sequencing strategy, including sample quality control, library construction, library quality control, mRNA sequencing and data quality control.

Clean RNA sequencing reads were achieved after removing reads containing adapters, reads containing more than 10% of bases that cannot be determined, and reads containing low quality (Qscore≤ 5) bases which is over 50% of the total read (Novogene). Q score represents the sequencing quality score of a given base, defined as a property that is logarithmically related to the base calling error probabilities (P), $Q=-10\log_{10}P$ (https://www.illumina.com/documents/products/technotes/technote_Q-Scores.pdf).

Clean RNA sequencing reads of three wild-type and *MIM159* tobacco plants were mapped to the reference *Nicotiana tabacum* TN90 genome (Sierro et al., 2014, Sol Genomics) by Tophat V2.1.1 software (Trapnell et al., 2009; https://ccb.jhu.edu/software/tophat/index.shtml). The alignment results were analysed by Cuffdiff software (cufflinks V2.2.1, Trapnell et al., 2012; http://cole-trapnell-

lab.github.io/cufflinks/cuffdiff/) to detect differentially expressed genes between wild-type and *MIM159* tobacco. The fold change >=2 and the false discovery rate (FDR) ⩽ 0.01 were used to define the significantly differentially expressed genes.

Genes differentially expressed at the level of more than five-fold or ten-fold in *MIM159* tobacco leaves compared to wild type, were next subjected to Gene Ontology (GO) enrichment analysis using agriGO (http://bioinfo.cau.edu.cn/agriGO/). The tobacco genome is not completely annotated but provides annotations from the closest Arabidopsis homologues, which were used to determine the GO enrichment of the biological pathways. Pathways with a p-value of ⩽ 0.05 were considered as significantly enriched and were ranked by p-value.

Chapter 3

Target RNA secondary structure is a major determinant of miR159 efficacy

3.1 Introduction

MicroRNAs (miRNAs) are small 20-24 nucleotide RNAs that guide the miRNA induced silencing complex (miRISC) to target mRNAs and mediate their silencing through a combination of transcript degradation and translational repression (reviewed in Axtell, 2013; Rogers and Chen, 2013). Central to understanding miRNA function is the identification of their target mRNAs (Sun et al., 2014). In plants, it clear that high sequence complementarity between a miRNA and its target mRNA(s) is an absolute requirement for a miRNA-target mRNA interaction (Schwab et al., 2005; Addo-Quaye et al., 2008; German et al., 2008), with most experimentally validated miRNA-target pairs having very few mis-matches (Schwab et al., 2005; Liu et al., 2014). Consequently, miRNA-target complementarity has been the cornerstone of plant miRNA biology, determining miRNA target prediction (Dai and Zhao, 2011; Dai et al., 2018), the design of artificial miRNAs (amiRNAs; Schwab et al., 2006), the design of artificial miRNA decoys, such as target *MIM*s (Todesco et al., 2010) or the identification of endogenous target *MIM*s (Karakülah et al., 2016). However, bioinformatic prediction of target genes often fails to accurately predict functionally relevant targets, where from numerous predicted targets, only a select few appear functionally significant (reviewed in Li et al., 2014a). Likewise, it has been reported that amiRNAs with high complementarity to their intended targets perform with considerable variability in plants (Li et al., 2013a; Deveson et al., 2013). Finally, different miRNA decoys that contain identical miRNA binding sites work with widely varying efficacies (Reichel et al., 2015). These observations argue that miRNA-target interaction is not simply a product of complementarity, but additional factors are required for functional miRNA-target recognition and interaction (Wang et al., 2015).

In animals, it has long been known that the contextual sequence features in which a miRNA-binding site resides can strongly impact silencing. For example, miRNA-binding site accessibility was shown to be important, where introduction of mutations to decrease predicted accessibility disrupted efficient regulation, with impacts being as strong as mutations within the binding site itself (Kertesz et al., 2007). Furthermore, it has been shown that for certain animal miRNA-target interactions, strong regulation only occurs when the binding sites are within specific sequence contexts (Bartel, 2009; Didiano and Hobert, 2006; Vella et al., 2004). Factors such as these have been studied less in plants, but evidence is accumulating that the sequence context of the miRNA binding sites may also be important in plants. Firstly, Gu and colleagues (2012) found a synonymous codon bias favouring AU-richness, and hence reduced RNA secondary structure, around predicted miRNA target sites in several plant species. Secondly, Li et al. (2012b) have found that miRNA binding sites in Arabidopsis are generally less structured than their flanking regions, indicating a preference for high accessibility. Indeed, Fei et al. (2015) found that target site accessibility may explain select regulation of only a few targets from a large number of predicted target genes. Therefore, it would be of interest to functionally test these potential factors on miRNA-target interactions.

In plants, the Arabidopsis miR159 family has been extensively studied as a model for plant miRNA-mediated gene expression regulation (Palatnik et al., 2003; Palatnik et al., 2007; Allen et al., 2007; Allen et al., 2010). The family has two major isoforms, miR159a and miR159b, which are strongly expressed throughout Arabidopsis (Palatnik et al., 2007; Li et al., 2016). Such expression is consistent with a loss-of-function *mir159ab* double mutant that displays strong pleiotropic developmental defects (Allen et al., 2007). In Arabidopsis, miR159 is bioinformatically predicted to regulate more than twenty targets, including eight genes encoding conserved R2R3 domain *MYB* transcription factors

(Palatnik et al., 2007). Despite this, genetic analysis revealed that miR159-mediated regulation of only two of the predicted target genes, *MYB33* and *MYB65*, accounted for the developmental defects of the *mir159ab* double mutant, as all defects are suppressed in the *myb33.myb65.mir159ab* quadruple mutant (Allen et al., 2007). This defined the functional specificity of Arabidopsis miR159 being restricted to *MYB33* and *MYB65*, but also raised the question of the functional significance of miR159-mediated regulation of the additional bioinformatically predicted targets, including those that have a strongly conserved miR159 binding site (Allen et al., 2007, Allen et al., 2010). Curiously, this narrower functional specificity as defined by genetic studies has also been found in other plant and animal miRNA systems, suggesting the functional scope of miRNA-mediated silencing is narrower than generally assumed (Li et al., 2014a; Seitz, 2009).

For miR159, it is likely that multiple factors contribute to this apparently narrow functional specificity, including non-overlapping transcriptional domains of the *MIR159* and *MYB* target genes (Allen et al., 2007). In addition, whether regulation of other targets is important under certain untested growth conditions remains unclear (Allen at al., 2010). However, one untested hypothesis is that *MYB33* and *MYB65* are more sensitive to miR159 regulation than the other *MYB* target genes. Recently, we have shown that factors beyond complementary govern the efficacy of the miR159-*MYB33* silencing outcome (Li et al., 2014b). This not only included the miR159:*MYB33* transcript stoichiometry, but also the sequence context of the miR159 binding site in *MYB33*. We showed that mutation of nucleotides (nts) that immediately flank the miR159-binding site attenuated silencing to a similar extent to mutating nts within the binding site itself (Li et al., 2014b). This is further evidence that sequence complementarity alone does not guarantee strong miRNA regulation and that additional factor(s) also influence miRNA-mediated regulation in plants.

Here, by carrying out *in planta* miR159 efficacy assays, we show that *MYB33* and *MYB65* are indeed much more sensitive to miR159 regulation than the other *MYB* genes with conserved miR159 binding sites. *MYB33* and *MYB65* are shown to share a predicted RNA secondary structure consisting of two stem-loops that partially overlap with the miR159 binding site. This chapter aims to investigate the strong stem-loop RNA structures in *MYB33* and whether they determine miR159-mediated silencing efficacy. I speculate that having strong RNA stem-loop structures adjacent to the miR159 binding site may facilitate accessibility of the binding site to the miR159, which in turn promotes efficient miR159 silencing of *MYB33/MYB65*.

3.2 Results

3.2.1 Conserved *MYB* gene family members have different sensitivities to miR159 silencing

Despite bioinformatics predicting miR159 to target approximately 20 genes in Arabidopsis, including eight *MYB* genes that contain conserved and highly complementary miR159 binding sites, miR159 appears functionally specific to only *MYB33* and *MYB65* (Figure 3.1A, Allen et al., 2007). Although many possible factors likely underlie this apparent functional specificity (Allen et al., 2010), here we investigate whether *MYB33* and *MYB65* are more sensitive to miR159-mediated regulation compared to these other conserved *MYB* targets. In order to investigate this, *in planta* miR159-silencing efficacy assays were performed. Here, each target was fused to the 35S promoter to constitutively transcribe each target in Arabidopsis. Multiple primary transformants were next obtained for each *35S-MYB* transgene and their phenotype scored. If the *MYB* transcripts are efficiently silenced by miR159, no developmental abnormalities are likely to be observed, as endogenous miR159 appears to be widely expressed throughout Arabidopsis (Palatnik et al., 2007; Allen et al., 2007; Li et al., 2016). Conversely, if the *MYB* transcripts are poorly silenced by miR159, constitutive *MYB* expression in the resulting transformant lines will develop phenotypic abnormalities (Figure 3.1B), with their frequency and severity in the primary transformant populations acting as a measure of silencing. As these *MYB* genes differ in their amino acid sequence, scoring for each *35S-MYB* (wild-type miRNA binding site) transgenic population was compared to a corresponding miR159-resistant *35S-mMYB* (mutated miRNA binding site) transgenic population, with the difference in frequency and severity of phenotypic abnormalities acting as a measure of miR159 silencing efficacy.

The *35S-MYB* and *35S-mMYB* transgenes were generated for *MYB33*, *MYB65*, *MYB81*, *MYB97*, *MYB101*, *MYB104* and *DUO1*, and then individually transformed into Arabidopsis. Multiple Arabidopsis primary transformants were selected for each construct, followed by phenotyping to assess developmental abnormalities. Phenotypic severities of *35S-MYB* rosettes were placed into three broad categories; No Phenotypic (N) abnormalities, where the plant appears indistinguishable from wild-type; Mild (M), where the plant contains one or more moderately curled leaves; and Severe (S) where the plant has strongly curled leaves with the abaxial (bottom) side of two or more leaves visible from an aerial view (Figure 3.1B). Multiple primary transformants (>35) were scored as to negate the impact of transgene position effects.

Firstly, the majority of *35S-MYB33* and all of *35S-MYB65* primary transformants were phenotypically indistinguishable from wild-type. By contrast, all *35S-mMYB33* and *35S-mMYB65* transformants resulted in developmental defects (Figure 3.1C). Additionally, *CYSTEINE PROTEINASE 1* (*CP1*), a marker gene of MYB protein activity (Alonso-Peral et al., 2010; Li et al., 2016) was measured in pooled primary transformant lines, and a strong difference was found for the *MYB33*/*mMYB33* and the *MYB65*/*mMYB65* pairs (Figure 3.1D). The stark difference in both developmental phenotypes and *CP1* levels demonstrated that both *MYB33* and *MYB65* are strongly silenced by miR159.

For the other four *GAMYB-like* genes, *MYB81*, *MYB97*, *MYB101* and *MYB104*, no such stark difference was seen for the *MYB*/*mMYB* pairs. Firstly, expression of all four of these *35S-MYB* genes could result in plants with severe phenotypic defects, resembling *mMYB33* and *mMYB65* transgenic lines or the *mir159ab* mutant (Figure 3.2). For *35S-MYB81*, *35S-MYB97* and *35S-MYB101* primary transformants, the majority displayed phenotypic defects (Figure 3.1C). For *35S-MYB104* primary transformants lines, the

majority displayed no phenotypic defects; however, some *35S-MYB104* transformants displayed severe phenotypes (Figure 3.1C). More importantly, for each *MYB* gene, the frequency and severity of phenotypes in the *35S-MYB* primary transformants populations were only slightly lower than the corresponding *35S-mMYB* populations. Although this difference implies that *MYB81, MYB97, MYB101* and *MYB104* are still miR159-regulated, such small differences would argue that these genes are poorly silenced by miR159. Supporting this notion for *MYB81*, *MYB97* and *MYB101* is the *CP1* transcript levels in these three populations, which were strongly up-regulated in the *35S-MYB* populations compared to wild-type (Col-0), and which had smaller differences between the *MYB*/m*MYB* pairs, relative to the *MYB33*/*mMYB33* and *MYB65/mMYB65* pairs (Figure 3.1D). For *MYB104*, the low frequency of severe phenotypes in the *35S-mMYB* populations and low *CP1* levels indicate that this gene does not trigger this pathway as strongly as the other genes. However, again the lack of difference between the *35S-MYB104* and *35S-mMYB104* populations strongly argues that this gene is poorly silenced by miR159 (Figure 3.1C, D). In 35S-*MYB81*/*97*/*101* transformants, *MYB33* and *MYB65* levels remained unchanged, confirming that the observed phenotypes were due to the poor miR159-mediated silencing of *MYB81*/*97*/*101* (Figure 3.3).

Finally, *DUO1* encodes an atypical R2R3 MYB transcription factor that does not belong in the *GAMYB-like* family clade, but nevertheless its miR159 binding site is highly conserved (Palatnik et al., 2007). All *35S-mDUO1* primary transformants (n = 55) exhibited phenotypic defects of reduced rosette size but with downward curled leaves (Figure 3.1E), suggesting *DUO1* activates different developmental pathways compared to the *GAMYB-like* family members. Interestingly, all *35S-DUO1* plants (n = 54) exhibited similar phenotypic defects to the *35S-mDUO1* plants and similar activation of a downstream gene, *MGH3* (Figure 3.1F). The lack of a large difference between the *35S-*

DUO1 and *35S-mDUO1* primary transformant populations again implies that *DUO1* is also poorly silenced by miR159 relative to *MYB33*/*MYB65*.

In summary, the seven *MYB* genes that contain conserved miR159 binding sites appear to fall into two distinct categories regarding the efficacy of miR159 silencing; *MYB33* and *MYB65* appear to be very efficiently silenced, whereas *MYB81*, *MYB97*, *MYB101*, *MYB104* and *DUO1* appear poorly silenced. Given that the latter five genes have highly conserved miR159 binding sites that appear of equivalent complementarity to the miR159 binding sites of *MYB33*/*MYB65*, the difference in silencing efficacies seems surprising.

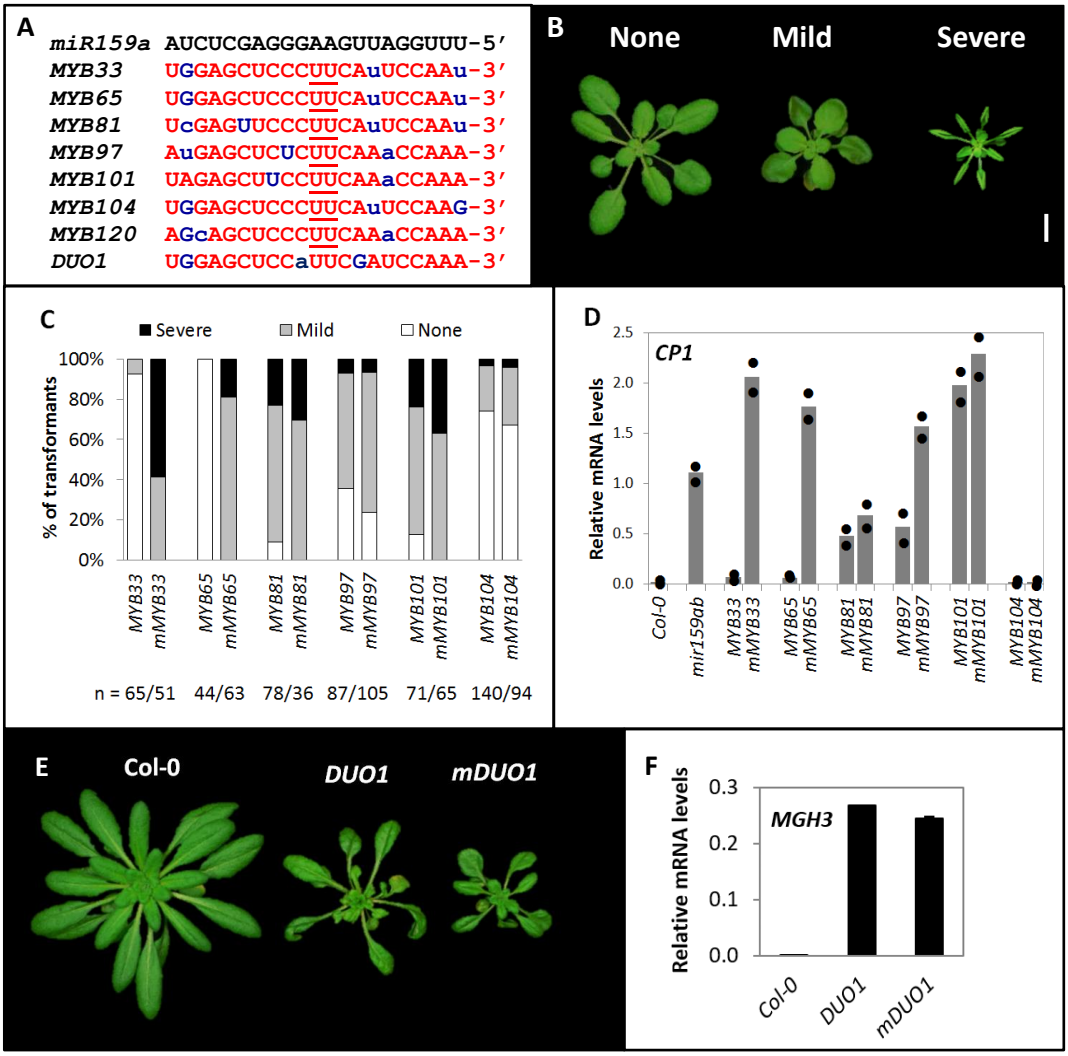


Figure 3.1. *MYB* target genes are differentially silenced by miR159. (A) Sequence of miR159-binding sites in *GAMYB-like* and *DUO1* transcripts. Nucleotides underlined represent the cleavage site, lower case blue letters represent mis-matches to miR159a, and G:U pairing is shown in uppercase blue letters. **(B)** Different phenotypic categories based on 35S-*GAMYB-like* expression; none (indistinguishable from wild-type), mild (display of some leaf curl) and severe (two or more leaves showing the abaxial side from an aerial view). Scale bar represents 1 cm. **(C)** Percentage of *35S-MYB* and *35S-mMYB* primary transformants showing None, Mild and Severe phenotypes. n = the number of primary transformants. **(D)** Transcript levels of the *MYB*-downstream gene *CP1* in *35S-MYB* and *35S-mMYB* primary transformants measured by qRT-PCR. RNA was extracted from two independent biological pools of 30-50 randomly selected, 15-day old primary transformants. Col-0 and *mir159ab* were used as controls, and mRNA levels were normalized to *CYCLOPHILIN*. The two measurements are shown as Dots, with the Bar representing the Mean. **(E)** Phenotypic characteristics of 40-day-old *35S-DUO1* and *35S-mDUO1* primary transformants. **(F)** Transcript levels of the *DUO1*-downstream gene *MGH3* in *35S-DUO1* and *35S-mDUO1* primary transformants. Measurements are the average of three technical replicates with error bars representing the standard error of the mean (SEM).

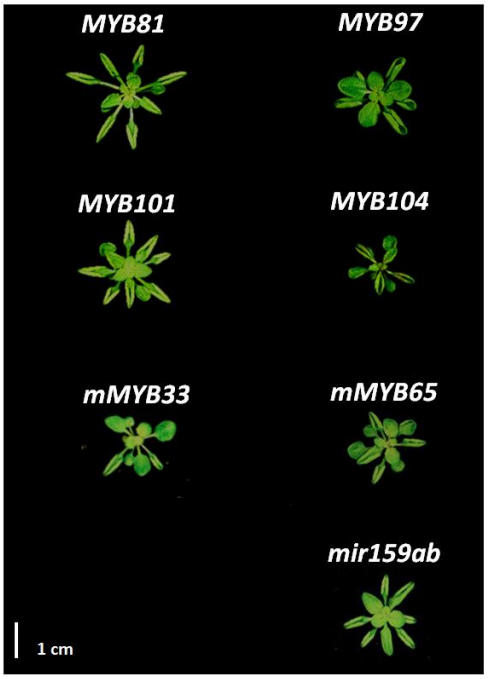


Figure 3.2. Expression of all *GAMYB-like* genes in rosettes results in similar phenotypic defects. Rosette phenotypes of 28-day old *35S-MYB* and *35S-mMYB* primary transformants grown under short-day conditions. Scale bar = 1 cm.

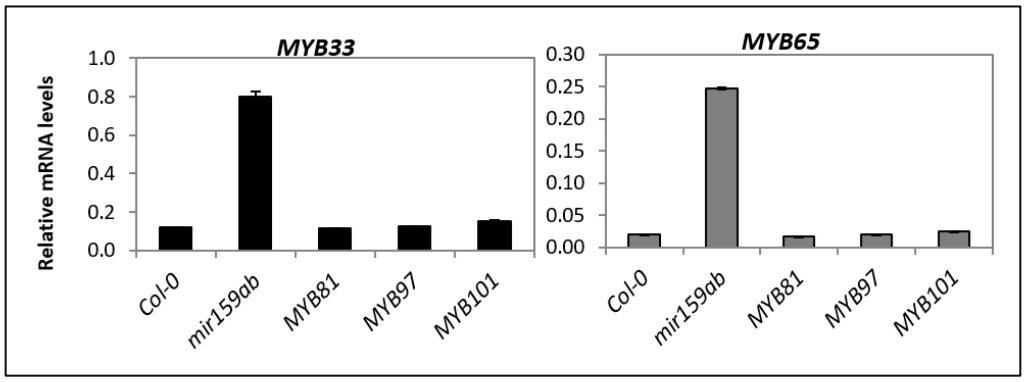


Figure 3.3. *MYB33* and *MYB65* transcript levels are unaltered in *MYB81, MYB97* and *MYB101* plants. qRT-PCR measurement of *MYB33* and *MYB65* mRNA levels in primary transformants of *35S-MYB81, 35S-MYB97* and *35S-MYB101*. The wild-type (Col-0) and *mir159ab* plants were used as controls. All measurements are relative to *CYCLOPHILIN*. RNA was extracted from pools of 30-50 randomly selected, 15-day old primary transformants. Measurements are the average of three technical replicates with error bars representing the SEM.

3.2.2 The ΔG free energy of a miRNA-target interaction is not an absolute determinant of silencing efficacy

Possibly explaining the stronger silencing of *MYB33*/*MYB65* is their stronger ΔG free energy interaction with miR159 compared to the other *MYB* genes (Figure 3.4), with the exception of *MYB104*, which has the strongest free energy, but is poorly silenced. To investigate this, we mutated the miR159 binding site of *MYB81* to make it identical to the miR159 binding site of *MYB33*, and investigated whether this modified *35S-MYB81* transgene (*MYB81-33*) was now as strongly silenced as *MYB33* by miR159. As the protein sequence of *MYB81-33* differs from *MYB81*, a miR159-resistant version (*MYB81-m33*) that was identical in amino acid sequence to *MYB81-33* was used as a control to ensure these amino acid changes did not impact protein activity (Figure 3.5A).

The incorporation of the miR159 binding site from *MYB33* into *MYB81* improved silencing, as over 60% of *MYB81-33* primary transformants had no phenotype as compared to less than 20% of *MYB81* primary transformants (Figure 3.5B). As the proportion of phenotypic categories was highly similar between *MYB81-m33* and *mMYB81* (Figure 3.5B), this difference between *MYB81-33* and *MYB81* cannot be attributed to an altered protein activity. However, silencing of *MYB81-33* was still not as strong as for *MYB33*, as a much greater proportion of *MYB81-33* plants exhibited mild and severe defects when compared to *MYB33* transformants (Figure 3.5B).

To support this data, we performed transcript profiling on the *MYB81* transgenes and the downstream marker gene of MYB protein activity, *CP1* (Alonso-Peral et al., 2010; Li et al., 2016). Consistent with miR159 having a strong translational repression component to its silencing of *MYB* targets (Li et al., 2014b), *MYB81* transcript levels do not reflect the differences in phenotypic severities of the different transgenic lines (Figure 3.5C). By

contrast, *CP1* levels tightly correlated with phenotypic severity, being highest in the lines carrying miR159-resistant versions (*mMYB81* and *MYB81-m33*), followed by *MYB81*, and lowest in *MYB81-33* (Figure 3.5C and Figure 3.6A). Therefore, consistent with the phenotypic severities, the high *MYB81* transcript levels in *MYB81-33* plant are being repressed to the greatest degree by miR159, as indicated by *CP1* levels.

Finally, *MYB33* transcript levels in *MYB81*, *mMYB81*, *MYB81-33* and *MYB81-m33* transformants remained unchanged, confirming that the observed phenotypes were due to miR159-mediated silencing of the respective *MYB81* transgenes (Figure 3.6B). Therefore, although incorporation of the *MYB33* miR159 binding site into *MYB81* has improved silencing, both morphological and molecular data indicated that *MYB81-33* is still not being as strongly silenced as *MYB33*, implying there are factors other than complementarity impacting silencing.

	ΔG hybrid
MYB33	-38.7
MYB65	-38.7
MYB81	-33.7
MYB97	-32.2
MYB101	-33.4
MYB104	-39.1
MYB120	-32.6
DUO1	-35.0

Figure 3.4. ΔG free energy of miR159-*MYB* target interactions. ΔG_{hybrid} was computed as the overall free energy of association for hybridisation of miR159a to the 21-nt miR159 binding site in different *MYB* target genes as calculated by the DINAMelt Web server (http://mfold.rna.albany.edu/?q=DINAMelt/Two-state-melting).

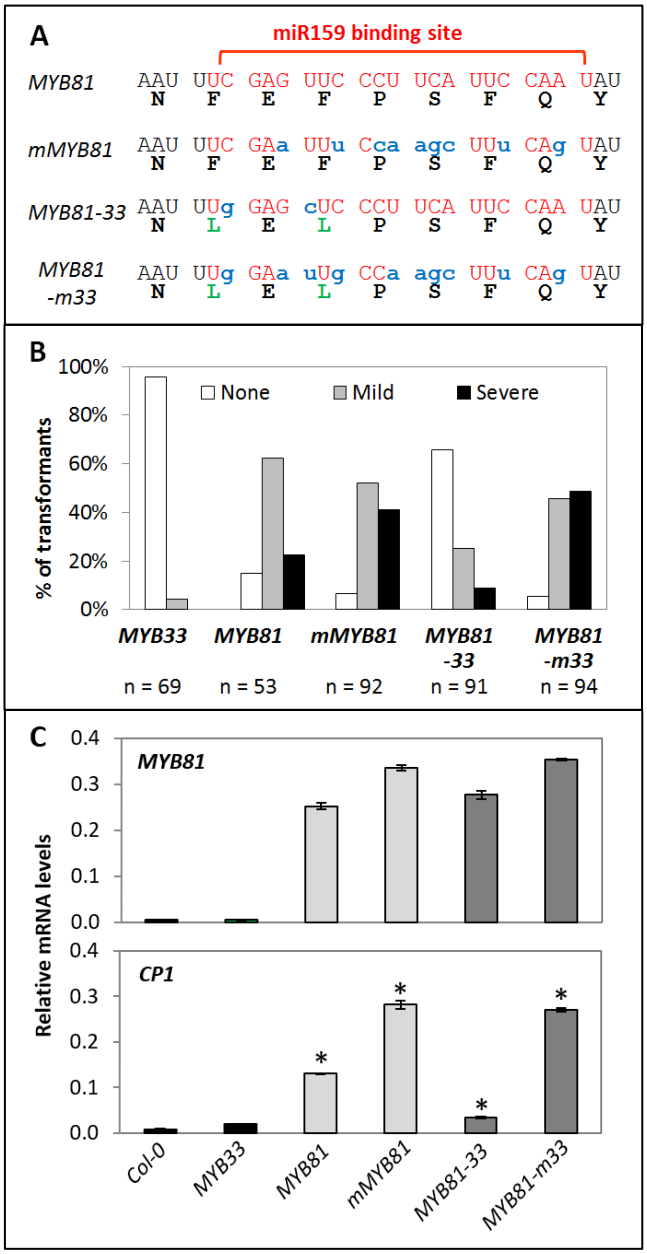


Figure 3.5. Lowering ΔG free energy of a miRNA-target interaction can improve silencing efficacy. (A) Nucleotide and amino acid changes introduced to generate *mMYB81*, *MYB81-33* and *MYB81-m33* compared to *MYB81*. The miR159 binding sequences are marked in red. Amino acid changes are marked in green; and nucleotide changes in blue. **(B)** The number of primary transformants of *MYB33*, *MYB81*, *mMYB81*, *MYB81-33* and *MYB81-m33* falling into each phenotypic category (none, mild or severe), as a percentage of the total number of transformants generated for each construct (n). **(C)** Transcript levels of *MYB81* and *CP1* in primary transformants of *MYB33*, *MYB81*, *mMYB81*, *MYB81-33* and *MYB81-m33* measured by qRT-PCR. RNA was extracted from pools of rosette tissue of 10-15 randomly selected, three-week-old primary transformants. Col-0 was used as wild-type control and mRNA levels were normalized to *CYCLOPHILIN*. Measurements are the average of three technical replicates with error bars representing the SEM. Values marked with * are significantly larger ($P < 0.05$) than corresponding measurements in *MYB33* plants as determined by *t*-test.

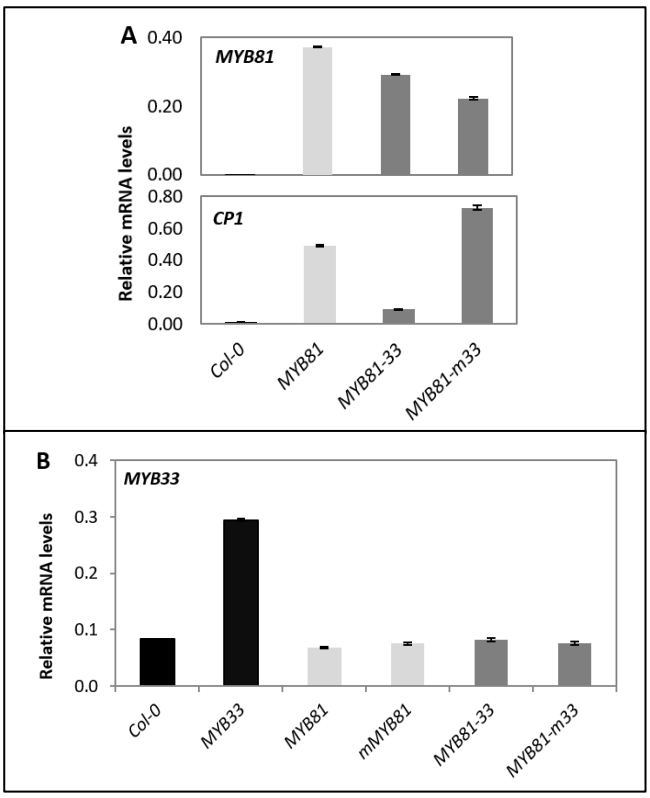


Figure 3.6. *MYB81*, *CP1* and *MYB33* transcript levels in various *MYB81* transgenic lines. (A) Transcript levels of *MYB81* and *CP1* in the second biological replicate of *MYB81*, *MYB81-33* and *MYB81-m33* transformants. **(B)** Transcript levels of *MYB33* in pools of primary transformants of *MYB33*, *MYB81*, *mMYB81*, *MYB81-33* and *MYB81-m33*. All measurements are relative to *CYCLOPHILIN* and wild-type (Col-0) was used as a control. RNA was extracted from pools of rosette tissue of 10-15 randomly selected, three-week-old primary transformants. Measurements are the average of three technical replicates with error bars representing the SEM.

3.2.3 The miR159 binding site of *MYB33* and *MYB65* abuts a strongly predicted RNA secondary structure

Another factor that is known to influence miRNA regulation in animals, but has not been given much attention to in plants is RNA secondary structure. Comparison of the predicted RNA secondary structure of *MYB33* and *MYB65* using the Vienna RNA fold server revealed that these transcripts have highly similar structures (Figure 3.7A). Of particular interest are the two stem-loop structures that are predicted to form immediately upstream the miR159 binding site (termed SL1 and SL2). These structures, which have

high base-pairing probabilities, are predicted to be present in *MYB* homologues of closely related species, such as *Arabidopsis lyrata,* and in that of more distantly related species such as wheat (*Triticum aestivum*), the basal angiosperm, Amborella, and the gymnosperm, Larix (Figure 3.7B). To investigate this further, my colleague performed nucleotide sequence alignments on *MYB33* homologues from dicotyledonous and monocotyledonous species separately, and then all species combined (generated by Gigi Wong). From these alignments, consensus sequences were determined for the region of the *MYB33* homologues that are predicted to form the stem-loop regions (Figure 3.7C). The nucleotides forming the stems of these structures were found to be the most strongly conserved nts in this region, and among the most highly conserved nucleotides in the entire transcript, with the sequences forming the stem of SL1, and the sequences forming the base of SL2, appearing very strongly conserved (Figure 3.7C). However, the consensus sequence of the upper part of SL2 varies between *MYB33* homologues from dicotyledonous and monocotyledonous species of plants, but in both instances, the sequences are still able to form stems, arguing that RNA secondary structure has being actively selected for in this region (Figure 3.7C). Supporting this, of the 25 conserved nts predicted to form stems in this predicted Arabidopsis *MYB33* RNA structure, 12 nts correspond to synonymous codon position regarding the amino acid sequence of the MYB33 protein (Figure 3.7D). Together, this argues that selection of these conserved nts is occurring mainly at the RNA secondary structure level, not at the protein level. Given their proximity to the miR159 binding site, it possibly suggests that these predicted structures have functional significance regarding miR159-mediated silencing. By contrast, SL1 and SL2 are not predicted to form in the other Arabidopsis *GAMYB-like* genes and *DUO1* (Figure 3.8), suggesting their presence maybe associated with strong silencing.

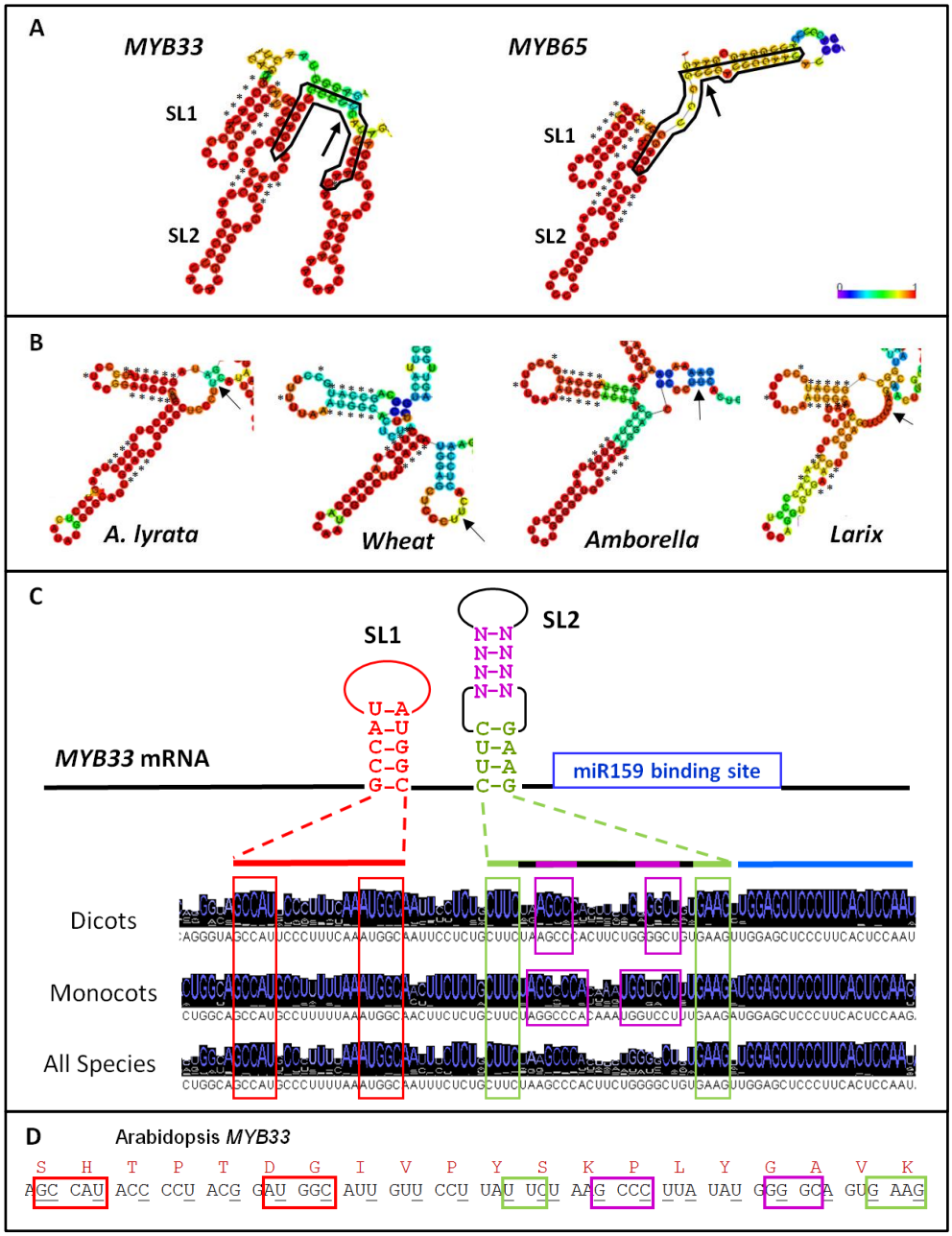


Figure 3.7. ***MYB33/MYB65*** **are predicted to contain highly similar RNA stem-loop structures abutting the miR159 binding site that appear strongly conserved. (A)** *MYB33* and *MYB65* transcripts contain two highly similar predicted stem-loop structures (SL1 and SL2) as determined by the Vienna RNAfold web server using the whole coding DNA sequences (CDS). Structure figure is a part of the whole structure of *MYB33/MYB65* CDS (Appendix Figure 1 and Appendix Figure 2). The miR159 binding sites are outlined in black and the cleavage sites indicated with arrows. The conserved nucleotides depending on the Figure 3.7C are labelled with asterisks. The heat maps indicate the probability of second structure formation, from low (purple) to high (red). **(B)** RNA secondary structures similar to SL1 and SL2 are also predicted to form in *MYB33* homologues of *Arabidopsis lyrata*, Wheat (*Triticum aestivum*), the basal angiosperm, *Amborella trichopoda*, and the gymnosperm, *Larix kaempferi* (Lamb.). The miR159 cleavage sites are indicated with arrows. Conserved nucleotides are marked with asterisks. **(C)**

MYB33/*MYB65* homologues from a diverse range of species contain conserved nucleotides corresponding to the predicted stem regions. Consensus sequences of *MYB33* homologues from dicots, monocots and all species combined (generated by Gigi Wong). Sequences which bind to form SL1 and SL2 are indicated in matching coloured boxes. The histogram indicates degree of conservation. Consensus sequence histogram and alignments are generated using Jalview (Waterhouse et al., 2009). **(D)** Many of the conserved nucleotide positions correspond to both synonymous and nonsynonymous nucleotides, implying conservation is not selected for at the protein level. Boxed nucleotides correspond to conserved stem positions as shown in Figure 3.7C, and underlined nucleotides correspond to synonymous positions.

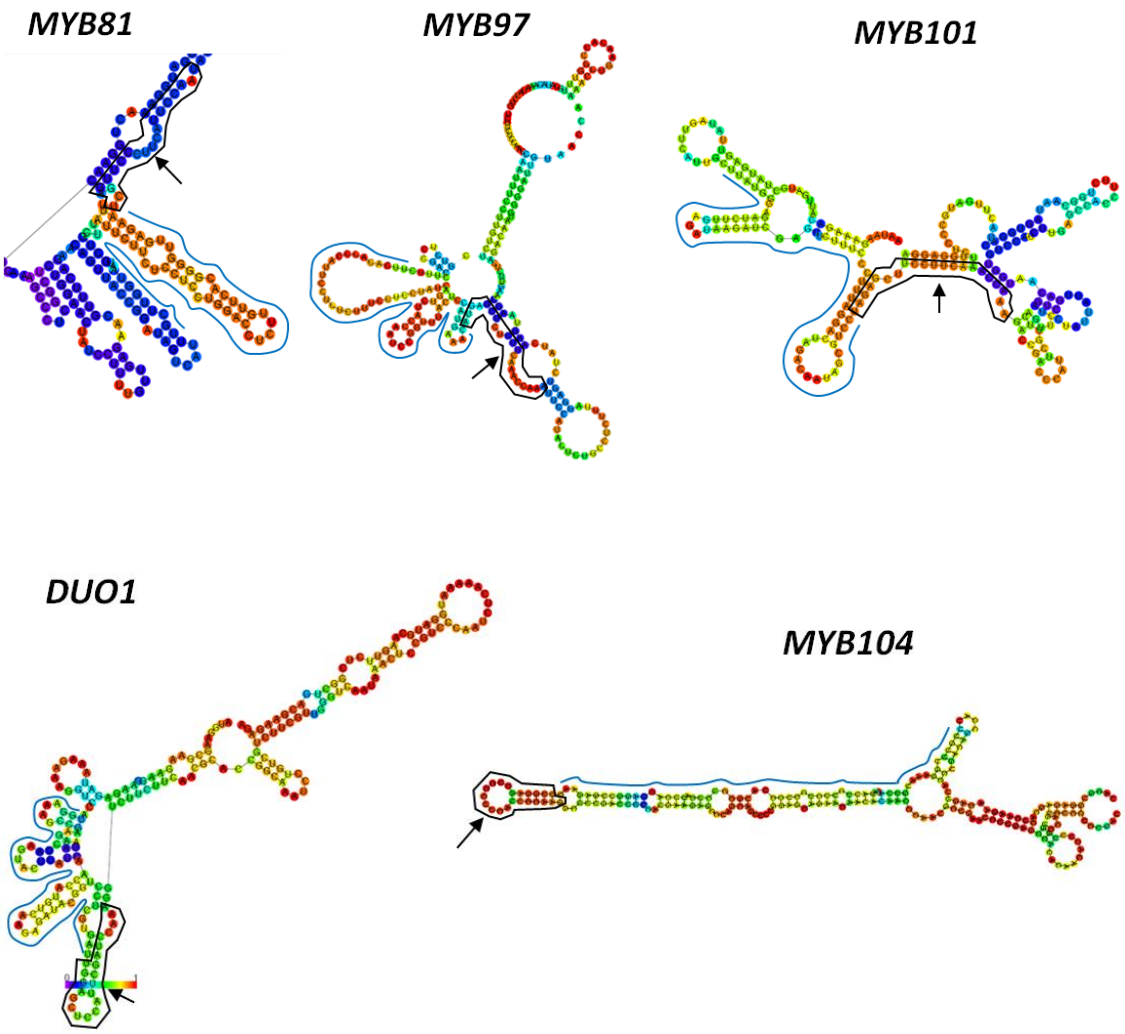


Figure 3.8. Predicted RNA secondary structures of *MYB* genes. Predicted RNA secondary structures of Arabidopsis *MYB81*, *MYB97*, *MYB101*, *MYB104* and *DUO1*. Structures were predicted via the Vienna RNAfold web server using a 221 bp window of the coding regions (miR159 binding site plus 100 bp immediately 5' and 3'). The miR159 binding site is outlined in black and the arrow indicates the miR159 cleavage site. The predicted secondary structures of SL1 and SL2 homologous, compared to *MYB33* and *MYB65*, is labelled with blue curve. The heat map indicates the probability of second structure formation, from low (purple) to high (red).

3.2.4 Mutations of either putative SL1 and SL2 structure attenuates strong *MYB33* silencing

To test whether, and potentially which of these predicted SL1 and SL2 structures are required for strong miR159-mediated silencing, synonymous mutations corresponding to SL1 or SL2 were introduced individually to disrupt these structures, but not changing the amino acid sequence. For SL1, the introduced mutations corresponded to seven nts in a region that was 40-60 nts upstream of the miR159 binding site, and the construct was termed *MYB33-mSL1* (Figure 3.9A). The mutations introduced into the SL2 forming corresponded to six nts that were more proximal, being 11-29 nts upstream of the miR159 binding site, and the construct was termed *MYB33-mSL2* (Figure 3.9A). Both *MYB33-mSL1* and *MYB33-mSL2* encoded a MYB33 protein identical to wild-type MYB33. Additionally, the mutations were predicted to result in a RNA secondary structure that does not disrupt the non-mutated SL (Figure 3.9B). These constructs, as well as a wild-type *MYB33* positive control, were individually transformed into *myb33* plants and multiple primary transformants were selected and classified according to their phenotypic defects shown as defined above [None (N), Mild (M) and Severe (S), Figure 3.9C].

Consistent with previous results, all of *MYB33* plants did not show any phenotypic abnormalities, demonstrating that *MYB33* is strongly silenced by miR159 (Figure 3.9D). However, for *MYB33-mSL1*, despite the distal nature of the mutations to the miR159 binding site, these mutations were able to perturb miR159 regulation. Compared to *MYB33* primary transformants that displayed no phenotypic defects, more than 60% of *MYB33-mSL1* primary transformants displayed mild or severe phenotypic defects (Figure 3.9D). For the more proximal mutations in SL2, a more severe perturbation of miR159 regulation was observed, where approximately 90% of transformants displayed mild or severe phenotypes (Figure 3.9D). Finally, *MYB33* and *CP1* mRNA levels were measured

in the *MYB33*, *MYB33-mSL1* and *MYB33-mSL2* plants. Levels of *MYB33* transcript will be dependent on at least two factors; firstly, the strength of transcription of the transgene, and secondly, the strength of miR159-mediated silencing, which includes transcript cleavage/degradation mechanism, as well as a translational repression mechanism (Li et al., 2014b). *MYB33* transcript levels were relatively higher in the *MYB33-mSL1* and *MYB33-mSL2* plants than in the *MYB33* plants (Figure 3.9E). In addition, *CP1* transcript levels were closely correlated with morphological phenotyping scoring, being most abundant in *MYB33-mSL2*, next in *MYB33-mSL1*, and lowest in *MYB33* and *myb33* plants (Figure 3.9E). Together, this morphological and molecular data argues that both SL1 and SL2 are required for strong miR159-mediated regulation of *MYB33*.

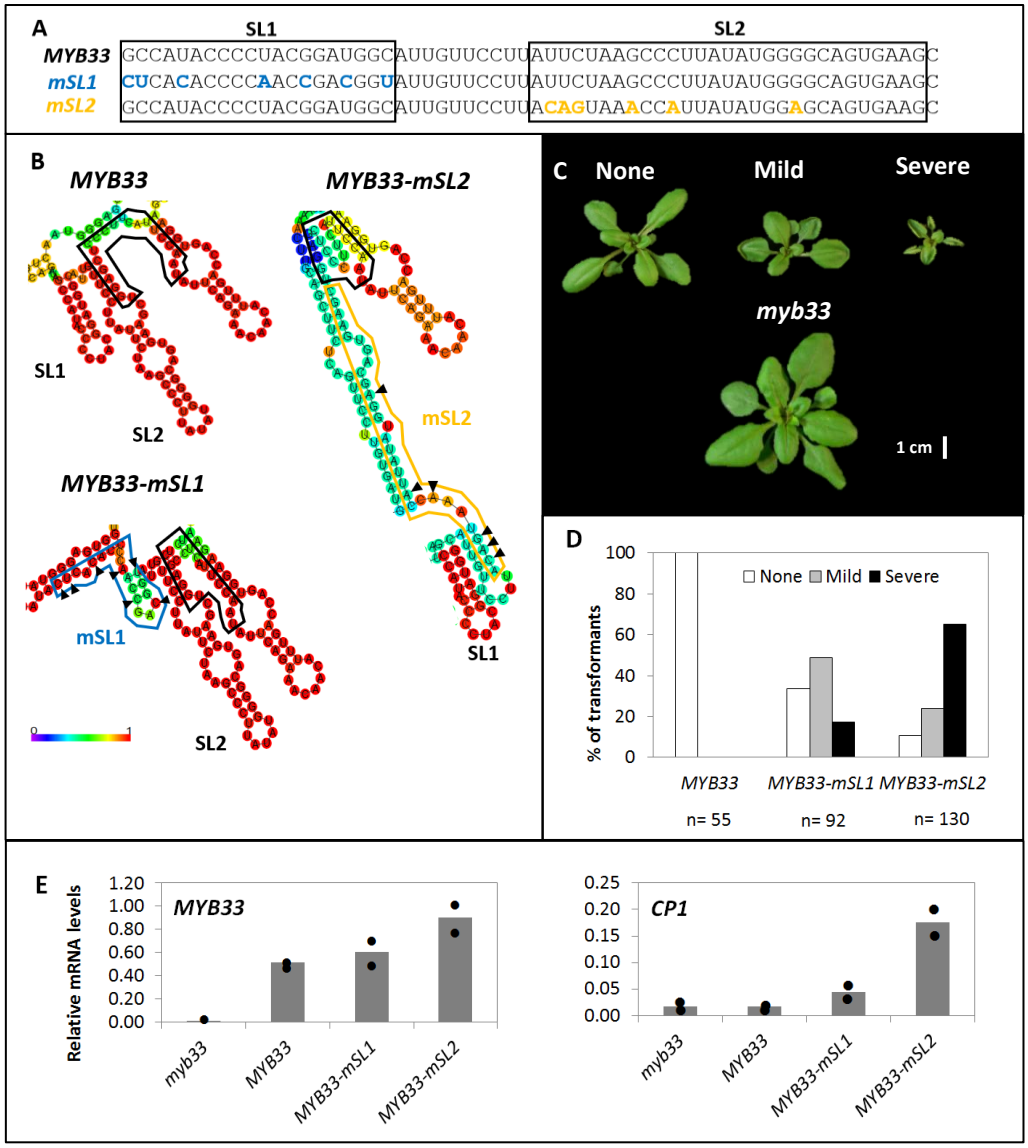


Figure 3.9. **Mutation of nucleotides corresponding to SL1 or SL2 attenuate miR159-mediated silencing of *MYB33*. (A)** Synonymous mutations were introduced to generate *MYB33-mSL1* (blue) and *MYB33-mSL2* (orange) compared to *MYB33*. The SL1 and SL2 sequences are marked in the black boxes, individually. **(B)** Predicted RNA secondary structure of *MYB33*, *MYB33-mSL1* and *MYB33-mSL2*. Structures were predicted via the Vienna RNAfold web server. The heat map indicates the probability of second structure formation, from low (purple) to high (red). The miR159 binding site is outlined in black. The altered mSL1 and mSL2 structure are outlined in blue and orange, individually. The mutated nucleotides in mSL1 and mSL2, regards to Figure 3.9A, are marked with "▲" on the structure figure. **(C)** Different phenotypic categories based on *MYB33-mSL* expression; none (indistinguishable from *myb33*), mild (display of some leaf curl) and severe (two or more leaves showing the abaxial side from an aerial view). The *myb33* plants served as control. Scale bar represents 1 cm. **(D)** Number of primary transformants falling into each phenotypic category, as a percentage of the total number of transformants analysed for each construct (n). **(E)** qRT-PCR measurement of *MYB33* and *CP1* mRNA levels in primary transformants of *MYB33*, *MYB33-mSL1* and *MYB33-*

mSL2. The *myb33* plants were used as a control. All measurements are relative to *CYCLOPHILIN*. RNA was extracted from two independent biological replicates, each being composed of rosette tissues sampled from 10-15 randomly selected, three-week-old primary transformants. The two measurements are shown as Dots, with the Bar representing the Mean.

3.2.5 An artificial predicted RNA stem-loop abutting the miR159 binding site promotes silencing

The above experiments raised the question as to what is the function of these sequences in promoting silencing. We speculate that if nts that are adjacent to a miRNA-binding site form a stem-loop structure, these nts are then unavailable to base-pair with nts of the miRNA-binding site, thus making the miRNA-binding site highly accessible for the miRNA. In an attempt to gain some insight into this possibility, we made an additional SL2 mutant, in which six nts were synonymously changed, but an alternative long stem-loop was predicted to be formed, which we called *MYB33-mSL2-2* (Figure 3.10A and B). Interestingly, despite the nt changes that alter the predicted RNA secondary structure, this transgene was found to be as strongly silenced as the wild-type *MYB33* transcript, as no phenotypic defects were seen in the more than 60 primary transformants screened (Figure 3.10C). Even though all six nt changes resulted in synonymous amino acid substitutions, a miR159-resistant version was also constructed (*mMYB33-SL2-2*) to ensure a functional protein is being made. As most *mMYB33-SL2-2* transformants had a severe phenotype, the lack of phenotypes seen in *MYB33-SL2-2* transformants is due to strong miR159-mediated silencing (Figure 3.10C). Supporting this strong silencing of *MYB33-SL2-2*, even though high levels of *MYB33* transcript accumulated, it appears to be totally silenced, as *CP1* mRNA levels remained low in *MYB33-SL2-2* plants, compared to the *CP1* levels in *mMYB33-SL2-2* plants (Figure 3.10D).

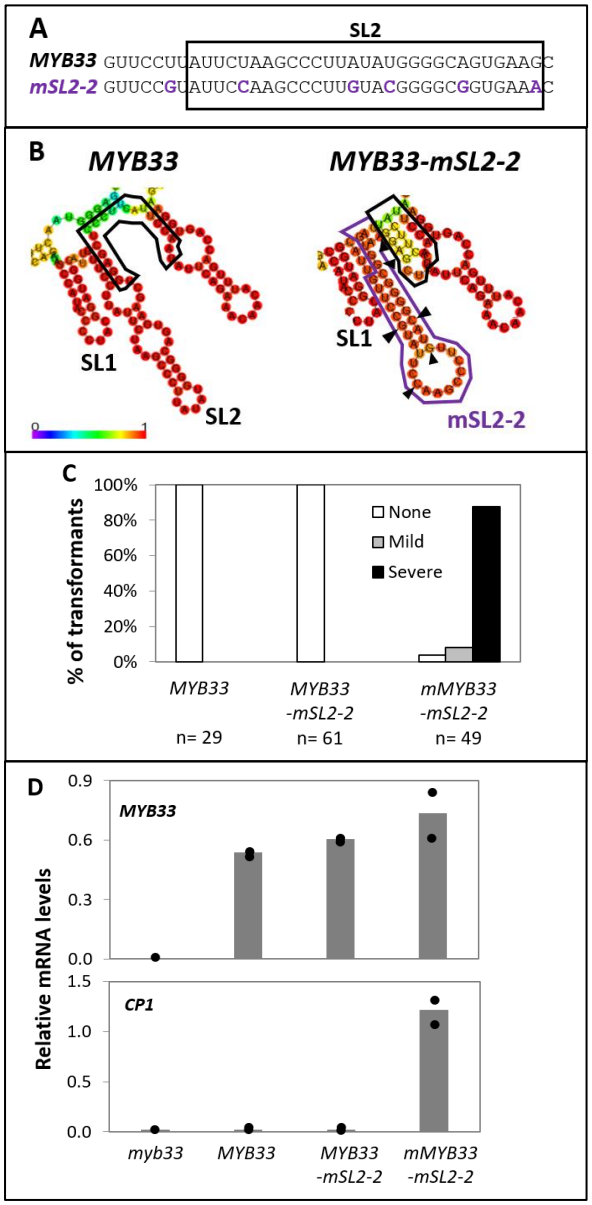


Figure 3.10. **An artificial stem-loop structure replacing mSL2 in *MYB33* promotes silencing. (A)** Synonymous mutations were introduced to generate *MYB33-mSL2-2* (purple) compared to *MYB33*. **(B)** Predicted RNA secondary structure of *MYB33* and *MYB33-mSL2-2*. Structures were predicted via the Vienna RNAfold web server. The heat map indicates the probability of second structure formation, from low (purple) to high (red). The miR159 binding site is outlined in black. The altered mSL2-2 structure is outlined in purple. The mutated nucleotides in mSL2-2, regards to Figure 3.10A, are marked with "▲" on the structure figure. **(C)** Number of primary transformants falling into each phenotypic category, as a percentage of the total number of transformants analysed for each construct (n). **(D)** qRT-PCR measurement of *MYB33* and *CP1* mRNA levels in primary transformants of *MYB33*, *MYB33-mSL2-2* and *mMYB33-mSL2-2*. The *myb33* plants were used as controls. All measurements are relative to *CYCLOPHILIN*. RNA was extracted from two independent biological replicates, each being composed of rosette tissue sampled from 10-15 randomly selected, three-week-old primary transformants. The two measurements are shown as Dots, with the Bar representing the Mean.

3.2.6 *MYB81* was engineered to have predicted SL structures, but not strongly silenced by miR159

As the conserved SL structures abutting the miR159 binding site facilitate the strong miR159-mediated silencing against *MYB33*, it is possible that engineering similar SL structures in the poorly-regulated miR159 targeted transcripts could promote their miR159-mediated silencing. As *MYB81* is not predicted to have these strong RNA secondary structures (Figure 3.11B), 13-nts of synonymous mutations were introduced to generate similar SL structures to those predicted for *MYB33*, with the resulting transgene termed, *MYB81-SL* (Figure 3.11A and B). However, this structure had an additional six nts compared to *MYB33*, therefore, another construct, *MYB81-SL-6nt*, was generated where these additional six nts between the two stem-loops were deleted, resulting in a predicted SL structures identical to that of *MYB33* (Figure 3.11B). Furthermore, the miR159 binding site of both *MYB81-SL* constructs was mutated to make it identical to the miR159 binding site of *MYB33*. As the protein sequence of *MYB81-SL* differs from *MYB81*, miR159-resistant versions *(mMYB81-SL* and *mMYB81-SL-6nt*) that were identical in amino acid sequence to *MYB81-SL* and *MYB81-SL-6nt* respectively, were used as controls to ensure these amino acid changes did not negatively impact on MYB81 protein activity.

Phenotypic scoring found that more than 50% of *MYB81-SL* and *MYB81-SL-6nt* transformants exhibited mild or severe mutant phenotypes, compared to the miR159-resistant versions (*mMYB81-SL* and *mMYB81-SL-6nt*), having around 80% of transformants with mutant phenotypes (Figure 3.11C). This indicated that the two *MYB81-SL* constructs were poorly regulated by miR159. Although a larger proportion of *MYB81-SL* and *MYB81-SL-6nt* transformants displayed a wild-type phenotype relative to *MYB81* transformants (Figure 3.11C), the slightly increased frequencies of “None” lines

were more likely to be caused by the incorporation of the *MYB33* miR159 binding site into *MYB81*, rather than effects of engineered SL structures. This can be supported by phenotyping scoring results of *MYB81*/*MYB81-33* transformants (Figure 3.5B). Therefore, the predicted SL structures engineered into *MYB81* are insufficient to improve miR159-mediated silencing efficacy.

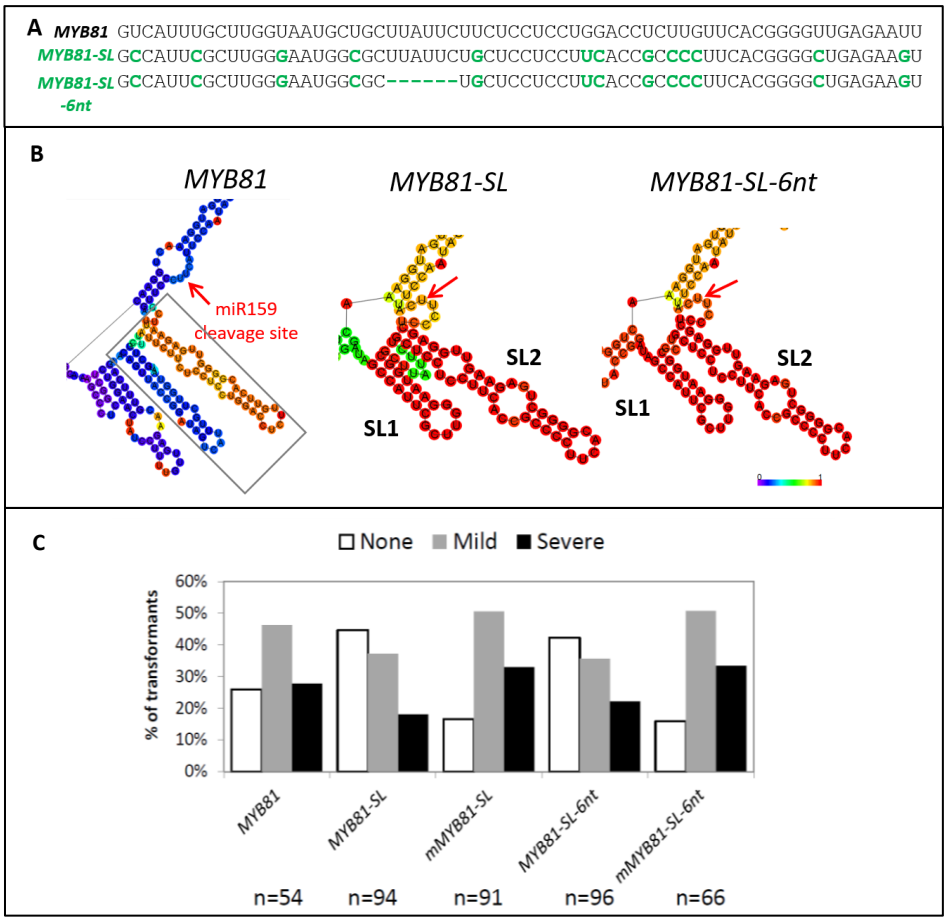


Figure 3.11. Engineering of predicted SL structures into *MYB81* does not increase sensitivity to miR159 regulation. (A) Mutations or deletions were introduced to generate *MYB81-SL* and *MYB81-SL-6nt* (green) compared to *MYB81*. **(B)** Predicted RNA secondary structure of *MYB81*, *MYB81-SL* and *MYB81-SL-6nt*. Structures were predicted via the Vienna RNAfold web server. The heat map indicates the probability of second structure formation, from low (purple) to high (red). The arrow indicates the miR159 cleavage site. The stem-loop RNA secondary structures in the low base-pair probability in *MYB81* is shown in the grey block. **(C)** Number of primary transformants falling into each phenotypic category, as a percentage of the total number of transformants analysed for each construct (n).

3.3 Discussion

Here, we have assessed the efficacy of miR159-mediated silencing *in planta* against multiple *MYB* target genes in Arabidopsis that contain conserved miR159 binding sites. Contrary to the possibility that miR159 regulates conserved *MYB* targets with equivalent efficacies, we show that there are differences in the degree of miR159-mediated silencing against its individual target genes.Although slight sequence composition of the individual miR159 binding sites can in part explain these differential efficacies, a predicted conserved RNA secondary structure, consisting of two stem-loops, that is associated with the miR159 binding site, was shown to confer efficient silencing of *MYB33*. This provides the rationale of why the nts flanking the miR159-binding site of *MYB33* are important in miRNA-mediated regulation (Li et al., 2014b), as the GAAG motif just upstream of the miR159 binding site was mutated, and this sequence composes the base of SL2 (Figure 3.7C). Moreover, this study represents the first evidence that RNA secondary structure can strongly impact the efficacy of miRNA-mediated target gene silencing in plants, and further implies the factor beyond complementary contributing to the narrow functional specificity of miR159 (Allen et al., 2007).

3.3.1 Differential miR159-mediated silencing of *MYB* family members.

Many plant miRNAs are predicted to target multiple paralogous members of gene families (Jones-Rhoades, 2012), including that of multiple *GAMYB-like* family members in many different species (Tsuji et al., 2006; Yang et al., 2014). However, how silencing varies amongst family members have not been extensively studied. Here, we have performed *in planta* efficacy assays on the *GAMYB-like* family members, as well as another *MYB* gene with a conserved miR159 binding site, *DUO1*. Despite each of the assessed genes containing a miR159 binding site that satisfies the widely accepted empirical parameters required for efficient miRNA-mediated regulation (Liu et al., 2014;

Schwab et al., 2005), we found that the strength of silencing varied dramatically, with *MYB33* and *MYB65* being much more sensitive to miR159-mediated regulation than the other *MYB* genes assessed. As miR159-guided cleavage products have been detected for *MYB81*, *MYB97*, *MYB101*, *MYB104* and *DUO1* (Allen et al., 2010; Alves-Junior et al, 2009; Palatnik et al., 2007), these are experimentally validated targets of miR159. However, given that only subtle differences were observed between the wild-type and miR159-resistant transgenes for all of these *MYB* targets, our *in planta* assays argue that these genes are weakly regulated by miR159, supporting genetic experiments that they are not major functional targets (Allen et al., 2007). Given that all these *MYB* targets have retained strong complementarity to miR159, the variability in silencing must be considered surprising.

Several studies have previously assessed the variation in silencing of related target genes. Firstly, quantitative transient assays in *Nicotiana benthamiana* to assess silencing efficacy of multiple family members of targets genes found that efficacy can vary dramatically amongst different family members, where subtle differences in complementarity can dramatically change efficacy (Liu et al., 2014). However, in this system, the miRNA binding sites were taken out of their endogenous sequence context and placed in the ORF or UTR of luciferase (LUC), therefore, what impact sequence context is having on these targets can no longer be accurately assessed. In another study, different *Medicago* miRNAs were over-expressed in Arabidopsis, with each miRNA potentially able to target a large number of *NBS-LRR* family members (Fei et al., 2015). However, all were found to only target a surprisingly few *NBS-LRR* genes (< 4%), to which differences in complementarity (or target score) could not explain and possibly resulting from target site accessibility determined by flanking sequences (Fei et al., 2015).

Such an observation adds further weight that factors in addition to complementarity have a large impact on a miRNA-target interaction.

This study has revealed that only *MYB33* and *MYB65* are sensitive to miR159 regulation, compared to other *MYB* targets, which can explain the genetic result that *MYB33* and *MYB65* account for the developmental defects displayed by the *mir159ab* double mutant. The findings presented here further suggest that the functional scope of miRNA-mediated silencing is narrower than generally assumed. Similar phenomenon appears to other plant and animal miRNA systems as well (Li et al., 2014a; Seitz, 2009). For example, the phenotypic defects in the loss-of-function *mir164abc* triple mutant are mainly caused by deregulation of two targets, *CUP-SHAPED COTYLEDON1* (*CUC1*) and *CUC2*, despite miR164 is validated to regulate five target genes with high sequence complementarity (Guo et al., 2005; Kasschau et al., 2003; Mallory et al., 2004a; Sieber et al, 2007).

Since *MYB33* and *MYB65* are widely transcribed throughout Arabidopsis development, and their expression has negative impacts on vegetative development (Allen et al., 2007; Millar and Gubler, 2005), this likely explains why they need to be sensitive to miR159 regulation at the post-transcriptional level. Meanwhile, *MYB97*, *MYB101* and *MYB120* are specifically transcribed in the pollen, where all three genes play important roles in pollen tube differentiation (Leydon et al., 2013; Liang et al., 2013). As these genes are not transcribed in vegetative tissues, the requirement for them to be strongly silenced by miR159 is absent, and therefore there is no selection for these genes to be sensitive targets of miR159. Therefore, the different transcriptional domains of the sensitive versus non-sensitive miR159 targets may underpin the differential silencing of miR159 towards the different *MYB* family members.

3.3.2 A conserved RNA secondary structure element promotes miR159 silencing of *MYB33*

We present two lines of evidence for the existence of a predicted RNA secondary structure which is important for miR159 silencing efficacy. Firstly, experimental, where we have carried out a structure/function analysis where strong miR159 regulation correlates with the predicted formation of the RNA secondary structure. Secondly, bioinformatic, where these structures, or variations thereof, are strongly predicted to reside in *MYB33* homologues of numerous angiosperm and gymnosperm plant species (Figure 3.7B, C), arguing that these structures have been integral in the miR159-*MYB* regulatory relationship over a long period of time. Interestingly, even though SL1 is more distal to the miR159 binding site, it appears more conserved than SL2, which varies in its consensus sequence between monocot and dicots species at the top of the predicted stem (Figure 3.7C). However, the consensus sequences both are still predicted to form stems (5'-AGCC and 5'-GGCU for dicots and 5'-AGGCCCA and 5'-UGGUCCU for monocots; Figure 3.7C). This variation again argues that a RNA secondary structure is being strongly selected for immediately adjacent to the miR159 binding site. Currently, no other miRNA binding site have been associated with a potential RNA structure, and very little is known about the role of RNA secondary structure in controlling plant gene expression.

Consistent with the conservation of nts corresponding to the stems of the RNA structures, we found that disrupting either predicted SL1 or SL2 attenuated efficient silencing of *MYB33*. Curiously, although more distal, the nts corresponding to SL1 appear more tightly conserved than the nts of the stem of SL2 (Figure 3.7C). The impact of the SL1 mutations must be considered surprising; seven synonymous mutations 40-60 nts upstream of the binding site would not be predicted to impact miRNA silencing with any accessibility program; for instance, the widely used target site accessibility program of

Kertesz et al., (2007) assesses 17 and 13 nts upstream and downstream of the binding site, respectively. Our result alone would suggest that a miRNA target site (binding site plus flanking nts) can be much larger than previously considered, and again highlights that current bioinformatic accessibility prediction programs likely need refinement in order to generate more accurate predictions.

3.3.3 Possible function of the predicted RNA structural element of *MYB33*.

We have not determined how these predicted structures are promoting silencing. One obvious possibility is that the stem-loop structures in the vicinity of the miR159 binding site maintain binding site accessibility for the miRNA, where if adjacent sequences are forming strong stem structures, they are less likely to base-pair with the binding site nts, thereby keeping the binding site accessible. Supporting such a notion is the *MYB33-mSL2-2* construct, which is predicted to form a different SL2, but appears as strongly silenced as the wild-type *MYB33* gene. This would argue that the presence of a stem *per se* is important, rather than the precise structure of the stem. This is consistent with the bioinformatic analysis of RNA structure of predicted miRNA targets in Arabidopsis, which has revealed less RNA structure in the miRNA binding sites, compared to the upstream and downstream flanking regions (Li et al., 2012b). Another supporting evidence is the STTM165/166 design, which contain two miRNA binding sites linked by a 48-nt spacer. This long length of spacer prefers to form a stable stem structure, which is used to stabilise the *STTM* transcript and make the two adjacent miRNA binding sites highly accessible (Yan et al., 2012).

However, this does not dismiss the possibility that RNA binding proteins (RBPs) interact with the stem-loop structures which potentially promotes silencing. In humans, the RBP Pumilio 1 (PUM1) binds to the stem RNA structure of the *p27* mRNA, which introduces

a local change in *p27* RNA structure and thereby enables efficient binding of miR-221 and miR-222 leading to *p27* silencing (Kedde et al, 2010). Whether such regulatory processes occur for miR159-mediated regulation of *MYB33* will require further investigation. However, it is interesting to note that the efficacy of miR159-mediated regulation of *MYB33* appears tissue dependent, which presents differential silencing efficacies between rosette and germinating seed (Alonso-Peral et al., 2012). This raises the possibility that a tissue-specific RBP-RNA interaction may mediate such differential regulation. Also uncertain is why engineering the SL RNA structure into the poorly regulated *MYB81* gene cannot improve its silencing, meaning these SL structures may be required for the promotion of strong silencing, but their presence *per se* is insufficient for strong silencing. This is supported by a poplar *GAMYB-like* homologue gene, whose transcript is predicted to have the stem-loop structures similar to Arabidopsis *MYB33*, was poorly silenced by miR159 when it was expressed in Arabidopsis (Kim et al., 2017). Therefore, how these RNA secondary structures impact miRNA silencing efficacy needs to be further investigated to understand the underlying mechanism. Understanding such principles may possibly have applications, where engineering specific RNA structures into a miRNA binding site may increase the sensitivity of the target to miRNA-mediated gene silencing.

3.3.4 Further investigating *in vivo* target RNA secondary structure is important for determining how to control miRNA silencing

This study provides the first direct *in vivo* evidence of target RNA secondary structure playing a major role in determining the efficacy of miRNA-mediated silencing in plants via a transgene-based approach. However, limitations still remain. Foremost, the design of RNA structures in *MYB33-mSL* constructs are dependent on bioinformatic prediction which may have no relation to what the real *in vivo* RNA secondary structures of *MYB33*

and *MYB33-mSL* transcripts are. This question can be in part resolved by the approach of dimethyl sulphate (DMS) sequencing, which has been established in plants to generate a map of *in vivo* RNA secondary structures with high experimental confidence (Ding *et al*., 2014; Kwok *et al*., 2013). Therefore, DMS sequencing could be performed on the *MYB33-mSL* transformants, exploring *in vivo* RNA secondary structures of the stem-loop region and the miR159 target site in *MYB33*. Furthermore, emerging DMS sequencing databases may provide opportunities for global analysis of *in vivo* RNA structures of miRNA targets, enabling the investigation of relationships between RNA secondary structures of target mRNAs and miRNA silencing efficacy.

3.3.5 MiRNA target recognition is still poorly understood

Presently, studies that have investigated the efficacy of miRNA-target mRNA interaction have focused solely on complementarity (Schwab et al., 2005; Iwakawa and Tomari, 2013; Liu et al., 2014). However, this study re-enforces that miRNA target recognition is a complex process involving multiple factors, which are far from fully appreciated. Consistent with strong sequence complementarity being a prerequisite for miRNA recognition, we have shown that replacing the miR159 binding site in *MYB81* with that of *MYB33*, which increases the free energy when bound to miR159, results in more efficient regulation. However, this change alone does not render *MYB81* as strongly silenced as *MYB33*, where it appears that both a favourable free energy of miR159 binding and the surrounding RNA landscape are required for strong silencing. Therefore, our study adds to the evidence that factors beyond complementarity, including target site structure, play a vital role in miRNA-mediated regulation of target gene expression. Just as miRNA-target complementarity or miRNA abundance can impact silencing efficacy and directly affect major plant traits (Todesco et al., 2012; Houston et al., 2013), our data would predict that sequence context likely has such impacts as well.

How common such potential RNA secondary structures are in controlling plant miRNA-mediated gene regulation is yet to be determined. One possible constraint on such structures arising is that most canonical miRNA binding sites reside in coding regions, where the amino acid sequence may limit the possibility of such structures occurring. Thereby, RNA secondary structures may be more likely to arise for the miRNA binding sites that reside in non-coding regions, or regions where the amino acid sequence conservation is not critical. Nevertheless, it is becoming clearer that the flanking sequences of miRNA binding sites can strongly impact miRNA-mediated silencing. Although algorithms used to predict targets and design amiRNAs take accessibility into account, it is clear that such analysis is rudimentary and will need much refinement to enable the development of more accurate miRNA target prediction programs.

Chapter 4

Investigating function of the miR159-*GAMYB* pathway in diverse plant species

4.1 Introduction

MiRNAs are a class of small RNAs that mediate silencing of target genes. In plants, many miRNA-target gene relationships are ancient, being conserved in multiple plant lineages of plant species. These target genes of highly conserved miRNAs mainly encode highly conserved regulatory proteins critical for a variety of biological processes, including plant development, stress response and plant defence response (reviewed in Chen, 2009; Rubio-Somoza and Weigel, 2011; Sunkar et al., 2012; Weiberg et al., 2014). Therefore, these ancient miRNA-target relationships are fundamental to plant biology.

One such miRNA-target relationship is the miR159-*GAMYB* pathway. The miR159 family in plants appears ancient, where highly conserved homologues are present in basal vascular plants to angiosperms (Axtell et al., 2007; Li *et al*., 2011). From sRNA sequencing experiments in a variety of plant species, it is apparent that miR159 is often highly abundant and widely expressed (Fahlgren et al., 2007; Jeong et al., 2011; Khatabi et al., 2016; Rajagopalan et al., 2006; Szittya et al., 2008). So far, all evidences demonstrate that *GAMYB* or *GAMYB-like* genes are the major targets of miR159, which encode conserved R2R3 MYB transcription factors (Csukasi et al., 2012; Li et al., 2013b; Millar and Gubler, 2005; Schwab et al., 2005; Tsuji *et al*., 2006). In Arabidopsis, it has been shown that miR159 mediates the post-transcriptional silencing of these *GAMYB-like* genes via two mechanisms: transcript cleavage and translational repression (Allen et al., 2010; Li et al., 2014b).

The functional role of these GAMYB transcription factors has been investigated in a number of different plant species. GAMYB was initially identified as a positive transcriptional regulator of the expression of GA-dependent α-amylase and hydrolytic enzymes in the barley aleurone cells (Gubler et al., 1995; Gubler et al., 1999).

Furthermore, GAMYB positively regulates GA-mediated programmed cell death (PCD) in the seed aleurone of both barley and Arabidopsis (Alonso-Peral et al., 2010; Guo and Ho, 2008). In Arabidopsis, the vacuolation of aleurone cells, a PCD-mediated process, is impaired in the aleurone of the *myb33.myb65.myb101* triple mutant (Alonso-Peral et al., 2010). Similarly in barley, RNA interference of *GAMYB* expression reduces PCD-associated vacuolation (Guo and Ho, 2008). Likewise, a conserved role of GAMYB in PCD of anther development exists in both dicot and monocot plant species. In the anthers of both the Arabidopsis *myb33.myb65* and rice *gamyb* loss-of-functional mutants, the tapetum fails to undergo PCD and degenerate, instead the tapetum abnormally expands to display tapetal hypertrophy, resulting in pollen abortion (Aya et al., 2009; Millar and Gubler, 2005). Therefore, GAMYB has conserved roles of PCD in both aleurone and tapetum development in diverse plant species.

By contrast, there is conflicting evidence of whether the miR159-*GAMYB* pathway controls flowering time. Expression of a 35S-miR159 transgene in Arabidopsis ecotype Landsberg delays flowering under short-day conditions (Achard et al., 2004), but does not change flowering time in Arabidopsis ecotype Columbia (Schwab et al., 2005). In another research, the *myb33.myb65* double mutant Arabidopsis displayed a similar flowering time as the wild type plants (Alonso-Peral et al., 2010). In the ornamental dicot species, Gloxinia, the expression of 35S-*MIR159* transgene was also reported to delay flowering time (Li et al., 2013c). Therefore, the role the miR159-*GAMYB* pathway plays in flowering time does not appear strongly conserved.

However, the functional role of the miR159-*GAMYB* pathway in vegetative tissues is not clear. In Arabidopsis, de-regulated expression of *MYB33* and *MYB65* in the *mir159ab* loss-of-functional mutant results in a smaller sized rosette with severe upward leaf curl

(Allen et al., 2007; Alonso-Peral et al., 2010). By contrast, a *myb33.myb65* double mutant displays a wild-type phenotype (Allen et al., 2007). This is curious, as although *MYB33* and *MYB65* appear widely transcribed throughout vegetative tissues, multiple lines of evidence readily suggest that they are strongly silenced by miR159 to levels that have very little phenotypic impact (Alonso-Peral et al., 2010). A similar scenario appears to be the case in rice. No phenotypic difference was observed between *gamyb* and wild-type rice plants during vegetative development (Kaneko et al., 2004). However, the deregulation of *GAMYB* in short tandem target *MIM159* (STTM159) rice contributed to reduced length of flag leaf and diameter of stems (Zhao et al., 2017). Therefore, in both Arabidopsis and rice, efficient miR159-mediated silencing of *GAMYB* is critical for normal vegetative development. This raises a question: why are *GAMYB-like* genes widely transcribed in vegetative tissues only to be strongly silenced by miR159? One possibility is that the GAMYB protein does play a subtle role in vegetative development. Indeed, Guo et al. (2017) found the Arabidopsis *myb33* mutant has a subtle but significant alteration in vegetative phase change, suggesting that the low level of MYB33 protein expression occurs in the rosette. Another possibility is that during stress, miR159 levels decrease, leading to elevated GAMYB accumulation. However, despite testing against multiple abiotic and biotic stresses, the miR159-*MYB33/MYB65* pathway in Arabidopsis had no apparent functional role in response to these stresses (Li et al., 2016). However, another report suggests the miR159-*MYB33/MYB65* pathway may play a role during heat stress response, but without any strong lines of evidence (Wang et al., 2012).

Therefore, during vegetative growth, *GAMYB* is widely transcribed only to be strongly silenced by miR159. The miR159-*GAMYB* pathway seems to be futile cycle. We are curious to determine whether the miR159-*MYB33*/*MYB65* relationship is specific to Arabidopsis, or if this relationship occurs in distant plant species. To investigate this, a

miR159 loss-of-function was developed. As most plant species have multiple functionally redundant miR159 family members (Table 1.3), a T-DNA loss-of-function approach is impractical. Previous reports have established molecular technologies to inhibit entire miRNA families in plants, including *SPONGE*s, target *MIMIC*s (*MIM*s) and STTMs (Reichel et al., 2015; Yan et al., 2012; Todesco *et al*., 2010). Considering that *MIM159* has a stronger efficacy than *SPONGE159* and STTM159 in Arabidopsis (Reichel et al., 2015), *MIM159* was chosen for this study. *MIM*s act as a decoy to sequester target miRNAs, due to a three- nucleotide loop at the cleavage site, a feature proposed to prevent cleavage and then subsequent sequestration of the miRNA from further target mRNAs (Todesco *et al*., 2010). In this chapter, *MIM159* transgenes were introduced into Arabidopsis, tobacco and rice, representative model species for the dicot and monocot plants; and further RNA sequencing analysis of *MIM159* transformants was used to determine *GAMYB* downstream events. This study aims to investigate the functional roles of the miR159-*GAMYB* pathway, and analyse the conserved nature or species-specific divergence in distant plant species.

4.2 Results

4.2.1 Expression of a 35S-*MIM159* transgene results in similar phenotypic outcomes in different *Arabidopsis thaliana* ecotypes

Currently, the majority of studies on the miR159-*GAMYB* pathway in Arabidopsis has been on the ecotype Columbia (Col-0), with only a single study performed on a second ecotype, Landsberg (Achard et al, 2004), generating conflicting data regarding the role of miR159-mediated silencing in flowering. Therefore, to determine whether miR159 function is conserved across different *Arabidopsis thaliana* ecotypes, a 35S-*MIM159* transgene was transformed into the ecotypes Col-0, C24 and Landsberg (Ler-1) to generate loss-of-function *mir159* phenotypes. All three Arabidopsis ecotypes transformed with the *MIM159* transgene displayed similar phenotypic defects, especially, reduced rosette size with upward leaf curl (Figure 4.1). On the basis of these data, the phenotypic impact of de-regulated *GAMYB* expression appears conserved across different Arabidopsis ecotypes.

Figure 4.1. Phenotypes of *MIM159* Arabidopsis Col-0, C24 and Ler-1 ecotypes. Three-week-old primary transformants that were grown under long-day conditions all displayed similar phenotypic defects: mild to severe up-ward leaf curl and reduce rosette size in Col-0, C24 and Landsberg (Ler-1) ecotypes. Scale bar represents 1 cm.

4.2.2 Investigating the miR159-*GAMYB* pathway in tobacco.

Next, the miR159-*GAMYB* pathway was investigated in a species from the Solanaceae family, *Nicotiana tabacum* (tobacco). According to miRBase (Kozomara and Griffiths-Jones, 2014), *Nicotiana tabacum* (*Nt*) has a single *Nt*-miR159 isoform, whose 21-nt sequence is identical to *Arabidopsis thaliana At*-miR159a (Figure 4.2A; Allen et al., 2010; Tang et al., 2012). Regarding target genes, there are three *GAMYB* or *GAMYB-like* homologues in *Nicotiana tabacum*, which are referred to here as *NtGAMYB1/2/3* (corresponding to the NCBI Reference accessions of XM_016589328, XM_016599515, XM_016629135, respectively). Based on nucleotide and amino acid sequence alignments, *NtGAMYB1* and *NtGAMYB2* are more closely related, with *NtGAMYB3* being the more divergent homologue (Figure 4.2B and C).

A

***At*-miR159a**	**AUCUCGAGGGAAGUUAGGUUU-5′**
***Nt*-miR159**	**AUCUCGAGGGAAGUUAGGUUU-5′**
NtGAMYB1/2	**UGGAGCUCCCUUCAcUCCAAA-3′**
NtGAMYB3	**UcGAGCUCCCUUCAcUCCAAA-3′**

B. Nucleotide identity	*NtGAMYB1*	*NtGAMYB2*	*NtGAMYB3*
NtGAMYB1		94%	82%
NtGAMYB2	94%		82%
NtGAMYB3	82%	82%	

C. Amino acid identity	*NtGAMYB1*	*NtGAMYB2*	*NtGAMYB3*
NtGAMYB1		93%	73%
NtGAMYB2	93%		73%
NtGAMYB3	73%	73%	

Figure 4.2. Sequence similarity of tobacco miR159 and *GAMYB* homologues. (A) Sequence alignments of tobacco mature miR159 and the miR159 binding sites in *NtGAMYB* transcripts. Nucleotides underlined represent the cleavage site, lower case blue letters represent mis-matches to Nt-miR159, and G:U pairing is shown in uppercase blue letters. *At*, *Arabidopsis thaliana*; *Nt*, *Nicotiana tabacum*. **(B)** Nucleotide and **(C)** amino acid sequence identity across three *NtGAMYB* homologues.

To determine whether miR159 and *GAMYB* are transcribed in vegetative and reproductive tissues of tobacco, mature miR159 and *GAMYB* levels were measured by quantitative real-time PCR (qRT-PCR). Due to high sequence similarity of the *NtGAMYB* homologues, primers were designed to amplify the 3'-UTR regions as to ensure gene-specific amplicons. As these primers do not span the miR159 cleavage site, these primers will measure both cleaved and uncleaved *GAMYB* transcripts. The miR159 was found to accumulate in both leaves and flowers of tobacco (Figure 4.3). *NtGAMYB1* and *NtGAMYB3* transcripts were more abundant in flowers than in leaves; while *NtGAMYB2* returned similar abundance in leaves and flowers (Figure 4.3). Therefore, similar to Arabidopsis, both mature miR159 and *GAMYB* mRNA are present in vegetative and reproductive tissues (Axtell and Bartel, 2005).

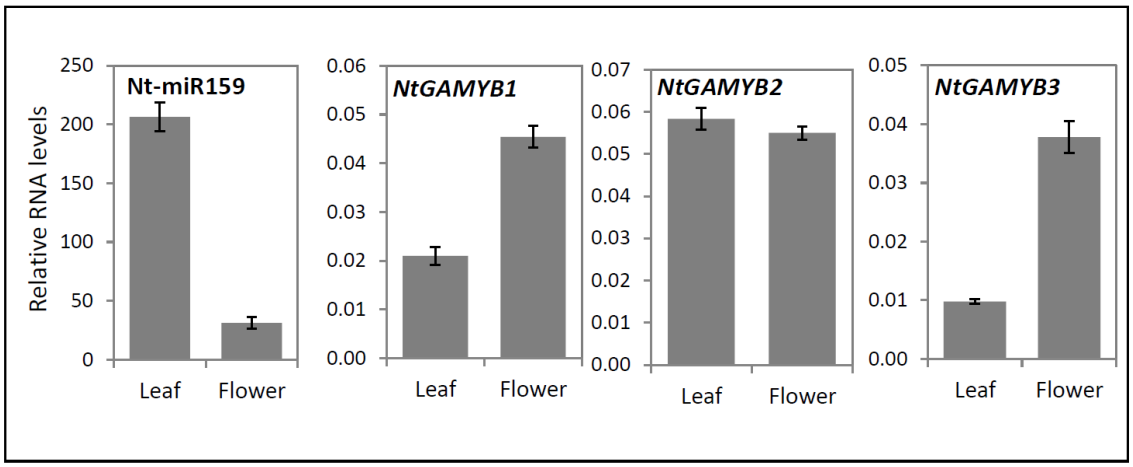


Figure 4.3. qRT-PCR measurement of miR159 and *GAMYB* transcript levels in tobacco. RNA was extracted from pools of two *Nicotiana tabacum* plants, each being composed of two young leaves and five mature open flowers. The mRNA levels of the *NtGAMYB* genes were normalized to *PROTEIN PHOSPHATASE 2A* (*PP2A*, Liu et al., 2012) and the miR159 RNA levels were normalized to snoR101. Measurements are the average of three technical replicates with error bars representing the standard error of the mean (SEM).

Next, since all tobacco *GAMYB* homologues contain a highly complementary miR159 binding site, it was investigated whether they undergo miR159-mediated cleavage. To investigate this, a 5' Rapid Amplification of Complementary DNA Ends (5'-RACE) miRNA cleavage assay was used to identify miR159-guided cleavage products. Here, the 3' end product of a miRNA-guided cleavage event is ligated with a 5' RNA adapter and then PCR products are amplified using adapter and *NtGAMYB* gene-specific primers. These PCR products are then examined by agarose gel electrophoresis; if they correspond to the expected size and then they are sequenced to examine whether they represent miR159-guided cleavage products (Figure 4.4A, Li et al., 2014b).

The miR159 cleavage assay in tobacco revealed that sizes of PCR products were as expected, being observed for all three *NtGAMYB* homologues. Sequencing of these PCR products determined that all three *NtGAMYB* homologues are cleaved by miR159, with the *NtGAMYB*-adapter DNA junction corresponding to the canonical miR159 cleavage site (Figure 4.4B, C and D). Although these cleavage assays are non-quantitative, they experimentally validate that these *NtGAMYB*s are targets of miR159-mediated regulation in both leaves and flowers of tobacco.

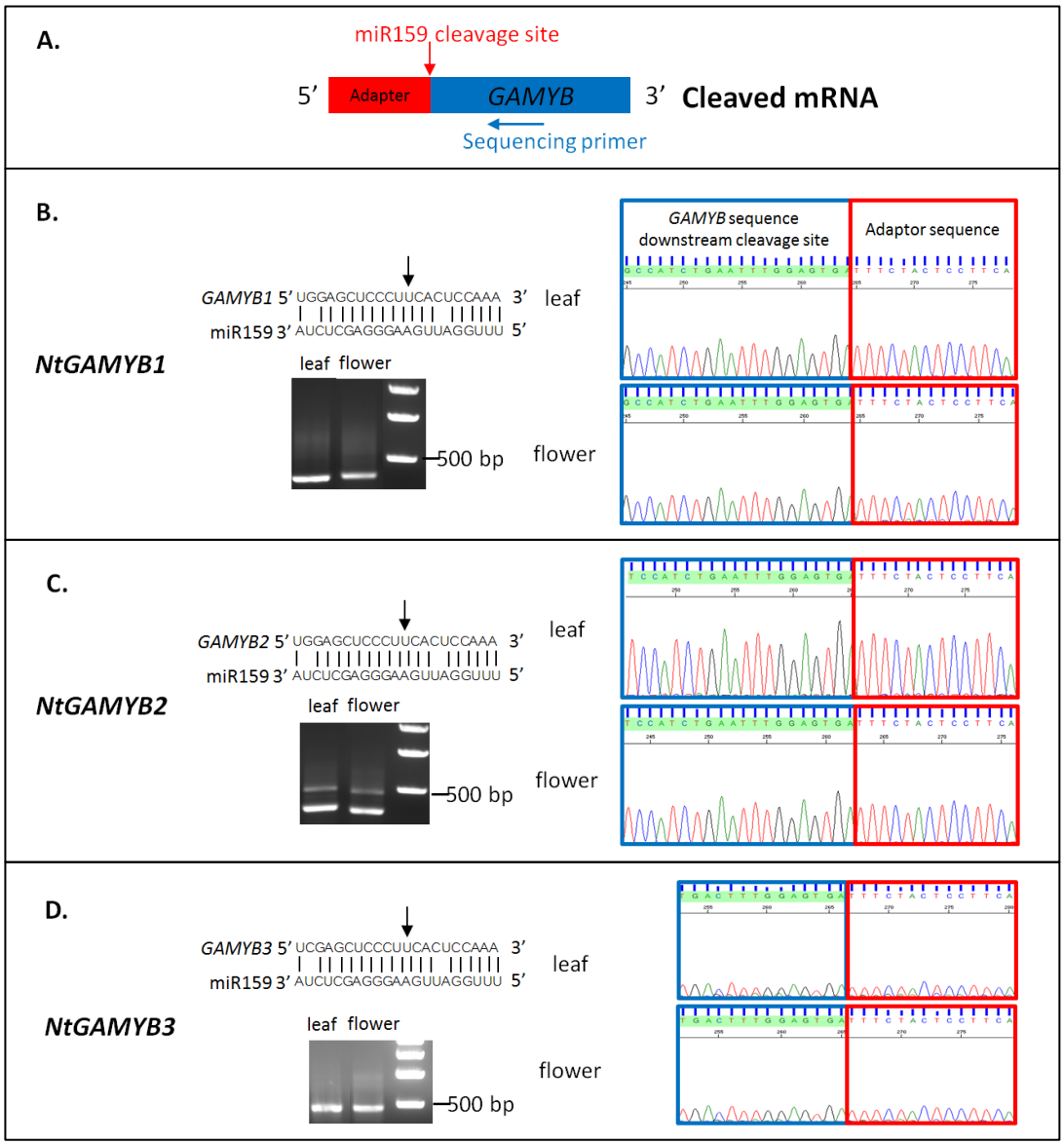


Figure 4.4. Investigating the miR159-mediated cleavage of *NtGAMYB* genes in tobacco. (A) Schematic representation of PCR amplicon of *GAMYB*-adapter DNA junction from the 5'-RACE miRNA cleavage assay (Li et al., 2014b). Red box represents RNA Oligo adaptor; blue box represents *GAMYB* sequences downstream the cleavage site; red arrow represents the miR159 cleavage site. Blue arrow represents the position of *GAMYB*-specific primer in the next DNA sequencing. **(B, C and D)** mRNAs isolated from tobacco young leaves and flowers were ligated with the 5'-RACE adapters and subjected to the 5'-RACE PCR. PCR products of *NtGAMYBs* have expected size of 430~450 bp. Next, PCR products were subjected to DNA sequencing with a *NtGAMYB*-specific primer. Sequencing chromatographs of *NtGAMYB*-adapter DNA junctions: peaks and sequences from *NtGAMYB* sequence downstream the miR159 cleavage site are highlighted in blue blocks and from adapter are in red blocks.

4.2.3 35S-*MIM159* tobacco displays pleiotropic developmental defects

To investigate the functional role of miR159 in tobacco, a 35S-*MIM159* transgene was stably transformed into *Nicotiana tabacum*. Of 15 independent T0 *MIM159* transformant tobacco plants, seven lines displayed similar developmental defects. These plants were still fertile, so seeds from three independent lines were sown and T1 progeny plants displayed similar phenotypic defects to their T0 parental plants, indicating that these defects are not artefacts due to the tissue culture aspects of the plant transformation.

Compared to wild-type, *MIM159* tobacco displayed stunted growth and reduced apical dominance with increased leaf numbers (Figure 4.5A, B and C). The leaves of *MIM159* tobacco were smaller and crinkled compared to wild type and displayed sectorized chlorosis (Figure 4.5C, D and E). In the reproductive organs, *MIM159* tobacco also developed smaller flowers with pale petals and shorter anther filaments (Figure 4.5F); and produced seeds that had reduced weight (Figure 4.5G). In addition, *MIM159* plants exhibited delayed flowering time compared to wild type (Figure 4.5H). Based on these observations, miR159 is required for correct growth and development at both the vegetative and reproductive stages.

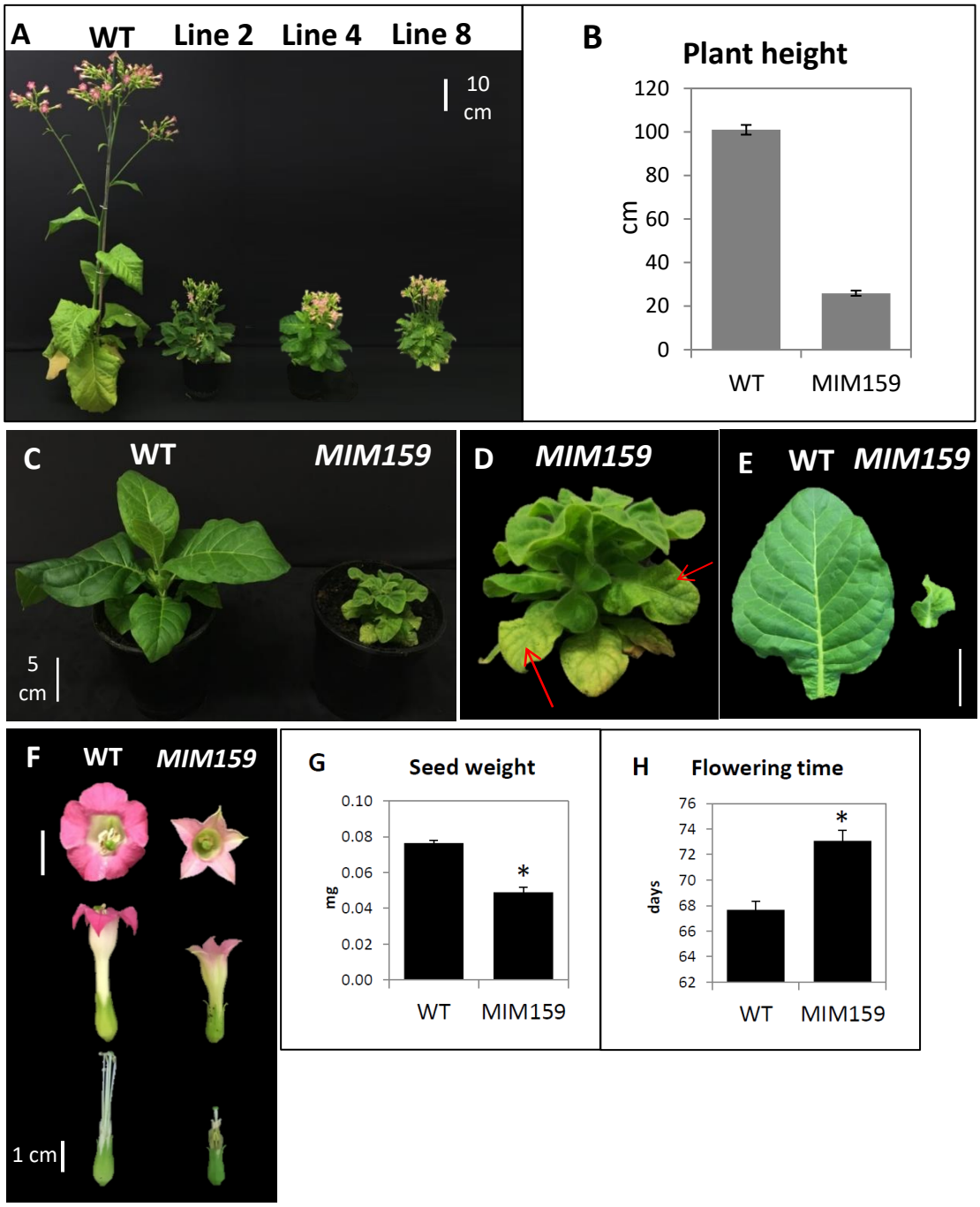


Figure 4.5. Expression of *MIM159* in tobacco results in strong phenotypic defects. (A) Three T1 independent lines of *MIM159* transformant *Nicotiana tabacum* displayed smaller growth stature compared to wild type. Scale bar represents 10 cm. **(B)** Average plant height of five 12-week-old wild-type plants and ten T1 *MIM159* plants from three independent lines. **(C)** Eight-week-old wild-type and *MIM159* tobacco grown under the identical conditions. Scale bar represents 5 cm. **(D)** Detailed view of the *MIM159* tobacco plant from the Figure (C). Arrows indicate sectorized chlorosis on leaves. **(E)** The 2nd leaf of eight-week-old wild-type and *MIM159* tobacco plants. Scale bar represents 5 cm. **(F)** *MIM159* tobacco produces smaller flowers, paler petals and shorter anther filaments, compared to the wild-type plants. Scale bar represents 1 cm. **(G)** Seed weight of wild-type and *MIM159* tobacco. The bar charts represent the average seed weight of more than one hundred seeds from five wild-type and five *MIM159* plants from three independent lines. Error bars represent the standard error of the

mean (SEM), and asterisk mark represents statistically significant difference analyzed from *t*-test. **(H)** Flowering time of wild-type and *MIM159* tobacco. All plants were grown without antibiotic selection. The bar charts represent the mean of flowering time of eight wild-type plants and 13 *MIM159* plants from three independent lines. Error bars represent SEM, and asterisk mark represents statistically significant difference analyzed from *t*-test.

4.2.4 *NtGAMYB* expression is de-regulated in *MIM159* tobacco.

Given the phenotypes of the *MIM159* tobacco plants, it would be predicted that expression of one or more of the *NtGAMYB* homologues may have become de-regulated due to inhibition of miR159. To determine this, transcript profiling was carried out on the *MIM159* transgene, mature miR159 and the *NtGAMYB* targets. qRT-PCR confirmed that *MIM159* had been successfully transformed into tobacco and was being highly transcribed (Figure 4.6A). Next, to identify whether miR159 was inhibited by *MIM159*, stem-loop qRT-PCR was performed to measure mature miR159 levels. Consistent with inhibition of miR159, mature miR159 levels were strongly reduced in both leaves and flowers of the *MIM159* transformant lines tested as compared to wild-type (Figure 4.6C). This confirms that miR159 is effectively suppressed by *MIM159* in tobacco.

The transcript levels of the three *NtGAMYB*s were next measured. Firstly, primers were used to generate amplicons that spanned the miR159 binding site, measuring uncleaved *NtGAMYB* mRNA levels. As these regions of the three *NtGAMYB* homologues are highly conserved, the primers will detect all three homologues, so the transcript levels determined would be a total of all three *NtGAMYB* homologues. qRT-PCR results revealed that the transcript levels of the total *NtGAMYB*s increased in both leaves and flowers of *MIM159* transformant lines compared to the wild-type plants (Figure 4.6B), with fold-change being greater in leaves. To investigate which of the three *NtGAMYB* homologues are de-regulated, gene-specific primers corresponding to the 3'-UTR of each gene were used to generate homologue specific amplicons. For *NtGAMYB1* and

NtGAMYB2, their mRNA levels were higher in both leaves and flowers of *MIM159* transformant lines compared to wild type (Figure 4.7A and B), with the only exception of the *NtGAMYB1* levels in the *MIM159* Line 8 which were lower than wild type in both leaves and flowers (Figure 4.7A). However, this distinct case of the low *NtGAMYB1* levels in the Line 8 was not corresponding to its morphological phenotype, which was shown in the severe phenotypes similar to the other *MIM159* lines. For *NtGAMYB3*, its transcript levels in leaves of *MIM159* transformant lines were higher than in wild type but were similar in flowers, suggesting this gene is not under miR159 regulation in floral tissues (Figure 4.7C). As miR159 can mediate strongly translational repression of *MYB33* in Arabidopsis (Li et al., 2014b), I cannot rule out the possibility that *NtGAMYB3* expression in tobacco is likely to be repressed at the translational level. Therefore, further assessment of NtGAMYB3 protein abundance in flower tissues between wide type and *MIM159* lines need to be performed. Taken together, on the basis of the qRT-PCR results, *NtGAMYB* expression appears to be de-regulated in both leaves and flowers of *MIM159* tobacco.

A previous report has claimed that expression of *MIM159* in Arabidopsis can result in off-target inhibition of miR319 (Reichel and Millar, 2015), as miR319 has high sequence identity to miR159 (identical at 17 of 21 nucleotides). In tobacco *MIM159* transformant lines, mature miR319 levels were decreased (Figure 4.6D and E), but not statistically significant in *MIM159* tobacco compared to wild type. Similarly, the mRNA levels of the miR319 target gene *TCP4*, were not statistically significant different between wild type and *MIM159* tobacco (Figure 4.7D and E). Therefore, significantly phenotypic difference in *MIM159* tobacco was due to inhibition of miR159 rather than miR319.

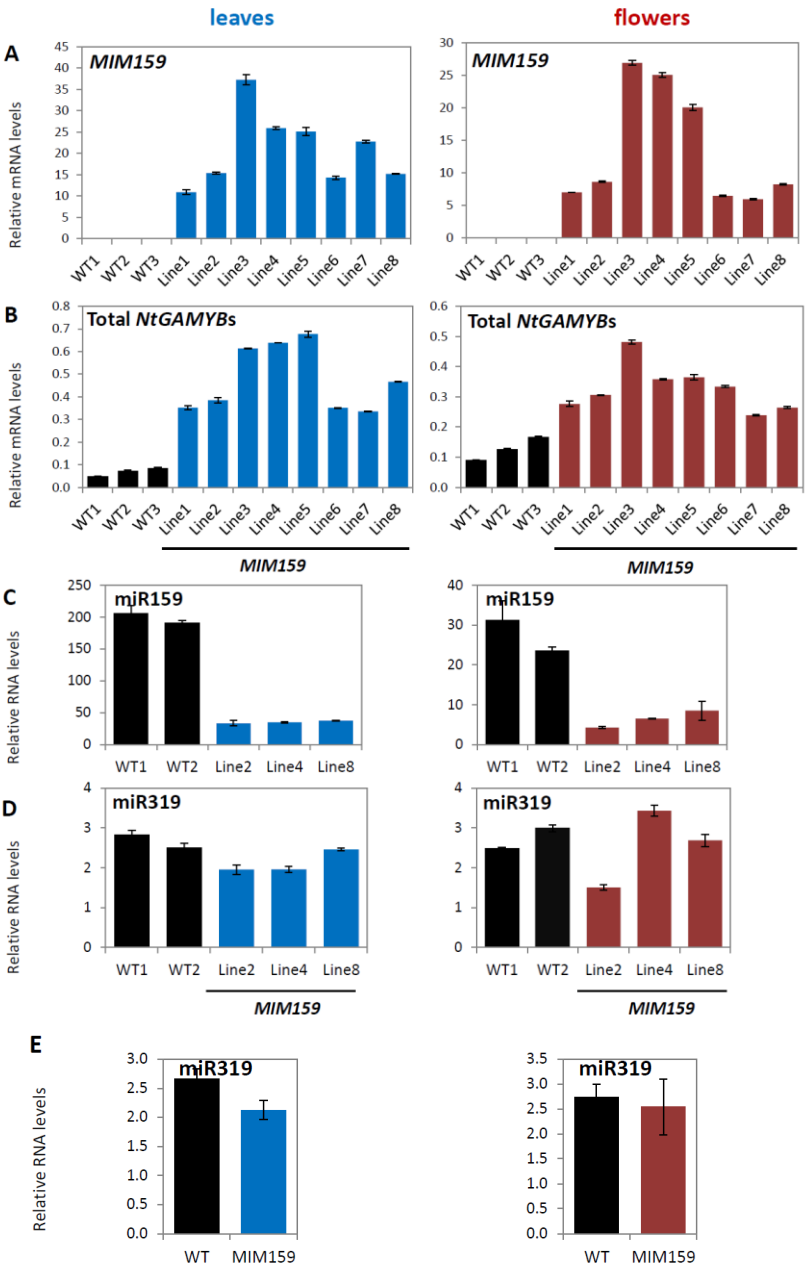


Figure 4.6. Transcript profiling in *MIM159* tobacco. qRT-PCR measurement of **(A)** *MIM159*; **(B)** Total *NtGAMYB*s; **(C)** miR159 and **(D)** miR319 in leaves (left panel) and flowers (right panel) of wild-type and independent T1 *MIM159* transformant lines. RNA was individually extracted from multiple tissues from each line; either the two young leaves of eight-week-old plants or five mature open flowers of 12-week-old plants. These *MIM159* transformant lines were coming from multiple T1 progeny plants from three independent T0 transformant lines. The transcript levels of total *NtGAMYB*s include all three *NtGAMYB* homologues, as *NtGAMYB* amplicons spanned the conserved sequences flanking to the miR159 binding site. The *MIM159* and *NtGAMYB*s mRNA levels were normalized to *PP2A*. The miR159 and miR319 levels were normalized to snoR101. Measurements are the average of three technical replicates with error bars representing the standard error of the mean (SEM). **(E)** Average levels of miR319 from Figure 4.6(D). The difference between wild type and *MIM159* is not statistically significant, analysed by t-test.

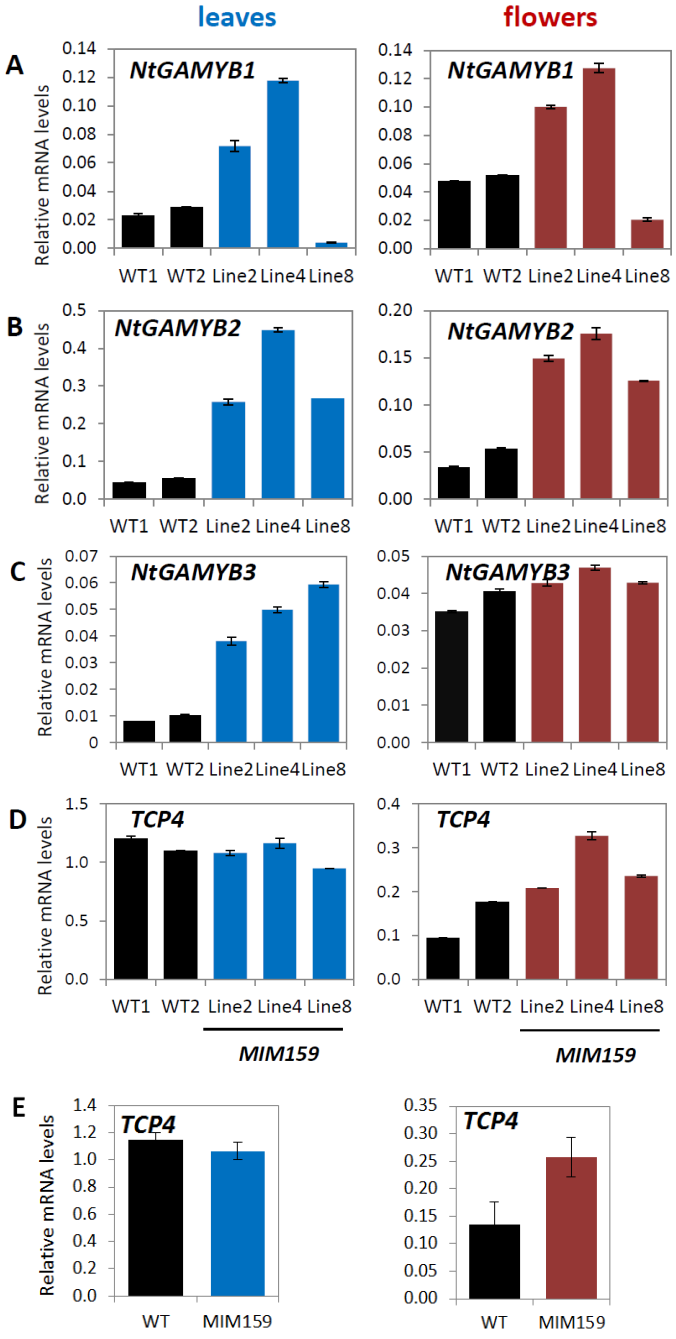


Figure 4.7. Transcript profiling of individual *GAMYB* homologue and *TCP4* in *MIM159* tobacco. qRT-PCR measurement of three *NtGAMYB* homologues **(A, B and C)**, as amplicons spanned the 3'-UTR of three different *NtGAMYB* homologues, and *TCP4* **(D)** mRNA levels in wild type and *MIM159* transformant lines. RNA was individually extracted from multiple tissues from each line; either the two young leaves of eight-week-old plants or five mature open flowers of 12-week-old plants. The *MIM159* Line 2, 4 and 8 were independent transformant lines. All measurements were normalized to *PP2A*. Measurements are the average of three technical replicates with error bars representing the standard error of the mean (SEM). **(E)** Average levels of *TCP4* from Figure 4.7(D). The difference between wild type and *MIM159* is not statistically significant, analysed by t-test.

4.2.5 *mNtGAMYB2* expression in tobacco exhibits phenotypic defects similar to *MIM159*.

To determine whether the phenotypic defects in *MIM159* tobacco were caused by de-regulated *NtGAMYB* expression, *mNtGAMYB2*, a miR159-resistant version of *NtGAMYB2* containing synonymous mutations within its miR159 binding site, was placed under the control of the 35S promoter and stably transformed into tobacco. Here, the *NtGAMYB2* transgene, but not *NtGAMYB1* or *NtGAMYB3*, was selected for *mNtGAMYB* over-expression in tobacco, due to its transcript abundance being the most relevant to the morphological phenotyopes of *MIM159* transformant lines (Figure 4.7A, B and C). Of 10 independent T0 *mNtGAMYB2* transformant tobacco lines, four displayed severe phenotypic defects (*mNtGAMYB2*-S lines) compared to six lines that appeared wild-type phenotypes (*mNtGAMYB2*-N lines). Consistent with de-regulated *NtGAMYB* expression resulting in the developmental defects in *MIM159* lines (Figure 4.5A, D and E), *mNtGAMYB2* severe lines displayed similar phenotypic defects, including stunted growth, upward leaf curl and leaf chlorosis (Figure 4.8B, D and E). Of the four *mNtGAMYB2*-S lines, only one line (*mNtGAMYB2*-S3) produced flowers (Figure 4.8C). Therefore, the phenotypic defects of the *mNtGAMYB2* plants were even more severe than that of the *MIM159* tobacco plants, which may reflect higher *mNtGAMYB2* expression due to the transgene positional effects within the tobacco genome. To investigate this, qRT-PCR was performed on leaves from the *mNtGAMYB2* lines and the phenotypic severity of the *mNtGAMYB2* plants was found to correlate with the *NtGAMYB2* transcript levels, which were considerably higher in the severe lines, S1, S2 and S3, compared to lines that did not display a phenotype, N1 and N2 (Figure 4.8F). The transcript levels of *mNtGAMYB2* were at least an order of magnitude higher than that of any of the *NtGAMYB* genes in the *MIM159* plants, potentially explaining the more severe phenotypes of the *mNtGAMYB2* plants. Therefore, this data argues the observed phenotypes in *MIM159*

tobacco were due to the deregulation of *NtGAMYB*, where its expression negatively impacts growth, as apparent in the *MIM159* and *mNtGAMYB2* tobacco plants.

Figure 4.8. Expression of *mNtGAMYB2* in tobacco results in severe phenotypic defects. (A) *mNtGAMYB2* tobacco presented stunted growth in the severe line (S) compared to the none line (N). Scale bar represents 10 cm. **(B)** The aerial view of the *mNtGAMYB2* severe line from Figure (A), which displayed up-ward leaf curl and leaf chlorosis (indicated by red arrows). **(C)** One of four severe lines (*mNtGAMYB2*-S3) produced flowers. **(D)** The aerial view of four-month-old *mNtGAMYB2*-S1 line. Scale bar represents 5cm. **(E)** The top part of the *mNtGAMYB2*-S1 tobacco plant from the Figure (D). The *mNtGAMYB2* tobacco presents up-ward leaf curl. **(F)** qRT-PCR measurement of *NtGAMYB2* transcript levels in two independent *mNtGAMYB2*-N lines and three independent *mNtGAMYB2*-S lines. RNA was individually extracted from two young leaves of 12-week-old plants. The *NtGAMYB2* mRNA levels were normalized to *PP2A*. Measurements are the average of three technical replicates with error bars representing the standard error of the mean (SEM).

4.2.6 RNA sequencing of *MIM159* tobacco leaves to determine pathways downstream of *GAMYB*

As *GAMYB* genes encode MYB R2R3 transcription factors, RNA sequencing was performed on leaves of wild-type and *MIM159* transformed tobacco to attempt to identify differentially expressed genes and downstream pathways possibly under the NtGAMYB control. RNA was isolated from young leaves of eight-week-old wild-type and *MIM159* plants. Each sample consisted of tissues from three individual plants; for the *MIM159* sample tissues came from multiple independent lines. The isolated RNAs were used to construct cDNA libraries on which Illumina deep sequencing was performed (Novogene). Sequencing reads were aligned to the reference *Nicotiana tabacum* TN90 genome (Sierro et al., 2014, Sol Genomics). The alignment results were analysed by Cuffdiff software (cufflinks V2.2.1) and sequence data from three libraries were statistically analysed to detect differentially expressed genes between the *MIM159* and wild type tobacco plants. Using a two-fold change cut-off, 12,418 genes were up-regulated and 9,431 genes were down-regulated in *MIM159* compared to wild type, suggesting that the loss of miR159 in tobacco facilitates global changes to the transcriptome.

Firstly, *NtGAMYB1*, *NtGAMYB2* and *NtGAMYB3* (*gene_47587*, *gene_56181*, *gene_42641*) were all found to be up-regulated in *MIM159* tobacco leaves (approximately three, nine and six times respectively), which are fold-level changes consistent with the qRT-PCR data (Figure 4.7A, B and C). Next, to gain insight into which biological pathways are altered in the *MIM159* plants, Gene Ontology (GO) analysis of the differentially expressed genes was performed. As the tobacco genome has not been completely annotated, the GO enrichment analysis was determined by using the closest homologues from Arabidopsis. Given that so many differentially expressed genes were detected at a two-fold level, the GO enrichment analysis was narrowed to only include

genes differentially expressed at the level of more than five-fold, narrowing the analysis to 1478 genes that were up-regulated in *MIM159*. Of the highest 20 enriched pathways ranked by p-value significance, three general themes emerged. Firstly, in addition to the highest enriched category, defence response, there were another four categories related to defence in the top 20 pathways, including defence response to fungus and bacterium (Table 4.1). The next highest category was programmed cell death (PCD), and again another four categories related to PCD were present in the most 20 enriched pathways (Table 4.1). Such a role is consistent with previous research which has demonstrated that GAMYB promotes PCD in the seed aleurone and the anther tapetum (Aya et al., 2009; Alonso-Peral et al., 2010; Millar and Gubler, 2005). Finally, another six pathways were related to responses to hormone and environmental stimulus (Table 4.1). Narrowing down the analysis to genes that are ten-fold up-regulated in *MIM159* tobacco (804 genes) largely resulted in the identification of the same pathways (16 out of 20 of the pathways were identical; Appendix Table 9).

Conversely, in *MIM159* tobacco leaves there were 2345 genes down-regulated more than five-fold. Identification of the 20 most enriched pathways found many pathways associated with cell cycle, size and morphogenesis (Table 4.2). The top pathway, "microtubule-based processes" is also a potential pathway related to cell division and development (Hamada, 2014). Down-regulation of such pathways could be consistent with the stunted growth and leaf phenotype of the *MIM159* tobacco plants. Again narrowing this analysis to ten-fold down-regulated genes in *MIM159* tobacco (1038 genes) largely identified the same pathways (Appendix Table 10).

Table 4.1. Top 20 enriched pathways from up-regulated genes in *MIM159* tobacco leaves. Based on the RNA sequencing data, 1478 genes up-regulated in *MIM159* tobacco leaves at the level of more than five-fold, compared to wild type. These genes were referred to the closest homologues from Arabidopsis and the Gene Ontology (GO) enrichment were analysed by agriGO (Du et al., 2010). The highest 20 enriched pathways were ranked by p-value. Pathway with a p-value of ≤0.05 was considered as significantly enriched.

Pathway	GO term	p-value
defense response	GO:0006952	1.20E-15
programmed cell death	GO:0012501	1.30E-10
phosphate metabolic process	GO:0006796	3.80E-10
response to carbohydrate stimulus	GO:0009743	7.60E-09
response to hormone stimulus	GO:0009725	2.10E-08
response to oxidative stress	GO:0006979	6.50E-08
apoptosis	GO:0006915	6.70E-08
response to chitin	GO:0010200	1.50E-07
defense response to fungus	GO:0050832	2.80E-07
immune system process	GO:0002376	1.40E-06
defense response, incompatible interaction	GO:0009814	6.80E-06
response to abscisic acid stimulus	GO:0009737	1.80E-05
aging	GO:0007568	3.90E-05
response to temperature stimulus	GO:0009266	5.80E-05
ion transport	GO:0006811	1.10E-04
cellular nitrogen compound metabolic process	GO:0034641	1.10E-04
senescence	GO:0010149	1.20E-04
host programmed cell death induced by symbiont	GO:0034050	1.30E-04
cellular amino acid derivative biosynthetic process	GO:0042398	1.30E-04
response to bacterium	GO:0009617	2.40E-04

Table 4.2. Top 20 enriched pathways from down-regulated genes in *MIM159* tobacco leaves. Based on the RNA sequencing data, 2345 genes down-regulated in *MIM159* tobacco leaves at the level of more than five-fold, compared to wild type. These genes were referred to the closest homologues from Arabidopsis and the Gene Ontology (GO) enrichment were analysed by agriGO (Du et al., 2010). The top 20 enriched pathways were ranked by p-value. Pathway with a p-value of ≤0.05 was considered as significantly enriched.

Pathway	GO term	p-value
microtubule-based process	GO:0007017	6.60E-13
cell cycle	GO:0007049	7.70E-12
lipid localisation	GO:0010876	4.90E-11
post-embryonic development	GO:0009791	1.50E-10
transmembrane receptor protein tyrosine kinase signalling pathway	GO:0007169	2.00E-10
cell surface receptor linked signalling pathway	GO:0007166	2.80E-09
cellular developmental process	GO:0048869	7.30E-09
post-embryonic morphogenesis	GO:0009886	1.80E-08
cellular component morphogenesis	GO:0032989	9.50E-08
tissue development	GO:0009888	2.10E-07
cell morphogenesis	GO:0000902	2.40E-07
response to radiation	GO:0009314	4.00E-07
regulation of cell size	GO:0008361	7.40E-07
regulation of cell cycle	GO:0051726	7.90E-07
cell wall polysaccharide metabolic process	GO:0010383	8.30E-07
regulation of cellular component size	GO:0032535	9.10E-07
cellular component organisation	GO:0016043	9.50E-07
lipid metabolic process	GO:0006629	1.60E-06
unidimensional cell growth	GO:0009826	1.70E-06
developmental growth involved in morphogenesis	GO:0060560	1.70E-06

Finally, a previous report proposed that GAMYB activates *MIR156* transcription, contributing to delayed vegetative phase transition and flowering time in Arabidopsis (Guo et al., 2017). Consistent with this, multiple miR156 predicted target *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) genes were down-regulated in *MIM159* tobacco leaves (Table 4.3). Additionally, a delayed flowering-time was observed in *MIM159* tobacco (Figure 4.5H). This suggests that the interplay between miR159-*GAMYB* and miR156-*SPL* could be conserved in Arabidopsis and tobacco.

Table 4.3. miR156 target *SPL* genes were down-regulated in *MIM159* tobacco leaves. (A) The tobacco *SPL* genes contain a conserved binding site highly complementary to miR156a. *Nt*, *Nicotiana tabacum*. The red letter represents mismatch. **(B)** RNA sequencing data reveals many predicted *SPL* genes, referred as *N. tabacum* gene ID, are down-regulated in fold changes of *MIM159* to wild type.

A

```
Nt-miR156a CACGAGUGAGAGAAGACAGU-5′
Nt-SPLs    GUGCUCUCUCUCUUCUGUCA-3′
```

B.

N. tabacum Gene ID	Fold change (*MIM159*/WT)	Description
gene_19223	0.327077782	*SPL16*
gene_22177	0.077335481	*SP3*
gene_43962	0.426240933	*SPL6*
gene_67009	0.035954508	*SPL6*
gene_12160	0.089387471	*SPL9*
gene_15333	0.217861314	*SPL16*

4.2.7 *PATHOGENESIS-RELATED* (*PR*) and other defence genes are strongly up-regulated in *MIM159* tobacco

Next, the 50 most up-regulated and down-regulated genes were identified (Table 4.4 and Appendix Table 11). BLAST analysis of these genes to the tobacco transcriptome predicted their protein descriptions. Of the top 50 most up-regulated genes, 22 genes putatively encode PR proteins. The mRNA levels of these *PR* genes were strongly up-regulated in *MIM159* tobacco leaves, being 360~2020 fold greater than wild-type (Table 4.5). According to their putative protein function, these up-regulated PR proteins include seven conserved PR sub-families, which have properties of β-1,3-glucanase, chitinase and osmotin. Of these, seven genes putatively encode PR-2 and two genes encode PR-Q'-like, which were identified as β-1,3-glucanase. Six genes putatively encode three PR families: PR-Q, PR-R major form and acidic endochitinase, which all have properties of chitinase. Another two genes putatively encode osmotin, which is a homologue of thaumatin and induced by osmotic stress. The thaumatin-like protein family has an antifungal function by disrupting the fungal membrane, resulting in osmotic rupture. Therefore, this family is identified as the PR-5 protein family (Edreva, 2005, Sinha et al., 2014). As previously determined, PR proteins are involved in plant defence against pathogens (Sinha et al., 2014). For example, PR proteins encoding the hydrolytic enzymes β-1,3-glucanase and chitinase defend against pathogens by degrading pathogen cell walls (Ebrahim et al., 2011). To validate the up-regulation of *PR* mRNA levels in *MIM159* tobacco leaves, qRT-PCR was performed on leaves from multiple *MIM159* transformant lines and wild-type plants. Compared to wild type, *PR-1b*, *PR-2* and *PR-Q* transcript levels were strongly increased in all *MIM159* transformant lines tested (Figure 4.9A), confirming the RNA sequencing data. Together, these results show that inhibition of miR159 in tobacco leaves leads to strong up-regulation of *PR* expression.

In addition to the *PR* genes, other genes associated with defence response were also strongly up-regulated in *MIM159* tobacco. For instance, *gene_46853* and *gene_13483*, which putatively encode PHOTOASSIMILATE-RESPONSIVE PROTEIN 1c (PAR-1c) homologs, were up-regulated approximately 500-fold in *MIM159* tobacco leaves, ranking in the list of top 50 most up-regulated genes (Table 4.4). *PAR* is identified as a distinctive class of *PR*, and its transcription is inducible by viruses, where its expression is coordinated with *PR* genes (Herbers et al., 1995). Furthermore, *gene_60737* and *gene_69711*, putatively encoding Sar8.2, were up-regulated 606-fold and 470-fold respectively in *MIM159* tobacco (Table 4.4). Expression of *Sar8.2* is induced by salicylic acid in tobacco via the systemic acquired resistance (SAR), which is involved in defence response against fungal, bacterial and viral diseases (Alexander et al., 1992; Song and Goodman, 2002).

Table 4.4. Top 50 most up-regulated genes in *MIM159* tobacco leaves. RNA sequencing results indicate top 50 genes, referred as *Nicotiana tabacum* gene ID and gene description, are highest up-regulated in fold changes of *MIM159* to wild type.

Gene ID	Description	Fold change
gene_58605	acetylajmalan esterase-like	2371
gene_40024	pathogenesis-related protein PR1c	2143
gene_56557	pathogenesis-related protein PR1a	2020
gene_5026	PR2, glucan endo-1,3-beta-glucosidase	1937
gene_14265	pathogenesis-related protein PR1c	1640
gene_1443	calcium-binding protein CML30	1555
gene_12559	pathogenesis-related protein PR1a	1451
gene_38599	xyloglucan endotransglucosylase/hydrolase protein 23	1292
gene_39617	PR2, glucan endo-1,3-beta-glucosidase	1262
gene_59805	PR2, glucan endo-1,3-beta-glucosidase	1069
gene_59471	osmotin	1056
gene_18980	AAA-ATPase	1001
gene_71672	AAA-ATPase	883
gene_556	5-epi-aristolochene synthase-like	843
gene_66775	PR2, glucan endo-1,3-beta-glucosidase	811
gene_68554	unknown	747
gene_34815	calcium-transporting ATPase 13, plasma membrane-type	745
gene_16878	pathogenesis-related protein R major form	731
gene_67890	protein SAR DEFICIENT 1-like	715
gene_32993	zingipain-2-like	632
gene_26188	protein yippee-like	627
gene_27639	somatic embryogenesis receptor kinase 2-like	623
gene_59815	PR2, glucan endo-1,3-beta-glucosidase	610
gene_60737	Partial Sar8.2	606
gene_12101	diacylglycerol kinase theta-like	595
gene_66475	ankyrin repeat-containing protein	588
gene_60502	LRR receptor-like serine/threonine-protein kinase	582
gene_8988	acidic chitinase PR-Q	578
gene_15771	PR2, glucan endo-1,3-beta-glucosidase	570
gene_26030	glucan endo-1,3-beta-glucosidase, acidic isoform PR-Q'-like	546
gene_46853	PAR-1c (photoassimilate-responsive)	542
gene_1355	glucan endo-1,3-beta-glucosidase, acidic isoform PR-Q'-like	516
gene_69452	suberisation-associated anionic peroxidase-like	503
gene_61952	pathogenesis-related protein 1b	500
gene_58140	osmotin	493
gene_13483	PAR-1c (photoassimilate-responsive)	491
gene_7291	protein P21-like	487
gene_69711	Partial Sar8.2	470
gene_71758	glucan endo-1,3-beta-glucosidase, basic vacuolar	469
gene_19816	unknown	465
gene_38916	abscisic acid and environmental stress-inducible protein TAS14-like	423
gene_54704	unknown	408
gene_27076	pathogenesis-related protein R major form	403
gene_18100	peroxidase 51-like	402
gene_46441	acidic chitinase PR-Q	400
gene_63189	bifunctional epoxide hydrolase 2-like	398
gene_59819	PR2, glucan endo-1,3-beta-glucosidase	392
gene_22847	E3 ubiquitin-protein ligase RNF217	379
gene_32773	acidic endochitinase	369
gene_4867	acidic endochitinase	360

Table 4.5. *PATHOGENESIS-RELATED (PR)* genes strongly up-regulated in *MIM159* tobacco leaves. *PR* genes refined from the Table 4.4, referred as *Nicotiana tabacum* gene ID and up-regulated in fold changes of *MIM159* to wild type, are predicted to encode seven PR subfamilies, which have a function of plant defence against pathogen (Sinha et al., 2014).

Gene ID	Fold change (*MIM159*/WT)	Description (Protein family)	Function
gene_56557	2020	**PR-1a**	**Antifungal**
gene_12559	1451		
gene_61952	500	**PR-1b**	
gene_40024	2143	**PR-1c**	
gene_14265	1640		
gene_5026	1937	**PR-2, β-1,3-glucanase**	**Cleaves β-1,3-glucans**
gene_39617	1262		
gene_59805	1069		
gene_66775	811		
gene_59815	610		
gene_15771	570		
gene_59819	392		
gene_26030	546	**PR-Q'-like, β-1,3-glucanase**	**Cleaves β-1,3-glucans**
gene_1355	516		
gene_8988	578	**PR-Q, acidic chitinase**	**Endochitinase**
gene_46441	400		
gene_32773	369	**acidic endochitinase**	**Endochitinase**
gene_4867	360		
gene_16878	731	**PR-R major form, chitinase**	**Antifungal and chitinase**
gene_27076	403		
gene_59471	1056	**osmotin**	**Respond to osmotic stress, antifungal**
gene_58140	493		

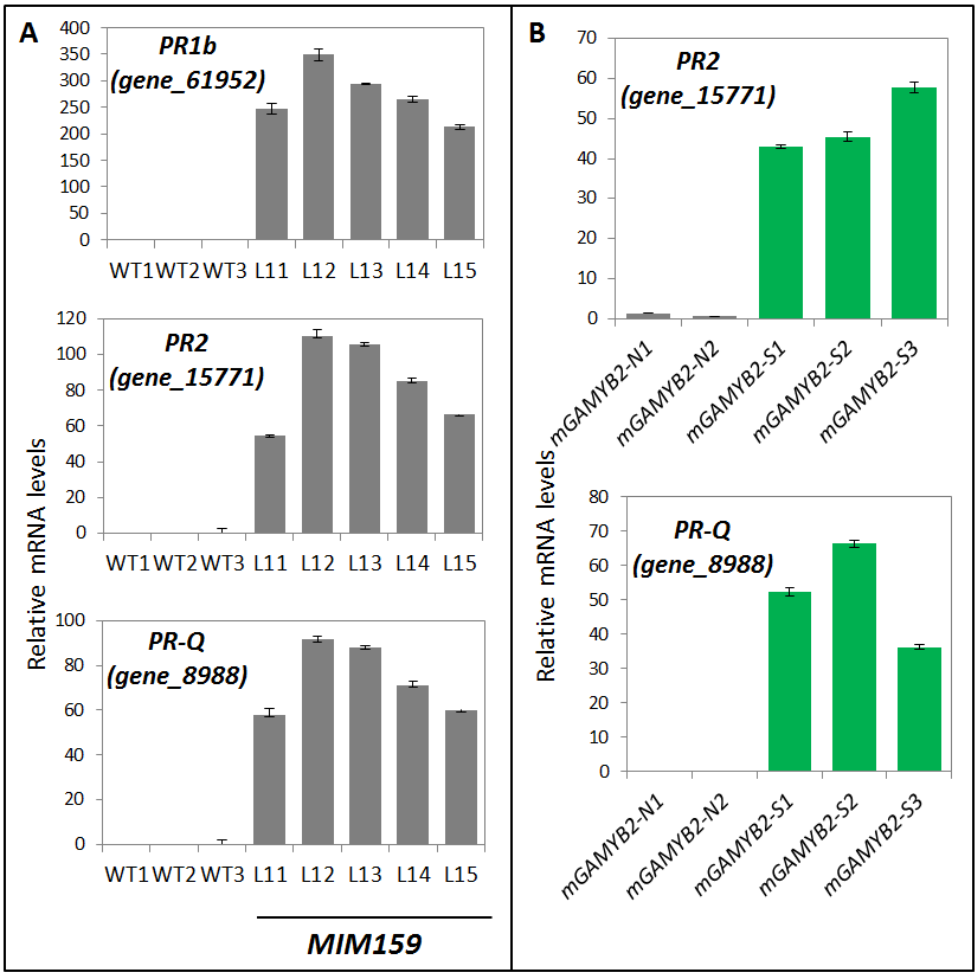


Figure 4.9. qRT-PCR measurement of *PR* mRNA levels in *MIM159* and *mNtGAMYB2* tobacco leaves. All measurements are relative to *PP2A*. **(A)** RNA was individually extracted from two young leaves of eight-week-old plants. These *MIM159* lines were coming from multiple independent transformant lines. **(B)** RNA was individually extracted from two young leaves of 12-week-old plants. These *mNtGAMYB2* lines were coming from two independent none (N) lines and three independent severe (S) lines. Measurements are the average of three technical replicates with error bars representing the standard error of the mean (SEM).

Finally, there are 27 predicted *NUCLEOTIDE-BINDING SITE AND LEUCINE-REPEAT* (*NBS-LRR*) and two *METACASPASE-1* genes that were more than five-fold up-regulated in *MIM159* tobacco leaves (Table 4.6). NBS-LRR and METACASPASE-1 are upstream elicitors of hypersensitive response (HR), a major type effector-triggered immunity (ETI) in plant defence response (Coll et al., 2011; Spoel and Dong 2012; Weiberg et al., 2014). Alignment of the sequences of these genes to the tobacco transcriptome revealed that they are predicted to encode resistance (R) proteins, belonging to two major classes: *Tobacco mosaic virus* (*TMV*) resistance N-like protein, containing the Toll and Interleukin-1 receptor homology/nucleotide binding site/leucine rich repeat (TIR-NBS-LRR); and late blight resistance protein homolog R1, containing NBS-LRR and a leucine zipper motif and being resistant to late blight disease resulting from *Phytophthora infestans* infection (Table 4.6; Ballvora et al., 2002; Loebenstein 2009; Marathe et al., 2002). In addition, all of these up-regulated *R* genes have highly related *NBS-LRR* homologues in Arabidopsis (Table 4.6), a finding that strongly suggests that this group of R proteins are conserved across multiple plant species. In conclusion, *PR* and many other genes related to plant defence response were strongly up-regulated in *MIM159* tobacco leaves, revealing that plant defence responses are universally enhanced.

Table 4.6. *NBS-LRR* genes up-regulated in *MIM159* tobacco leaves. The *NUCLEOTIDE-BINDING SITE AND LEUCINE-REPEAT* (*NBS-LRR*) and *METACASPASE-1* genes, referred as *Nicotiana tabacum* gene ID, are up-regulated in fold changes of *MIM159* to wild type and are predicted to encode resistance (R) proteins. In addition, these tobacco R proteins have the closest homologues in Arabidopsis, which are NBS-LRR resistance protein, including TIR-NBS-LRR, coiled-coil/nucleotide binding site/leucine rich repeat (CC-NBS-LRR) and other NBS-LRR sub-families (Li et al., 2012a).

Gene ID	Fold change	Description	Arabidopsis homologue
gene_19271	350	TMV resistance protein N-like	TIR-NB-LRR (AT4G12010)
gene_52683	19	TMV resistance protein N-like	TIR-NBS-LRR (AT5G36930)
gene_356	17		
gene_50778	8		
gene_19766	6		
gene_23647	6		
gene_49683	18	TMV resistance protein N-like	TIR-NBS-LRR (AT5G17680)
gene_19771	12		
gene_23091	7		
gene_22802	7		
gene_21302	6		
gene_5620	15	TMV resistance protein N-like	TIR domain-containing protein (AT1G72940)
gene_68433	111	TMV resistance protein N-like	transmembrane receptors / ATP binding protein (AT1G27170)
gene_73371	31	late blight resistance protein homolog R1A-3	CC-NBS-LRR (AT5G35450)
gene_3520	15	late blight resistance protein homolog R1B-16	
gene_4942	15		
gene_6485	12		
gene_65347	13	late blight resistance protein homolog R1C-3	
gene_72190	12	late blight resistance protein homolog R1A-10	
gene_12902	17	late blight resistance protein homolog R1A-3	NBS-LRR type R protein, RECOGNITION OF PERONOSPORA PARASITICA 13 (AT3G46530)
gene_30232	13		
gene_18983	11	late blight resistance protein R1-A-like	
gene_59049	15	late blight resistance protein homolog R1B-17	NB-ARC domain-containing disease resistance protein (AT1G59780)
gene_38070	14	probable disease resistance protein	CC-NBS-LRR (AT5G66900)
gene_188	18	probable disease resistance protein	NBS-LRR immune receptors, the ADR1 family (AT4G33300)
gene_34031	11	probable disease resistance protein	
gene_13131	15	disease resistance protein RPM1-like	NB-ARC domain disease resistance protein (AT3G46710)
gene_35884	90	metacaspase-1 -like	metacaspase-1 (AT1G02170)
gene_13554	15	metacaspase-1 -like	

4.2.8 *PR* up-regulation in *MIM159* tobacco leaves due to *GAMYB* de-regulation

If the up-regulation of *PR* expression in the *MIM159* tobacco plants is due to de-regulation of *GAMYB* expression, these genes should also be up-regulated in the *mNtGAMYB2* plants. To investigate this, qRT-PCR of *PR* transcripts was performed on *mNtGAMYB2* tobacco leaves. It was found that the transcript levels of *PR2* and *PR-Q* were strongly correlated with *mNtGAMYB2* expression, being negligible in the *mNtGAMYB2-N* plants (low *mNtGAMYB2* transcript levels, wild-type phenotype) and high in the *mNtGAMYB2-S* plants (high *mNtGAMYB2* transcript levels, severe phenotypic defects) (Figure 4.8F and Figure 4.9B). Given this positive relationship between *GAMYB* and *PR* transcript levels, argues that *PR* up-regulation in *MIM159* tobacco leaves is caused by de-regulated *GAMYB* expression.

4.2.9 Investigating the miR159-*GAMYB* pathway in rice.

As the miR159-*GAMYB* pathway plays important roles in Arabidopsis and tobacco, next *Oryza sativa* (rice) was chosen to extend the investigation of the miR159-*GAMYB* pathway to a monocot species. Based on miRBase (Kozomara and Griffiths-Jones, 2014), mature miR159a and miR159b in rice (*Os*-miR159a and *Os*-miR159b) have identical sequences and correspond to the most abundant isoform in rice, and only differ to the mature Arabidopsis miR159a sequence by one nucleotide (Figure 4.10A). In rice, one *GAMYB* and two *GAMYB-like* genes have been identified, all of which have a conserved miR159 binding site that is highly complementary to Os-miR159a/b (Figure 4.10A, Tsuji et al., 2006).

To begin to determine the spatial transcriptional domains of miR159 and *GAMYB* in rice, qRT-PCR was performed to measure mature *Os*-miR159 levels and *OsGAMYB* mRNA levels in leaves and flowers. For the *OsGAMYB* genes, primers were designed to amplify

their 3'-UTR regions to ensure gene-specific amplicons. As these primers do not span the miR159 cleavage site, these primers will measure both cleaved and uncleaved *OsGAMYB* transcripts. *Os*-miR159a/b was found to be expressed in both leaves and flowers of rice (Figure 4.10B), consistent with the previous report (Tsuji et al., 2006). The transcript levels of all three *OsGAMYB* homologues were substantially higher in leaves than in flowers (Figure 4.10B). This goes against a previous report that found *OsGAMYB* and *OsGAMYBL1* were not transcribed in vegetative tissues, as determined by *OsGAMYB* promoter-GUS reporter genes (Tsuji et al., 2006). Of the three *GAMYB* homologues, *OsGAMYBL2* had the lowest transcript levels (Figure 4.10B). In conclusion, similar to Arabidopsis and tobacco, mature miR159 and *GAMYB* mRNA are present in both rice leaves and flowers.

As the *OsGAMYB* homologues all contain a highly complementary miR159 binding site, they are predicted to be regulated by miR159. To verify this, 5'-RACE miRNA cleavage assays on RNA from rice leaves and flowers were performed. PCR products consistent with the expected sizes were observed for *OsGAMYB* and *OsGAMYBL1*, but not for *OsGAMYBL2* (Figure 4.10C and D). Sequencing of these PCR products found they corresponded to the miR159-guided cleavage products, as the *OsGAMYB*-adapter boundary corresponds to the canonical miR159 cleavage site (Figure 4.10C and D). These results indicate that *OsGAMYB* and *OsGAMYBL1* are regulated by miR159 in both rice leaves and flowers, confirming previous results (Tsuji et al., 2006).

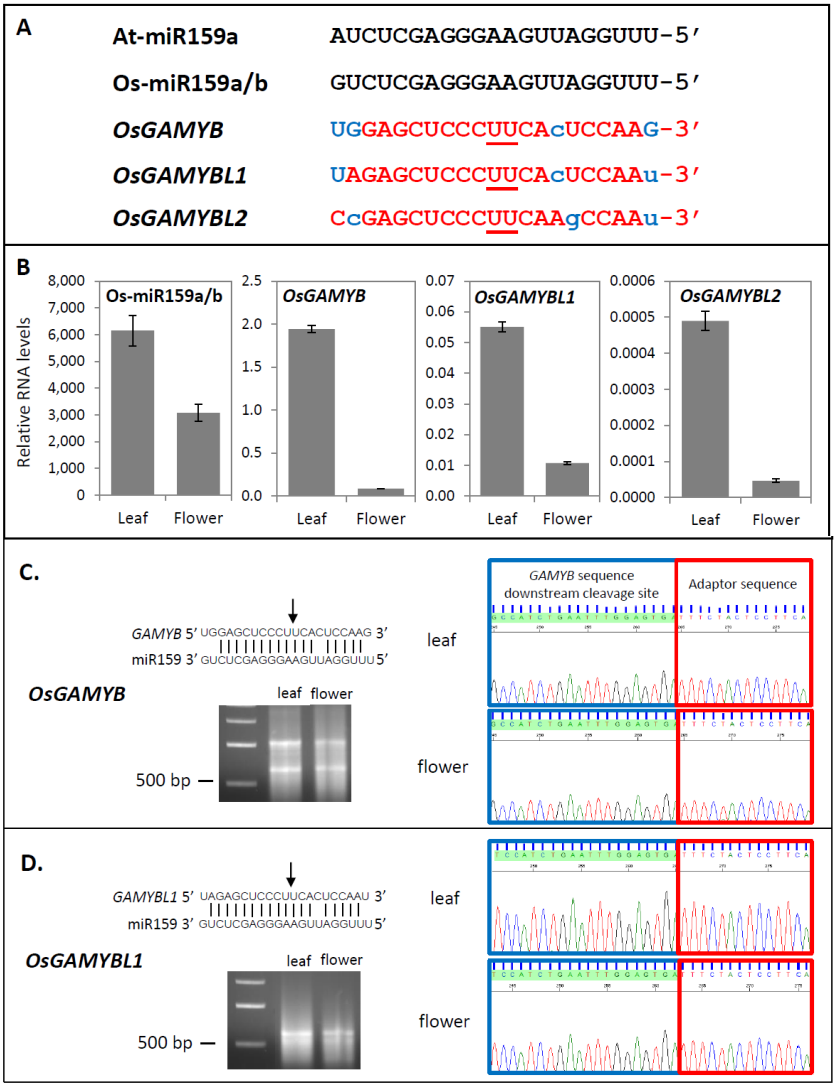


Figure 4.10. Investigating the miR159-*GAMYB* regulation in rice. (A) Sequence alignments of rice miR159 and the miR159-binding site in *OsGAMYB* homologues. Nucleotides underlined represent the cleavage site, lower case blue letters represent mis-matches to *Os*-miR159a/b, and G:U pairings are shown in upper case blue letters. *At*, *Arbidopsis thaliana*; *Os*, *Oryza sativa*. **(B)** qRT-PCR measurement of miR159 and *GAMYB* transcript levels in rice. RNA was extracted from pools of two *Oryza sativa* plants, each being composed of three flag leaves and 30 florets. The mRNA levels of the *OsGAMYB* genes were normalized to *ACTIN* and the miR159 RNA levels were normalized to snoR14. Measurements are the average of three technical replicates with error bars representing the standard error of the mean (SEM). Investigating miR159-mediated cleavage of *OsGAMYB* **(C)** and *OsGAMYBL1* **(D)** in rice flag leaves and flowers. The mRNA isolated from rice flag leaves or flowers was ligated with 5'-RACE adapters and subjected to 5'-RACE PCR. PCR products of *OsGAMYB* and *OsGAMYBL1* have expected size of 1000 and 650 bp, respectively. Next, PCR products were subjected to DNA sequencing with *OsGAMYB*-specific primer. Sequencing chromatographs of *OsGAMYB*-adapter DNA junction: peaks and sequences from *OsGAMYB* sequence downstream the miR159 cleavage site are highlighted in blue block and from adapter are in red.

4.2.10 Investigating the roles of miR159-*GAMYB* in rice

To investigate the functional role of miR159 in rice, a rice *MIM159* construct was placed under the control of ubiquitin promoter and transformed into *Oryza sativa* Japonica. An empty vector control (VC) harbouring ubiquitin promoter and Hygromycin-resistant gene was also transformed into rice. Of 24 independent T0 *MIM159* transformants, more than half of the lines displayed a smaller growth stature, compared to the wild-type and VC plants (Figure 4.11A). Although average height of VC plants was significantly different to wild type (Figure 4.11B), this could be due to the T0 transformants being grown on antibiotic selection or tissue culture artefacts.

Next, qRT-PCR measurements of *MIM159*, mature miR159, *OsGAMYB* and *OsGAMYBL1* were performed on wild type, VC Line 5, and *MIM159* Lines, 1, 23 and 25. Lines 23 and 25 were two of the smallest *MIM159* lines, whereas Line 1 displayed the highest stature among *MIM159* lines (Figure 4.11A). Consistent with this, the transcript levels of *MIM159* were higher in the *MIM159* Line 23 and 25 compared to that of Line 1 (Figure 4.11C). These *MIM159* transcript levels inversely correlated with miR159, with Lines 23 and 25 having the lowest miR159 levels (Figure 4.11D). Furthermore, *OsGAMYB* and *OsGAMYBL1* levels were higher in *MIM159* Line23 and Line25 compared to wild-type, VC5 and Line 1 (Figure 4.11 E and F). On the basis of these data, it appears that the RNA levels of *MIM159,* miR159, *OsGAMYB* and *OsGAMYBL1* were generally correlated with phenotypic severity.

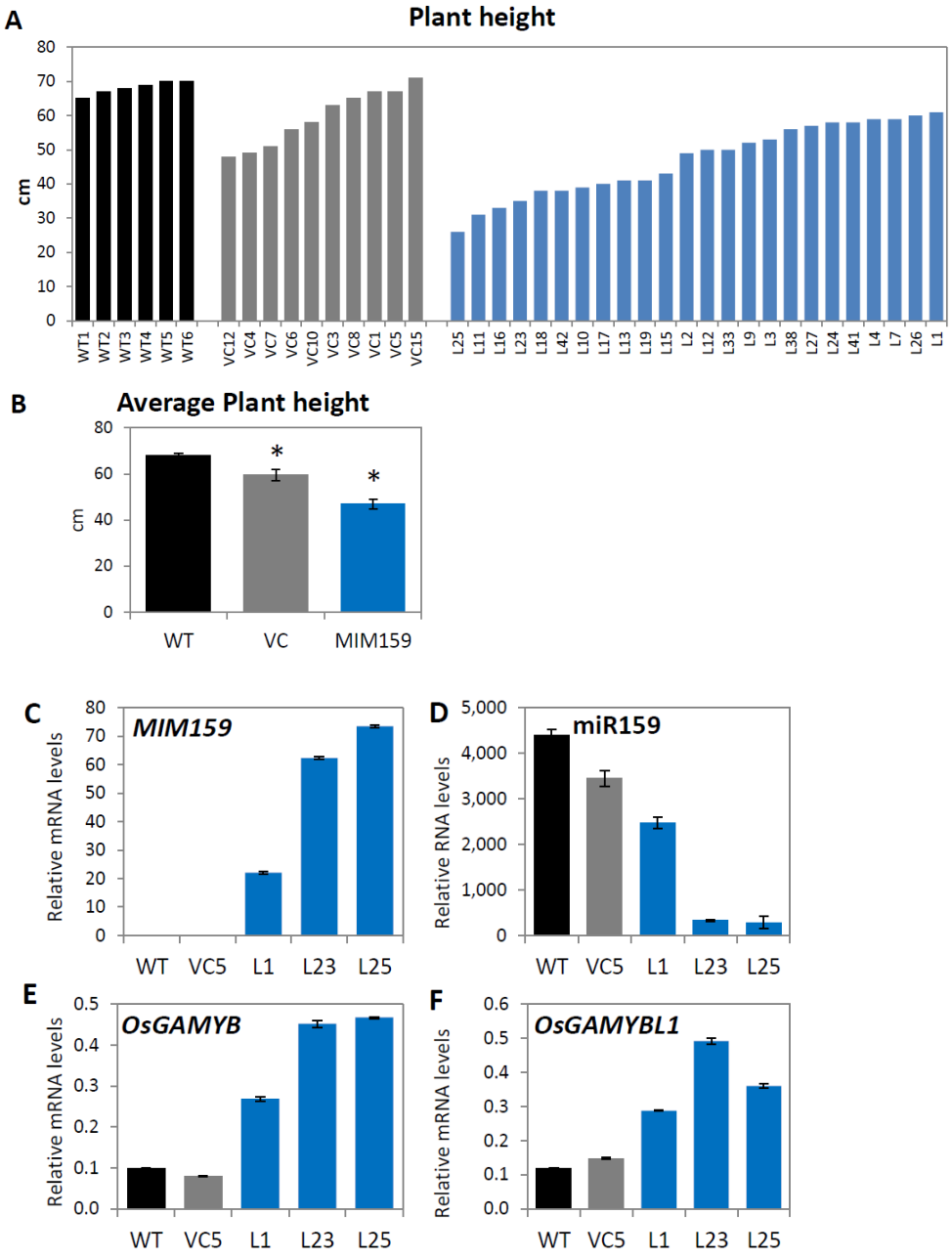


Figure 4.11. Phenotyping of T0 *MIM159* rice. (A) Plant heights of multiple wild-type plants (black bar), independent T0 transformant lines of empty vector control (VC) (grey bar) and *MIM159* (blue bar) grown in nine-week-old. The rice transformants were grown under Hygromycin selected medium. **(B)** Average plant heights of wild-type, independent VC transformant lines and independent *MIM159* transformant lines from Figure 4.11(A). Error bars represent the standard error of the mean (SEM), and asterisk mark represents statistically significant difference compared to wild type, statistically analyzed from *t*-test. Transcript profiling of *MIM159* **(C)**, miR159 **(D)**, *OsGAMYB* **(E)** and *OsGAMYBL1* **(F)** measured by qRT-PCR. RNA was individually extracted from wild type, one VC transformant line and three independent *MIM159* transformant lines, each being composed of three flag leaves. The *MIM159*, *GAMYB* and *GAMYBL1* mRNA levels were normalized to *ACTIN*. The miR159 RNA levels were normalized to snoR14. Measurements are the average of three technical replicates with error bars representing SEM.

To determine whether this stunted growth phenotype was inherited, seeds from three T0 independent lines (L23, L25 and L42), which all had stunted growth, were sown and T1 progeny plants were planted out and different phenotypic traits were measured. Compared to the wild-type plants, these *MIM159* transformant lines displayed pleiotropic phenotypic defects throughout development, including smaller growth stature (Figure 4.12A, B and D), shorter flag leaf (Figure 4.12C), and smaller florets containing relatively shorter anther filaments, but normal pistils (Figure 4.12E). These demonstrate the loss of miR159 has negative impacts on both vegetative and reproductive development in rice.

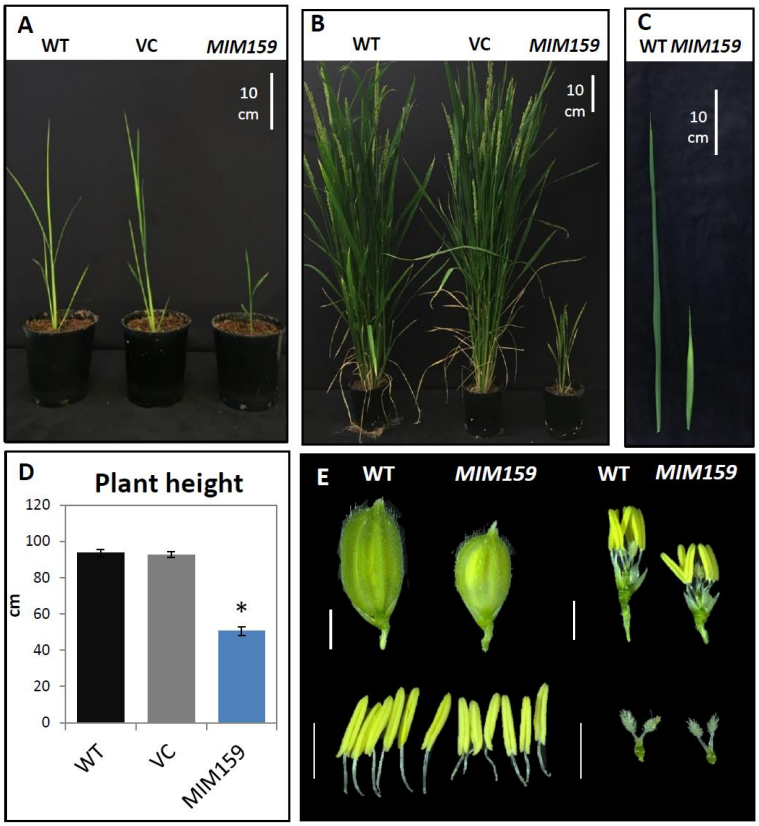


Figure 4.12. Expression of *MIM159* in rice results in phenotypic defects. (A) Three-week-old wild type, VC and *MIM159* transformant rice. Scale bar represents 10 cm. **(B)** 11-week-old *MIM159* transformant rice displayed smaller growth stature compared to wild type and VC. Scale bar represents 10 cm. **(C)** The flag leaf of wild type and *MIM159* rice. Scale bar represents 10 cm. **(D)** Average plant heights of ten wild-type plants, ten VC T1 plants and twenty *MIM159* T1 plants. These VC and *MIM159* T1 plants were coming from multiple independent transformant lines. Error bars represent the standard error of the mean (SEM), and asterisk mark represents statistically significant difference compared to wild type, statistically analyzed from *t*-test. **(E)** Phenotypes of floret, anther and pistil of wild type and *MIM159* rice. Scale bar represents 2mm.

Next, to confirm molecular effects were stably inherited in *MIM159* T1 progeny and correlated with phenotypes, transcript profiling of the *MIM159* transgene, miR159 and the *OsGAMYB/OsGAMYBL1* was performed on multiple T1 progeny plants of VC and *MIM159* rice. High expression of *MIM159* resulted in a drastic repression of miR159 in both leaves and florets of *MIM159* transformant lines, compared to wild type and VC lines (Figure 4.13A and D). Meanwhile, *OsGAMYB* and *OsGAMYBL1* transcript levels were elevated in both leaves and florets of *MIM159* transformant lines (Figure 4.13B and C), supporting the observation that *GAMYB* and *GAMYBL1* were released from miR159 repression in *MIM159* rice. These data revealed, as de-regulated *GAMYB* expression had negative impacts in vegetative growth of Arabidopsis, tobacco and rice, miR159 strong silencing of *GAMYB* seems to be conserved across monocots and dicots.

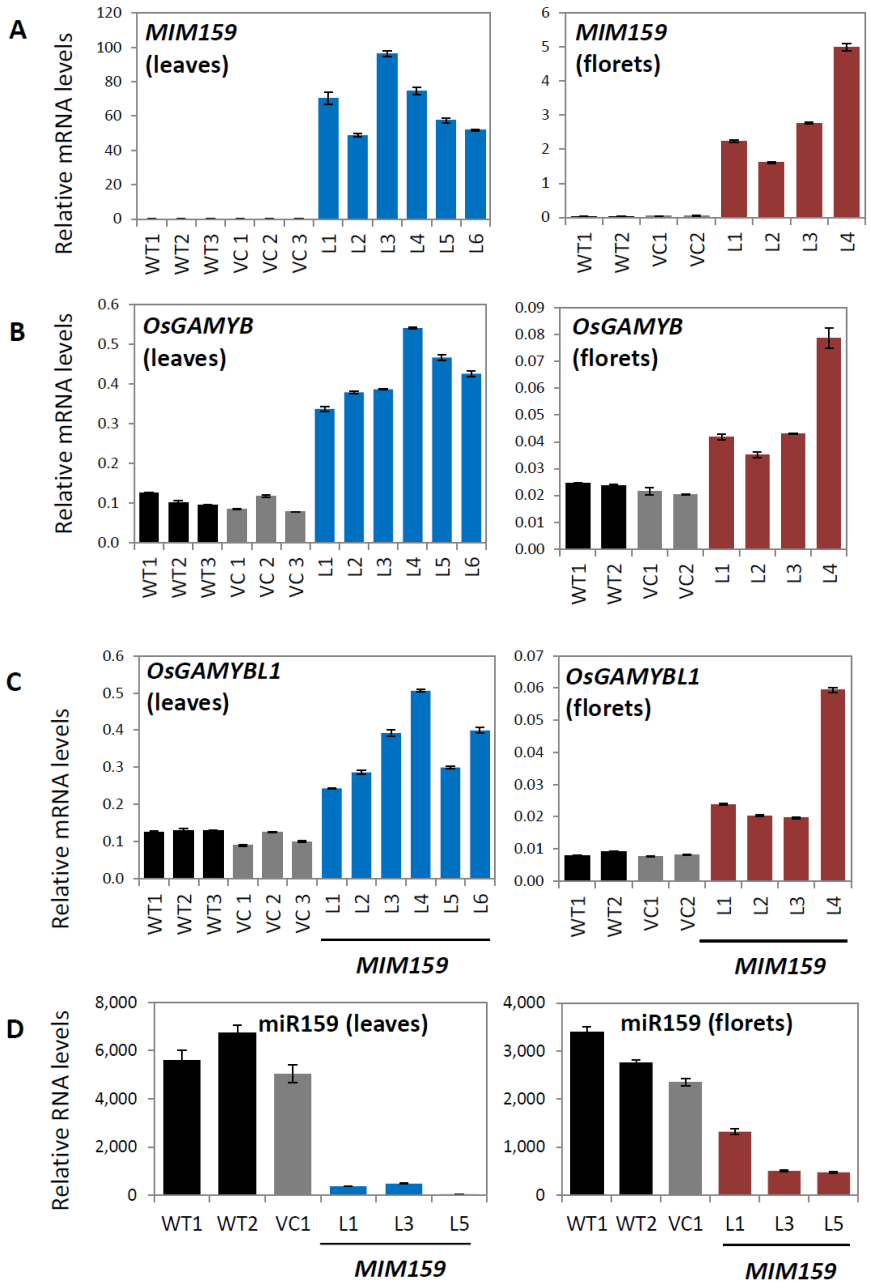


Figure 4.13. Transcript profiling in T1 *MIM159* rice. qRT-PCR measurement of *MIM159* **(A)**, *OsGAMYB* **(B)**, *OsGAMYBL1* **(C)** and miR159 **(D)** in flag leaves (left panel) and florets (right panel). RNA was individually extracted from multiple wild type plants, VC T1 plants and *MIM159* T1 plants, each being composed of three flag leaves and 30 florets. These VC and *MIM159* T1 plants were coming from multiple independent transformant lines. The *MIM159*, *GAMYB* and *GAMYBL1* mRNA levels were normalized to *ACTIN*. The miR159 RNA levels were normalized to snoR14. Measurements are the average of three technical replicates with error bars representing the standard error of the mean (SEM).

4.2.11 *PR* transcript levels were not up-regulated in miR159 loss-of-function Arabidopsis and rice

As *PR* mRNA levels were dramatically up-regulated in *MIM159* tobacco leaves (Table 4.5 and Figure 4.9A), it was investigated whether *PR* mRNA levels are also up-regulated in the Arabidopsis *mir159ab* mutant and in *MIM159* rice. However, qRT-PCR analysis found that the corresponding *PR* homologues in Arabidopsis and rice were not up-regulated and even decreased in these miR159 loss-of-function plants (Figure 4.14A and B). This suggests that up-regulation of the *PR* genes due to loss-of-function of miR159 appears to be restricted specifically to tobacco, rather than being broadly conserved in different plant species such as Arabidopsis and rice.

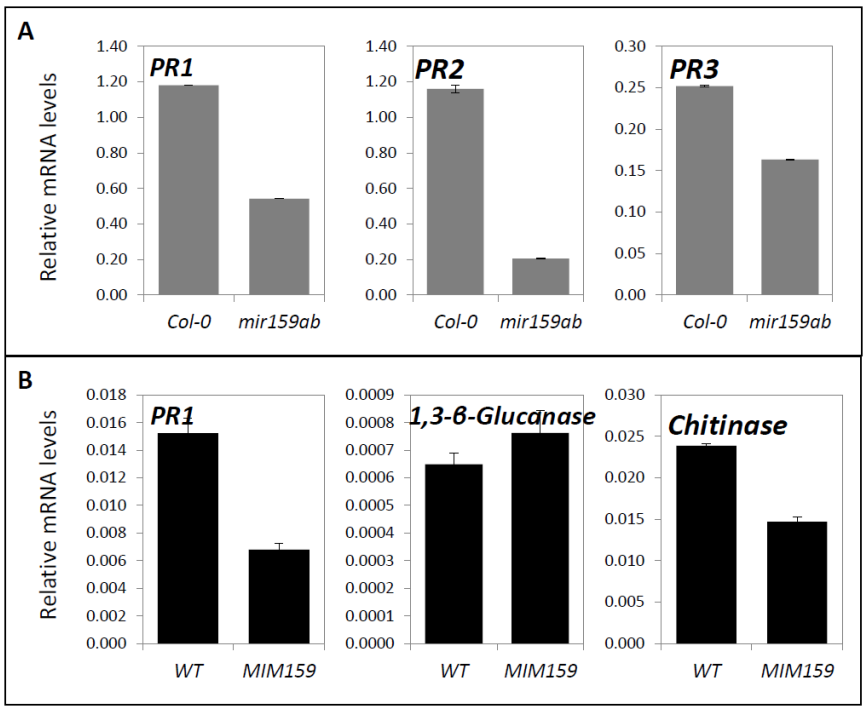


Figure 4.14. qRT-PCR measurement of *PR* mRNA levels in Arabidopsis and rice. (A) The mRNA levels of *PR* homologues in Col-0 and Arabidopsis *mir159ab* mutant. RNA was extracted from sample pools of rosette tissue of 15 randomly selected, three-week-old Arabidopsis. Measurements are relative to *CYCLOPHILIN*. **(B)** The mRNA levels of *PR* homologues in wild type and *MIM159* rice. RNA was individually extracted from three flag leaves of one wild type plant and the *MIM159* transformant Line 3. Measurements are relative to *ACTIN*. Measurements are the average of three technical replicates with error bars representing the standard error of the mean (SEM).

4.2.12 GAMYB proteins from diverse plant species activate similar pathways when expressed in Arabidopsis.

As *PR* expression was differentially regulated in tobacco and Arabidopsis/rice with respect to *GAMYB* expression, this raises the question of whether the tobacco GAMYB protein has a divergent function compared to the Arabidopsis and rice GAMYB proteins. To investigate this, homologues of *GAMYB* (or *GAMYB-like*) genes from representative plant species were aligned to analyse amino acid sequence identity. The proteins were found to be strongly conserved within the R2R3 region, but highly divergent in all other regions, especially within their respective C-terminal domains (Figure 4.15). Conserved BOX1, BOX2 and BOX3 found in Arabidopsis and rice GAMYB-like proteins (Gocal *et al*., 2001) have low similarity in moss, lycophytes and gymnosperm homologues (Figure 4.15). This raises the possibility that the divergent protein sequences in GAMYB homologues could activate different downstream pathways. To determine this, the coding regions of *GAMYB-like* genes from tobacco (*Nt*, *Nicotiana tabacum*), rice (*Os*, *Oryza sativa*) and pinetree (*La*, *Larix kaempferi*) were isolated and constitutively expressed in Arabidopsis via the 35S promoter. As these three *GAMYB-like* genes all contained a highly conserved miR159 binding site (Figure 4.2A and Figure 4.10A, Li et al., 2013b), synonymous mutations were introduced to generate miR159-resistant *mGAMYB* transgenes. These 35S-*mGAMYB* transgenes were individually transformed into Arabidopsis Col-0 plants. Multiple primary transformants were obtained for each construct and phenotypically analysed.

Expression of *mNtGAMYB2*, *mOsGAMYB* and *mLaMYB33* in Arabidopsis led to morphological defects in rosettes, all showing mild and severe upward leaf curl, similar to *mAtMYB33* transformants (Figure 4.16A). Supporting this, *CYSTEIN PROTEINASE 1* (*CP1*), a marker gene of Arabidopsis MYB protein activity (Alonso-Peral et al., 2010; Li

et al., 2016) was measured in pooled transformant lines. The *CP1* transcript levels in *mNtGAMYB2*, *mOsGAMYB* and *mLaMYB33* transformants were at similar levels compared to the *mAtMYB33* transformants, and dramatically higher than in wild type (Figure 4.16B). This suggests GAMYB proteins from diverse plant species activate similar pathways when expressed in Arabidopsis. However, regarding *PR* expression in Arabidopsis, *PR* mRNA levels in Arabidopsis *mNtGAMYB2* transformants were similar to that in Col-0 and transformant Arabidopsis expressing the *GAMYB* homologues from the other plant species (Figure 4.16C). Therefore, it appears that *NtGAMYB* only up-regulates *PR* expression in the context of tobacco.

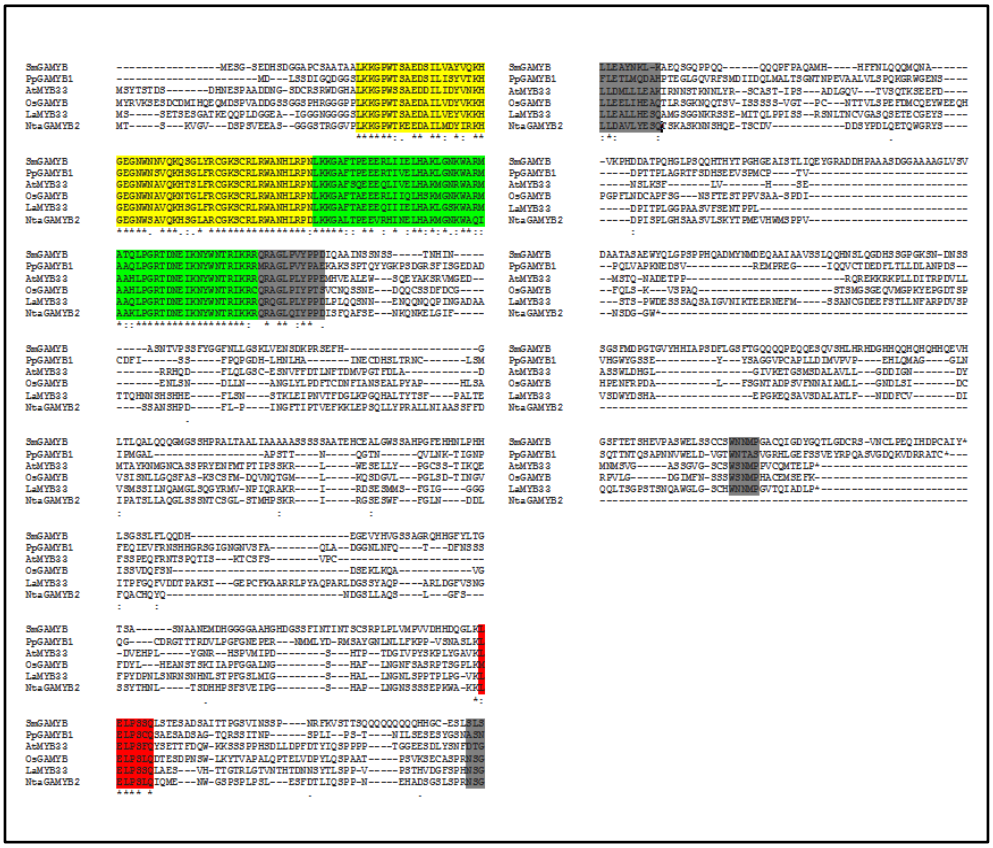


Figure 4.15. Sequence analysis of GAMYB proteins. Protein sequence alignment of GAMYB amino acid sequences from diverse plant species. Conserved sequences (asterisks at the bottom) are only in the R2 (shaded in yellow) and R3 (shaded in green) domains, and the region which the miR159 binding site (shaded in red). Conserved BOX1, BOX2 and BOX3 of GAMYB-like proteins that have been identified in Arabidopsis and rice (Gocal et al., 2001) are shaded in grey. *Sm, Selaginella moellendorffii; Pp, Physcomitrella patens; At, Arabidopsis thaliana; Os, Oryza sativa; La, Larix kaempferi; Nta, Nicotiana tabacum.*

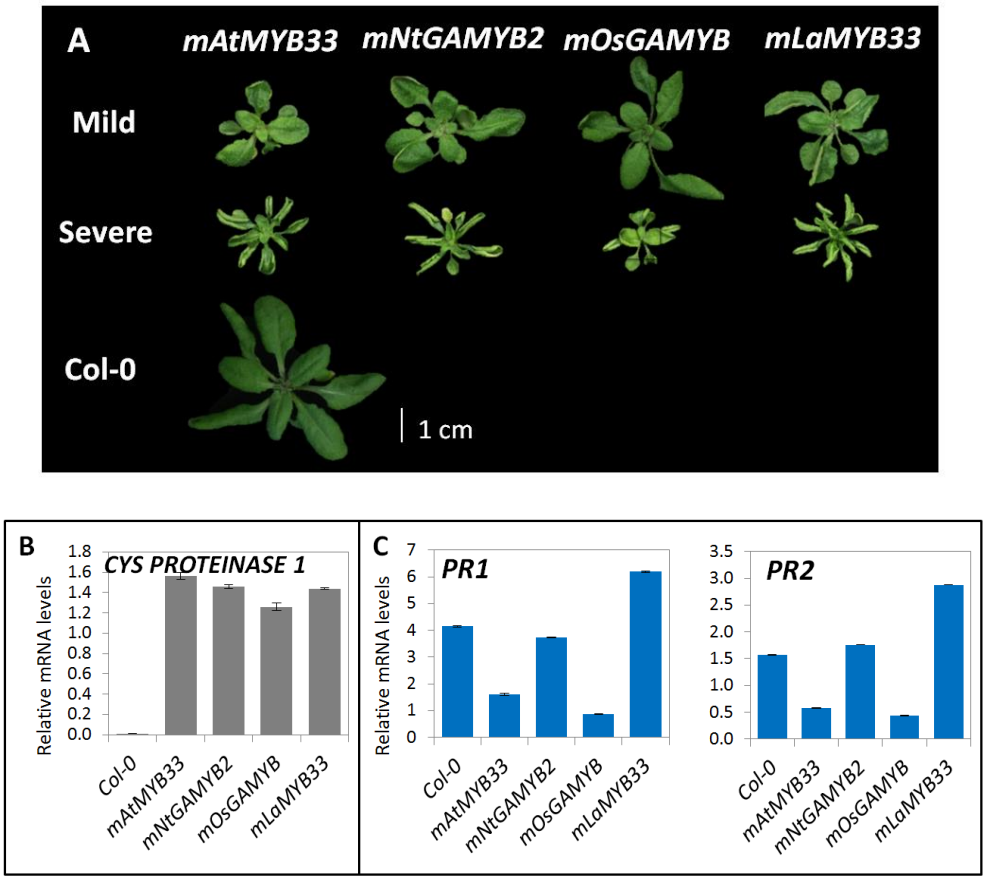


Figure 4.16. GAMYB proteins from diverse plant species activate similar pathways when expressed in Arabidopsis. (A) Three-week-old Arabidopsis transformants constitutively expressing *mAtMYB33* (*Arabidopsis thaliana*), *mNtGAMYB2* (*Nicotiana tabacum*), *mOsGAMYB* (*Oryza sativa*) and *mLaMYB33* (*Larix kaempferi*) all display similar mild and severe phenotypes. Scale bar represents 1 cm. Transcript levels of the *CYS PROTEINASE 1* (*CP1*) **(B)** and *PR*s **(C)** in 35S-*mGAMYB* primary transformants measured by qRT-PCR. RNA was individually extracted from sample pools of rosette tissue of 15 randomly selected, three-week-old primary transformants. Col-0 was used as wild-type control and mRNA levels were normalized to *CYCLOPHILIN*. Measurements are the average of three technical replicates with error bars representing the standard error of the mean (SEM).

4.3 Discussion

In this chapter, the function of the miR159-*GAMYB* pathway has been investigated in distantly related plant species, including dicotyledonous (Arabidopsis and tobacco) and monocotyledonous (rice) plants. Inhibition of miR159 resulted in similar phenotypic defects in Arabidopsis, tobacco and rice *MIM159* transformant lines, implying multiple aspects of the miR159-*GAMYB* pathway are strongly conserved. Foremost, *GAMYB* expression contributes to stunted vegetative development. Therefore, miR159 appears widely expressed throughout these plants, which is necessary to strongly silence *GAMYB* expression, enabling normal growth. This raises several questions: (1) why is *GAMYB* widely transcribed if its expression is detrimental to the plant, and; (2) why has this been strongly conserved across multiple species, given that *GAMYB* expression is strongly silenced by miR159 throughout the plant resulting in little to no net phenotypic impact of this miR159-*GAMYB* pathway.

To gain insights into these questions, RNA sequencing of *MIM159* tobacco leaves was performed to investigate downstream events of the miR159-*GAMYB* pathway. GO analyses of *MIM159* up-regulated genes found a strong enrichment of plant defence pathways. This included 22 of the 50 most up-regulated genes corresponding to *PR* genes. Confirming that these *PR* genes are downstream of *GAMYB* expression, they were strongly up-regulated in *mNtGAMYB2* tobacco leaves. However, up-regulation of *PR* expression was not observed in *mir159ab* Arabidopsis or *MIM159* rice, suggesting that GAMYB only up-regulates *PR* genes in the context of tobacco. Given the ubiquitous presence of this highly conserved miR159-*GAMYB* pathway and its ability to trigger PCD, it is proposed that this pathway may be involved in defence response, which is activated upon deregulated *GAMYB* expression via miR159 inhibition.

4.3.1 *GAMYB* expression inhibits plant growth, widely conserved in flowering plants

This chapter has elucidated the functional roles of miR159-*GAMYB* in distant plant species via expression of the *MIM159* transgene to inhibit miR159 activity. In general, *MIM159* Arabidopsis, tobacco and rice displayed similar phenotypic defects in both vegetative and reproductive tissues, including stunted growth, abnormal leaf development and smaller reproductive organs. This demonstrates deregulated *GAMYB* expression from miR159 has negative impact on plant growth, which is widely conserved across dicot and monocot species.

Other studies have inhibited miR159 activity with *MIM159* being expressed in Gloxinia and also *STTM159* in rice (Li et al., 2013c; Zhao et al, 2017). For Gloxinia, a dicotyledonous species, expressing *MIM159* did not inhibit vegetative growth, but instead promoted flowering (Li et al., 2013c). For *STTM159*-expressing rice, the plants displayed very similar morphological defects to the *MIM159* rice described here, including reduced stature, shorter leaf and smaller florets (Figure 4.12). In addition, the *MIM159* rice showed a more severe phenotype than the *STTM159* rice with regards to the plant growth (Figure 4.12, Zhao et al, 2017). Based on this, it is possible that *MIM159* has an efficacy stronger than the *STTM159* transgene in rice, which is the case in Arabidopsis (Reichel et al., 2015).

In our studies, consistent with smaller size and crinkle leaves displayed by *MIM159* tobacco (Figure 4.5E) and smaller and upward curling leaves seen in *mNtGAMYB2* tobacco (Figure 4.8E), GO analysis found that pathways related to cell cycle, cell size and cell morphogenesis, in *MIM159* tobacco leaves were significantly down-regulated (Table 4.2). Likewise, leaves of *mir159ab* Arabidopsis have decreased cell numbers and altered cell morphology in rosette leaves (Alonso-Peral et al., 2010). Also, STTM159 rice plants

had decreased cell layers in parenchymatous tissue in stems and reduced numbers of small veins in the flag leaf (Zhao et al., 2017), and a less well established vascular network was also reported in Arabidopsis plants (Alonso-Peral et al., 2010). Therefore, these smaller phenotypes of vegetative tissues are speculated to be due to the repressed pathways of cell cycle and cell morphogenesis mediated by *GAMYB* expression. On the basis of vegetative defects from miR159 loss-of-function effects, it reveals that *GAMYB* expression inhibits vegetative growth, conserved in multiple plant species.

Furthermore, the miR159-*GAMYB* pathway may play a role in the developmental transition from vegetative growth to flowering. Supporting this, *MIM159* tobacco plants have a delayed flowering time compared to wild type (Figure 4.5H), and of the four *mNtGAMYB2* tobacco lines that displayed severe phenotypic defects, only one line eventually flowered (Figure 4.8C). However, it cannot be distinguished whether the miR159-*GAMYB* pathway plays a direct role in flowering, or whether the defect in flowering of these *MIM159* and *mNtGAMYB2* tobacco lines is a secondary effect from the deleterious impacts of *GAMYB* expression, which has an overall negative effect on growth. One possible mechanism that *GAMYB* expression could up-regulate *MIR156* levels, further resulting in a delayed vegetative phase transition, which has been demonstrated in Arabidopsis (Guo et al., 2017). Therefore, a coordinated action between the miR159-*MYB33* and miR156-*SPL* pathways in modulating phase transition and flowering time has been proposed in Arabidopsis (Guo et al., 2017; Teotia and Tang, 2015). Supporting this, the RNA sequencing of *MIM159* tobacco leaves found down-regulated expression of multiple members of the *SPL* gene family, targets of miR156 (Table 4.3).

However, previous reports have claimed that GAMYB acts as a positive regulator of flowering time via activation of *LEAFY* expression in Arabidopsis ecotype Landsberg and in Gloxinia (Achard et al., 2004; Li et al., 2013c). These seem to be inconsistent with our tobacco data. Therefore, the roles of miR159-*GAMYB* in controlling plant flowering need to be further investigated.

4.3.2 Possible roles of miR159-*GAMYB* in vegetative tissues

Although *GAMYB* genes are widely transcribed in vegetative tissues, miR159 accumulates to very high levels in the same tissues to ensure strong silencing of *GAMYB* expression, resulting in the miR159-*GAMYB* pathway having little to no physiological impact. In this regard, continual *GAMYB* transcription/silencing could be considered a largely dispensable pathway in vegetative tissues. This raises the question: what is the selective pressure to conserve the miR159-*GAMYB* pathway in vegetative tissues, especially given that *GAMYB* expression has such negative impacts on vegetative growth. In Arabidopsis, most of the *GAMYB* homologues are not expressed in vegetative tissues, but rather, their transcriptional domains are restricted to seeds and anthers (Allen et al., 2007; Allen et al., 2010). Given this, there must be strong selection for this pathway to be retained in vegetative tissues, as this *GAMYB* transcription occurs in vegetative tissues of both monocot and dicot species.

Previous reports have proposed some speculations. First possibility is that the GAMYB protein does play a tissue-specific or timing-specific role in vegetative growth. Guo et al. (2017) found the Arabidopsis *myb33* mutant has a subtle but significant alteration to leaf trichome production and vegetative phase change, suggesting that the low level of MYB33 protein expression occurs in the rosette. In this paper, Arabidopsis MYB33 protein has been determined to promote *MIR156* transcription by binding directly to the

MYB33 gene promoter, which results in plants remaining in vegetative stage (Guo et al., 2017). However, it is worthy to note that the subtle alteration of vegetative phase change in the *myb33* mutant did not impact primary vegetative growth or contribute to any defects in leaf morphogenesis (Guo et al., 2017).

Secondly, the miR159-*GAMYB* pathway is suggested to be involved in heat stress response (Wang et al., 2012); however, this report did not provide any firm evidence of this proposed role for the miR159-*GAMYB* pathway. In addition, Li et al., (2016) subjected the Arabidopsis *mir159ab.myb33.myb65* quadruple mutant (appears wild-type phenotype) to a number of different abiotic stresses. However, no phenotypic differences were found between the wild type and *mir159ab.myb33.myb65* plants regarding their abiotic stress responses, arguing that the miR159-*MYB33/MYB65* pathway has no major role during abiotic stresses (Li et al., 2016).

Based on the previous observations, *GAMYB* function in vegetative growth remains to experimentally validated. In this tobacco study, *GAMYB* expression promotes universal plant defence responses, including strong up-regulation of many *PR* families, in *MIM159* and *mNtGAMYB2* tobacco leaves. This provides a new insight that the miR159-*GAMYB* pathway might be in response to plant-pathogen interactions.

4.3.3 *GAMYB* expression in tobacco promotes downstream gene expression invloved in plant defence responses

As *GAMYB* genes encode conserved R2R3 MYB transcription factors, they potentially regulate a variety of downstream events. Of these, the pathways related to plant defence response are most up-regulated in *MIM159* tobacco leaves (Table 4.1). Consistently, RNA sequencing and qRT-PCR results have demonstrated *PR* mRNA levels are strongly

correlated with *GAMYB* expression in both *MIM159* and *mNtGAMYB2* tobacco (Table 4.5 and Figure 4.9). Furthermore, dramatic up-regulation of *PR* expression is universal for many *PR* families. Previous reports have been identified that these *PR* families have various functions in plant resistance against pathogens (Ebrahim et al., 2011; Edreva, 2005; Sinha et al., 2014). In addition to the *PR* genes, other genes, such as *PAR-1c* and *Sar8.2*, which are implicated in the plant defence response pathways (Herbers et al., 1995; Song and Goodman, 2002), are also strongly up-regulated in *MIM159* tobacco leaves. Therefore, *GAMYB* expression in tobacco appears to promote broad-spectrum plant defence responses.

In complex signalling pathways of plant defence responses, hypersensitive response (HR) is a major type effector-triggered immunity (ETI). HR signals can modulate expression of defence-related genes, such as *PR*s, and result in HR cell death, which is a form of programmed cell death (PCD) at the site of pathogen invasion and has a major feature of chloroplast disruption (Coll et al., 2011; Mur et al., 2007; Teh and Hofius, 2014). In the *MIM159* and *mNtGAMYB2* tobacco plants, leaf chlorosis was observed (Figure 4.5D and Figure 4.8B), consistent with significantly up-regulated PCD related pathways (Table 4.1). Furthermore, many *NUCLEOTIDE-BINDING SITE AND LEUCINE-REPEAT* (*NBS-LRR*) and *METACASPASE-1* gene homologues were strongly up-regulated in *MIM159* tobacco leaves (Table 4.6). These two protein families are major upstream elicitors of HR cell death. In plants, NBS-LRR proteins, served as plant immune receptors, can specifically recognise pathogen effectors and mediate activation of metacaspase-1, which positively regulates HR cell death (Coll et al., 2011; Teh and Hofius, 2014). Therefore, I speculate that leaf chlorosis in the *MIM159* and *mNtGAMYB2* tobacco plants could be a result of HR, triggered by *GAMYB* expression.

4.3.4 GAMYB promotes differential downstream responses between plant species

Transcript profiling results reveal that *PR* mRNA levels were strongly induced in the *MIM159* and *mNtGAMYB2* tobacco lines (Figure 4.9), but *PR* levels were unchanged in the Arabidopsis *mir159ab* mutant or *MIM159* rice (Figure 4.14). In addition, *PR* levels were not up-regulated when tobacco *GAMYB* was expressed in Arabidopsis (Figure 4.16C). Therefore, GAMYB appears to up-regulate *PR* levels only when expressed in the context of tobacco (Figure 4.17). To explain this differential response between species, it is possible that some tobacco-specific intermediate(s) exist. When *GAMYB* is expressed in tobacco, these intermediate(s) are activated and then promote defence response pathways. It may be that the intermediate(s) are only present in tobacco, or that additional factors/cues are required to activate the defence response pathways in Arabidopsis and rice. Despite the conserved nature of the miR159-*GAMYB* pathway, these results highlight that species-specific differences exist.

However, some doubts remain in the above speculation. Firstly, as many cases of dwarfing plants have the auto-immune responses that promote expression of defence-related genes (Gou and Hua, 2012), we cannot rule out a possibility that the strong up-regulation of defence-related gene expression in *MIM159* tobacco was due to the secondary effects of the stunted growth. Secondly, Nicotiana species is likely to be more sensitive to pathogen responses, which renders Nicotiana as a good model of plant disease research (Jube and Borthakur, 2007; Sinha et al., 2014). Therefore, how the miR159-*GAMYB* pathway controls tobacco plant defence requires further investigation via pathogen infection of *MIM159* tobacco leaves. It is possible that some particular factors/biotic stresses could inhibit miR159 activity, resulting in deregulated *GAMYB* expression that induces plant defence response. Therefore, it is necessary to perform further studies of biotic stresses, may be conserved or divergent between species,

investigating a clear role of the miR159-*GAMYB* pathway in vegetative growth. This is significant to explain why the miR159-*GAMYB* pathway constantly exists under a strong evolutionary force; and may provide a new insight of how the miR159-*GAMYB* pathway diverged during evolution.

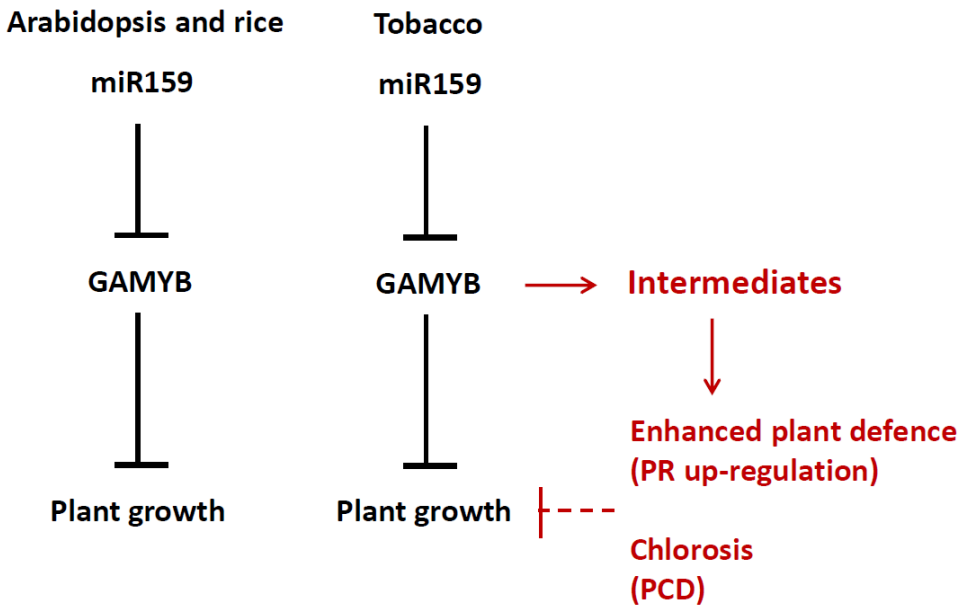


Figure 4.17. The miR159-*GAMYB* mystery: does it play a role in plant defence response? In vegetative tissues, miR159 strongly silences *GAMYB* to levels that have very little phenotypic impact. When *GAMYB*s are deregulated from miR159, *GAMYB* expression is detrimental to plant growth, widely conserved in Arabidopsis, tobacco and rice. However, in tobacco, *GAMYB* expression might activate the tobacco-specific intermediate(s), which promote plant defence response, confirmed by universal *PR* up-regulation; and contribute to leaf chlorosis, potentially resulted from the increased PCD.

Chapter 5

General discussion

5.1 Why have *GAMYB*s evolved to contain conserved stem-loop RNA structures?

In Arabidopsis, miR159 displays differential silencing efficacies against eight *GAMYB-like* homologues, despite each transcript containing a highly complementary miR159 binding site. Regarding the degree to which they are silenced by miR159, there are two classes: *MYB33* and *MYB65* are strongly silenced, whereas the other *MYB*s are poorly silenced. Interestingly, these two classes of *MYB*s are characterised by their RNA secondary structures, where *MYB33* and *MYB65* transcripts harbouring stem-loop RNA structures abutting the miR159 binding site. These structures are present in *MYB33/MYB65* that are widely transcribed in vegetative tissues; whereas they are absent in other *GAMYB-like* homologues that are only transcribed in seed and pollen where miR159 appears absent or of low activity (Alonso-Peral et al., 2010; Leydon et al., 2013; Li et al., 2016; Liang et al., 2013; Reyes and Chua, 2007). Taken together, the results presented in Chapter 3 strongly suggest that selection for such structures only occurs for homologues transcribed in vegetative tissues as to confer strong silencing (Figure 5.3).

We are curious whether this trend is specific for Arabidopsis *GAMYB-like* members, or whether a similar scenario occurs in other plant species. In numerous angiosperm and gymnosperm plant species, the stem-loop RNA structures are strongly predicted to reside in *GAMYB-like* genes, which are homologous to Arabidopsis *MYB33*. Of these, tobacco and rice *GAMYB-like* homologues have been verified to be transcribed in vegetative tissues. We speculate that the stem-loop RNA structures have arisen to render these *GAMYB-like* homologues that are transcribed in vegetative tissues sensitive to miR159 regulation (Figure 5.3). For example, *NtGAMYB2* over-expression in tobacco strongly silenced, due to the result that 35S-*NtGAMYB2* transformant tobacco displayed a wild-type phenotype (data not shown). However, it is unknown whether there are other *GAMYB-like* members that are only transcribed in seeds or reproductive tissues, in species

beyond Arabidopsis. The *AtMYB81*, *AtMYB101* and *AtMYB104* genes are predicted to have close homologues in *Brassicaceae* species but not in other distant species. Of these, the *MYB104-like* homologue in *Brassica napus* has the conserved nts in the 5' region abutting the miR159 binding site and is predicted to have loosely conserved stem-loop RNA structures, whereas the *MYB101-like* homologue is not (Figure 5.1). However, their transcriptional domains have not been characterised and how strongly they are silenced by miR159 *in planta* is unknown. Furthermore, another example is the poplar *GAMYB-like* homologue gene, *PtrMYB012*, whose transcript is predicted to have the conserved stem-loop RNA structures abutting the miR159 binding site; this gene can be strongly silenced by miR159 in poplar, as 35S-*PtrMYB012* poplar displayed a wild-type phenotype (Kim et al., 2017). However, the *PtrMYB012* mRNA levels are much higher in flowers relative to other vegetative tissues, suggesting it strongly transcribed in flowers (Kim et al., 2017). The case of poplar *PtrMYB012* seems to be conflicting with my speculation; but it cannot rule out a possibility that *PtrMYB012* would be still transcribed but strongly silenced by miR159 in poplar vegetative tissues; whereas it is highly expressed in flowers due to the low miR159 activity. This needs to be further proven in the miR159 loss-of-function poplar.

Therefore, it would be of interest to further explore whether miR159 displaying differential silencing efficacies against the *GAMYB-like* members occurs in other species and whether silencing efficacies of these *GAMYB*s are correlated with existence of the stem-loop RNA structures and their transcription in vegetative tissues.

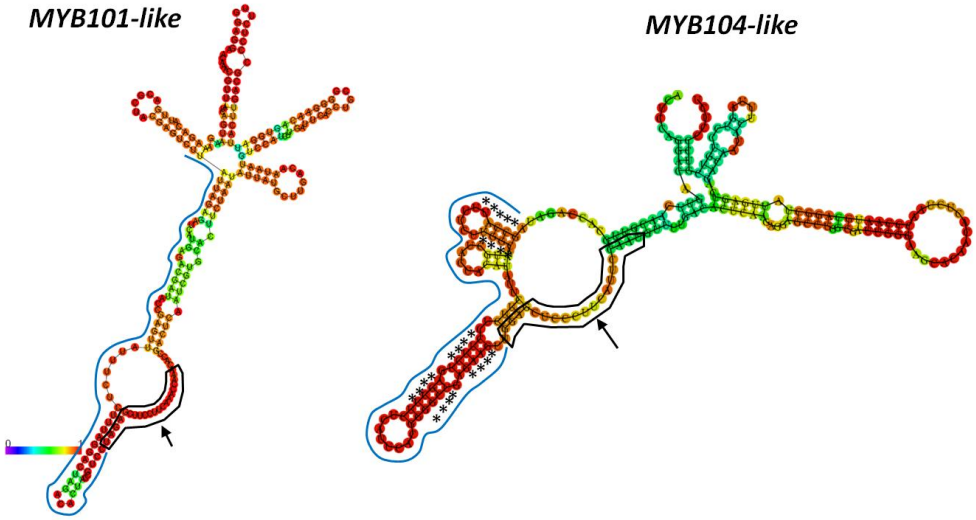


Figure 5.1. Prediction of RNA secondary structures of *MYB101* and *MYB104* homologues in *Brassica napus*. Arabidopsis *MYB101* has a close *MYB101-like* homologue in *Brassica napus*. Arabidopsis *MYB81* and *MYB104* have the same *MYB104-like* homologue in *Brassica napus*. RNA structures are predicted by the Vienna RNAfold web server. The miR159 binding sites are outlined in black and the cleavage sites indicated with arrows. The predicted secondary structures of SL1 and SL2 homologous, compared to Arabidopsis *MYB33*, is labelled with blue curve. The conserved nucleotides depending on the Figure 3.7C are labelled with asterisks, only present in *MYB104-like*, but not in *MYB101-like*. The heat maps indicate the probability of second structure formation, from low (purple) to high (red).

5.2 Target RNA secondary structure is an important but not absolute determinant of miRNA efficacy

This study has proposed that target RNA secondary structure is an important determinant of miRNA efficacy, demonstrated in the *MYB33-mSL* systems. I found that the RNA structure in *MYB33* correlated with strong silencing efficacy; introducing mutations to disrupt either SL attenuated miR159 efficacy, while introducing mutations to form an artificial stem-loop structure adjacent to a miRNA-binding site restored strong miR159-mediated silencing.

However, the similar stem-loop structures engineered into *MYB81* did not improve miR159 silencing efficacy. In addition, the poplar *PtrMYB012*, homologue to

Arabidopsis *MYB33*/*MYB65*, whose transcript is predicted to have the stem-loop structures, similar to *MYB33*. However, it was poorly silenced by miR159 when expressed in Arabidopsis; despite it harbouring a conserved binding site highly complementary to *At*-miR159a (Kim et al., 2017). Interestingly, *PtrMYB012* over-expression in poplar did not contribute to any phenotypic differences compared to wild type (Kim et al., 2017); and introduction of the poplar *MIR159* transgene into the 35S-*PtrMYB012* transformant Arabidopsis rescued the mutant phenotypes (Kim et al., unpublished data). On the basis of Kim Lab's results, poplar *PtrMYB012* is only strongly silenced by poplar miR159 but not Arabidopsis miR159, possibly due to the sequence of poplar miR159 different to that of Arabidopsis miR159 in the 7th nucleotide (Kim et al., 2017). This reveals that RNA secondary structure is an important but not absolute factor in determining efficacies of miRNA-mediated silencing. In addition, RNA secondary structure may interplay with other factors, like RNA binding proteins (RBPs), mutually determining miRNA silencing efficacy. This has been demonstrated in animal systems, where a RBP binding on the target mRNA results in a change of the local RNA secondary structure to further promote accessibility of the miRNA target site (Kedde et al., 2010). If the RBP expression occurs in a spatial-temporal manner, there may be differential silencing efficacies in different tissues or at different developmental stages. For example, *MYB33* is strongly silenced by miR159 in vegetative tissues, but surprisingly poorly silenced in germinating seeds, despite miR159 being expressed at similar levels in the two tissues (Alonso-Peral et al., 2012). Although what factors contributing to the differential silencing outcomes for *MYB33* remain unknown, I speculate there is a tissue-specific RBP impacting on the *MYB33* RNA secondary structures, rendering differential silencing efficacies by miR159. These potential factors that can determine miRNA-target interactions are evolving under selective pressure, rendering the optimal miRNA regulatory action on plant fitness.

5.3 Difficulties in investigating how target RNA structure determines miRNA silencing efficacy

Although this study has revealed target RNA structure is a major determinant of miRNA silencing efficacy, how target RNA structure determines miRNA silencing efficacy remains unclear. This will be likely a difficult question to resolve. Firstly, bioinformatic prediction of RNA structures depending on thermodynamics is not very reliable. For example, all of three tobacco *NtGAMYB* homologues have conserved nts in the 5' region abutting the miR159 binding site, which enable being base-paring and form the stems of SL structures, similar to Arabidopsis *MYB33*. However, these conserved nts of *NtGAMYB*s are not predicted to form the conserved SL1, using the Vienna RNAfold software (Figure 5.2). Subtle nucleotide changes may interfere with thermodynamic folding of RNA secondary structures, and *in silico* prediction would be much different to the *in vivo* RNA structures. Therefore, the *in vivo* target RNA structures need to be established instead of relying on bioinformatic prediction. One such approach is dimethyl sulphate (DMS) sequencing which can provide a map of *in vivo* RNA secondary structures with high experimental confidence in plants (Ding et al., 2014; Kwok et al., 2013). Such a methodology could be used to determine *in vivo* RNA structures of miRNA targets. However, the transcript levels of many endogenous miRNA targets are low, a situation likely caused by being strongly silenced by miRNAs. Although DMS sequencing has successfully identified the RNA structures of a low-abundance mRNA (*GPR3S*) (Kwok et al., 2013), Arabidopsis eFP browser reveals its abundance is still ten-fold more than the abundance of many miRNA targets, like *MYB33* (Waese et al., 2017). Therefore, determining *in vivo* RNA structures of miRNA targets by DMS sequencing is still challenging due to the low signals hard to be detected. In addition, RNA structures are dynamic, with their folding likely being affected by many factors, resulting in

differential miRNA-target regulatory efficacies. Therefore, as these *in vivo* RNA structures are difficult to determine, investigating the role of target RNA structure in miRNA silencing efficacy will remain predictive and inaccurate.

Although studies of how target RNA structure determines miRNA silencing efficacy will be time- and labour-consuming, the mechanism is significant to be incorporated in applications. First important application is the prediction of miRNA targets, still being largely dependent on sequence complementarity after development in many years (Dai et al., 2011; Dai et al., 2018). As complementarity-based *in silico* prediction of target genes often fails to accurately predict functionally relevant targets (Li et al., 2014a), target RNA structure incorporated in the predicted programs will be expecting to narrow down functional miRNA targets. In addition, the particular RNA secondary structures should be considered in design of the artificial miRNA decoys, promoting efficacy. For example, the STTM165/166 contains a 48-nt linker between two miR165/166 binding sites, which can fold into a strong stem RNA structure to stabilize the *STTM165/166* transcript, increasing accessibility of the two miRNA binding sites (Yan et al., 2012). Although STTM decoys for other miRNA families display varying efficacies (Reichel et al., 2015), this still provides a possibility that RNA structure incorporated in the miRNA decoy system increases its efficacy.

A.

```
AtMYB33         GCCATACCCCTACGGATGGCATTGTTCCTTATTCTAAGCCCTTATATGGGGCAGTGAAGC
NtGAMYB1        GCCATGCCCCTTTAAATGGCAACTCTTCTTCTTCAGAGCC---CAAATGGGCTAAGAAGC
NtGAMYB2        GCCATGCCCCTTTAAATGGCAACTCTTCTTCTTCAGAGCC---CAAATGGGCTAAGAAGC
NtGAMYB3        GCCATGCCCATTTAAATGGCAACTCCTCTTCTTCAGAGCC---CTCACGGGCAAAGAAGC
                ***** *** *    ******      *** ***  ****        ****   *****

AtMYB33         TGGAGCTCCCTTCATTCCAAT
NtGAMYB1        TGGAGCTCCCTTCACTCCAAA
NtGAMYB2        TGGAGCTCCCTTCACTCCAAA
NtGAMYB3        TCGAGCTCCCTTCACTCCAAA
                * ************ *****
```

B.

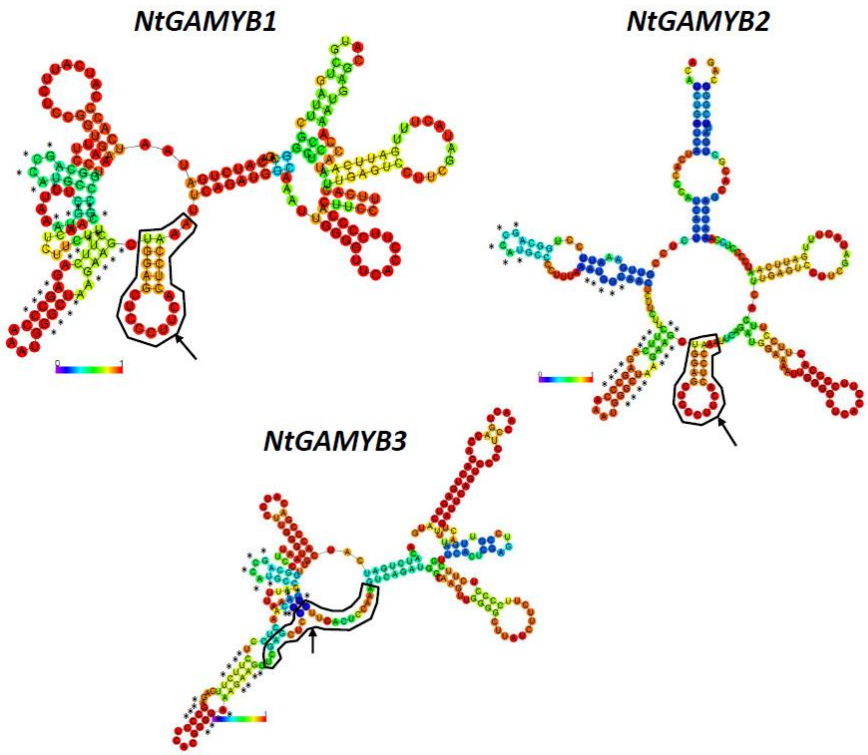


Figure 5.2. Prediction of RNA secondary structures of *NtGAMYB* homologues. (A) Sequence alignment of three tobacco *NtGAMYB*s and Arabidopsis *MYB33* nucleotides (nts). Conserved nts are marked with asterisk. Red, green and purple labelling nts are conserved nts that expect to be base-paring in the stem of SL RNA structures, referring to Figure 3.7C. Blue labeling nts are the miR159 binding site. **(B)** RNA structures of *NtGAMYB*s predicted by the Vienna RNAfold web server. The miR159 binding sites are outlined in black and the cleavage sites indicated with arrows. The conserved nucleotides depending on the Figure 3.7C are labelled with asterisks. The heat maps indicate the probability of second structure formation, from low (purple) to high (red).

5.4 What is the functional role(s) of the ancient miR159-*GAMYB* pathway in vegetative tissues?

In plants, many miRNA-target relationships are ancient and they appear to play fundamental roles in plant growth and development. For example, the miR156-*SPL* pathway controls vegetative phase transitions and floral transition (Teotia and Tang, 2015; Zhang et al., 2015); the miR165/miR166-*PHV*/ *PHB*/ *REV* pathway controls leaf polarity (Chen, 2009; Emery et al., 2003; Rubio-Somoza and Weigel, 2011); and the miR319-*TCP* pathway controls leaf morphogenesis (Palatnik et al., 2007; Rubio-Somoza and Weigel, 2011). This regulation results in complex temporal and spatial expression of the miRNA target genes. By contrast, regarding the ancient miR159-*GAMYB* regulatory pathway, where *GAMYB* are widely transcribed in vegetative tissues, appear strongly silenced by miR159; and as a result, there has been no major functional role ascribed to this regulatory pathway. The *GAMYB* genes were named "GA" because they were found to be involved in the gibberellin (GA) signalling pathway and positively regulated by GA in tissues, such as in barley seed aleurone (Gubler et al., 1995), rice anther tapetum (Aya et al., 2009) and strawberry fruit receptacle (Csukasi et al., 2012). However, in vegetative tissues, the levels of miR159 and *GAMYB* remain unaltered after GA treatment (Alonso-Peral et al., 2010). When miR159 is inhibited, multiple aspects of plant vegetative growth are altered; but what biological processes are being interfered remains unclear (Figure 5.3). Therefore, it is crucial to figure out what factors inhibit miR159 function, rendering GAMYB protein expression under particular developmental or environmental conditions.

5.4.1 Does the miR159-*GAMYB* pathway play a role in plant defence response?

In the tobacco context, when miR159-mediated regulation is inhibited, *GAMYB* expression promotes broad-spectrum plant defence responses, confirmed by the strong up-regulation of *PATHOGENESIS-RELATED* (*PR*) families that encode PR proteins as

the final products of plant immunity that directly act on pathogens (Figure 5.3). This finding provides a new insight that the miR159-*GAMYB* pathway may be respond to biotic stresses, such as plant-pathogen interactions. I speculate that in response to pathogen attack miR159 function is inhibited, resulting in *GAMYB* expression that promotes various plant defence responses to protect it from further pathogen attack. Interestingly, inhibition of miR159 throughout the *MIM159* and *mNtGAMYB2* tobacco plants displaying dwarf phenotypes is reminiscent of the auto-immune mutants, which similarly display dwarf and increased cell death but also high levels of *PR* expression without pathogen triggers (Gou and Hua, 2012; Spoel and Dong, 2012).

Data so far is insufficient to demonstrate whether the miR159-*GAMYB* pathway plays a role in plant defence. Therefore, further studies are required to investigate this and determine what biotic factors can inhibit miR159 silencing of *GAMYB*. We are performing experiments of pathogen infection of *MIM159* tobacco leaves, such as *Phytophthora* infection. The *PR1a* and *Sar8.2* expression in tobacco increases resistance to *Phytophthora parasitica* (Alexander et al., 1993; Song and Goodman, 2002), and these two genes were strongly up-regulated in *MIM159* tobacco (Table 4.4 and Table 4.5). Therefore, we speculate that the *MIM159* tobacco plants will be more resistant to *Phytophthora*. If demonstrated, the factors that inhibit the miR159-mediated silencing of *GAMYB* will need to be identified, and this will go a long way in determining the functional role of the miR159-*GAMYB* pathway in vegetative tissues.

5.4.2 Is *GAMYB*-mediated plant defence response specific in Solanaceae?

Although the phenomenon that miR159 strongly silences *GAMYB* is widely conserved in distant plant species, however, when miR159 is absent, GAMYB downstream events are activated, which are divergent in different species. For example, *PR*s were only up-

regulated in tobacco, but not in Arabidopsis and rice. In addition, tomato infected by *Phytophthora infestans* contributed to the decreased miR159 levels and increased *MYB* levels (Luan et al., 2015), suggesting that miR159-*GAMYB* might be involved in *Phytophthora* resistance in tomato. Therefore, we speculate whether the miR159-*GAMYB* pathway responding to plant defence is only specific in Solanaceae species. To determine this, *MIM159* will be expected to be transformed into tomato and potato, investigating whether transformants activate similar *GAMYB* downstream events and whether they are resistant to pathogen infection.

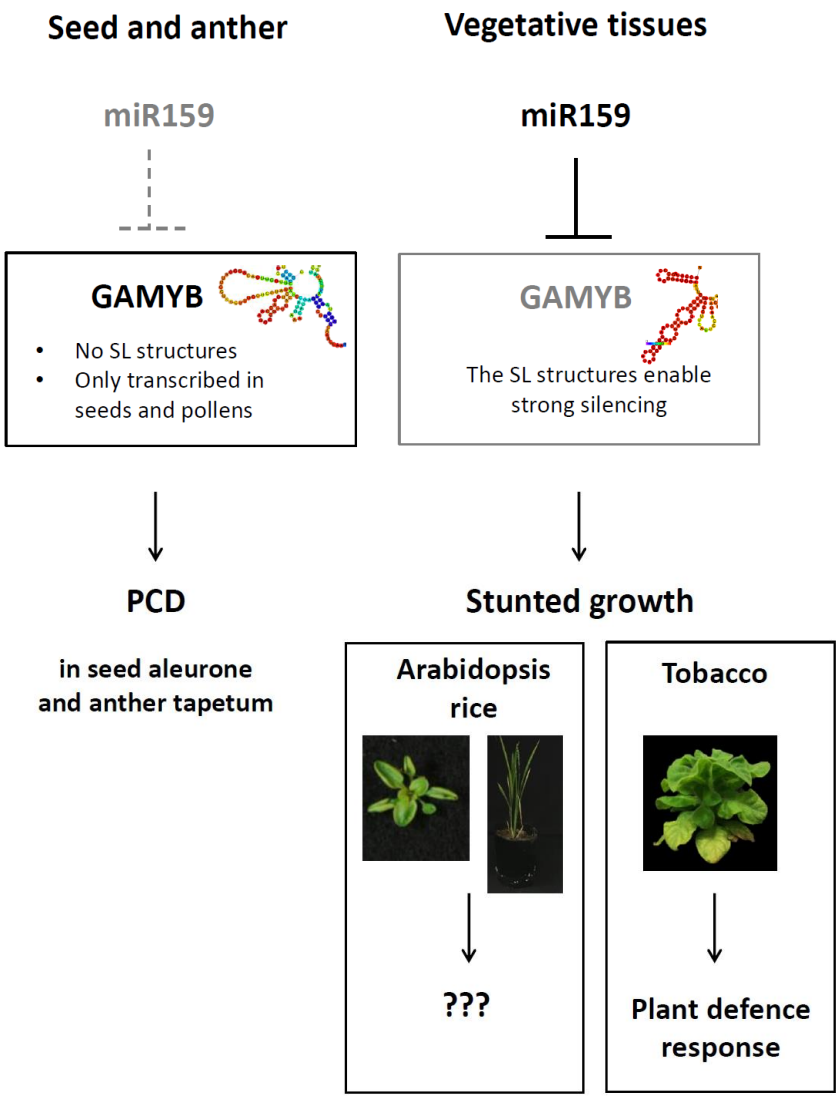


Figure 5.3. Proposed model of the miR159-*GAMYB* pathway in plants. The ancient miR159-*GAMYB* pathway is present from basal land plants to flowering plants. GAMYB-like protein expression occurs in germinating seeds and anthers, where the miR159 is absent. In Arabidopsis, these seed/pollen-specific *GAMYB-like* homologues are not predicted to have conserved RNA structures; they play a conserved role of promoting programmed cell death (PCD) in the seed aleurone and anther tapetum, essential for seed germination and anther fertility. By contrast, in vegetative tissues, miR159 is widely present and strongly silence *GAMYB* to little physiological impact. These *GAMYB*s are predicted to have the conserved stem-loop RNA structures abutting the miR159 binding site, required for the sensitive regulation of miR159. When miR159 is inhibited, deregulated *GAMYB* expression contributes to the stunted growth, widely conserved in Arabidopsis, rice and tobacco. In tobacco, *GAMYB* expression seems to promote universal plant defence responses, whereas this scenario appears not in Arabidopsis and rice, suggesting divergent function of the miR159-*GAMYB* pathway between species.

References

Achard, P., Herr, A., Baulcombe, D. C., & Harberd, N. P. (2004). Modulation of floral development by a gibberellin-regulated microRNA. *Development*, **131**(14): 3357-3365.

Addo-Quaye, C., Eshoo, T. W., Bartel, D. P., & Axtell, M. J. (2008). Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. *Current Biology*, **18**(10): 758-762.

Akdogan, G., Tufekci, E. D., Uranbey, S., & Unver, T. (2016). miRNA-based drought regulation in wheat. *Functional & integrative genomics*, **16**(3): 221-233.

Alexander, D., Goodman, R. M., Gut-Rella, M., Glascock, C., Weymann, K., Friedrich, L., Maddoxi, D., Ahl-Goy, P., Luntz, T., Ward, E. & Ryals, J. (1993). Increased tolerance to two oomycete pathogens in transgenic tobacco expressing pathogenesis-related protein 1a. *Proceedings of the National Academy of Sciences*, **90**(15): 7327-7331.

Alexander, D., Stinson, J., Pear, J., Glascock, C., Ward, E., Goodman, R. M., & Ryals, J. (1992). A new multigene family inducible by tobacco mosaic virus or salicylic acid in tobacco. *Molecular plant-microbe interactions: MPMI*, **5**(6): 513-515.

Allen, R. S., Li, J., Alonso-Peral, M. M., White, R. G., Gubler, F., & Millar, A. A. (2010). MicroR159 regulation of most conserved targets in Arabidopsis has negligible phenotypic effects. *Silence*, **1**(1): 18.

Allen, R. S., Li, J., Stahle, M. I., Dubroué, A., Gubler, F.; & Millar, A. A. (2007). Genetic analysis reveals functional redundancy and the major target genes of the Arabidopsis miR159 family. *PNAS.* **104:** 16371-16376.

Alonso-Peral, M. M., Li, J., Li, Y., Allen, R. S., Schnippenkoetter, W., Ohms, S., White, R. G., & Millar, A. A. (2010). The microRNA159-regulated GAMYB-like genes inhibit growth and promote programmed cell death in Arabidopsis. *Plant physiology.* **154:** 757-771.

Alonso-Peral, M. M., Sun, C., & Millar, A. A. (2012). MicroRNA159 can act as a switch or tuning microRNA independently of its abundance in Arabidopsis. *PLoS One*, **7**(4): e34751.

Alves-Junior, L., Niemeier, S., Hauenschild, A., Rehmsmeier, M., & Merkle, T. (2009). Comprehensive prediction of novel microRNA targets in Arabidopsis thaliana. *Nucleic acids research*, **37**(12): 4010-4021.

Ameres, S. L., Martinez, J., & Schroeder, R. (2007). Molecular basis for target RNA recognition and cleavage by human RISC. *Cell*, **130**(1): 101-112.

Aukerman, M. J., & Sakai, H. (2003). Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes. *The Plant Cell*, **15**(11): 2730-2741.

Axtell, M. J. (2013). Classification and comparison of small RNAs from plants. *Annual review of plant biology*, **64**: 137-159.

Axtell, M. J., & Bartel, D. P. (2005). Antiquity of microRNAs and their targets in land plants. *The Plant Cell*, **17**(6): 1658-1673.

Axtell, M. J., Snyder, J. A., & Bartel, D. P. (2007). Common functions for diverse small RNAs of land plants. *The Plant Cell*, **19**(6): 1750-1769.

Aya, K., Ueguchi-Tanaka, M., Kondo, M., Hamada, K., Yano, K., Nishimura, M., & Matsuoka, M. (2009). Gibberellin modulates anther development in rice via the transcriptional regulation of GAMYB. *The Plant Cell*, **21**(5): 1453-1472.

Ballvora, A., Ercolano, M. R., Weiß, J., Meksem, K., Bormann, C. A., Oberhagemann, P., Salamini, F., & Gebhardt, C. (2002). The R1 gene for potato resistance to late blight (Phytophthora infestans) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. *The Plant Journal*, **30**(3): 361-371.

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *cell*, **136**(2): 215-233.

Beauclair, L., Yu, A., & Bouché, N. (2010). microRNA-directed cleavage and translational repression of the copper chaperone for superoxide dismutase mRNA in Arabidopsis. *The Plant Journal*, **62**(3): 454-462.

Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I., & Filipowicz, W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell.* **125**(6): 1111–1124.

Bologna, N. G., & Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. *Annual review of plant biology*, **65**(1): 473-503.

Borges, F., & Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nature Reviews Molecular Cell Biology*, **16**(12): 727.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L., & Voinnet, O. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. *Science*, **320**(5880): 1185-1190.

Budak, H., & Akpinar, B. A. (2015). Plant miRNAs: biogenesis, organization and origins. *Functional & integrative genomics*, *15*(5), 523-531.

Caldana, C., Scheible, W. R., Mueller-Roeber, B., & Ruzicic, S. (2007). A quantitative RT-PCR platform for high-throughput expression profiling of 2500 rice transcription factors. *Plant Methods*, **3**(1): 7.

Chen, X. (2009). Small RNAs and Their Roles in Plant Development. *Annual Review of Cell and Developmental Biology*. **25**: 21-44.

Chen, X. (2012). Small RNAs in development–insights from plants. *Current opinion in genetics & development*, **22**(4): 361-367.

Chen, X., Liu, J., Cheng, Y., & Jia, D. (2002). HEN1 functions pleiotropically in Arabidopsis development and acts in C function in the flower. *Development*, **129**(5): 1085-1094.

Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *The plant journal*, **16**(6): 735-743.

Coll, N. S., Epple, P., & Dangl, J. L. (2011). Programmed cell death in the plant immune system. *Cell death and differentiation*, **18**(8): 1247.

Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of p-values. *Royal Society open science*, **1**(3): 140216.

Csukasi, F., Donaire, L., Casañal, A., Martínez-Priego, L., Botella, M. A., Medina-Escobar, N., Llave, C., & Valpuesta, V. (2012). Two strawberry miR159 family members display developmental-specific expression patterns in the fruit receptacle and cooperatively regulate Fa-GAMYB. *New phytologist*, **195**(1): 47-57.

Curtis, M. D., & Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant physiology*, **133**(2): 462-469.

da Silva, E. M, Silva, G. F. F. E., Bidoia, D. B., da Silva Azevedo, M., de Jesus, F. A., Pino, L. E., Peres, L. E. P., Carrera, E, López-Díaz, I., & Nogueira, F. T. S. (2017). microRNA159-targeted SlGAMYB transcription factors are required for fruit set in tomato. *The Plant Journal*. **92**(1): 95-109.

Dai, X., and Zhao, P. X. (2011). psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res*. **39**: W155-159.

Dai, X., Zhuang, Z., & Zhao, P. X. (2018). psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic acids research*.

Deveson, I., Li, J., & Millar, A. A. (2013). MicroRNAs with analogous target complementarities perform with highly variable efficacies in Arabidopsis. *FEBS letters*, **587**(22): 3703-3708.

Didiano, D., & Hobert, O. (2006). Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nature Structural and Molecular Biology*, **13**(9): 849.

Ding, Y., Tang, Y., Kwok, C. K., Zhang, Y., Bevilacqua, P. C., & Assmann, S. M. (2014). In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature*, **505**(7485): 696-700.

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

Ebrahim, S., Usha, K., & Singh, B. (2011). Pathogenesis related (PR) proteins in plant defense mechanism. *Sci Against Microb Pathog*, **2**: 1043-1054.

Edreva, A. (2005). Pathogenesis-related proteins: research progress in the last 15 years. *Gen Appl Plant Physiol*, **31(1-2)**: 105-24.

Emery, J. F., Floyd, S. K., Alvarez, J., Eshed, Y., Hawker, N. P., Izhaki, A., Baum, S., F., & Bowman, J. L. (2003). Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. *Current Biology*, **13**(20): 1768-1774.

Fahlgren, N., Howell, M. D., Kasschau, K. D., Chapman, E. J., Sullivan, C. M., Cumbie, J. S., Givan, S. A., Law, T.F., Grant, S. R., Dang, J. L., & Carrington, J. C. (2007). High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. *PloS one*, **2**(2): e219.

Fang, Y., Xie, K., & Xiong, L. (2014). Conserved miR164-targeted NAC genes negatively regulate drought resistance in rice. *Journal of experimental botany*, *65*(8), 2119-2135.

Fei, Q., Li, P., Teng, C., & Meyers, B. C. (2015). Secondary siRNAs from Medicago NB-LRRs modulated via miRNA-target interactions and their abundances. *The Plant Journal*, **83**(3): 451-465.

Flynt, A. S., & Lai, E. C. (2008). Biological principles of microRNA-mediated regulation: shared themes amid diversity. *Nature Reviews Genetics*, **9**(11): 831.

Franco-Zorrilla, J. M., Valli, A., Todesco, M., Mateos, I., Puga, M. I., Rubio-Somoza, I., Leyva, A., Weigel, D., Garcia, J. A., & Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature genetics*, **39**(8): 1033-1037.

Garcia, D. (2008). A miRacle in plant development: role of microRNAs in cell differentiation and patterning. *Semin Cell Dev Biol*, **19**(6): 586-595.

German, M. A., Pillay, M., Jeong, D. H., Hetawal, A., Luo, S., Janardhanan, P., Kannan, V., Rymarquis, L. A., Nobuta, K., German, R., De Paoli, E., Lu, C., Schroth, G., Meyers, B. C., & Green, P. J. (2008). Global identification of microRNA–target RNA pairs by parallel analysis of RNA ends. *Nature biotechnology*, **26**(8): 941.

Gocal, G. F., Sheldon, C. C., Gubler, F., Moritz, T., Bagnall, D. J., MacMillan, C. P., Li, S. F., Parish, R. W., Dennis, E. S., Weigel, D. & King, R. W. (2001). GAMYB-like genes, flowering, and gibberellin signaling in Arabidopsis. *Plant Physiology*, **127**(4): 1682-1693.

Gou, M., & Hua, J. (2012). Complex regulation of an R gene SNC1 revealed by autoimmune mutants. *Plant signaling & behavior*, **7**(2): 213-216.

Grimson, A., Farh, K. K. H., Johnston, W. K., Garrett-Engele, P., Lim, L. P., & Bartel, D. P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Molecular cell*, **27**(1): 91-105.

Gu, W., Wang, X., Zhai, C., Xie, X., & Zhou, T. (2012). Selection on synonymous sites for increased accessibility around miRNA binding sites in plants. *Molecular biology and evolution*, **29**(10): 3037-3044.

Guan, Q., Lu, X., Zeng, H., Zhang, Y., & Zhu, J. (2013). Heat stress induction of miR398 triggers a regulatory loop that is critical for thermotolerance in Arabidopsis. *The Plant Journal*, **74**(5): 840-851.

Gubler, F., Kalla, R., Roberts, J. K., & Jacobsen, J. V. (1995). Gibberellin-regulated expression of a myb gene in barley aleurone cells: evidence for Myb transactivation of a high-pI alpha-amylase gene promoter. *The Plant Cell*, **7**(11): 1879-1891.

Gubler, F., Raventos, D., Keys, M., Watts, R., Mundy, J., & Jacobsen, J. V. (1999). Target genes and regulatory domains of the GAMYB transcriptional activator in cereal aleurone. *The Plant Journal*, **17**(1): 1-9.

Guo, C., Xu, Y., Shi, M., Lai, Y., Wu, X., Wang, H., Zhu, Z., Poethig, R.S. & Wu, G. (2017). Repression of miR156 by miR159 Regulates the Timing of the Juvenile-to-Adult Transition in Arabidopsis. *The Plant Cell*, **29(**6): 1293-1304.

Guo, H. S., Xie, Q., Fei, J. F., & Chua, N. H. (2005). MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for Arabidopsis lateral root development. *The Plant Cell*, **17**(5): 1376-1386.

Guo, W. J., & Ho, T. H. D. (2008). An abscisic acid-induced protein, HVA22, inhibits gibberellin-mediated programmed cell death in cereal aleurone cells. *Plant physiology*, **147**(4): 1710-1722.

Hamada, T. (2014). Microtubule organization and microtubule-associated proteins in plant cells. In *International review of cell and molecular biology*. **312**:1-52. Academic Press.

Hellens, R., Mullineaux, P., & Klee, H. (2000). Technical focus: a guide to Agrobacterium binary Ti vectors. *Trends in plant science*, **5**(10): 446-451.

Herbers, K., Mönke, G., Badur, R., & Sonnewald, U. (1995). A simplified procedure for the subtractive cDNA cloning of photoassimilate-responding genes: isolation of cDNAs encoding a new class of pathogenesis-related proteins. *Plant molecular biology*, **29**(5): 1027-1038.

Hofius, D., Schultz-Larsen, T., Joensen, J., Tsitsigiannis, D. I., Petersen, N. H., Mattsson, O., Jorgensen, L. B., Jones, J., Mundy, J., & Petersen, M. (2009). Autophagic components contribute to hypersensitive cell death in Arabidopsis. *Cell*, **137**(4): 773-783.

Houston, K., McKim, S. M., Comadran, J., Bonar, N., Druka, I., Uzrek, N., Cirilloc, E., Guzy-Wrobelskad, J., Collinse, N. C., Halpinb, C., Hanssonf, M., Dockterf, C., Drukaa. A., & Waugha, R. (2013). Variation in the interaction between alleles of HvAPETALA2 and microRNA172 determines the density of grains on the barley inflorescence. *Proceedings of the National Academy of Sciences*, **110**(41): 16675-16680.

Huntzinger, E., & Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nature Reviews Genetics*, **12**(2): 99-110.

Ito, H. (2012). Small RNAs and transposon silencing in plants. *Development, growth & differentiation*, **54**(1): 100-107.

Iwakawa, H. O., & Tomari, Y. (2013). Molecular insights into microRNA-mediated translational repression in plants. *Molecular cell*, **52**(4): 591-601.

Jeong, D. H., Park, S., Zhai, J., Gurazada, S. G. R., De Paoli, E., Meyers, B. C., & Green, P. J. (2011). Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. *The Plant Cell*, **23**(12): 4185-4207.

Jones-Rhoades, M. W. (2012). Conservation and divergence in plant microRNAs. *Plant molecular biology*, **80**(1): 3-16.

Jones-Rhoades, M. W., & Bartel, D. P. (2004). Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Molecular cell*, **14**(6): 787-799.

Jones-Rhoades, M. W., Bartel, D. P., & Bartel, B. (2006). MicroRNAs and their regulatory roles in plants. *Annu. Rev. Plant Biol.*, **57**: 19-53.

Jube, S. and Borthakur, D. (2007). Expression of bacterial genes in transgenic tobacco: methods, applications and future prospects. *Electronic journal of biotechnology : EJB*. **10**(3):452-467.

Jung, J. H., Seo, Y. H., Seo, P. J., Reyes, J. L., Yun, J., Chua, N. H., & Park, C. M. (2007). The GIGANTEA-regulated microRNA172 mediates photoperiodic flowering independent of CONSTANS in Arabidopsis. *The Plant Cell*, **19**(9): 2736-2748.

Kaneko, M., Inukai, Y., Ueguchi-Tanaka, M., Itoh, H., Izawa, T., Kobayashi, Y., Hattori, T., Miyao, A., Hirochika, H., Ashikari, M., & Matsuoka, M. (2004). Loss-of-function mutations of the rice GAMYB gene impair α-amylase expression in aleurone and flower development. *The Plant Cell*, **16**(1): 33-44.

Karakülah, G., Kurtoğlu, K. Y., & Unver, T. (2016). PeTMbase: a database of plant endogenous target mimics (eTMs). *PLoS One*, **11**(12): e0167698.

Kasschau, K. D., Xie, Z., Allen, E., Llave, C., Chapman, E. J., Krizan, K. A., & Carrington, J. C. (2003). P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA function. *Developmental cell*, **4**(2): 205-217.

Kedde, M., Strasser, M. J., Boldajipour, B., Vrielink, J., Slanchev, K., Le Sage, C., Nagel, R., Voorhoeve , P.M., van Duijse, J., Andersson Ørom, U., Lund, A. H., Perrakis, A., Raz, E., & Agami, R. (2007). RNA binding protein Dnd1 inhibits microRNA access to target mRNA. *Cell*. **131:** 1273–1286.

Kedde, M., Van Kouwenhove, M., Zwart, W., Vrielink, J. A. O., Elkon, R., & Agami, R. (2010). A Pumilio-induced RNA structure switch in p27-3 [prime] UTR controls miR-221 and miR-222 accessibility. *Nature cell biology*, **12**(10): 1014-1020.

Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., & Segal, E. (2007). The role of site accessibility in microRNA target recognition. *Nat Genet.* **39**(10): 1278-1284.

Khatabi, B., Arikit, S., Xia, R., Winter, S., Oumar, D., Mongomake, K., Meyers, B. C., & Fondong, V. N. (2016). High-resolution identification and abundance profiling of cassava (Manihot esculenta Crantz) microRNAs. *BMC genomics*, **17**(1): 85.

Kim, M. H., Cho, J. S., Lee, J. H., Bae, S. Y., Choi, Y. I., Park, E. J., Lee, H., & Ko, J. H. (2017). Poplar MYB transcription factor PtrMYB012 and its Arabidopsis AtGAMYB orthologs are differentially repressed by the Arabidopsis miR159 family. *Tree physiology*.

Koyama, T., Furutani, M., Tasaka, M., & Ohme-Takagi, M. (2007). TCP transcription factors control the morphology of shoot lateral organs via negative regulation of the expression of boundary-specific genes in Arabidopsis. *The Plant Cell*, **19**(2): 473-484.

Koyama, T., Mitsuda, N., Seki, M., Shinozaki, K., & Ohme-Takagi, M. (2010). TCP transcription factors regulate the activities of ASYMMETRIC LEAVES1 and miR164, as well as the auxin response, during differentiation of leaves in Arabidopsis. *The Plant Cell*, **22**(11): 3574-3588.

Kozomara, A., & Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic acids research*, **42**(D1): D68-D73.

Kwok, C. K., Ding, Y., Tang, Y., Assmann, S. M., & Bevilacqua, P. C. (2013). Determination of in vivo RNA structure in low-abundance transcripts. *Nature communications*, **4**: 2971.

Lanet, E., Delannoy, E., Sormani, R., Floris, M., Brodersen, P., Crété, P., Voinnet, O., and Robaglia, C. (2009). Biochemical evidence for translational repression by Arabidopsis microRNAs. *Plant Cell.* **21**: 1762–1768.

Leydon, A. R., Beale, K. M., Woroniecka, K., Castner, E., Chen, J., Horgan, C., Palanivelu, R., & Johnson, M. A. (2013). Three MYB transcription factors control pollen tube differentiation required for sperm release. *Current Biology*, **23**(13): 1209-1214.

Li, F., Pignatta, D., Bendix, C., Brunkard, J. O., Cohn, M. M., Tung, J., Sun, H., Kumar, P., & Baker, B. (2012a). MicroRNA regulation of plant innate immune receptors. *Proceedings of the National Academy of Sciences*, **109**(5): 1790-1795.

Li, F., Zheng, Q., Vandivier, L. E., Willmann, M. R., Chen, Y., & Gregory, B. D. (2012b). Regulatory impact of RNA secondary structure across the Arabidopsis transcriptome. *The Plant Cell*, **24**(11): 4346-4359.

Li, J., & Millar, A. A. (2013). Expression of a microRNA-resistant target transgene misrepresents the functional significance of the endogenous microRNA: target gene relationship. *Molecular plant*, **6**(2): 577-580.

Li, J., Reichel, M., Li, Y., & Millar, A. A. (2014a). The functional scope of plant microRNA-mediated silencing. *Trends in plant science*, **19**(12): 750-756.

Li, J., Reichel, M., & Millar, A. A. (2014b). Determinants beyond both complementarity and cleavage govern microR159 efficacy in Arabidopsis. *PLoS genetics*, **10**(3): e1004232.

Li, J. F., Chung, H. S., Niu, Y., Bush, J., McCormack, M., & Sheen, J. (2013a). Comprehensive protein-based artificial microRNA screens for effective gene silencing in plants. *The Plant Cell*, **25**(5): 1507-1522.

Li, W. F., Zhang, S. G., Han, S. Y., Wu, T., Zhang, J. H., & Qi, L. W. (2013b). Regulation of LaMYB33 by miR159 during maintenance of embryogenic potential and somatic embryo maturation in Larix kaempferi (Lamb.) Carr. *Plant Cell, Tissue and Organ Culture (PCTOC)*, **113**(1): 131-136.

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. *The Plant Cell*, *20*(8), 2238-2251.

Li, X., Bian, H., Song, D., Ma, S., Han, N., Wang, J., & Zhu, M. (2013c). Flowering time control in ornamental gloxinia (Sinningia speciosa) by manipulation of miR159 expression. *Annals of botany*, **111**(5): 791-799.

Li, Y., Alonso-Peral, M., Wong, G., Wang, M. B., & Millar, A. A. (2016). Ubiquitous miR159 repression of MYB33/65 in Arabidopsis rosettes is robust and is not perturbed by a wide range of stresses. *BMC plant biology*, **16**(1): 179.

Li, Y., Li, C., Ding, G., & Jin, Y. (2011). Evolution of MIR159/319 microRNA genes and their post-transcriptional regulatory link to siRNA pathways. *BMC evolutionary biology*, **11**(1): 122.

Li, Y., Lu, Y. G., Shi, Y., Wu, L., Xu, Y. J., Huang, F., Guo, X., Zhang, Y., Fan, J., Zhao, J., Zhang, H., Xu, P., Zhou, J., Wu, X., Wang, P., & Wang, W. (2014c). Multiple rice microRNAs are involved in immunity against the blast fungus Magnaporthe oryzae. *Plant physiology*, **164**(2): 1077-1092.

Li, Y., Zhang, Q., Zhang, J., Wu, L., Qi, Y., & Zhou, J. M. (2010). Identification of microRNAs involved in pathogen-associated molecular pattern-triggered plant innate immunity. *Plant physiology*, **152**(4): 2222-2231.

Liang, G., He, H., Li, Y., Wang, F., & Yu, D. (2014). Molecular mechanism of microRNA396 mediating pistil development in Arabidopsis. *Plant physiology*, **164**(1): 249-258.

Liang, Y., Tan, Z. M., Zhu, L., Niu, Q. K., Zhou, J. J., Li, M., Chen, Li., Zhang, X. & Ye, D. (2013). MYB97, MYB101 and MYB120 function as male factors that control pollen tube-synergid interaction in Arabidopsis thaliana fertilization. *PLoS genetics*, **9**(11): e1003933.

Liu, D., Shi, L., Han, C., Yu, J., Li, D., & Zhang, Y. (2012). Validation of reference genes for gene expression studies in virus-infected Nicotiana benthamiana using quantitative real-time PCR. *PLoS One*, **7**(9): e46451.

Liu, H., and Naismith, J.H. (2008). An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol*. **8:** 91.

Liu, Q., Wang, F., & Axtell, M. J. (2014). Analysis of complementarity requirements for plant microRNA targeting using a Nicotiana benthamiana quantitative transient assay. *The Plant Cell*, **26**(2): 741-753.

Llave, C., Xie, Z., Kasschau, K. D., & Carrington, J. C. (2002). Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science*, **297**(5589): 2053-2056.

Loebenstein, G. (2009). Local lesions and induced resistance. In *Advances in virus research.* **75**: 73-117). Academic Press.

Luan, Y., Cui, J., Zhai, J., Li, J., Han, L., & Meng, J. (2015). High-throughput sequencing reveals differential expression of miRNAs in tomato inoculated with Phytophthora infestans. *Planta*, **241**(6): 1405-1416.

Luo, Y., Guo, Z., & Li, L. (2013). Evolutionary conservation of microRNA regulatory programs in plant flower development. *Developmental biology*, **380**(2): 133-144.

Ma, X., Tang, Z., Qin, J., & Meng, Y. (2015). The use of high-throughput sequencing methods for plant microRNA research. *RNA biology*, **12**(7): 709-719.

Mallory, A., & Vaucheret, H. (2010). Form, function, and regulation of ARGONAUTE proteins. *The Plant Cell*, **22**(12): 3879-3889.

Mallory, A. C., Bartel, D. P., & 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.

Mallory, A. C., Dugas, D. V., Bartel, D. P., & Bartel, B. (2004a). MicroRNA regulation of NAC-domain targets is required for proper formation and separation of adjacent embryonic, vegetative, and floral organs. *Current Biology*, **14**(12): 1035-1046.

Mallory, A. C., Reinhart, B. J., Jones-Rhoades, M. W., Tang, G., Zamore, P. D., Barton, M. K., & Bartel, D. P. (2004b). MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5′ region. *The EMBO journal*, **23**(16): 3356-3364.

Marathe, R., Anandalakshmi, R., Liu, Y., & Dinesh-Kumar, S. P. (2002). The tobacco mosaic virus resistance gene, N. *Molecular Plant Pathology*, **3**(3): 167-172.

McConnell, J. R., Emery, J., Eshed, Y., Bao, N., Bowman, J., & Barton, M. K. (2001). Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots. *Nature*, **411**(6838): 709-713.

Millar, A. A., and Gubler, F. (2005). The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. *The Plant Cell*, **17**(3): 705-721.

Mur, L. A., Kenton, P., Lloyd, A. J., Ougham, H., & Prats, E. (2007). The hypersensitive response; the centenary is upon us but how much do we know?. *Journal of experimental Botany*, **59**(3): 501-520.

Nakamura, S., Mano, S., Tanaka, Y., Ohnishi, M., Nakamori, C., Araki, M., Niwa, T., Nishimura, M., Kaminaka, H., Nakagawa, T., Sato, Y. & Ishiguro, S. (2010). Gateway binary vectors with the bialaphos resistance gene, bar, as a selection marker for plant transformation. *Bioscience, biotechnology, and biochemistry*, **74**(6): 1315-1319.

Padmanabhan, C., Zhang, X., & Jin, H. (2009). Host small RNAs are big contributors to plant innate immunity. *Current opinion in plant biology*, **12**(4): 465-472.

Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J. C., & Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. *Nature.* **425:** 257-263.

Palatnik, J. F., Wollmann, H., Schommer, C., Schwab, R., Boisbouvier, J., Rodriguez, R., Warthmann, N., Allen, E., Dezulian, T., Huson, D., Carrington, J., Carrington, J. C. & Weigel, D. (2007). Sequence and expression differences underlie functional specialization of Arabidopsis microRNAs miR159 and miR319. *Developmental cell*, **13**(1): 115-125.

Rajagopalan, R., Vaucheret, H., Trejo, J., & Bartel, D. P. (2006). A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. *Genes & development*, **20**(24): 3407-3425.

Reichel, M., Li, Y., Li, J., & Millar, A. A. (2015). Inhibiting plant microRNA activity: molecular SPONGEs, target MIMICs and STTMs all display variable efficacies against target microRNAs. *Plant biotechnology journal*, **13**(7): 915-926.

Reichel, M., & Millar, A. A. (2015). Specificity of plant microRNA target MIMICs: Cross-targeting of miR159 and miR319. *Journal of plant physiology*, **180**: 45-48.

Reyes, J. L., & Chua, N. H. (2007). ABA induction of miR159 controls transcript levels of two MYB factors during Arabidopsis seed germination. *The Plant Journal*, **49**(4): 592-606.

Rodriguez, R. E., Mecchia, M. A., Debernardi, J. M., Schommer, C., Weigel, D., & Palatnik, J. F. (2010). Control of cell proliferation in Arabidopsis thaliana by microRNA miR396. *Development*, **137**(1): 103-112.

Rogers, K., & Chen, X. (2013). Biogenesis, turnover, and mode of action of plant microRNAs. *The Plant Cell*, **25**(7): 2383-2399.

Rubio-Somoza, I., & Weigel, D. (2011). MicroRNA networks and developmental plasticity in plants. *Trends in plant science*, **16**(5): 258-264.

Schauer, S. E., Jacobsen, S. E., Meinke, D. W., & Ray, A. (2002). DICER-LIKE1: blind men and elephants in Arabidopsis development. *Trends in plant science*, **7**(11): 487-491.

Schwab, R., Maizel, A., Ruiz-Ferrer, V., Garcia, D., Bayer, M., Crespi, M., Voinnet, O., & Martienssen, R. A. (2009). Endogenous TasiRNAs mediate non-cell autonomous effects on gene regulation in Arabidopsis thaliana. *PloS one*, **4**(6): e5980.

Schwab, R., Ossowski, S., Riester, M., Warthmann, N., & Weigel, D. (2006). Highly specific gene silencing by artificial microRNAs in Arabidopsis. *The Plant cell*. 18: 1121-1133.

Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M., & Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Developmental cell*, **8**(4): 517-527.

Seitz, H. (2009). Redefining microRNA targets. *Current biology*, **19**(10): 870-873.

Shivaprasad, P. V., Chen, H. M., Patel, K., Bond, D. M., Santos, B. A., & Baulcombe, D. C. (2012). A microRNA superfamily regulates nucleotide binding site–leucine-rich repeats and other mRNAs. *The Plant Cell*, **24**(3): 859-874.

Sieber, P., Wellmer, F., Gheyselinck, J., Riechmann, J. L., & Meyerowitz, E. M. (2007). Redundancy and specialization among plant microRNAs: role of the MIR164 family in developmental robustness. *Development*, **134**(6): 1051-1060.

Sierro, N., Battey, J. N., Ouadi, S., Bakaher, N., Bovet, L., Willig, A., Goepfert, S., Peitsch, M. C. & Ivanov, N. V. (2014). The tobacco genome sequence and its comparison with those of tomato and potato. *Nature communications*, **5**: 3833.

Sinha, M., Singh, R. P., Kushwaha, G. S., Iqbal, N., Singh, A., Kaushik, S., Kaur, P., Sharma, S. & Singh, T. P. (2014). Current overview of allergens of plant pathogenesis related protein families. *The Scientific World Journal*, **2014**.

Song, C., Jia, Q., Fang, J., Li, F., Wang, C., & Zhang, Z. (2010). Computational identification of citrus microRNAs and target analysis in citrus expressed sequence tags. *Plant Biology*, *12*(6), 927-934.

Song, F., & Goodman, R. M. (2002). Cloning and identification of the promoter of the tobacco Sar8. 2b gene, a gene involved in systemic acquired resistance. *Gene*, **290**(1): 115-124.

Spoel, S. H., & Dong, X. (2012). How do plants achieve immunity? Defence without specialized immune cells. *Nature reviews immunology*, **12**(2): 89.

Subramanian, S., Fu, Y., Sunkar, R., Barbazuk, W. B., Zhu, J. K., & Yu, O. (2008). Novel and nodulation-regulated microRNAs in soybean roots. *BMC genomics*, **9**(1): 160.

Sun, X., Zhang, Y., Zhu, X., Korir, N. K., Tao, R., Wang, C., & Fang, J. (2014). Advances in identification and validation of plant microRNAs and their target genes. *Physiologia plantarum*, **152(**2): 203-218.

Sunkar, R., Li, Y. F., & Jagadeeswaran, G. (2012). Functions of microRNAs in plant stress responses. *Trends in plant science*, **17**(4): 196-203.

Szittya, G., Moxon, S., Santos, D. M., Jing, R., Fevereiro, M. P., Moulton, V., & Dalmay, T. (2008). High-throughput sequencing of Medicago truncatula short RNAs identifies eight new miRNA families. *BMC genomics*, **9**(1): 593.

Tang, S., Wang, Y., Li, Z., Gui, Y., Xiao, B., Xie, J., Zhu, Q., & Fan, L. (2012). Identification of wounding and topping responsive small RNAs in tobacco (Nicotiana tabacum). *BMC plant biology*, **12**(1): 28.

Teh, O. K., & Hofius, D. (2014). Membrane trafficking and autophagy in pathogen-triggered cell death and immunity. *Journal of experimental botany*, **65**(5): 1297-1312.

Teotia, S., & Tang, G. (2015). To bloom or not to bloom: role of microRNAs in plant flowering. *Molecular plant*, **8**(3): 359-377.

Todesco, M., Balasubramanian, S., Cao, J., Ott, F., Sureshkumar, S., Schneeberger, K., Meyer, R. C., Altmann, T. & Weigel, D. (2012). Natural variation in biogenesis efficiency of individual Arabidopsis thaliana microRNAs. *Current Biology*, **22**(2): 166-170.

Todesco, M., Rubio-Somoza, I., Paz-Ares, J., & Weigel, D. (2010). A collection of target mimics for comprehensive analysis of microRNA function in Arabidopsis thaliana. *PLoS Genet*, **6**(7): e1001031.

Trapnell, C., Pachter, L., & Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, **25**(9): 1105-1111.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L., & Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols*, **7**(3): 562.

Tsuji, H., Aya, K., Ueguchi-Tanaka, M., Shimada, Y., Nakazono, M., Watanabe, R., Nishizawa, N. K., Gomi, K., Shimada, A., Kitano, H., Ashikari, M. & Matsuoka, M. (2006). GAMYB controls different sets of genes and is differentially regulated by microRNA in aleurone cells and anthers. *The Plant Journal*, **47**(3): 427-444.

Vallarino, J. G., Osorio, S., Bombarely, A., Casañal, A., Cruz-Rus, E., Sánchez-Sevilla, J. F., Amaya, I., Giavalisco, P., Fernie, A. R., Botella, M. A. & Valpuesta, V. (2015). Central role of FaGAMYB in the transition of the strawberry receptacle from development to ripening. *New Phytologist*, **208**(2): 482-496.

Vaucheret, H., Vazquez, F., Crété, P., & Bartel, D. P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes & development*, **18**(10): 1187-1197.

Vella, M. C., Choi, E. Y., Lin, S. Y., Reinert, K., & Slack, F. J. (2004). The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3′ UTR. *Genes & development*, **18**(2): 132-137.

Waese, J., Fan, J., Pasha, A., Yu, H., Fucile, G., Shi, R., Cumming, M., Kelley, L., Sternberg, M., Krishnakumar, V., Ferlanti, E., Miller, J., Town, C., Stuerzlinger, W., & Provart, N. J. (2017). ePlant: visualizing and exploring multiple levels of data for hypothesis generation in plant biology. *The Plant Cell*, tpc-00073.

Wan, Y., Qu, K., Zhang, Q. C., Flynn, R. A., Manor, O., Ouyang, Z., Zhang, J., Spitale, R. C., Snyder, M. P., Segal, E. & Chang, H. Y. (2014). Landscape and variation of RNA secondary structure across the human transcriptome. *Nature*, **505**(7485): 706-709.

Wang, F., Polydore, S., & Axtell, M. J. (2015). More than meets the eye? Factors that affect target selection by plant miRNAs and heterochromatic siRNAs. *Current opinion in plant biology*, **27**: 118-124.

Wang, J., Meng, X., Dobrovolskaya, O. B., Orlov, Y. L., & Chen, M. (2017). Non-coding RNAs and Their Roles in Stress Response in Plants. *Genomics, proteomics & bioinformatics*.**15(**5): 301-312.

Wang, J. W., Wang, L. J., Mao, Y. B., Cai, W. J., Xue, H. W., & Chen, X. Y. (2005). Control of root cap formation by MicroRNA-targeted auxin response factors in Arabidopsis. *Plant Cell*, **17:** 2204-2216.

Wang, Y., Sun, F., Cao, H., Peng, H., Ni, Z., Sun, Q., & Yao, Y. (2012). TamiR159 directed wheat TaGAMYB cleavage and its involvement in anther development and heat response. *PloS one*, **7**(11): e48445.

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**(9): 1189-1191.

Weiberg, A., Wang, M., Bellinger, M., & Jin, H. (2014). Small RNAs: a new paradigm in plant-microbe interactions. *Annual review of phytopathology*, **52**(1): 495-516.

Wu, G. (2013). Plant microRNAs and development. *Journal of Genetics and Genomics*, **40**(5): 217-230.

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

Xie, Z., Allen, E., Fahlgren, N., Calamar, A., Givan, S. A., & Carrington, J. C. (2005). Expression of Arabidopsis MIRNA genes. *Plant physiology*, **138**(4): 2145-2154.

Yan, J., Gu, Y., Jia, X., Kang, W., Pan, S., Tang, X., Chen, X., & Tang, G. (2012). Effective small RNA destruction by the expression of a short tandem target mimic in Arabidopsis. *The Plant Cell*, **24**(2): 415-427.

Yang, J., Zhang, N., Mi, X., Wu, L., Ma, R., Zhu, X., Yao, L., Jin, X., Si, H., & Wang, D. (2014). Identification of miR159s and their target genes and expression analysis under drought stress in potato. *Computational biology and chemistry*, **53**: 204-213.

Yang, L., Wu, G., and Poethig, R.S. (2012). Mutations in the GW-repeat protein SUO reveal a developmental function for microRNA mediated translational repression in Arabidopsis. Proc. *Natl. Acad. Sci.* USA **109**: 315–320.

Yao, Y., Guo, G., Ni, Z., Sunkar, R., Du, J., Zhu, J. K., & Sun, Q. (2007). Cloning and characterization of microRNAs from wheat (Triticum aestivum L.). *Genome biology*, **8**(6): R96.

Zhai, J., Jeong, D. H., De Paoli, E., Park, S., Rosen, B. D., Li, Y., Gonzalez, A. J., Yan, Z., Kitto, S. L., Grusak, M. A., Jackson, S. A., Stacey, G., Cook, D. R., Green, P. J., Sherrier, D., J., & Meyers, B. C (2011). MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes & development*, **25**(23): 2540-2553.

Zhang, B. (2015). MicroRNA: a new target for improving plant tolerance to abiotic stress. *Journal of experimental botany*, **66**(7): 1749-1761.

Zhang, B., & Wang, Q. (2015). MicroRNA-based biotechnology for plant improvement. *Journal of cellular physiology*, **230**(1): 1-15.

Zhang, L., Chia, J. M., Kumari, S., Stein, J. C., Liu, Z., Narechania, A., Maher, C., Guill, K., McMullen, M., & Ware, D. (2009). A genome-wide characterization of microRNA genes in maize. *PLoS genetics*, **5**(11): e1000716.

Zhang, T., Wang, J., & Zhou, C. (2015). The role of miR156 in developmental transitions in Nicotiana tabacum. *Science China Life Sciences*, **58**(3): 253-260.

Zhang, W., Gao, S., Zhou, X., Chellappan, P., Chen, Z., Zhou, X., Zhang, X., Fromuth, N., Coutino, G., Coffey, M., & Jin, H. (2011a). Bacteria-responsive microRNAs regulate plant innate immunity by modulating plant hormone networks. *Plant molecular biology*, **75**(1-2): 93-105.

Zhang, X., Zou, Z., Gong, P., Zhang, J., Ziaf, K., Li, H., Xiao, F., & Ye, Z. (2011b). Over-expression of microRNA169 confers enhanced drought tolerance to tomato. *Biotechnology letters*, **33**(2): 403-409.

Zhao, Y., Wen, H., Teotia, S., Du, Y., Zhang, J., Li, J., Sun, H., Tang, G., Peng, T., & Zhao, Q. (2017). Suppression of microRNA159 impacts multiple agronomic traits in rice (Oryza sativa L.). *BMC plant biology*, **17**(1): 215.

Zhu, Q. H., Fan, L., Liu, Y., Xu, H., Llewellyn, D., & Wilson, I. (2013). miR482 regulation of NBS-LRR defense genes during fungal pathogen infection in cotton. *PLoS One*, **8**(12): e84390.

Appendix

Primer Table 1

Name	Template	Sequence 5'-3'	Purpose
attB-MYB33-F	*mir159abc* cDNA	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAGAAAAGAGATGAGTTACACGAGC	*MYB33* CDS gateway construct
attB-MYB33-R	*mir159abc* cDNA	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCGTTAGGGTAGTTCTGTCATTTG	*MYB33* CDS gateway construct
attB-MYB65-F	*mir159abc* cDNA	GGGGACAAGTTTGTACAAAAAAGCAGGCTGTAGAAAGAGATGAGTTACACGACGG	*MYB65* CDS gateway construct
attB-MYB65-R	*mir159abc* cDNA	GGGGACCACTTTGTACAAGAAAGCTGGGTTTACAGCGACCAAAACAGGAGGC	*MYB65* CDS gateway construct
attB-MYB81-F	Col-0 genomic	GGGGACAAGTTTGTACAAAAAAGCAGGCTTACAGATGACACAAGATGGGGAAAG	*MYB81* genomic gateway construct
attB-MYB81-R	Col-0 genomic	GGGGACCACTTTGTACAAGAAAGCTGGGTCAATCACCATCAAAGGATGTGTGTC	*MYB81* genomic gateway construct
attB-MYB97-F	*mir159abc* cDNA	GGGGACAAGTTTGTACAAAAAAGCAGGCTAGAGGGGGCAATGATCGTGTAC	*MYB97* CDS gateway construct
attB-MYB97-R	*mir159abc* cDNA	GGGGACCACTTTGTACAAGAAAGCTGGGTCTAGCAGATCCCTGGCAAGTTG	*MYB97* CDS gateway construct
attB-MYB101-F	*mir159abc* cDNA	GGGGACAAGTTTGTACAAAAAAGCAGGCTGTGTTGAAAAGGATGGATGGTG	*MYB101* CDS gateway construct
attB-MYB101-R	*mir159abc* cDNA	GGGGACCACTTTGTACAAGAAAGCTGGGTGCAGGTTCCAAACTAACAGATGC	*MYB101* CDS gateway construct
attB-MYB104-F	Col-0 genomic	GGGGACAAGTTTGTACAAAAAAGCAGGCTAAAATTTTCCATGATTCAGGATCAAG C	*MYB104* genomic gateway construct
attB-MYB104-R	Col-0 genomic	GGGGACCACTTTGTACAAGAAAGCTGGGTTTGGTGTTATATGAATCCCTCATCG	*MYB104* genomic gateway construct
attB-DUO1-F	*mir159abc* cDNA	GGGGACAAGTTTGTACAAAAAAGCAGGCTGATGAGGAAAATGGAAGCGAAG	*DUO1* CDS gateway construct
attB-DUO1-R	*mir159abc* cDNA	GGGGACCACTTTGTACAAGAAAGCTGGGTAATCTAAGGACTTGGGATTGGATC	*DUO1* CDS gateway construct

Primer Table 2

Name	Template	Sequence 5'-3'	Purpose
mMYB33-F	*MYB33* CDS entry vector	TCGAATTGCCAAGCTTTCAGTATTCAGAAACAACATTTGACCAGTGGAAGAAATCG	Synonymous mutation of *MYB33* entry vector
mMYB33-R	*MYB33* CDS entry vector	ACTGAAAGCTTGGCAATTCGAGCTTCACTGCCCCATATAAGGGCTTAGAATAAGG	Synonymous mutation of *MYB33* entry vector
mMYBb65-F	*MYB65* CDS entry vector	TCGAATTGCCAAGCTTTCAGTATTCAGAAACTAGTGCATTTGATCAGTGGAAGACG	Synonymous mutation of *MYB65* entry vector
mMYB65-F	*MYB65* CDS entry vector	ACTGAAAGCTTGGCAATTCGAGCTTCACTGCCCCAAACAAGGGCTTAGAAGTAG	Synonymous mutation of *MYB65* entry vector
mMYB81-F	*MYB81* genomic entry vector	GAGAATTTCGAATTTCCAAGCTTTCAGTATCATGAAGAGCCTGGC	Synonymous mutation of *MYB81* entry vector
mMYB81-R	*MYB81* genomic entry vector	CTTCATGATACTGAAAGCTTGGAAATTCGAAATTCTCAACCCCGTG	Synonymous mutation of *MYB81* entry vector
mMYB97-F	*MYB97* CDS entry vector	ATGAACTTAGCAGCAATCAGATTCCATACTCTGCCTCTTTAGGAGTCTATC	Synonymous mutation of *MYB97* entry vector
mMYB97-R	*MYB97* CDS entry vector	AATCTGATTGCTGCTAAGTTCATGTTTCATGGAGTAAACCGGATTGGTAG	Synonymous mutation of *MYB97* entry vector
mMYB101-F	*MYB101* CDS entry vector	CAATAGCGTCCTCGAGCTTCCAAGCAATCAGCGACCGACCCATTCG	Synonymous mutation of *MYB101* entry vector
mMYB101-R	*MYB101* CDS entry vector	CGGTCGCTGATTGCTTGGAAGCTCGAGGACGCTATTGTCTAGTCCTAAAG	Synonymous mutation of *MYB101* entry vector
mMYB104-F	*MYB104* genomic entry vector	TAGAATTGCCAAGCTTTCAGGGCTTTGATTTTCATGAAGAGCCTAGCAG	Synonymous mutation of *MYB104* entry vector
mMYB104-R	*MYB104* genomic entry vector	CCTGAAAGCTTGGCAATTCTAAGTTCTCAACCTCGTGAATGAGAGGC	Synonymous mutation of *MYB104* entry vector
mDUO1-F	*DUO1* CDS entry vector	GATACGGTCCTCGTGATTGGTCGTCGATCCGGTCGAGGGTCTTCTTCAACGCAC	Synonymous mutation of *DUO1* entry vector
mDUO1-R	*DUO1* CDS entry vector	GTGCGTTGAAGAAGACCCTTCGACCGGATCGACGACCAATCGAGGACCGTATC	Synonymous mutation of *DUO1* entry vector

Primer Table 3

Name	*Template*	*Sequence 5'-3'*	*Purpose*
MYB33-q-F	*35S:MYB33*	CCAGATAGCCATACCCCTACG	qRT-PCR of un-cleaved MYB33
MYB33-q-R	*35S:MYB33*	CCTCGGATTTAGTTTGGGATAC	qRT-PCR of un-cleaved MYB33
MYB65-q-F	*35S:MYB65*	ATATGATGATGGTTCCTGATAGCC	qRT-PCR of un-cleaved MYB65
MYB65-q-R	*35S:MYB65*	AGGCATCAACAGAGTCAAGGAG	qRT-PCR of un-cleaved MYB65
MYB81-q-F	*35S:MYB81*	GATAGTCATTTGCTTGGTAATGC	qRT-PCR of un-cleaved MYB81
MYB81-q-R	*35S:MYB81*	AGTGTTATCTGACTCATGTTCCG	qRT-PCR of un-cleaved MYB81
MYB97-q-F	*35S:MYB97*	CCTTTCTTCACACCCATTCCTC	qRT-PCR of un-cleaved MYB97
MYB97-q-R	*35S:MYB97*	TGTTCAGGTTTTGGTTACAATCC	qRT-PCR of un-cleaved MYB97
MYB101-q-F	*35S:MYB101*	CGAGTTCTTTCCCTTTAGGACT	qRT-PCR of un-cleaved MYB101
MYB101-q-R	*35S:MYB101*	TGGCTCATTGTACTTGTTGTG	qRT-PCR of un-cleaved MYB101
MYB104-q-F	*35S:MYB104*	AGCACTCAAGAACAGTTACCTGAC	qRT-PCR of un-cleaved MYB104
MYB104-q-R	*35S:MYB104*	CAGCATTGGGTTATATTGTTCTG	qRT-PCR of un-cleaved MYB104
MYB120-q-F	*35S:MYB120*	CTTCCACCTCATTTCATACCAAC	qRT-PCR of un-cleaved MYB120
MYB120-q-R	*35S:MYB120*	CTGTCTCTCGTTGTCTGTGAAGTT	qRT-PCR of un-cleaved MYB120
DUO1-q-F	*35S:DUO1*	TGGAAAGCCGAAGAAGACG	qRT-PCR of un-cleaved DUO1
DUO1-q-R	*35S:DUO1*	TTGGGACGGAGTTTATTGACC	qRT-PCR of un-cleaved DUO1
CP1-F	*35S:MYB*	ACGCTCGATAAACATTTGGG	qRT-PCR of CP1
CP1-R	*35S:MYB*	CAAGCAAACAATCTTTCTTTGAA	qRT-PCR of CP1
Cyclo-F	*35S:MYB*	TGGACCAGGTGTACTTTAATGG	qRT-PCR of reference gene
Cyclo-R	*35S:MYB*	CCACTGTCTGCAATTACGACTTTG	qRT-PCR of reference gene
MGH3-F	*35S:DUO1*	GATTTAAGGTTCCAAAGCCATG	qRT-PCR of MGH3
MGH3-R	*35S:DUO1*	GCCAACTGAATGTCCTTGGAC	qRT-PCR of MGH3

Primer Table 4

Name	***Template***	***Sequence 5'-3'***	***Purpose***
MYB81-33-F	*MYB81* genomic entry vector	TGGAGCTCCCTTCATTCCAATATCATGAAGAGCCTGGCGGTTGG	mutation of *MYB81* entry vector to *MYB81-33* entry vector
MYB81-33-R	*MYB81* genomic entry vector	ATTGGAATGAAGGGAGCTCCAAATTCTCAACCCCGTGAACAAGAGGTC	
MYB81-m33-F	*MYB81* genomic entry vector	TGGAATTGCCAAGCTTTCAGTATCATGAAGAGCCTGGCGGTTGG	mutation of *MYB81* entry vector to *MYB81-m33* entry vector
MYB81-m33-R	*MYB81* genomic entry vector	ACTGAAAGCTTGGCAATTCCAAATTCTCAACCCCGTGAACAAGAGGTC	
MYB81-q-F	*MYB81*	GATAGTCATTTGCTTGGTAATGC	qRT-PCR of un-cleaved *MYB81*
MYB81-q-R	*MYB81*	AGTGTTATCTGACTCATGTTCCG	
CP1-F	*CP1*	ACGCTCGATAAACATTTGGG	qRT-PCR of *CP1*
CP1-R	*CP1*	CAAGCAAACAATCTTTCTTTGAA	

Primer Table 5

Name	***Template***	***Sequence 5'-3'***	***Purpose***
MYB33-mSL1-F	*MYB33* genomic entry vector	TTCTCACACCCCAACCGACGGTATTGTTCCTTATTCTAAGCCCTTAT	mutation of *MYB33* entry vector to *MYB33-mSL1* entry vector
MYB33-mSL1-R	*MYB33* genomic entry vector	TACCGTCGGTTGGGGTGTGAGAATCTGGAATCATAACAGGTGAATGTCGGTTTCCATAG	
MYB33-mSL2-F	*MYB33* genomic entry vector	TACAGTAAACCATTATATGGAGCAGTGAAGCTGGAGCTCCCTTC	mutation of *MYB33* entry vector to *MYB33-mSL2* entry vector
MYB33-mSL2-R	*MYB33* genomic entry vector	GCTCCATATAATGGTTTACTGTAAGGAACAATGCCATCCGTAGG	
MYB33-mSL2-2-F	*MYB33* genomic entry vector	ATTCCAAGCCCTTGTACGGGGCGGTGAAACTGGAGCTCCCTTCAT	mutation of *MYB33* entry vector to *MYB33-mSL2-2* entry vector
MYB33-mSL2-2-R	*MYB33* genomic entry vector	CCCGTACAAGGGCTTGGAATACGGAACAATGCCATCCGTAGGG	
mMYB33-mSL2-2-F	*MYB33-mSL2-2 entry* vector	TCGAATTGCCAAGCTTTCAGTATTCAGAAACAACATTTGACCAGTGGAAG	mutation of *MYB33-mSL2-2* entry vector to *mMYB33-mSL2-2* entry vector
mMYB33-mSL2-2-R	*MYB33-mSL2-2 entry* vector	ACTGAAAGCTTGGCAATTCGAGTTTCACCGCCCCGTACAAGG	
MYB33-1F	*MYB33* genomic entry vector	TTCCTTGGGACATTTTCG	Sequencing of *MYB33-mSL* entry vector
MYB33-2F		GTTGTGGGCTTTTCTTTAG	
MYB33-3F		GTATTTTCCTTAGCCTGG	
MYB33-4F		GCTGCTGATGATAATGGAAG	
MYB33-5F		CTGGGGAGTTGTGAATCT	
MYB33-6F		GGGGAGAAGAGTCAGATTT	
MYB33-7F		CTTGGCGATGACATAGGA	
MYB33-8F		AGACAGAGAGCATAAATAAAA	
MYB33-mSL-q-F	*MYB33-mSL*	CGACATTCACCTGTTATGATTC	qRT-PCR of *MYB33-mSL*
MYB33-mSL-q-R	*MYB33-mSL*	AGTGTGGAGGAGATGACGATT	
CP1-F	*CP1*	ACGCTCGATAAACATTTGGG	qRT-PCR of *CP1*
CP1-R	*CP1*	CAAGCAAACAATCTTTCTTTGAA	

Primer Table 6

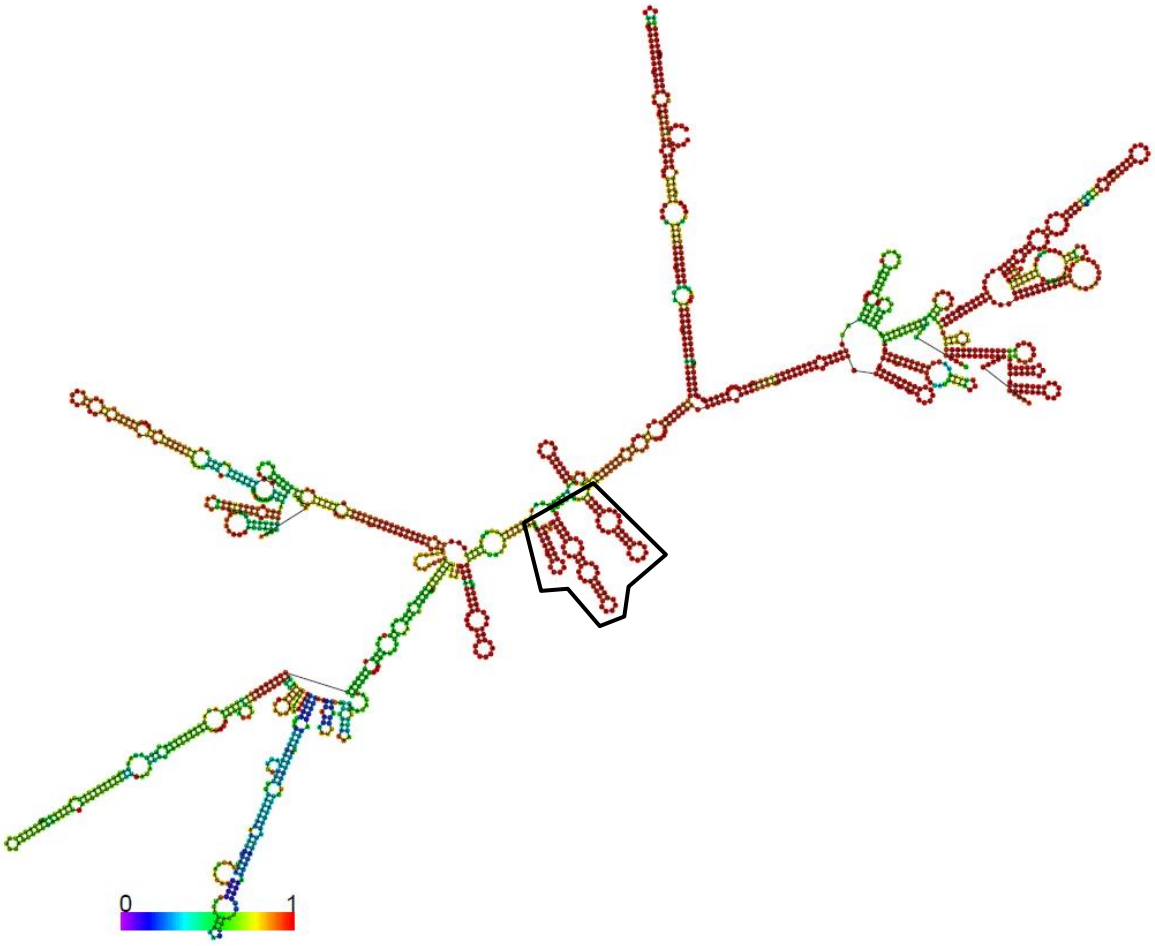
Name	***Template***	***Sequence 5'-3'***	***Purpose***
MYB81-SL-6nt-F	*MYB81-SL* entry vector	AATGGCGCTGCTCCTCCTTCACCGCCCCTTCA	mutation of *MYB81-SL* entry vector to *MYB81-SL-6nt* entry vector
MYB81-SL-6nt-R	*MYB81-SL* entry vector	AGGAGCAGCGCCATTCCCAAGCGAATGGCTATCC	
mMYB81-SL-F	*MYB81-SL and MYB81-SL-6nt* entry vector	TCGAATTGCCAAGCTTTCAGTATCATGAAGAGCCTGGCGGTTGG	mutation of *MYB81-SL(-6nt)* entry vector to *mMYB81-SL(-6nt)* entry vector
mMYB81-SL-R	*MYB81-SL and MYB81-SL-6nt* entry vector	ACTGAAAGCTTGGCAATTCGAACTTCTCAGCCCCGTGAAGGG	
MYB81-SL-q-F	*MYB81-SL*	CGTTTCAGAACAGTTACCGGA	qRT-PCR of uncleaved *MYB81-SL*
MYB81-SL-q-R	*MYB81-SL*	AGTGTTATCTGACTCATGTTCCG	

Primer Table 7

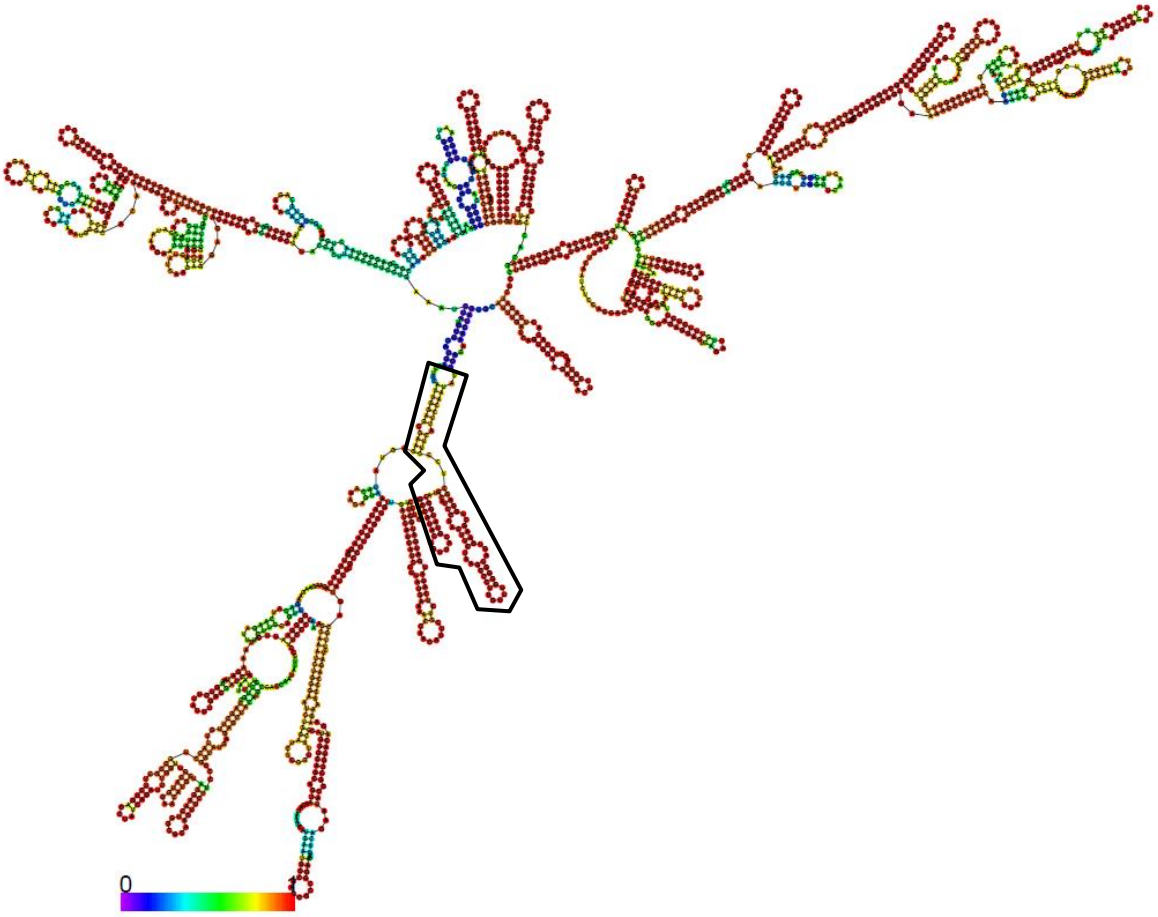
Name	***Template***	***Sequence 5'-3'***	***Purpose***
attB-NtGAMYB2-F	*tobacco cDNA*	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGACTTCAAAAGTTGGTGTGGA	*NtGAMYB2* gateway construct
attB-NtGAMYB2-R	*tobacco cDNA*	GGGGACCACTTTGTACAAGAAAGCTGGGTTCACCAACCACCATCGCT	
NtGAMYB2-1	*NtGAMYB2* entry vector	ATGACTTCAAAAGTTGGTGTGGA	Sequencing of *NtGAMYB2* entry vector
NtGAMYB2-2	*NtGAMYB2* entry vector	GGGGACCACTTTGTACAAGAA	
mNtGAMYB2-F	*NtGAMYB2 entry vector*	TCGAATTGCCAAGCCTTCAGATTCAGATGGAAAATTGGGGTTCACCTTC	mutation of *NtGAMYB2* entry vector to *mNtGAMYB2* entry vector
mNtGAMYB2-R	*NtGAMYB2 entry vector*	TCTGAAGGCTTGGCAATTCGAGCTTCTTAGCCCATTTGGGCTCTGAAG	
NtGAMYB1-3'q-F	*wild-type tobacco cDNA*	GGTGAGAATGGTTATCGATTTGAC	qRT-PCR of 3' UTR of *NtGAMYB1* transcripts
NtGAMYB1-3'q-R		TGATAACTAACTCAACCTTCTTCTGCT	
NtGAMYB2-3'q-F		GGTGAGAATGGTTATCGATTTGAC	qRT-PCR of 3' UTR of *NtGAMYB2* transcripts
NtGAMYB2-3'q-R		CATCCTTCCTATCAACCAGAACTAAGT	
NtGAMYB3-3'q-F		TTGTGGACATGGCTCTTGCA	qRT-PCR of 3' UTR of *NtGAMYB3* transcripts
NtGAMYB3-3'q-R		TCAACACAATCTAACATGCAGAAGTC	
NtGAMYB-adapter-F	*5'-adapter-NtGAMYB cleaved products*	ATGGACTGAAGGAGTAGAAATCACT	PCR of *NtGAMYB* cleaved products
NtGAMYB1-RACE-R	*5'-adapter-NtGAMYB1 cleaved products*	TGATAACTAACTCAACCTTCTTCTGCT	
NtGAMYB2-RACE-R	*5'-adapter-NtGAMYB2 cleaved products*	CATCCTTCCTATCAACCAGAACTAAGT	
NtGAMYB3-RACE-R	*5'-adapter-NtGAMYB3 cleaved products*	AAGATCGGGTTTGAGGCGTCAT	
NtGAMYB-Seq	*5'-adapter-NtGAMYB PCR products*	CAGAATGTCCTAAAGGAGAGATTGG	*Sequencing of 5'-adapter-NtGAMYB PCR products*
MIM159-q-F	*tobacco MIM159*	AGAAGGCTGATTCAGACTGCG	*qRT-PCR of Arabidopsis and tobacco MIM159*
MIM159-q-R		GATACATAATCTGCTGATTCCGAGG	
NtGAMYB-q-F	*three NtGAMYB isoforms*	AATTCCTGGCAGCCATGC	*qRT-PCR of total NtGAMYBs*
NtGAMYB-q-R		GATTCATAAAGTACAGCGTCCAA	
NtTCP4-q-F	*NtTCP4*	TTGGCGGCCAAAATCAGT	*qRT-PCR of NtTCP4*
NtTCP4-q-R		GGATCAATCCATGCTCGAA	
NtPR1b-q-F	*N. tabacum gene_61952*	GATGTGGGTCGATGAGAAACAG	*qRT-PCR of NtPR1b*
NtPR1b-q-R		CTCCATTGTTGCACTTAACCCTA	
NtaPR2-q-F	*N. tabacum gene_15771*	TATCTGAAAGTGGCTGGCCT	*qRT-PCR of NtPR2*
NtaPR2-q-R		TCCAGGTTTCTTTGGAGTTCC	
NtaPR-Q-q-F	*N. tabacum gene_8988*	AAGCCATTGGCAGGACATC	*qRT-PCR of NtPR-Q, acidic chitinase*
NtaPR-Q-q-R		CTGTTCTTAACAGGGAGCTTGG	
NtPP2A-F	*NtPP2A*	GACCCTGATGTTGATGTTCGCT	*qRT-PCR of N. tabacum reference gene*
NtPP2A-R		GAGGGATTTGAAGAGAGATTTC	

Primer Table 8

Name	***Template***	***Sequence 5'-3'***	***Purpose***
attB-OsGAMYB-F	*rice cDNA*	GGGGACAAGTTTGTACAAAAAAGCA GGCTATGTATCGGGTGAAGAGCGA	*OsGAMYB* gateway construct
attB-OsGAMYB-R	*rice cDNA*	GGGGACCACTTTGTACAAGAAAGCTG GGTCCAAGAATCTGCTGTAGCGTCT	
mOsGAMYB-F	*OsGAMYB entry vector*	TGGAATTGCCAAGCTTGCAGGATACT GAATCTGATCCAAACAGCTGGCTCAA GTA	mutation of *OsGAMYB* entry vector to *mOsGAMYB* entry vector
mOsGAMYB-R	*OsGAMYB entry vector*	CCTGCAAGCTTGGCAATTCCATCTTCA AAGGACCACTTGTGGGCCTAGAAG	
OsGAMYB-3'q-F	*Wild-type rice cDNA*	GCATGGAGATAAGCCGGTTA	qRT-PCR of 3' UTR of *OsGAMYB* transcripts
OsGAMYB-3'q-R		GCTCAACAAGATAGTACACCACAAA	
OsGAMYBL1-3'q-F		AGATGACGCTTTCGGTCATG	qRT-PCR of 3' UTR of *OsGAMYBL1* transcripts
OsGAMYBL1-3'q-R		AATCAGGTATGTCACAGGCACC	
OsGAMYBL2-3'q-F		AACCACTCCAGCTCCGTCT	qRT-PCR of 3' UTR of *OsGAMYBL2* transcripts
OsGAMYBL2-3'q-R		AGATGTTGTCCCACTGGCA	
OsGAMYB-adapter-F	*5'-adapter OsGAMYB cleaved products*	ATGGACTGAAGGAGTAGAAATCACT	PCR of *OsGAMYB* cleaved products
OsGAMYB-RACE-R	*5'-adapter-OsGAMYB cleaved products*	GCTCAACAAGATAGTACACCACAAA	
OsGAMYBL1-RACE-R	*5'-adapter-OsGAMYBL1 cleaved products*	AATCAGGTATGTCACAGGCACC	
OsGAMYBL2-RACE-R	*5'-adapter-OsGAMYBL2 cleaved products*	AGATGTTGTCCCACTGGCA	
OsGAMYB-Seq	*5'-adapter-OsGAMYB PCR products*	TCTGGGCTAAGAACCGTAGTAT	*Sequencing of 5'-adapter-OsGAMYB PCR products*
OsGAMYBL1-Seq	*5'-adapter-OsGAMYBL1 PCR products*	CTCAGAGAGAAAGCATTCGATT	*Sequencing of 5'-adapter-OsGAMYBL1 PCR products*
OsGAMYBL2-Seq	*5'-adapter-OsGAMYBL2 PCR products*	AGGTCGAAGAGTGTCTCGATGT	*Sequencing of 5'-adapter-OsGAMYBL2 PCR products*
OsMIM159-F	*OsIPS1 (AY568759*)	TACAGAGCTCCCTCAATCAATCCAAA TTATTCGGTGGATGT	mutation of *OsIPS1* entry vector to *OsMIM159* entry vector
OsMIM159-R		TGATTGAGGGAGCTCTGTACCTTAGT AGAGGTAAAAGTCGTCG	
Hygromycin-F	hygromycin resistant gene in binary vector	GTCTGTCGAGAAGTTTCTGATCGA	*Genotyping of OsMIM159 transformation*
Hygromycin-R		CACTATCGGCGAGTACTTCTACACA	
OsMIM159-q-F	*OsMIM159*	GTCGACGACTTTTACCTCTACTAAGG	qRT-PCR of *MIM159*
OsMIM159-q-R		TGTCTAACATCATACAGCCAAGATC	
OsGAMYB-q-F	*OsGAMYB*	TAGGCCCACAAGTGGTCCTT	qRT-PCR of *OsGAMYB*
OsGAMYB-q-R		TGCAACGCAGGAGCTACAGT	
OsGAMYBL1-q-F	*OsGAMYBL1*	TCAGTCAGTACCTTTTGGTAGTGC	*qRT-PCR of OsGAMYBL1*
OsGAMYBL1-q-R		ATGTCCTAAGGGGTCGTGATAT	
OsPR1a-F OsPR1a-R	*OsPR1a*	TCGGAGAAGCAGTGGTACGA CGAGTAGTTGCAGGTGATGAAGA	*qRT-PCR of OsPR1a*
Os-1,3-glucanase-F	*Os-1,3-glucanase*	GTTCTTGACGTTGTCGATCACG	*qRT-PCR of Os-1,3-glucanase*
Os-1,3-glucanase-R		CACCTACACCAACCTCTTCAACG	
Os-chitinase 8-F	*Os-chitinase 8*	CTCCTTCAAGACGGCCATCT	*qRT-PCR of Os-chitinase 8*
Os-chitinase 8-R		GATGTTGGTGATGACGCCATA	
OsActin-F	OsActin, Os03g50890	CTCCCCCATGCTATCCTTCG	*qRT-PCR of rice reference gene*
OsActin-R		TGAATGAGTAACCACGCTCCG	
mLaMYB33-F	*LaMYB33*	TAGAATTGCCAAGCAGTCAGCTTGCG GAGTCTGTGCATACCACA	mutation of *LaMYB33* entry vector to *mLaMYB33* entry vector
mLaMYB33-R		GCTGACTGCTTGGCAATTCTAACTTCA CACCTGGTAGGGGTGTAGGA	



Appendix Figure 1. Predicted RNA secondary structure of *MYB33*. *MYB33* structure is predicted by the Vienna RNAfold web server using the whole coding DNA sequence (CDS). Black outlined part is shown in Figure 3.7A. The heat maps indicate the probability of second structure formation, from low (purple) to high (red).



Appendix Figure 2. Predicted RNA secondary structure of *MYB65*. *MYB65* structure is predicted by the Vienna RNAfold web server using the whole coding DNA sequence (CDS). Black outlined part is shown in Figure 3.7A. The heat maps indicate the probability of second structure formation, from low (purple) to high (red).

Appendix Table 9. Top 20 enriched pathways from 10-fold up-regulated genes in *MIM159* tobacco leaves. Based on the RNA sequencing data, 804 genes up-regulated in *MIM159* tobacco leaves at the level of more than ten-fold, compared to wild type. These genes were referred to the closest homologues from Arabidopsis and the Gene Ontology (GO) enrichment were analysed by agriGO (Du et al., 2010). The highest 20 enriched pathways were ranked by p-value. Pathway with a p-value of ≤0.05 was considered as significantly enriched.

Pathway	GO term	p-value
defense response	GO:0006952	4.3E-15
multi-organism process	GO:0051704	2.8E-11
response to carbohydrate stimulus	GO:0009743	8.4E-10
response to chitin	GO:0010200	1.8E-09
protein amino acid phosphorylation	GO:0006468	6E-09
defense response to fungus	GO:0050832	1.2E-07
response to endogenous stimulus	GO:0009719	1.3E-07
response to oxidative stress	GO:0006979	1.4E-07
immune system process	GO:0002376	1.6E-07
phosphorylation	GO:0016310	1.7E-07
cellular response to stimulus	GO:0051716	2.6E-07
innate immune response	GO:0045087	2.7E-07
cell death	GO:0008219	3.1E-07
post-translational protein modification	GO:0043687	3.5E-07
programmed cell death	GO:0012501	8.5E-07
localisation	GO:0051179	1.8E-06
defense response, incompatible interaction	GO:0009814	1.9E-06
response to hormone stimulus	GO:0009725	2.6E-06
transport	GO:0006810	2.6E-06
protein modification process	GO:0006464	8.80E-06

Appendix Table 10. Top 20 enriched pathways from 10-fold down-regulated genes in *MIM159* tobacco leaves. Based on the RNA sequencing data, 1038 genes down-regulated in *MIM159* tobacco leaves at the level of more than ten-fold, compared to wild type. These genes were referred to the closest homologues from Arabidopsis and the Gene Ontology (GO) enrichment were analysed by agriGO (Du et al., 2010). The top 20 enriched pathways were ranked by p-value. Pathway with a p-value of ≤0.05 was considered as significantly enriched.

Pathway	GO term	p-value
enzyme linked receptor protein signaling pathway	GO:0007167	2.30E-11
cell surface receptor linked signaling pathway	GO:0007166	1.00E-10
microtubule-based process	GO:0007017	4.00E-10
post-embryonic development	GO:0009791	2.90E-09
regulation of cell cycle	GO:0051726	7.10E-08
organ development	GO:0048513	9.30E-08
lipid localisation	GO:0010876	9.60E-08
anatomical structure development	GO:0048856	1.20E-07
anatomical structure morphogenesis	GO:0009653	1.80E-07
cell cycle	GO:0007049	2.10E-07
post-embryonic morphogenesis	GO:0009886	1.10E-06
multicellular organismal process	GO:0032501	1.40E-06
cell wall modification	GO:0042545	1.70E-06
phenylpropanoid metabolic process	GO:0009698	3.00E-06
regulation of cellular process	GO:0050794	4.40E-06
tissue development	GO:0009888	3.40E-05
regulation of biosynthetic process	GO:0009889	4.90E-05
cellular amino acid derivative metabolic process	GO:0006575	6.10E-05
stomatal complex development	GO:0010374	6.70E-05
secondary cell wall biogenesis	GO:0009834	6.90E-05

Appendix Table 11. Top 50 most down-regulated genes in *MIM159* tobacco leaves. RNA sequencing results indicate top 50 genes, referred as *Nicotiana tabacum* gene ID and gene description, are most down-regulated in fold changes of *MIM159* to wild type.

Gene ID	Description	Fold change
gene_71053	(E,E)-geranyllinalool synthase-like	0.0009
gene_12247	UDP-glycosyltransferase 74E2-like	0.0011
gene_41674	vacuolar cation/proton exchanger 3-like	0.0015
gene_15917	vacuolar cation/proton exchanger 3-like	0.0017
gene_22273	anthocyanidin 3-O-glucosyltransferase 5-like	0.0018
gene_14124	lactoylglutathione lyase-like	0.0025
gene_55704	CASP-like protein 4D2	0.0026
gene_64057	transmembrane protein 45A-like	0.0026
gene_21475	PIN-LIKES 3-like	0.0030
gene_70089	chalcone synthase A	0.0032
gene_64613	anthocyanidin 3-O-glucosyltransferase-like	0.0035
gene_15080	chalcone synthase 2	0.0040
gene_5374	aspartic proteinase A1-like	0.0050
gene_69940	unknown	0.0055
gene_8364	expansin-B3-like	0.0055
gene_64852	endoglucanase 18-like	0.0062
gene_16711	3-ketodihydrosphingosine reductase-like	0.0062
gene_64623	anthocyanin 5-O-glucoside-6'-O-malonyltransferase-like	0.0063
gene_76159	zeatin O-xylosyltransferase-like	0.0065
gene_42936	antimicrobial peptide (SN1a)	0.0070
gene_44680	squamosa promoter-binding-like protein 9	0.0071
gene_50474	aspartic proteinase A1-like	0.0073
gene_10894	glutamine synthetase-like	0.0074
gene_1011	cysteine-rich extensin-like protein	0.0080
gene_6668	auxin-induced protein AUX22-like	0.0081
gene_8731	calvin cycle protein CP12-2, chloroplastic-like	0.0081
gene_20582	hybrid proline-rich protein (HyPRP1)	0.0083
gene_563	chalcone synthase A	0.0083
gene_43615	sulfate transporter 3.1-like	0.0089
gene_36409	GDSL esterase/lipase At5g45950-like	0.0091
gene_46573	aspartic proteinase A1-like	0.0091
gene_1010	keratinocyte proline-rich protein-like	0.0092
gene_73264	agamous-like MADS-box protein AGL8 homolog	0.0093
gene_25683	expansin-B3-like	0.0094
gene_58432	non-specific lipid-transfer protein 1-like	0.0097
gene_18067	ornithine decarboxylase-like	0.0100
gene_22611	proline dehydrogenase 2, mitochondrial-like	0.0102
gene_28437	aspartic proteinase A1-like	0.0103
gene_100	probable pectinesterase 68	0.0103
gene_14682	unknown	0.0105
gene_45615	21 kDa protein-like	0.0109
gene_61081	cytochrome P450 CYP736A12-like	0.0110
gene_53863	zinc-finger homeodomain protein 5-like	0.0111
gene_16624	protodermal factor 1-like	0.0112
gene_33281	isoaspartyl peptidase/L-asparaginase 2	0.0114
gene_63135	zinc-finger homeodomain protein 5-like	0.0114
gene_61083	cytochrome P450 CYP736A12-like	0.0116
gene_9127	bifunctional purple acid phosphatase 26-like	0.0123
gene_15133	expansin-B3-like	0.0125
gene_12954	cytochrome P450 CYP736A12-like	0.0127