ABSTRACT

Title of Dissertation: THE ROLE OF SMALL RNAS IN REGULATING *APETALA2* AND *CAULIFLLOWER* DURING FLOWER DEVELOPMENT

Boyana Grigorova, Master of Science, 2010

Thesis Directed By: Professor Zhongchi Liu
Department of Cell Biology and Molecular Genetics

The discovery of small RNAs, including miRNAs and siRNAs, has added new layers of complexity to the numerous pathways that direct plant development. These molecules play a fundamental role as negative regulators of gene expression in a variety of developmental processes, including meristem initiation and differentiation, light responses, and proper formation of leaves, roots and inflorescences. My work provides deeper understanding into the function of *miR172,* and characterizes a novel interfering RNA, encoded by the first intron of the meristem-specific gene *CAULIFLOWER.*

The first part of my thesis focuses on the transcriptional regulation of two *miR172* genes by the LUG, SEUSS and AP2 co-repressor complex, which binds to the microRNA promoter to negatively and directly regulate its expression. My study provides evidence that a negative regulatory feedback loop exists between *miR172* and AP2, where *miR172* restricts AP2 function to the outer two floral whorls, while AP2 limits *miR172* expression to the inner two floral whorls.

Additionally, *lug* loss-of-function mutation causes a dramatic decrease in the transcript level of *AGO1,* an essential component of the RISC complex, suggesting that LUG acts as a regulator of *AGO1* as well. My thesis work highlights the importance of LUG as a major regulator of the miRNA pathway and further elucidates the molecular mechanisms underlying the antagonistic interactions between class A and class C genes during flower development.

The second part of my thesis addresses the mechanisms of intron-mediated gene silencing. My project provides data that the intron of the MADS-box transcription factor *CAULIFLOWER* can silence the expression of its host gene. Specifically, I identified a novel siRNA encoded by the first intron of the *CAULIFLOWER* gene, which transcriptionally inhibits *CAL* and restricts its expression domain. Moreover, my results indicate that the intron-derived siRNA leads to heterochromatin repression of the whole *CAL* gene locus. This silenced epigenetic pattern is stable across generations and can be inherited without the presence of the transgene. Conceivably, my thesis work on the novel intronic small RNA can be used as an effective tool to generate transgenic plants for research and agricultural purposes.

The Role of small RNAs in Regulating *APETALA2* and *CAULIFLOWER* during Flower Development

By

Boyana Grigorova

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment of
the requirements for the degree of
Master of Science
2010

Advisory Committee:
Associate Professor Zhongchi Liu, Chair
Associate Professor Caren Chang
Associate Professor Stephen M. Mount

Acknowledgements

I am truly grateful to my supervisor, Dr. Zhongchi Liu, whose encouragement, guidance and support from the initial to the final stages of my projects enabled me to develop an understanding of the subject. She has been a great mentor to me for the past four years, and this thesis work would not have been possible without her patience, her enthusiasm, her inspiration and sound advice.

I would also like to thank my committee members, Dr. Caren Chang and Dr. Stephen Mount, for their critical thinking and wise advice regarding the development of my thesis work.

I wish to thank Dr. Chloe Mara who deserves my deepest gratitude not only for patiently teaching me new techniques and collaborating with me on many of the experiments, but also for helping me get through the difficult times, and for all the emotional support, friendship, and caring she provided.

Lastly, and most importantly, I wish to thank my family (my mother, father and sister) and my boyfriend for providing a loving network for me. Without them I would have been lost. To them I dedicate this thesis.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Literature Review

1.1 Introduction

1.2 Flower Development

1.2.1 The ABCE model of flower development

Although flowers appear in numerous varieties of colours, shapes and forms, floral development in angiosperms is based on variations of the same fundamental pattern. Comprehensive analyses of the small flowering plant *Arabidopsis thaliana* have provided some of the most in-depth information on floral pattern formation. The flowers of dicots originate as a small group of undifferentiated cells, called the floral meristem, on the edges of the inflorescence (shoot) apical meristem. These lateral primordia eventually give rise to four concentric whorls of floral organs, each whorl being a separate domain that produces a single type of floral organ. In *Arabidopsis*, the outer two whorls consist of sterile organs, where the sepals occupy the outermost first whorl and the colorful petals take up the second whorl. The third and fourth whorls contain the reproductive organs- six stamens, and two fused carpels, respectively.

There are numerous genes that are required for the initiation and development of flowers and their role has been largely elucidated by the study of floral homeotic mutants. In these mutants a flower develops normally, but floral organs that normally occupy a different whorl replace floral organs at one position of the flower. Typically, the identity of two adjacent whorls is affected, but not the position or the number of floral organs. Genetic analyses of these floral homeotic mutants have led up to the postulation of the ABCE model of flower

development, which has become a milestone in the proper understanding of the mechanisms underlying floral patterning (Bowman et.al., 1991; Coen and Meyerowitz, 1991). The original ABC model proposed that three classes of transcription factors act alone or in combination with another class to specify the identity of one or more of the four organ whorls. Class A genes are required for the development of sepals and petals (whorls 1 and 2), B class genes specify petals and stamens (whorls 2 and 3), and C class genes determine the formation of stamens and carpels (whorls 3 and 4). In addition to the genes described in the initial ABC model, later work established the importance of class E genes, which are necessary to allow class A, B and C genes their flower specific activity (Honma and Goto, 2001; Ditta et.al., 2004). In Arabidopsis, *APETALA1/APETALA2* (*AP1/AP2*) are class A genes, *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) are class B genes, *AGAMOUS* (*AG*) is a class C gene and *SEPALLATA1/2/3/4* (*SEP1/2/3/4*) are class E genes. Nearly all genes of the ABCE model belong to a large family encoding transcription factors containing the conserved MADS domain, a protein domain that binds to the consensus CArG-box sequence in promoters of prospective target genes (Nam et.al., 2003). The only exception is *APETALA2 (AP2),* an A class gene that belongs to a different family of transcription factors with a novel AP2 binding domain (Okamuro et.al., 1997).

Class	*Arabidopsis thaliana* Genes
A	*APETALA1* (*AP1*) /*APETALA2* (*AP2*)
B	*APETALA3* (*AP3*) /*PISTILLATA* (*PI*)
C	*AGAMOUS* (*AG*)
E	*SEPALLATA1,2,3,4* (*SEP1,2,3,4*)

Table 1.1: Floral Organ Identity Genes in *Arabidopsis thaliana*

Mutations in the different classes of genes lead to a variety of floral organ transformations. In class A mutants the first and second whorls are affected, with carpels replacing sepals and stamens replacing petals. Mutants of the B-class genes exhibit identical phenotype, namely the conversion of petals in the second whorl to sepals and of stamens in the third whorl to carpels. Finally, mutations in the class C gene result in the conversion of stamens to petals in whorl 3 and sepals rather than carpels in whorl 4 (or alternatively the initiation of the first whorl of an inner flower) (Coen and Meyerowitz, 1991). The existence of these mutants confirms that floral organ identity is specified by the combinatorial action of more than one type of gene activity in the different regions of the floral meristem.

Genotype	Phenotype of Whorl 1	Phenotype of Whorl 2	Phenotype of Whorl 3	Phenotype of Whorl 4
Wild type	Sepal	Petal	Stamen	Carpel
ap2	Carpel	Stamen	Stamen	Carpel
ap3/pi	Sepal	Sepal	Carpel	Carpel
ag	Sepal	Petal	Petal	New Flower

Table 1.2: Phenotype of some organ identity mutants in *Arabidopsis thaliana*

To explain the homeotic phenotypes fully the ABCE model contains two important precepts, namely that class A and class C genes mutually inhibit each other's activity and that C class gene *AGAMOUS* is required for determinate flower development. In class A mutants, A function is lost and the mutual antagonism between the two classes of genes is removed. Class C gene activity spreads to all four whorls, resulting in the replacement of sepals with carpels, and petals by stamens (Bowman et.al., 1989). Conversely, in class C mutants, C function is lost and A class activity spreads to all four whorls giving rise to a pattern of sepals, petals, petals, new flower. Inactivation of B class genes causes second whorl organs to convert to first whorl organs, and third whorl organs to fourth whorl organs, giving rise to flowers consisting of sepals, sepals, carpels, carpels. Throughout most of a flower's development, the transcription patterns of *AP2*, *AP3, PI* and *AG* corresponds to the domains affected by their respective mutations although the molecular mechanisms behind this pattern and the antagonism between the A and C class genes are not fully understood. However, experiments with double mutants and transgenic plants overexpressing the ABC

genes support the idea that the regulation of their activity occurs at the transcription level. For example, ectopic expression of *AP3* and *PI* leads to flowers in which the first whorl adopts the same fate as the second whorl and is occupied by petals, and the fourth whorl carpels are replaced by stamens (Krizek and Meyerowitz, 1996). Quadruple transgenic plants, overexpressing class B genes *APETALA3* and *PISTILLATA*, class C gene *AGAMOUS*, and class E gene *SEPALLATA3* produce cauline leaves transformed into stamens or staminoid organs, supporting the hypothesis that floral organs are modified leaves (Honma and Goto, 2001).

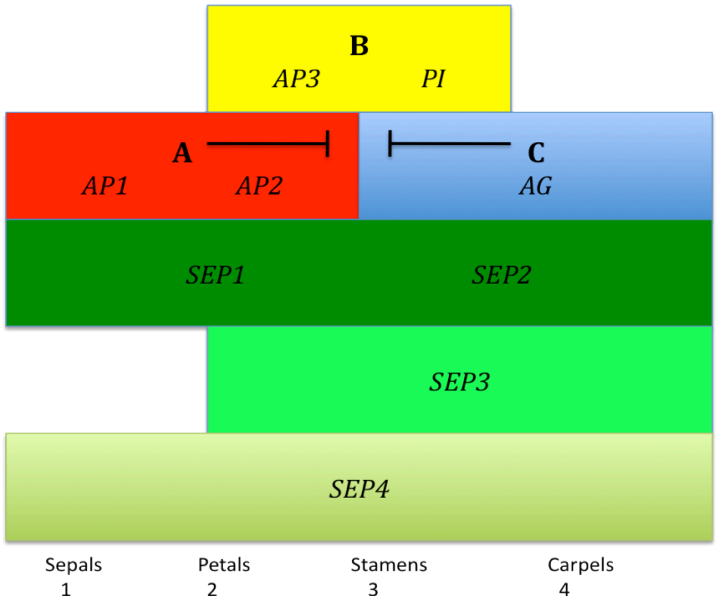

Fig. 1.1: The ABC and SEP genes specify floral organ identity. The ABCE model postulates that in whorl 1, A-class activity specifies sepals; in whorl 2, A+B+SEP activities specify petals; in whorl 3, B+C+SEP activities specify stamens; and, in whorl 4, C+SEP activities specify carpels (Modified from Jack, 2001).

1.2.2 Beyond the ABCE model: establishing the pattern of A, B and C activity

The ABC model successfully explains how floral organ identity is established through the combinatorial action of several transcription factors. However, it does not demonstrate how the expression patterns and activities of the ABC genes are set up. Central to the ABC model is the mutual inhibition between the A and C class genes, which suggests that they might be involved in regulating each other's activity. Flowers of plants with ectopic *AGAMOUS* expression resemble those of *ap2* mutants, indicating that AG is able to repress AP2 activity in the outer two whorls through an unknown mechanism (Mizukami, 1992). More recently, a plant microRNA, miR172, has been identified as sharing a high degree of sequence complementarity to a coding region outside of the AP2 domain of *APETALA2,* which results in its post-transcriptional regulation (Xuemei Chen, 2003; Aukerman and Sakai, 2003; Scwab et.al., 2005). Studies in *ap2-2* and *ap2-2 ag-1* mutants indicate that *AP2* promotes *AP3* and *PI* expression while the antagonistic interaction between the A and C class genes determine the functional domain of the B class genes (Zhao et.al., 2008). Thus, indirectly, by repressing *AP2* activity to the first two whorls, miR172 also plays an important role in establishing the inner boundary of the B class gene expression domain.

In addition to *AG*, *AP2* and miR172, the expression of B class genes *AP3* and *PI* is directly regulated by several other transcription factors, many of which are central to the transition from vegetative to reproductive stage in flower development. Early acting genes controlling the identity of floral meristems are

LEAFY, *AP1* and *CAULIFLOWER* and all functions of *CAULIFLOWER* are redundant with those of *AP1* (Weigel et.al., 1992; Bowman et.al., 1993). *LEAFY* and *AP1* gene products are necessary for the correct transition of an inflorescence meristem to a floral meristem (Weigel et.al., 1992). After the identity of the floral meristem is specified, AP1 interacts with AP2 to specify the fate of the outer two floral organs. Thus, plants homozygous for the *ap1-1* mutation exhibit homeotic conversion of sepals into bracts and lack petals (Irish and Sussex, 1990). *AP1* is negatively regulated by two factors: the floral homeotic gene *AGAMOUS* suppresses its expression in the inner two whorls, while the meristem identity gene *TERMINAL FLOWER* prevents *AP1* RNA accumulation in the inflorescence meristem (Gustafson-Brown et.al., 1994). LEAFY is a direct transcriptional activator of *AP1*, as its expression is significantly delayed and reduced in *lfy* mutants (Wagner et.al., 1999; Liljegren et.al., 1999). Furthermore, the *ap1* phenotype is significantly enhanced by a mutation at the *CAULIFLOWER* gene locus (Bowman et.al., 1993). *CAULIFLOWER* is expressed in young floral primordia and it encodes a MADS-box transcription factor, homologous to *AP1* (Kempin et.al., 1995). It has been proposed that CAULIFLOWER may act to positively regulate *APETALA1* and *LEAFY* expression in early floral primordia. Alternatively, *CAULIFLOWER* and *AP1* may have completely overlapping roles in flower development, as the *cal-1* mutant phenotype is only revealed in an *ap1* background (Bowman et.al., 1993).

In addition to promoting class A (*AP1*) and class B (*AP3* and *PI*) expression, LEAFY also acts as a direct upstream activator of C class gene

AGAMOUS (Parcy et.al., 1998; Busch et.al., 1999). In strong *lfy* mutants there is a delay in the onset of *AG* expression (Weigel and Meyerowitz, 1993). LFY has been shown to directly bind to a transcriptional enhancer located in the first intron of *AG* (Busch et.al., 1999), thus promoting its expression in the third and fourth floral whorls. While *LEAFY's* role is to activate *AG*, A class gene AP2 negatively regulates AG expression. In *ap2* mutants, *AG* RNA is present in all four whorls, consistent with the mutual inhibition between class A and class C activity (Drews et.al., 1991). Moreover, *AG* expression is ectopic also in *leunig* mutants (Liu and Meyerowitz, 1995). LEUNIG encodes a transcriptional co-repressor that belongs to the Gro/Tup1-like family of transcriptional co-repressors in *Arabidopsis* (Liu and Karmarkar, 2008). Narrow floral organs, sepals transformed into stamens and carpels, and petals that are either staminoid or absent characterize the *leunig* mutant phenotype. Thus, the combinatorial action of various molecular components defines the correct expression pattern of the genes involved in regulating floral organ development.

1.2.3 Role of *LEUNIG* in flower development

LEUNIG (LUG) was initially identified in a genetic screen as an enhancer of the floral homeotic gene *APETALA2* (Liu and Meyerowitz, 1995). LUG is a transcriptional co-repressor, member of the Gro/Tup1 family of co-repressors (Liu and Karmarkar, 2008). At the N-terminus of LUG is the LUFS domain, which contains a LisH and a conserved PFAM:SSDP domain, followed by a Q-rich domain and then C-terminal WD repeats. The LUFS-domain is a protein-protein

interaction domain, which interacts with another transcriptional co-regulator SEUSS (Franks et al, 2002). The 7 WD repeats at the C-terminus are protein-protein interaction domains and may mediate diverse interactions between LUG and various transcription factors resulting in the repression of numerous gene targets.

lug mutants are characterized by distinct phenotypic abnormalities, including narrow leaves and floral organs, homeotic transformations in floral organ identity, reduction in floral organs and split carpels (Liu and Meyerowitz, 1995). In *lug* mutants, class C gene *AG* expression expands into the first two whorls. Double mutants of *lug* and *ag* display a single *ag* mutant phenotype, suggesting that *AG* is epistatic to *LUG*. The ectopic *AG* expression in *lug* mutants shows that *LUG* prevents *AG* expression in outer two whorls. In addition, *lug* enhances the defects of class A mutant, *ap2*. *ap2 lug-1* double mutants develop filaments in the lateral positions and carpels in whorl 1. Whorl 2 organs are completely missing. there is a reduced number of stamens in whorl 3 and the carpel in whorl 4 ends in horn-like protrusions. In contrast, whorls 3 and 4 are largely normal in single *ap2* mutants.

How does LEUNIG repress *AGAMOUS* ? *SEUSS (SEU),* a Q-rich protein with a putative dimerization domain, acts as the bridging partner between LUG and a DNA biding transcription factor (Franks et.al., 2002). *AG* is ectopically expressed in single *seu* and double *seu lug* mutants suggesting that SEU acts as a co-regulator for LUG in order to negatively regulate *AG* expression in *Arabidopsis* flowers. Because neither LUG nor SEU possess a DNA-binding

motif they need to interact with a DNA-binding protein in order to downregulate their target genes. In this manner, a complex of LUG, SEU and A-class gene AP1 binds to a regulatory element in the second intron of *AG* to repress *AG* expression (Sridhar et.al., 2006). It is also suspected that LUG, SEU, and the second class A gene AP2 are similarly involved in *AG* repression. LUG is proposed to silence its target genes via two distinct mechanisms: reversible histone modifications and interaction with the transcription Mediator complex (Sridhar et.al., 2004; Gonzalez et.al., 2007). In the first scenario LUG recruits Histone Deacetylase (HAD) 19 which may result in histone deacetylation and consequently gene silencing. On the other hand, the LUG-SEU complex also interacts with the Mediator complex to modify RNA polymerase II activity and thus regulate gene expression.

LEUNIG and SEUSS are global plant regulators controlling many aspects of plant development. Identifying additional DNA binding partners in the LUG-SEU complex will provide further insight into the identity of other downstream targets and will facilitate the understanding of the functional specificity for different pathways for similar large protein complexes in other organisms. Future research into the functions of the LUG-SEU regulatory complex is crucial for elucidating important aspects of plant development, including body patterning, fertility and pathogen resistance.

1.3 microRNAs: Small Molecules with a Big Impact on Flower Development

1.3.1 MicroRNA identification, origin and conservation in *Arabidopsis*

MicroRNAs (miRNAs) are approximately 21-bp- long single-stranded RNA molecules that play a key regulatory role in cellular gene expression patterns. After their discovery in 2001 (Lee and Ambros, 2001; Lau et.al. 2001; Lagos-Qunitana et.al., 2001) small RNAs have been shown to regulate key aspects of plant growth, development and stress responses. Much research has been focused on understanding the function of this diverse class of gene regulators; however genetic studies still fall short of fully comprehending how microRNA expression itself is regulated. The availability of a sequenced genome in *Arabidopsis* has facilitated the identification of numerous microRNAs; however the discovery of microRNAs in economically important crop species remains an ongoing process. Three main methods have been employed in identifying microRNAs: forward genetics, direct cloning and sequencing, and bioinformatics (Lee et.al., 1993; Lu et.al., 2005; Zhang et.al., 2005). Usually a microRNA is also conserved among plant species starting from ferns to eudicots and monocots (Reinhart et.al., 2002).

A total of 959 microRNAs from 10 different plant species can currently be found in the microRNA database *MirBase* (http://microrna.sanger.ac.uk/sequences/index.shtml). 117 microRNAs have been identified in *Arabidopsis* and grouped into 42 gene families, many of which are fully or partially conserved in crop species such as rice and maize (Griffiths-

Jones, 2004). Studies suggest that microRNAs originate from the duplication of protein-coding sequences. Allen et.al. propose a model where plant miRNAs arose from the their target genes by formation of inverted duplications which have been transcribed but not modified further. Some miRNA genes are contained in parts of the genome considered to contain no coding sequences, such as introns and UTR regions, implying that miRNA origin may be a complex interplay of many mechanisms involving inversion and duplication.

1.3.2 Biogenesis of plant microRNAs

MicroRNA biogenesis involves multiple steps in order to covert a transcribed miRNA gene sequence into a mature miRNA. Much like regular genes microRNAs are transcribed from their own transcriptional unit, which may contain various binding motifs for known transcription factors (Megraw et. al., 2006). Initially, RNA polymerase II transcribes a microRNA into a long primary RNA (pri-miRNA) transcript (Bartel et.al., 2004; Lee et.al., 2002). Most pri-miRNAs begin with an adenosine, which is located 40 nucleotides upstream of a conserved TATA-box-like sequence. Next, the microRNA precursor, called a pre-miRNA, is released from the pri-miRNA transcript in two sequential steps involving the RNAse III endonuclease Dicer-like 1 (DCL1) (Kurihara and Watanabe, 2004). DCL1, assisted by the double-stranded RNA binding protein HYPONASTIC LEAVES1 (HYL1), catalyzes the cleavage of the long pri-miRNA transcript into a long microRNA precursor, followed by cleavage of the long pre-miRNA into a short pre-miRNA. Finally, in the nucleus DCL1 cleaves

the short pre-miRNA into a miRNA:miRNA∗ duplex and the mature miRNA is translocated into the cytoplasm by HASTY (Bartel, 2004; Park et.al., 2005). In plants, there is also an additional step in microRNA biogenesis where HEN1, a microRNA methyltransferase, deposits a methyl group on the 3' terminal nucleotide of the mature miRNA before its export to the cytoplasm (Yang et.al., 2006). In the cytoplasm, the miRNA is unwound into a single-stranded mature miRNA by a helicase and loaded onto the RNA-induced silencing complex (RISC), whose major protein component is an Argonaute protein (AGO1) (Bartel, 2004). AGO proteins have key catalytic residues that cleave the target mRNA in the middle of the complementary region between the mRNA and the miRNA (Vaucheret et.al., 2004).

1.3.3 Plant Argonaute

ARGONAUTE1 (AGO1) was first identified in a series of *Arabidopsis* mutants with severely compromised general plant architecture and leaf development (Bohmert et.al., 1998). In an *ago1* mutant axillary meristems rarely develop, the leaves lack adaxial/abaxial differentiation, filamentous structures without adaxial/abaxial differentiation develop instead of cauline leaves and the inflorescence consists of infertile filamentous structures. Extensive studies of the *ARGONAUTE* gene family in *Arabidopsis* later showed that ARGONAUTE proteins are involved in the regulation of gene expression via the RNAi silencing complex (RISC) (Baumberger and Baulcombe, 2005). The number of ARGONAUTE proteins in *Arabidopsis* is ten. AGO proteins possess one variable

N-terminal domain and 3 conserved C-terminal domains, named PAZ, MID and PIWI domains. The MID domain recognizes and binds to the 5' end of small RNAs, while the PAZ domain binds to the 3' end of small RNAs. The PIWI domain is similar to an RNAseH enzyme and has an endonuclease activity (Hutvagner et.al., 2008). Not all AGO proteins possess a slicer activity and thus mediate RNA interference. In *Arabidopsis* only AGO1, AGO4 and AGO7 have been demonstrated to cleave a target mRNA in the middle of their miRNA or siRNA complementary sequence (Qi et.al., 2005; Qi et.al., 2006; Baumberger and Baulcombe, 2005). In addition, AGO1 itself is regulated by a microRNA, miR168, illustrating that AGO1 is subject to a negative feedback regulation through the action of the small RNA silencing pathway. This mechanism ensures that AGO1 activity in the cells is maintained at a constant critical level required for the proper function of the RNAi machinery (Vaucheret et.al., 2004). Other direct regulators of AGO1 expression and function remain to be identified.

AGO1 is necessary for maintaining the integrity of the RNAi silencing pathway. AGO1 preferentially associates with miRNAs and siRNAs to cleave miRNA-targeted mRNAs. The severe developmental effects of *ago* mutants indicate that ARGONAUTE proteins are the most important constituents of the small-RNA mediated regulatory pathways in the cell. Plant *ago* mutants are very susceptible to viral infections suggesting that AGO proteins and the other components of the small RNA pathway silence viruses and mediate host defense (Morel et.al., 2002). Because miRNAs play roles in control of flowering time,

floral organ identity, cell division patterns, stem cell function and organ polarity the action of *AGO1* in the miRNA pathway is crucial for plant development.

1.3.4 Molecular mechanisms of mode of action of plant microRNAs

MicroRNAs generally silence genes via two types of interference: translational repression and reduction in the mRNA level. In the RISC complex, miRNAs bind to their target mRNA and inhibit gene expression through an imperfect (in animals) or close to perfect (in plants) complementarity. Most plant miRNA exhibit an almost precise match to a 21bp long sequence in their target mRNAs, and lead to mRNA cleavage. In animals, miRNAs usually bind to the 3'UTR region of the target mRNA and prevent ribosome movement along the mRNA and thus repress translation (Carrington and Ambros, 2003). However, certain plant miRNAs, such as miR172 and miR834, have also been shown to cause translational repression of their target mRNA (Chen, 2004; Aukerman et.al., 2003; Brodersen et.al., 2008). Translational silencing is initiated when there is an imperfect pairing with central mismatches in small RNA-target hybrids which impairs slicing and therefore cannot result in target mRNA degradation (Hutvagner et.al., 2002).

However, in some cases such as miR172, the same miRNA can be regulated via both translational and transcriptional inhibition (Chen 2004; Aukerman et al, 2003; Schwab et al., 2005). miR172 is initially present throughout the floral meristem, but is limited to the inner two whorls after floral

stage 7. miR172 negatively regulates the expression of the floral-homeotic gene *APETALA2* through translation inhibition (Chen, 2003). Additionally, overexpression of *miR172* in *Arabidopsis* causes early flowering through mRNA cleavage and protein downregulation of the *AP2*-like genes *TARGET OF EAT1 (TOE1)* and *TARGET OF EAT2 (TOE2)* (Aukerman and Sakai, 2003; Scwab et.al. 2005).

MicroRNAs play diverse functions in plant development as important regulators of gene expression. Several studies indicate that microRNAs regulate various developmental aspects including leaf morphogenesis and polarity, floral differentiation, root initiation and development, vascular development and transition from the vegetative to reproductive stage (Emery et.al., 2003; Chen, 2004; Mallory et.al. 2004; Kim et.al., 2005). Many of the microRNAs affect plant development through the regulation of crucial transcription factors that control various cellular processes. In addition, miRNAs are essential mediators of plant disease resistance and environmental stress responses through pathogen and stress-induced post-transcriptional gene silencing (Ding, 2000; Sunkar and Zhu, 2004). Thus, miRNAs appear to be involved in various aspects of plant development ranging from basic physiological processes to vital biotic and abiotic plant stress responses. It is therefore imperative that studies into miRNA function, mode of action and regulation remain the frontier of future plant research. In depth understanding of how miRNAs themselves are regulated and designing novel strategies for miRNA-mediated gene silencing could enhance plant resistance to environmental stresses and increase crop yields.

1.4 *Intron-Mediated Gene Regulation*

1.4.1 Introns

Most eukaryotic genes are interrupted by the presence of long, non-protein coding, apparently functionless sequences of DNA that are transcribed by RNA polymerase II and subsequently spliced out from the primary messenger RNA transcript. The function of introns remains to be fully understood; however, studies in recent years suggest that they may be involved in several types of gene regulation, both positive and negative. Alternative splicing of introns may also allow an increase in the complexity of some eukaryotic genomes as well as facilitate genome evolution. Intron sequences are sometimes short, while sometimes within a coding region the total length of the introns maybe much longer than that of the actual gene. Traditionally, introns have been looked at as deleterious sequences, whose insertion in a bad place may interfere with normal gene expression, and whose transcription consumes vast amounts of metabolic resources and energy. However, introns have been retained within eukaryotic genomes throughout evolution, which suggests that they must play some profound and previously overlooked role in the cell. One vital role for introns may be their regulation of gene expression.

1.4.2 Intron-Mediated Direct and Indirect Gene Regulation

Non-sense-mediated mRNA decay (NMD) is a process that facilitates the degradation of truncated mRNA transcripts that contain a premature stop codon. Studies in various plant species, including *Arabidopsis,* and mammals indicate that introns may play a pivotal role in the process, acting as an NMD signal located downstream of the termination codon (Maquat et.al., 2004). Introns can act as NMD *cis* elements in plants. In *Arabidopsis,* an intron, located in the 3'-UTR region of a transcript can trigger NMD in a position-dependant manner (Kertesz et.al., 2006). Sometimes, introns located upstream, rather than downstream of the transcript stop codon, can activate NMD (Isshiki et.al., 2001). On the other hand, it has been shown that introns can also raise mRNA accumulation via stimulation of transcription and 3'-end processing through a largely unknown mechanism (Lu and Cullen, 2003). Studies of the *Arabidopsis* TRYPTOPHAN BIOSYNTHESIS1 (TRP1) and POLYUBIQUITIN10 (UBQ10) gene introns indicate that there exists an intron- and position-dependant mechanism that stimulates gene transcription (Rose et.al., 2004). This enhancement of gene expression is possibly due to intronic U-rich sequences that render the nascent transcript more stable and leave RNA Polymerase II the more likely to elongate through to the 3' end of the gene. Additionally, increased association between the mRNA and the ribosome, due to an interaction with the exon junction comlex (EJC) proteins marking sites upstream of the introns can result in increased translational efficiency (Nott et.al., 2004).

Studies in various organisms have shown that intron-mediated enhancement of gene expression is ubiquitous and found throughout a diverse and wide range of organisms (plants, mammals, fungi, nematodes, insects). Taken together, these findings imply that introns play a fundamental role in regulating eukaryotic gene expression; however, their role remains yet to be fully understood. Some introns influence the expression of genes much more than the promoter of a gene. For example, the first intron of *PROFILIN1 (PRF1)*, a gene that encodes an *Arabidopsis* vegetative profilin gene, is required for the strong constitutive and tissue specific expression of the gene (Jeong et.al., 2006). In addition, the first intron of *ACT2*, a target gene of profilin, is necessary to direct the correct expression of vegetative actin within the shoot apical meristem, suggesting a positive role for introns in regulating actin gene family members (Jeong et.al., 2008).

The ability of introns to positively modulate translational efficiency suggests a more important function of introns than their traditionally accepted role to increase the coding capacity of genes via alternative splicing. Reduced transcription of intron-devoid areas may be a mechanism to prevent wasteful transcription of intergenic regions, as well as avoid the transcription of areas that might contain potentially harmful transcripts such as antisense transcripts. Intron-dependent interaction between the EJC proteins and a new transcript may lead to a more efficient loading of the transcript onto the translational machinery thus allowing a quicker response to changes in gene expression (Rose 2002). In

addition, functionally related genes may use introns as a common mechanism to ensure proper spatial and temporal expression.

1.4.3 How do Introns Promote Gene Expression?

The mechanisms underlying intron-mediated gene enhancement still remain to be characterized. In general introns can increase the expression either through the action of a transcriptional enhancer or an alternative promoter, located within the intron. In recent years, studies in maize have shown that there exists a third method for introns to elevate gene expression, called intron-mediated enhancement (IME) (Callis et.al., 1987). An intron falls into the last category if it can increase gene expression outside of transcribed sequences or in either orientation, as well as when it stimulated the expression of a gene that has a minimal or no promoter.

One of the most well characterized introns is the second intron of the floral homeotic gene *AGAMOUS (AG)*, which specifies the expression of *AG* in stamen and carpel whorls in *Arabidopsis*. The second *AG* intron drives an inner whorl-specific expression of a transgene, which contains translational fusion of *AG* to a *GUS*-reporter gene (Deyholos and Sieburth, 2000). The second intron of *AGAMOUS* has been shown to contain enhancer elements both in the 5' and the 3' end of the intron (Busch et.al., 1999). Fusing the intron enhancer element to a reporter gene results in a stamen- and carpel-specific expression pattern, typical of *AGAMOUS* itself, and such constructs may prove a reliable transgene

containment strategy in a variety of plant species (Liu and Liu, 2008). In addition to enhancer elements, introns can also contain alternative promoter and be able to drive or increase the expression of promoterless genes. In *Sesamum indicum* and *Arabidopsis* the *FAD2* gene encodes a desaturase, which catalyses the conversion of oleic acid to linoleic acid (Kim et.al., 2006). In both species the *FAD2* gene contains a large intron within the 5' untranslated region, which is able to increase the expression of a *GUS*-reporter construct up to a 100 fold in transgenic *Arabidopsis* plants.

A potentially significant, but poorly characterized, method by which introns elevate gene expression is intron-mediated gene ehancement (IME). The most defining feature of IME is the ability of an intron to stimulate gene expression without containing discrete enhancer or promoter elements. The mechanism underlying how IME occurs remains elusive; however, studies suggest that intron-mediated increases in mRNA levels may be a consequence of splicing or variations in intron position. Studies of the *Arabidopsis* tryptophan pathway biosynthetic gene *PAT1* show that the ability of the *PAT1* intron to enhance expression is only 50% diminished in the absence of proper splicing, implying that a complete splicing process is not absolutely required for IME to occur. IME can be completely abolished only if all the branchpoints are eliminated at the same time, which suggests that IME might be dependent upon an association with the spliceosome if not for splicing to be completed (Rose, 2004).

Another element that may control the extent to which an intron can enhance gene expression is intron position. The IME efficiency of an intron may

be dependent upon its location within the gene, or its length. Many introns that can stimulate gene expression are located within the 5' UTR of their native genes and are longer than other introns within the same gene (Chung et.al., 2006). In general, introns, located within the 5' UTR are longer than the ones within the 3' UTR or the coding sequence, and they preferentially are located close to the initiating ATG codon of the gene. The proximity of the 5' UTR introns to the transcription start site might be relevant to achieve maximum enhancement of gene expression. The further an intron is moved away from the transcription start site of a *GUS* reporter gene, the lesser its ability to stimulate gene transcription becomes (Rose, 2004).

An intron can also effect gene expression in a negative, rather than a positive, manner. For example, the *cis*-regulatory elements within *AG* intron contain tethering sites for transcriptional repressors such as the LUG-SEU-AP1 protein complex, which acts to restrict *AG* expression to the third and fourth whorl (Sridhar et.al., 2006). There are other genes that are also negatively controlled by regulatory elements located within one of their introns. The ovule and septum-specific expression of the MADS-box gene *SEEDSTICK (STK)* is controlled by sequences comprising its first intron, located within the 5' UTR region (Kooiker et.ak., 2005). *GUS*-reporter gene constructs containing deletions of the first *STK* intron led to ectopic expression of the gene within the flower, suggesting that the intron is necessary to restrict gene expression to the ovule and septum only. Both the introns of *AG* and *STK* are large, which may facilitate the presence of numerous regulatory sequences.

1.4.4 How do Intron-encoded small RNAs Inhibit Gene Expression?

Small RNA molecules of about 20-30 nucleotides have been shown powerful regulators of gene expression both on the transcriptional and translational level. Plant genomes contain several types of small regulatory RNAs, including microRNAs and small interfering RNAs (siRNAs). Early studies of small RNA molecules indicated that they are located in the noncoding regions between genes and are transcribed by unidentified promoters. However, recent advances in small RNA research have reported that there exists a new class of tiny noncoding RNAs, called intron-derived microRNAs (Id-miRNAs), transcribed from the introns of genes (Ambros et.al., 2003).

Several different types of Id-miRNAs have been identified in human, mouse and *C.elegans* cells; however, only 10 have been mapped in *Arabidopsis* genome and they still remain to be characterized (Rodriguez et.al., 2004; Lin et.al., 2004; Llave et.al., 2002). In difference to regular miRNAs and siRNAs, intronic microRNAs share the same promoter as their encoded gene target and are encoded in the gene transcript precursors, although in some cases an intron may contain a microRNA in the opposite direction of the gene transcript and with its own separate promoter. Such a microRNA, although contained within a gene intron, is not classified as an Id-miRNA and due to its antisense orientation is proposed to target the transcript of its host gene. The biogenesis of intronic miRNAs is identical to the biogenesis of regular *miRNAs*.

Intron-derived microRNAs can mediate gene silencing not only by using the traditional RNAi-silencing machinery, but also through RNAi-related chromatin remodeling events. It has been proposed there might exists a correlation between human disease and intronic microRNAs, as numerous introns containing microRNAs seem to be involved in RNAi-related chromatin silencing mechanisms (Jin et.al., 2004). For instance, fragile X syndrome occurs as a result of an erroneous intronic expansion, resulting in dysregulation of a specific 3'-UTR intronic microRNA that leads to heterochromatin repression of the whole gene locus. Such an event alters the condition-specific and time-specific manner of expression of the Id-miRNA encoding gene and results in the genetically inherited mental retardation that characterizes fragile X syndrome (Jin et.al., 2004). Thorough understanding of such conditions is crucial to the successful design of miRNA-based drugs for future gene therapies, aimed to cure or alleviate the symptoms of genetic diseases. Moreover, the construction of an artificial intron-derived miRNA system has recently become a successful strategy for knockdown of selected oncogenes and viral genome replication (Lin et.al., 2004). Man-made introns carrying miRNA precursors have already been successfully used in triggering RNAi-like gene silencing in human prostate cancer cells. Conceivably, intronic miRNAs can be also used as an effective tool to generate not only transgenic mammalian cells, but also transgenic plants for agricultural purposes. Similar constructs can potentially be used to enhance crop disease resistance and yield through the long-term and efficient suppression of specific plant genes.

1.4.5 RNA-mediated DNA methylation

RNA-directed DNA methylation, an important type of RNAi-like mechanism for gene silencing, has been recognized not only in mammalian cells, but in plant cells as well. Previously, only small interfering RNAs were believed to exert gene regulation through this mechanism; however recent studies into the regulation of class III HD-Zip transcription factors provide compelling evidence that microRNAs such as *miR166/miR165* can also mediate epigenetic silencing of their gene targets (Bao et.al., 2004). In contrast to miRNAs, which are encoded by specific miRNA genes as short hairpin pri-miRNAs in the nucleus there are no genes that encode for siRNAs. siRNAs are synthesized from double-stranded segments of matched mRNA via RNA-dependent RNA polymerase, while miRNAs are synthesized from an unmatched segment of RNA precursor featuring a hairpin turn. miRNAs are entirely endogenous to the cell, while siRNAs can either be derived endogenously from repetitive DNA sequences and associated transposons and centromers or exogenously from viruses. In plants, siRNAs interact with a homologous DNA locus to induce DNA chromatin modifications that may result in transcriptional silencing of the target gene (Matzke et.al., 2005). Similar to siRNAs, *miR166/miR165* during transcription guide the RISC complex to the PHB/ PHV gene locus and initiate chromatin methylation and subsequent epigenetic silencing, thus uncovering another layer of miRNA-mediated gene regulation.

It is possible that plant intronic small RNA molecules can also direct epigenetic modifications that repress gene expression. A novel plant-specific

protein named RNA polymerase IV (RNAP IV) is required for this process. Initially, RNAP IV was identified as a key component in the biogenesis of more than 90% of plant siRNAs, and only subsequently shown to be involved in siRNA-mediated chromatin condensation (Zhang et.al., 2007; Huettel et.al., 2007). There are several distinct features that characterize plant siRNA-directed DNA methylation. First, methylation occurs primarily at the region of RNA-DNA sequence similarity, suggesting that RNA-DNA base pairing acts as a substrate for methylation (Pellisier et.al., 1999). Another distinctive feature of plant RNA-directed DNA methylation is that cytosine methylation is the initial epigenetic mark as sequences as short as 30bp can be methylated (Pellisier et.al., 2000). Furthermore, cytosines in all sequence variants become modified (CG, CNG and CNN where N is A, T or C).

There are several conserved components of the RNA-directed DNA methylation in plants. First, DNA methylation requires at least two DNA methylatransferases, MET1 and the plant-specific CMT3, which are necessary to maintain continuous CG and CNG methylation during DNA replication (Matzke M., 2005; Chan S.W., 2005). In addition, DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2), plays an important role in *de novo* methylation (Cao et.al., 2003). Histone deacetylases such as SUVH4/KRYPTONITE (KYP) are also necessary for the enhancement of DNA-methylation and the reinforcement of CG and CNG methylation induced by double-stranded RNA (Aufsatz et.al., 2002). Furthermore, RNA-directed DNA methylation is dependent upon core RNAi proteins that generate and process

small RNA's in the RNA-methylation pathway. RNA DEPENDENT RNA POLYMERASE 2 (RDR2), DICER-LIKE 3 (DCL3) and ARGONAUTE 4 (AGO4) play a crucial role in small RNA biogenesis and target modification. RDR2 processes various single stranded siRNA precursors into double-stranded RNA precursors, from which the final siRNA molecules are produced. DCL3 catalyzes the enzymatic cleavage of the long dsRNA precursors into 21-24bp duplex siRNAs. The siRNA is finally loaded onto AGO4 which recruits the methylation machinery at the target chromatin site and mediates the DNA methylation (Qi et.al., 2006).

Small RNA-mediated chromatin modification has been shown as an advantageous mechanism for protecting plant genomes from accumulation of transgenerational loss of correct DNA methylation patterns (Teixeira et.al., 2009). In *Arabidopsis*, there exists a complex network of proteins that coordinates the interplay between *de novo* cytosine methylation and "maintenance" of a preexisting methylation pattern. The RNAi machinery plays a crucial role in restoring WT-methylation in *Arabidopsis* plants following loss of methylation in subsequent generations, suggesting that siRNA-mediated DNA methylation is an effective mechanism to prevent genome instability due to reactivation of naturally silenced gene sequences and transposable elements. siRNAs may direct inheritance of chromatin states and their associated gene-expression patterns, but the expression of specific siRNAs can also be a subject to epigenetic regulation . Studies in *Arabidopsis* show that in developing seeds there exists a novel class of maternally expressed small RNAs, called p4-siRNAs whose expression is

dependant upon epigenetic marks initiated in the maternal gametophyte and carried over to the endosperm of the fertilized seeds (Mosher et.al., 2009). Expression of p4-siRNAs in the developing endosperm are not produced by the maternally derived seed coat, but rather their expression is specified by the maternal chromosomes. Although the exact mechanism underlying the uniparental expression of p4-siRNAs is unclear, the study shows that it is due to an effect at the maternal loci that specify siRNAs and not at loci encoding proteins in the siRNA pathways.

Understanding various small RNA -mediated gene methylation may provide further insights into how epigenetic inheritance and genomic imprinting occur. Studies in *S.pombe* and plant epigenetics propose that siRNAs corresponding to a specific chromatin domain may act to recruit H3K9 methylation machinery to that domain and thus ensure maintenance of parental histone methylation pattern during cell division (Moazed, 2009).

Deep sequencing of the *Arabidopsis* genome has revealed that it contains large and diverse populations of small RNAs, including siRNAs and microRNAs. A new frontier for research would be the discovery of intronic small RNAs in the plant genome and revealing their regulatory role in the cellular processes. Intronic small RNA molecules might play a central role in more than one gene-silencing pathways both at the transcriptional and post-transcriptional level. Plant intronic siRNAs or microRNAs may potentially direct DNA methylation or histone modifications to homologous DNA sequences. Promoter-directed double-stranded RNAs have already been shown to induce methylation and silencing of a

transgene promoter in tobacco, as well as pigmentation genes in petunia (Mette et.al., 2000; Sijen et.al., 2001). In the same manner constructs containing promoter-driven introns, containing small RNA molecules, can potentially be exploited in a unique fashion to generate transgenic *Arabidopsis* and crop plants carrying silenced genes of interest without necessarily being mutants.

Chapter 2: Regulation of *miR172* by *APETALA2* requires the LEUNIG and SEUSS co-repressors and contributes to the A and C antagonism in flowers

2.1 Abstract

Central to the ABCE model of flower development is the antagonistic interactions between class A and class C activities that are responsible for restricting perianth organ development to the outer two whorls and sexual organ development to the inner two whorls. The molecular mechanisms underlying the A-C antagonism are not completely understood. In *Arabidopsis thaliana*, *miR172* is expressed largely in the inner two whorls and down-regulates class A gene *APETALA2 (AP2*). However, what controls this predominantly inner whorl-specific expression of *miR172* is not known. We show that the LEUNIG (LUG) and SEUSS (SEU) co-repressors in flowers negatively and directly regulate two miR172 genes. The recruitment of LUG/SEU to the *miR172* promoters is dependent on AP2, suggesting that *AP2* positively autoregulates in the outer floral whorls by repressing the expression of *miR172.* Such mutual inhibition between *miR172* and *AP2* underlies the A-C antagonism. Further*, lug* loss-of-function mutations cause a dramatic decrease in the transcript level of *AGO1*, an essential component of the RISC complex. Restoring *AGO1* transcript level in *lug* mutants significantly enhances *lug-3* flower and leaf phenotypes, indicating that mis-regulated microRNAs in *lug-3* are exerting greater negative effects on their target

genes when *AGO1* expression level is restored in *lug-3*. This highlights that the importance of *LUG* in the transcriptional regulation of the microRNA pathway is not limited to *miR172*. Together, this study provides novel insights into the molecular mechanisms underlying the A-C antagonism and sheds light on the transcriptional regulation of two important regulatory molecules, *AGO1* and *miR172*.

2.2 Introduction

Although flowers appear in numerous varieties of colors, shapes and forms, floral development in angiosperms is based on variations of the same fundamental pattern. In *Arabidopsis thaliana*, the flowers originate as a small group of undifferentiated cells, called the floral meristem. The floral meristem eventually gives rise to four concentric whorls of floral organs, each whorl being a separate domain that produces a single type of floral organs. The outer two whorls consist of sterile organs, where the sepals occupy the outermost whorl and the petals take up the second whorl. The third and fourth whorls contain the reproductive organs- six stamens and two fused carpels, respectively.

A large number of genes have been reported to regulate the initiation and development of flowers. Genetic analyses of floral homeotic mutants have led to the ABCE model of flower development, which has become a milestone in the understanding of floral organ identity specification, floral patterning, and floral pattern evolution (Coen and Meyerowitz, 1991; Theissen and Saedler, 2001;

Krizek and Fletcher, 2005). The ABCE model successfully explains how floral organ identity is established through the combinatorial action of A, B, C, and E classes of genes, all of which encode DNA-binding transcription factors. Central to the ABCE model is the mutual inhibition, or the antagonistic interaction between the A and C class genes. The repression of A by C in the inner two whorls is responsible for restricting A gene activities to the outer two whorls to specify sepal and petal identity, while the repression of C by A in the outer two whorls limits C function to the inner two whorls to specify stamen and carpel development (Coen and Meyerowitz, 1991; Bowman et.al., 1991; Drews et.al., 1991).

What are the molecular mechanisms underlying this antagonism between A and C? The class C gene *AGAMOUS (AG)* was found to be expressed in all four whorls in the strong class A mutant, *apetala2* (*ap2*), leading to reproductive organs in place of perianth organs in the outer two whorls, thus implicating *AP2* as a repressor of *AG* transcription (Bowman et.al., 1991; Drews et.al., 1991). However, the molecular mechanism of how AP2 represses *AG* is not yet demonstrated. In addition to *ap2* mutants, *AG* is ectopically expressed in the outer two whorls of *leunig (lug)* and *seuss (seu)* mutants (Liu and Meyerowitz, 1995; Franks et.al., 2002). LUG is a transcriptional co-repressor homologous to the Gro/Tup1 type family of transcriptional co-repressors in animals and fungi (Conner and Liu, 2000), while SEU encodes a plant specific transcription co-regulator with a conserved dimerization domain also found in animals (Franks

et.al., 2002). Neither LUG nor SEU contains a known DNA-binding domain. LUG interacts with SEU both genetically and physically (Franks et.al., 2002; Sridhar et.al., 2004) and SEU bridges the interaction between LUG and DNA-binding factors. These DNA-binding factors specify the regulatory targets of LUG/SEU. The class A ternary complex components APETALA1 (AP1) and SEPALLATA3 (SEP3) were shown to recruit LUG/SEU to repress the expression of class C gene *AG* by directly interacting with SEU (Sridhar et.al., 2006; Gregis et.al., 2006; Gregis et.al., 2009). Therefore, while LUG and SEU are not technically class A genes, they are integral to the A to C antagonism and essential for restricting the expression of *AG* to the inner two whorls.

How do the class C genes in turn antagonize class A genes? In *ag* mutants, petals develop in place of stamens and a new flower forms in place of carpels (Bowman et.al., 1991), indicating expansion of class A activities into inner whorls. Indeed, class A gene *AP1* mRNA was detected in inner two whorls of *ag* mutants (Mandel. et.al., 1992; Gustafson-Brown, 1994). The second class A gene *AP2* is unique among ABCE genes in that it does not encode a MADS box protein and its mRNA is detected in all floral whorls despite its class A function (Jofuku et.al., 1994). A sequence located near the 3' end of the *AP2* coding region is complementary to a microRNA, *miR172*, which cleaves *AP2* mRNA as well as inhibits *AP2* translation (Chen 2004; Aukerman and Sakai, 2003). *35S::AP2m* transgenic plants expressing an *AP2* mutated in *miR172* binding sites and immune to *miR172* regulation developed flowers that resemble *ag* loss-of-function mutants (Chen 2004; Scwab et.al., 2005), probably due to ectopic AP2 activity in the inner

two whorls. *In situ* hybridization revealed that *miR172* is expressed at highest levels in inner two whorls of wild type flowers at stage 7, but what determines this spatial and temporal regulation of *miR172* remains unknown.

MicroRNAs regulate their target mRNA via the RISC (RNA-induced silencing complex). An important component of the RISC complex is the slicer encoded by the *ARGONAUTE (AGO)* family of genes (Mallory et.al., 2008; Baumberger and Baulcombe, 2005). In *Arabidopsis*, ten *AGO* genes have been identified. AGO1 catalyzes broad miRNA- and siRNA-guided mRNA cleavage and translation inhibition (Baumberger and Baulcombe, 2005). *AGO1* mRNA itself is the target of a miRNA *miR168,* constituting a homeostatic *AGO1* regulatory loop (Vaucheret et.al., 2006; Vaucheret et.al., 2004; Mallory and Vaucheret, 2009). *ago1* null mutants exhibited severe developmental defects, including abnormal inflorescences of infertile flowers with filamentous organs, a lack of axillary and shoot apical meristems, and narrow pointy or filamentous leaves (Bohmert et.al., 1998; Kidner and Martienssen, 2005). Interestingly, *lug* mutants exhibited many of the similar developmental phenotypes as *ago1* mutants, albeit to a lesser degree of severity. This phenotypic similarity prompted us to investigate the possibility of *LUG* in the regulation of *AGO1*.

We showed that *LUG* regulates *AGO1* positively and thus is a general regulator of the microRNA pathway. Further, LUG directly and negatively regulates *miR172* in the outer two floral whorls. This direct repression of *miR172*

by LUG also requires SEU and AP2, suggesting that *AP2* may recruit the LUG/SEU co-repressor to repress *miR172* in the outer two whorls of a flower. While much has been learned about microRNA biogenesis, less is known about the transcriptional regulation of miRNAs. Our study provides important insights into *microRNA* regulation and reveals a positive feedback loop, where *AP2* maintains its own activity by negatively regulating the expression of its cognate microRNA, providing novel insights into the molecular mechanisms underlying A-C antagonism.

2.3 Materials and Methods

Plant materials, growth conditions and transformation

Arabidopsis thaliana plants were grown on Metromix soil (Griffin) at 22°C under 16h light/8h dark conditions. *lug-3, seu-1,* and *ap2-2* mutants and the transgenic lines, *35S:: AGO1* in WT, *35S:: AGO1* in *lug-3, 35S::GFP-LUG* in *lug-16*, *pSEU:: GFP-SEU* in WT, and *pSEU::GFP-SEU* in *ap2-2* are all in Landsberg erecta (Ler) background. *pSEU::GFP-SEU; seu-1*, a gift from Robert Franks (31), was crossed into *ap2-2+/+ant-9* plants to generate *pSEU::GFP-SEU* in WT and *pSEU::GFP-SEU* in *ap2-2*.

To construct *35S:: AGO1*, an *AGO1* cDNA clone was obtained from ABRC (C105223). Primers 5'ATGGTGAGAAAGAGAAGAACG3' and 5'TCAGCAGTAGAACATGACACG 3' were used to amplify *AGO1* cDNA, which was cloned into pCR8/GW/TOPO using TA cloning kit (Invitrogen, Carlsbad, CA, USA) and then recombined into the pEarleyGate100 plant

transformation vector (Earley et.al., 2006) using the Gateway® technology (Invitrogen, Carlsbad, CA, USA). The construct was introduced into *Agrobacterium tumefaciens* GV3101 that was used to transform *Arabidopsis thaliana* Ler and *lug-3* plants via floral dip. T1 transgenic plants were selected on soil using BASTA.

For constructing *35S:: GFP-LUG, p*AVA393 containing full length LUG fused to the C-terminal end of GFP (8) was cut with HindIII and SacI. The fragment was cloned into HindIII and SacI in pCAMBIA2300 vector (Cambia). The final construct was introduced into *Agrobacterium tumefaciens* GV3101 and transformed into *lug-16* plants via floral dip. Transgenic plants were selected on kanamycin (50 mg/ml) plates and analyzed in T2 for the presence of GFP by PCR. 40 T1 plants were initially recovered after selection and 11 lines with a near wild-type phenotype were further analyzed.

For constructing *mir172cPromoter::GUS* we amplified 1kb upstream of the *miR172c* coding sequence using the following end primers: F 5' GAGCTGAACAGAGTGGAA 3', R 5' GGTTGATGATAGGGATGTAT 3'. To construct *mir172c*^{m}*Promoter::GUS,* the forward and reverse central primers used here contain mutated second AP2 binding site and were paired with end primers described above to generate 5' and 3' promoter PCR fragments, which were mixed and served as the template for a third PCR using the end F and R primers. The mutated central primers are F 5'

ATCGAAGAAAAATGAAAAATTGGGCTTTAAAG; 3' R 5' CTTTTAAGCCCAATTTTTCATTTTTCTTCGAT 3'. The promoter fragments were cloned into pMDC164 gateway system (Stock CD3-762) (Curtis and Grossniklaus, 2003). 34 *mir172cPromoter::GUS (WT),* 12 *mir172cPromoter::GUS (lug-3),* 21 *mir172c^{m}Promoter::GUS (WT)* and 10 *mir172c^{m}Promoter::GUS (lug-3)* transgenic lines were obtained. 10 of each kind of transgenic flowers were analyzed for *GUS* expression.

Northern blot analysis

Total RNA was isolated from inflorescences of wild type, *lug-3*, and *35S::AGO1* transgenic plants using TRI® Reagent (Sigma, St. Louis, MO, USA). For small RNA blots, a 30μg total RNA was separated on a 15% acrylamide gel and then transferred to the Hybond-N nylon membrane (Amersham, Little Chalfont, Buckinghamshire, UK) by electroblotting. RNA was cross-linked onto the membrane using EDC (N-(3-Dimethylamionopropyl)-N'-ethylcarbodiimide hydrochloride) (Sigma, St. Louis, MO, USA) according to a published protocol (45). P^{32}-ATP and mirVANA™ Probe & Marker Kit were used to label the oligo probe according to manufacturer's instructions (Ambion Inc., Austin, TX, USA). Probe sequences for *miR172, miR168, miR165,* and *miR166* are listed in Table S3. The small RNA blot was hybridized and washed as previously described (Lee et.al.,2001).

For Northern blots, a PCR fragment of the first exon of *AP2* was labeled

with alpha P^{32}-dCTP (Perkin-Elmer) with Ready-To-Go DNA Labeling Beads according to the manufacturer's instruction (Amersham, Little Chalfont, Buckinghamshire, UK). The *AGO1* and 5SRNA probes were similarly made. Sequences of primers used to amplify *AP2, AGO1* and *5SRNA* are provided in Table S3. A 15μg total RNA was separated on a 1% agarose gel, transferred onto BrightStar-Plus membrane, hybridized and washed using the Northern Max-Gly kit according to the accompanying manual (Ambion Inc., Austin, TX, USA).

Chromatin immunoprecipitation

ChIP procedure is essentially the same as previously described (Sridhar et.al., 2006) except for a few changes described below. Nuclear extracts were prepared from about 1 gram inflorescence tissues (with flowers older than stage 12 removed) using MC, M1, M2 and M3 buffers (Ito et.al., 1997). Following sonication, 40μl of Protein A Dynabeads (Dynal) was added to the chromatin solution and incubated for 1 hour at 4 °C to eliminate nonspecific binding. After removing the Dynabeads, 5μl of a polyclonal anti-GFP antibody (catalog number AB290-50; Abcam, Cambridge, MA, USA) was added to the purified chromatin and then incubated overnight at 4 °C. 40μl fresh Protein A Dynabeads was added to the antibody-chromatin solution, and incubated for 1 hr at 4 °C. The bound chromatin was then washed and eluted as previously described (Sridhar et.al., 2006).

Semi-quantitative PCR was first used to scan larger and more putative AP2-binding sites of *miR172c, miR172e,* and *AGO1* promoters for possible enrichment. The primer sequences and results are summarized in Table S2. Subsequently, qPCR was used to quantify fold enrichment reported in Fig. 3. The AmpliFX program (http://ifrjr.nord.univ-mrs.fr/AmplifX-Home-page?lang=en) was used for the primer design to give amplicon size equal or less than 150bp. Percent efficiency for each primer is shown in Table S1. From 100 ml resuspended ChIP pellets, the input was diluted 1:100 and immunoprecipitated chromatin was diluted 1:10. 1ml of each dilution was used in 25 ml qPCR reactions with the QuantiTect SYBR Green PCR kit (Qiagen, Valencia, CA, USA) and run on BioRad CFX96 machine. PCR program consists of 35 cycles at 95°C for 15m, 56°C for 45s, 72°C for 45s.

To calculate fold enrichment, Mukhopadhyay et al. 2008 (Mukhopahhyay et.al., 2008) and the BioRad Real Time PCR Application Guide were followed. For each primer pair, Ct value for +Antibody (+AB) and for –AB is subtracted from the Ct value of input (reference Ct) to yield ΔCT. Then ΔCT^{+AB} is subtracted from ΔCT^{-AB} to yield ΔΔCT for each sample. Each ΔΔCT number is entered into the Pfaffl formula ($2^{-\Delta\Delta Ct}$) to yield "fold enrichment". Error bar in Fig. 2.3A represents standard deviation of two biological replicates (each has three technical replicates). Fig 2.3B is derived from three technical replicates of one biological experiment using similar normalization and calculation methods.

In-situ hybridization

A direct tandem oligonucleotide concatamer (4X) of the sense *miR172* strand, 5' AGAATCTTGATGATGCTGTAG 3', containing a T7 RNA polymerase-binding site, 5' TAATACGACTCACTATAGGGAGA 3', at the 3' end was synthesized (Bioneer) and serves as the template for T7 labeling reactions. T7 RNA polymerase (New England Biolabs, Ipswich, MA, USA) and the digoxygenin (DIG) RNA Labeling mix (Roche, USA, catalog #11277073910) were used to make DIG-labeled antisense *miR172* RNA probe. 500 pmol of the template was used per labeling reaction. Procedures for tissue fixation, embedding, section, *in situ* hybridization, and detection were performed as described previously (Carr and Irish, 1997) with the following modifications. First, hybridization was carried out at 42°C overnight. For each slide, 1 ml probe (from the 50 ml T7 transcription reaction) was mixed with 19 ml 50% Formamide, heated to 80 °C, and then mixed with 80 ml 1X hybridization solution containing 1x Salt, 50% Formamide, 10% Dextran Sulfate, 1X Denhardt's to yield 100 ml hybridization solution per slide. 10X salt contains 3M NaCl, 0.1 M Tris buffer PH 6.5, 0.1 M Sodium Phosphate buffer PH. 6.8, and 50 mM EDTA. After hybridization the slides were washed with 0.2X SSC twice at 40°C for 30 minutes each wash.

Yeast two-hybrid assay

Clontech Matchmaker system was used. LUH (Sitaraman et.al., 2008), ANT (Krizek and Sulli, 2006), and AP2*delta* (residues 124-394) are each fused in frame to the Activation Domain (AD) in pGAD424. Both ANT-pGAD424 and

AP2delta-pGAD424 are gifts from Beth Krizek. SEU-BD in pGBT9 contains a truncated SEU (residues 1-563) with its C-terminal self-activating domain removed (Sridhar et.al., 2006). Full length AP2 was excised from pGG30 (a gift from Detlef Weigel) with NcoI and EcoRI and inserted at the same restriction sites in pGADT7 to create AP2-AD. Constructs were introduced one at a time into *Saccharomyces* strain PJ694A according to the Yeast Protocols Handbook (Clontech). Media testing for interaction was –Leu, -Trp, - His, - Ade, and 3-Amino-1,2,4-triazole at 0.126g/500ml.

Bimolecular fluorescence complementation

Full length cDNAs of AP2 and SEU were PCR amplified with following 5' and 3' primers containing SalI and XmaI sites, respectively, and cloned into pGEM-T (Promega). AP2: 5' AT GTC GAC ATG TGG GAT CTA AAC GAC GCA 3' and 5' CCC GGG TCC AGA AGG TCT CAT GAG AGG AG 3'. SEU: 5' AT GTC GAC ATG GTA CCA TCA GAG CCG CCT AAT 3' and 5' CCC GGG TCC CGC GTT CCA ATC AAA ATT 3'. LUH was PCR amplified with SalI and XmaI restriction sites in the forward and reverse primers, respectively, and then cloned using the TOPO TA cloning kit (Invitrogen). LUH: 5' AT GTC GAC ATG GCT CAG AGT AAT TGG GAA 3' and 5' CCC GGG CTT CCA AAT CTT TAC GGA TTT GT 3'. AP2, SEU, and LUH were then excised with SalI and XmaI from pGEN-T or pTOPO generated above and cloned into pSPYNE (for fusion to the N-terminal fragment of YFP) and pSPYCE (for fusion to the C-terminal fragment of YFP) (33) at the SalI and XmaI restriction sites.

The above BiFC plasmids were bombarded into onion epidermal cells using the Helios Gene Gun (BioRad) system according to published procedures (Hollender and Liu, 2010). Respective recombinant pSPYNE and pSPYCE plasmids were mixed in equal quantity and loaded into cartridges of the gene gun. SEU-N was tested against AP2-C, and SEU-C was tested against AP2-N, yielding similar results that were observed under a Zeiss inverted fluorescent microscope

2.4 Results

ARGONAUTE1 (AGO1) expression is reduced in lug-3 mutant flowers

As *ago1* mutants (Bohmert et.al., 1998; Kidner and Martienssen, 2005) and *lug-3* (Liu and Meyerowitz, 1995; Conner and Liu, 2000) mutant flowers similarly exhibit split carpels, narrow pointy petals and leaves, and filamentous floral organs and leaves, we examined *AGO1* expression in a microarray data set (Gonzalez et.al., 2007) (http://affymetrix.arabidopsis.info/narrays/experimentbrowse.pl; NASCARRAYS-327) that compares gene expression between wild type (WT) and *lug-3*. *AGO1* transcript level is reduced in *lug-3* flowers compared to wild-type flowers, suggesting that *AGO1* transcription and accumulation may require wild type *LUG*. Northern analysis confirmed that *AGO1* mRNA decreases significantly in *lug-3* flowers (Fig. 2.1A), suggesting that certain aspects of the *lug-3* phenotype may be mediated by a reduction of *AGO1* in *lug-3*.

To determine which of the *lug-3* phenotypes are mediated by a reduction of *AGO1*, we restored *AGO1* mRNA levels in *lug-3* mutants by introducing a *35S::AGO1* construct into *lug-3*. The same construct was also transformed into wild type as a control. Since *AGO1* transcript level is also regulated by a microRNA, *miR168* (Vaucheret et.al., 2004), and *miR168* level remains unchanged in *lug-3* (Fig. 2.1B), *35S::AGO1* will likely lead to normal levels of *AGO1* mRNA in *lug-3* due to *miR168*-mediated post-transcriptional regulation. This is confirmed by Northern blots showing that *35S::AGO1* (WT) and *35S::AGO1 (lug-3)* flowers possess *AGO1* mRNA at a level similar to non-transgenic wild type plants (Fig. 2.1A). This indicates that the 35S promoter driven *AGO1* can bypass the requirement of *LUG* for proper *AGO1* expression and that *AGO1* is regulated by *LUG* at transcriptional rather than post-transcriptional level.

Ninety-six independent T1 transgenic lines of *35S::AGO1* in wild type background were generated and analyzed, all of which are phenotypically wild type except for a subtle increase in floral bud number per inflorescences (Fig. 2.1C, G) and slightly irregular phyllotaxis. The lack of any prominent gain-of-function phenotype in these *35S::AGO* (WT) transgenic plants is consistent with the wild type level of *AGO1* mRNA in these transgenic plants (Fig. 2.1A).

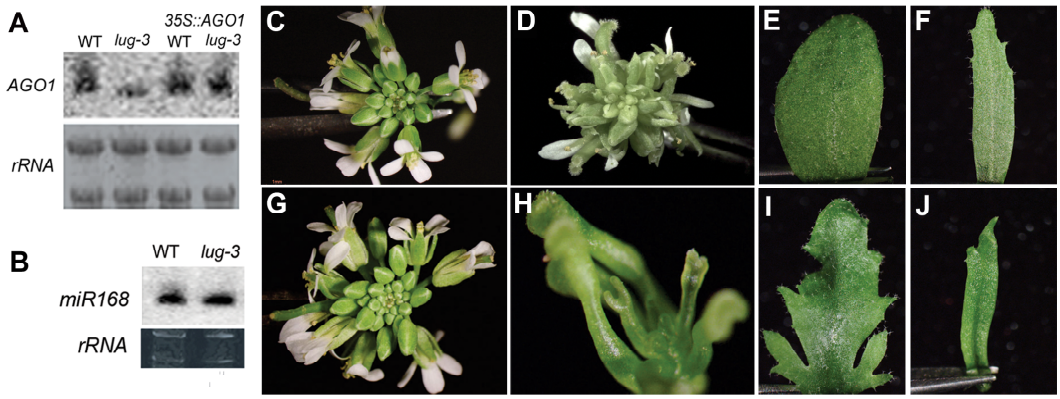

Fig. 2.1: *LUG* is required for wild type level *AGO1* expression.
(A). Northern analysis of *AGO1* mRNA in inflorescence tissues. The rRNA bands served as the loading control. (B). A small RNA blot of *miR168* comparing wild type (WT) and *lug-3* mutant inflorescences. The gel photo shows *rRNA* bands that serve as the loading control. (C). A wild type inflorescence. (D). A *lug-3* inflorescence. (E). A *lug-3* rosette leaf. (F) A *lug-3* cauline leaf. (G). An inflorescence of a *35S::AGO1* transgenic plant showing wild type phenotype except for a slight increase in the number of flower bud. (H). Shoot tip of what appeared to be an inflorescence of a *35S::AGO1; lug-3* transgenic plant, showing about five flowers each consisting of a single gynoecium. (I). A *35S::AGO1; lug-3* rosette leaf. (J). A *35S::AGO1; lug-3* cauline leaf.

The restoration of *AGO1* mRNA to wild type level in *lug-3* mutants should rescue some of the *lug-3* defects, allowing us to distinguish *lug-3* phenotypes mediated by a reduction of *AGO1* from those that are *AGO1*-independent. Forty-three independent T1 lines of *35S:: AGO1* in the *lug-3* background were generated and analyzed. Unexpectedly, these *35S::AGO1; lug-3* plants show an enhanced rather than rescued phenotype in leaves and flowers (compare Fig. 2.1D-F with 2.1H-J), and the plants are small and short in size. The extent of enhancement varies among different *35S::AGO1; lug-3* transgenic lines. An example is shown in Fig. 2.1H, where an inflorescence is comprised of several abnormal gynoecia. The majority of *35S::AGO1; lug-3* lines develop leaves that are either serrated (Fig. 2.1E, I) or narrow (Fig. 2.1F, J). The fertility defect of

lug-3 is also enhanced. *lug-3* normally gives rise to a few seeds per silique; the *35S::AGO1; lug-3* plants are completely sterile, limiting our ability to analyze T2 generation.

miR172 and AP2 expression are altered in lug-3 and 35S::AGO1; lug-3 flowers

What might underlie the dramatic phenotypic differences between *35S::AGO1* in wild type vs. *lug-3* background? One interpretation is that *lug-3* may cause several microRNAs to be expressed at a higher than normal level, but may have little effect on respective microRNA targets due to a simultaneous reduction of *AGO1* in *lug-3*. When *AGO1* expression is restored to wild type level in *lug-3*, the increased microRNAs in *lug-3* are now better able to downregulate their target mRNAs.

The severe carpelloidy and reduction of floral organ numbers observed in flowers of *35S::AGO1; lug-3* plants resemble flowers of strong *ap2* mutants (Bowman et.al., 1991). The *ap2*-like flower phenotype of *35S::AGO1; lug-3* could be caused by a decreased *AP2* function due to an increase in *miR172* that targets *AP2* mRNA for degradation or translational inhibition. To test this possibility, a small RNA Northern was used to analyze *miR172* expression. It revealed significantly increased *miR172* levels in *lug-3* flowers and equally increased *miR172* levels in *35S::AGO1; lug-3* flowers (Fig. 2.2A). The probe used does not distinguish among different *miR172* genes; at least five of which have been described in *Arabidopsis*. Northern blots were used to investigate the

effect of increased *miR172* on *AP2* transcript levels. *AP2* mRNA decreases slightly in *lug-3* but more significantly in *35S::AGO; lug-3* (Fig. 2.2B). Therefore, restoring *AGO1* levels in *lug-3* may lead to a more efficient degradation of *AP2* transcripts and thus a more severe floral phenotype (Fig. 2.1D, H).

Ectopic miR172 expression in lug-3 and seu-1 flowers

miR172 was previously shown to be expressed predominantly in the inner two floral whorls at stage 7 flowers (Chen 2004). Based on the A-C antagonism, an expansion of *miR172* expression to the outer two floral whorls could result in repression of *AP2* in the outer two whorls, leading to carpelloid sepals and loss of petals and stamens, a phenotype observed in *lug-3* and *35S::AGO1; lug-3* plants. *In situ* hybridization with a *miR172* antisense RNA probe was used to examine the *miR172* expression domain in *lug-3* as well as *seu-1* mutants for *SEU* is known to be a partner of *LUG* (Sridhar et.al., 2006). At early floral stages (stage 1-4) in wild type, *miR172* expression appears in both the sepals as well as the inner whorl primordia (Fig. 2. 2C). Soon afterwards, *miR172* RNA starts to abate in the abaxial side of the sepal. At stage 7 or later of wild type flowers, *miR172* RNA is absent from the sepal and petal primordia (Fig. 2.2C). In *lug-3* and *seu-1* flowers, while early stage expression of *miR172* appears similar to wild type, later stage expression appears altered. *miR172* is detected in the sepals and petals of *lug-3* and *seu-1* mutant flowers at stage 7 or later (Fig. 2.2D, E). Therefore, *LUG* and *SEU* are both needed to repress *miR172* in outer two floral whorls at stage 7 or later.

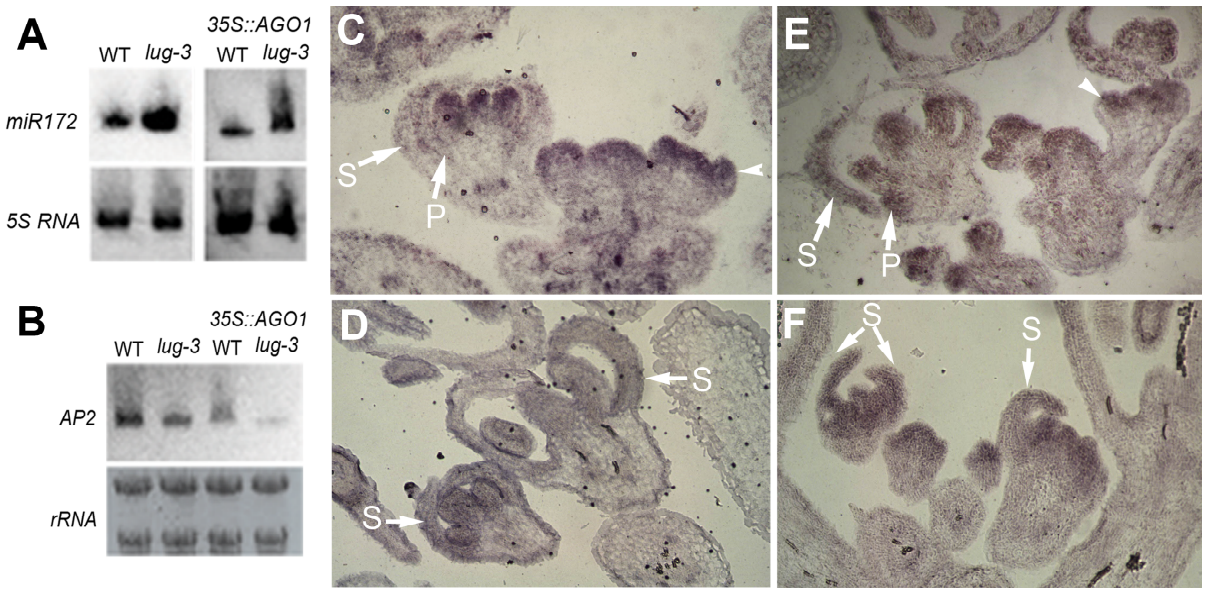

Fig. 2.2: Increased and ectopic *miR172* expression in *lug-3*, *seu-1,* and *ap2-2* mutants revealed by RNA blots (A-B) and *in situ* hybridization (C-F). (A). A small RNA blot showing *miR172* RNA levels in inflorescences. 5SRNA served as the loading control. (B). A Northern blot showing *AP2* mRNA levels. The same Northern blot for *AGO1* (Fig. 1A) was stripped and probed with an *AP2* probe shown here. rRNA bands served as the loading control. (C) *in situ* hybridization showing *miR172* RNA expression in wild type flowers. *miR172* signal is absent from the sepal (S) and petal (P) of a stage 7-8 flower (left). *miR172* signal is detected in the entire floral primordium of a stage 4 flower (right) but appears to start to abate in the abaxial side of the sepal (arrowhead). (D). *miR172* expression in *lug-3* flowers. Ectopic *miR172* RNA is detected in sepals (S) of a stage 6 and a stage 10 flower. (E) *miR172* RNA expression in *seu-1* flowers. Ectopic *miR172* expression is detected in the sepal (S) and petal (P) of a stage 7-8 flower (left). Also shown is a stage 5 flower, which is similar to wild type in that the *miR172* signal is absent in the abaxial side of the sepal (arrowhead). (F) *miR172* expression in *ap2-2*. Ectopic *miR172* RNA expression is detected in sepals (S) of stage 6 and older flowers.

LUG directly regulates miR172

The regulation of *miR172* and *AGO1* by *LUG* could be either direct or indirect. For example, the increased and ectopic *miR172* expression observed above could result from ectopic carpelloid organ development in the outer floral whorls of *lug-3* because carpels are known to express *miR172*. This indirect scenario is however not supported when *miR172* levels were shown similarly elevated in *lug-3* and *35S::AGO1; lug-3* (Fig. 2.2A), the later of which shows more carpelloidy.

In order to determine if LUG directly associates and thus regulates *miR172*, we conducted chromatin immunoprecipitation (ChIP) experiments. A *35S::GFP-LUG* construct, shown to rescue *lug-16* mutants, was used. Nuclear extracts of inflorescence tissues from *35S::LUG-GFP; lug-16* transgenic plants that are phenotypically wild type and negative control plants (non-transgenic wild type) were immunoprecipitated with an anti-GFP antibody. The immunoprecipitated DNA was quantified with quantitative real time PCR (qPCR).

We searched *miR172c* and *miR172e* promoters for conserved binding sites for class A genes (ie. MADS box and AP2). No MADS box binding site (CArG box) was found in the promoters of *miR172c* or *miR172*e.

Gene Name	qPCR primer sequence	AP2 binding site∗	Enrichment (LUG-GFP) (*lug-16)*	Enrichment (SEU-GFP) (*WT)*	Slope	Primer Percent Efficiency
miR172c-F1	gagattacgagaatccgcactca	1.6kb	No	No	-3.246	103.2
miR172c-R1	ggttttaggcttttagcccaagga					
miR172c-F2	ccacatgtgcccatattgat	1.4kb	Yes	Yes	-3.518	92.3
miR172c-R2	gaagatccacttttaaagcccaat					
miR172e-F	gctgtctgaatcctcttgctttcctc	118	Yes	Yes	-3.155	107.4
miR172e-R	cggtttcgaggtctaaagttgtga					
AGO1-F	ccgtaacttactctaaccacagaacct	125	No	No	-3.474	93.6
AGO1-R	aatccgtacgaaacaccaaccct					

Table 2. 1: qPCR primer sequences and summary of ChIP results
∗Approximate location of each AP2-binding site either from the start of cDNA (*miR172*) or from translational start (*AGO1*) based on sequence information from www. Arabidopsis.org.

Four putative AP2 binding sites, TTTGTT, were found in the promoter of *miR172c*, and one in the promoter of *miR172e*. The putative AP2-binding site sequence was determined by *in vitro* and *in vivo* approaches (T. Dinh and X. Chen unpublished results). qPCR primers were designed to flank these putative AP2-binding sites and generate amplicons of 150 bp or less (Table 2.1). For *miR172c*, only one of the four putative AP2-binding sites, *miR172c-2*, showed over 4-fold enrichment with the *GFP-LUG; lug-16* sample (Fig. 2.3A). The other three AP2-binding sites, including *miR172c-1*, showed no enrichment with the *LUG-GFP; lug-16* sample (Fig. 2.3A; Table 2.2). For *miR172e*, the single putative AP2-binding site was enriched over 7 fold with the *GFP-LUG* sample *(*Fig. 2.3A). Neither *miR172c* nor *miR172e* showed enrichment with the negative control

sample (WT) (Fig. 2.3A; Table 2. 2). These data suggest that AP2 or other AP2 family members might be involved in recruiting LUG/SEU to the promoters of *miR172c* and *miR172e.*

.

AGO1 promoter region contains no CArG box but possesses four putative AP2-binding sites. Semi-quantitative PCR showed no enrichment with any of the four AP2-binding sites in the 4kb promoter region of *AGO1* (Table 2.2). qPCR was performed to test the AP2-binding site closest to the transcriptional initiation, which showed no enrichment (Fig. 2.3A). Thus, *LUG* may regulate *AGO1* indirectly.

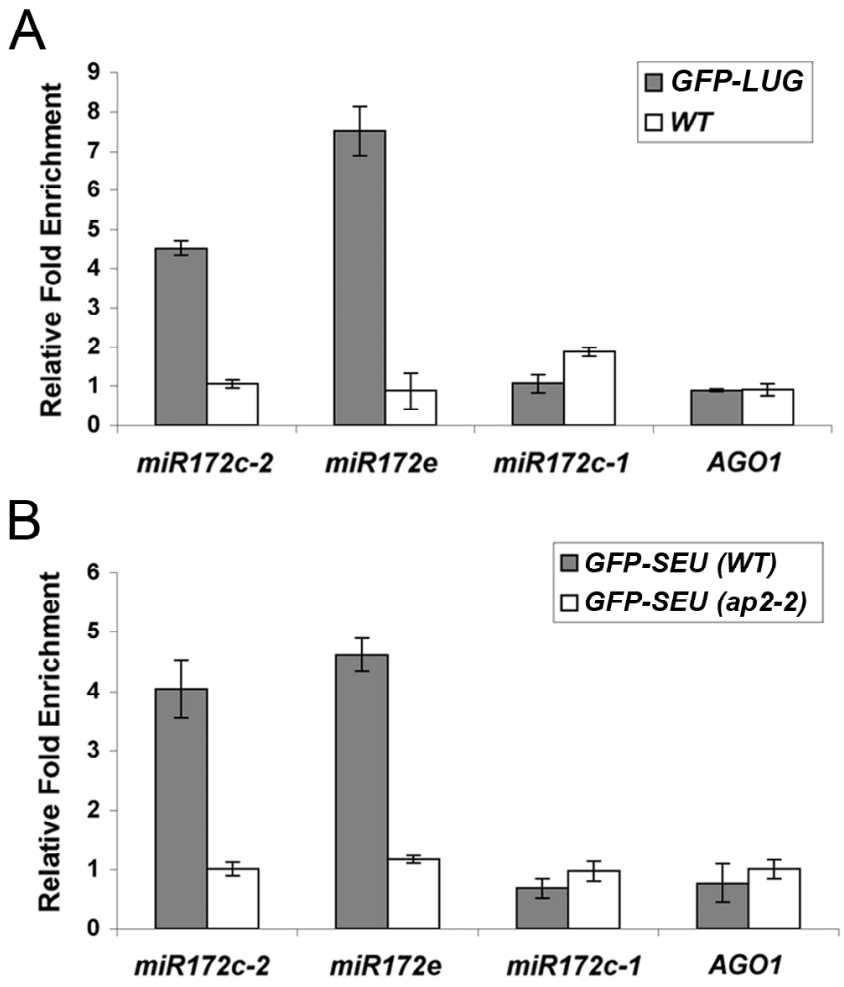

Fig. 2.3: ChIP detects *in vivo* association of GFP-LUG and GFP-SEU to *miR172* promoter sequences. (A). Anti-GFP antibody was used to immunoprecipitate chromatin from *35S::GFP-LUG; lug-16* inflorescences (Grey bar) and non-transgenic WT inflorescences (white bars). Fold enrichment of immunoprecipitated chromatin was quantified by qPCR with primers flanking putative AP2-binding sites at the promoters of *miR172c (miR172c-1* and *miR172c-2), miR172e* and *AGO1* (Table 2.1). Error bar represents standard deviation (s.d.) based on three technical replicates of two biological repeats. (B) Anti-GFP antibody was used to immunoprecipitate chromatin from *pSEU::GFP-SEU* in wild type background (Grey bar) and *pSEU::GFP-SEU* in *ap2-2* background (white bars). Fold enrichment for the same primer sets described in (A) is shown. Standard deviation is based on three technical replicates of one biological experiment.

Gene Name	PCR primer sequence	AP2 binding site*	Enrichment (LUG-GFP) (lug-16)	Enrichment (SEU-GFP) (WT)
miR172c-F1	GAGATTACGAGAATCCGCACTCA	1.6kb	No	No
mir172c-R1	GGTTTTAGGCTTTTAGCCCAAGGA			
miR172c-F2	AGGATCCACATGTGCCCATA	1.4kb	Yes	Yes
miR172c-R2	CGTGAGCTATTCAGGATCACGA			
miR172c-F3	TCGCTGACTAAATAGTTGGA	366	No	No
miR172c-R3	GAGGCCAATTTTGTCATCC			
miR172c-F4	ACCTGAGTATCTGAGATCTCAGT	225	No	No
miR172c-R4	AGGGATGTATGTAGTGATTTGG			
mir172e-F	AGCCTTTGGCTTCTGTTCCTGA	118	Yes	Yes
miR172e-R	CGGTTTCGAGGTCTAAAGTTGGGA			
AGO1-F1	AGCATTCCAAAGACCCAACCCTG	125	No	No
AGO1-R1	CTGAAAAGCAATCTCTCGCCGGAG			
AGO1-F2	TCTAAACGTTTTCCCCAAGATTTA	1.2kb	No	No
AGO1-R2	TTCACTAGTTGATTTGCTTCTGT			
AGO1-F3	ACGGCTTCCTACAATATCTCG	1.8kb	No	No
AGO1-R3	GGTGCGGTCTTGACGAGAAACG			
AGO1-F4	CGTAAAATAACTGAGTCATAACC	3.2kb	No	No
AGO1-R4	CTATCTTGATTCTCATAGGTCATT			
EIF4a1-F	CGCATCCTATCGGATTGTCT	NA	No	No
EIF4a1-R	CTCAGATGATGTGCGGAGAA.			

Table 2.2: Semi-quantitative PCR primers and summary of ChIP results.

Further evidence of AP2 in mediating the repression of miR172 by LUG/SEU

If AP2 recruits LUG/SEU to repress *miR172*, *ap2* mutants should also exhibit an increased or expanded expression of *miR172*. *In situ* hybridization revealed that *miR172* is ectopically expressed in sepals of stage 7 or older *ap2-2* flowers (Fig. 2.2F).

To further test the role of AP2 in the repression of *miR172c* and *miR172e*, ChIP was performed to compare *SEU-GFP (WT)* with *SEU-GFP (ap2-2)* inflorescence tissues (Fig. 2.3B). *ap2-2* is a splice acceptor site mutation leading to a stop codon in the first AP2 domain and no AP2 protein product (Chen 2004).

The *pSEU::GFP-SEU,* previously shown to rescue the strong *seu-1* mutants (31), was introduced into *ap2-2*. The same primer pairs used in Fig. 2.3A were used for quantifying fold enrichment. *miR172c-2* and *miR172e* were enriched 4 and 4.8 fold, respectively, with the *GFP-SEU* (WT) chromatin but not with the *GFP-SEU (ap2-2)* chromatin (Fig. 2.3B). Similar to *GFP-LUG* (Fig. 2.3A), *GFP-SEU* failed to associate with *miR172c-1* or *AGO1* (Fig. 2.3B). The data supports that *AP2* is required for mediating the association of SEU with the *miR172c-2* and *miR172e* promoters. Since SEU and LUG exhibit similar specificity for associating with specific AP2-binding sites at the tested promoters (Fig. 2.3 and Table 2.2), SEU and LUG may work in the same complex to regulate *miR172*.

Additionally, we attempted to confirm that the *LUG/SEU* complex regulates *miR172* through an AP2 binding site by monitoring *GUS* expression in *miR172Promoter::GUS* transgenic plants and plants containing a mutation in the second AP2 binding site, *miR172^{m}Promoter::GUS.* 10 transgenic flowers were analyzed for each construct both in WT and *lug-3* background but no staining was observed.

SEU but not LUG directly interacts with AP2

Results from above suggest that AP2 may directly interact with SEU and/or LUG to regulate *miR172* expression in flowers. Yeast two-hybrid assay was conducted and revealed a direct protein-protein interaction between AP2 and SEU (Fig. 2.4A.2). A truncated AP2 (AP2 delta) containing two AP2 domains (residues 124-394) and missing both N- and C- terminal domains failed to interact

with SEU (Fig. 2.4A.3). Thus the two AP2 domains are not sufficient for the interaction with SEU. Another AP2 domain-containing protein AINTEGUMANTA (ANT) also failed to interact with SEU (Fig. 2.4A.4), indicating that SEU specifically interacts with AP2. In contrast to SEU, LUG failed to interact with AP2 (Fig. 2.4A.5). Since LUG is known to interact with SEU, SEU may bridge the interaction between LUG and AP2. LUH and SEU were previously shown to interact in yeast (Sitaraman et.al., 2008) and served as a positive control (Fig. 2.4A.1).

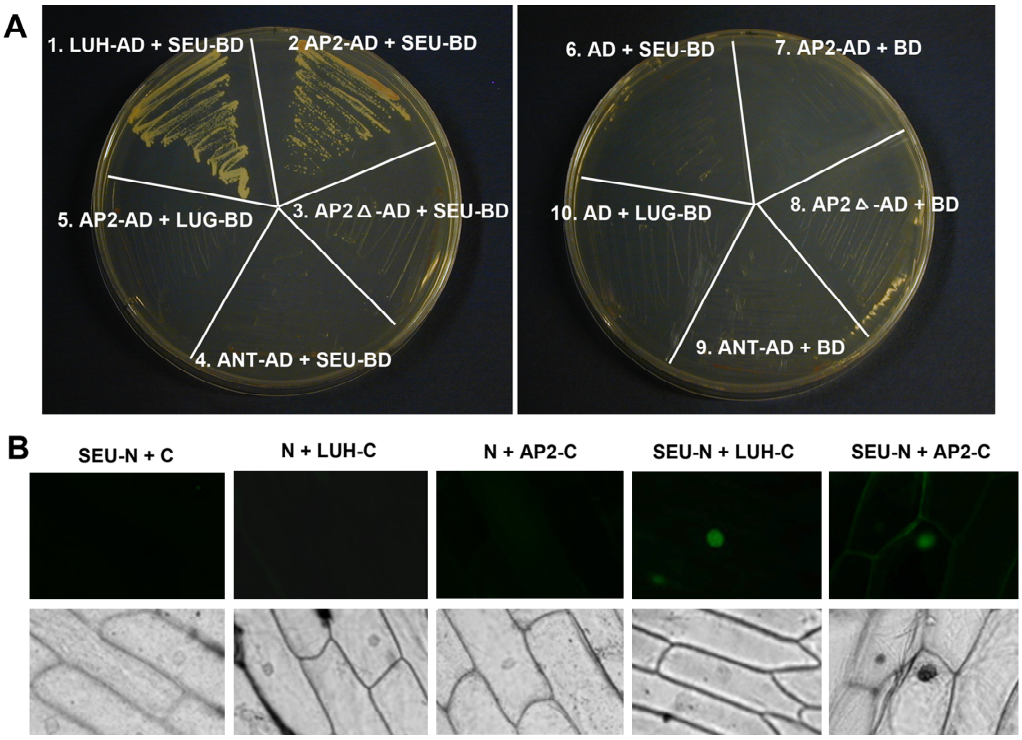

Figure 2.4: AP2 interacts with SEU but not LUG (Figure courtesy of Courtney Hollender). (A). Yeast two hybrid assay on selection media. A positive interaction between a bait protein fused to the DNA-binding domain (BD) and a prey protein fused to the Activation Domain (AD) is indicated by the formation of colony. SEU-BD contains a truncated SEU (residues 1-563) with its C-terminal domain removed to avoid self-activation (Sridhar et.al., 2006). The plate on the right shows various negative control combinations. (B). BiFC assay showing an interaction between AP2 and SEU *in planta*. Fluorescent (A-E) and bright field (F-J) images of onions epidermal peels bombarded with BiFC plasmids. N and C represent N- and C- fragment of YFP, respectively. White arrows indicate fluorescent nuclei and black arrows indicate nuclei in bright field.

The interaction between SEU and AP2 was further confirmed in onion epidermal cells via particle bombardment and Bimolecular Fluorescent Complementation or BiFC (Walter et.al., 2004). SEU and AP2 were each fused to the N-terminal and C-terminal fragment of YFP using the pSPYNE and pSPYCE vectors, respectively (Walter et.al., 2004). The fusion proteins SEU-N and AP2-C were shown to interact (Fig. 4B). SEU-N was previously shown to

interact with LUH-C via BiFC (Hollender and Liu, 2010) and served as a positive control (Fig. 4B).

2.5 Discussion

AP2 mediates the transcription repression of miR172 by LUG/SEU

miR172 regulates many important developmental processes in diverse plant species. In addition to regulating flowering, juvenile to adult phase change, and floral organ identity (Chen 2004; Aukerman and Sakai, 2003; Wu et.al., 2009; Lauter et.al., 2005; Jung et.al., 2007), *miR172* regulates sex determination in maize and tuberization in potato (Martin et.al., 2009; Chuck et.al., 2007). Understanding the regulatory mechanisms of *miR172* has profound impact for basic science as well as for agriculture. Whereas much attention has been focused on finding the targets of miRNAs and the biogenesis and metabolism of miRNAs, the transcriptional regulation of miRNAs remains largely unknown.

Our data supports that *AP2*, the *miR172* target, likely recruits the LUG/SEU co-repressor to repress *miR172* transcription in the outer two whorls of flowers. First, *lug-3, seu-1* and *ap2-2* mutants all showed similar ectopic *miR172* expression in the outer whorls of stage 7 or older flowers (Fig. 2.2). Second, anti-GFP antibodies preferentially precipitate *miR172c* and *miR172e* promoter regions that contain putative AP2-binding sites from chromatins isolated from *GFP-LUG* as well as *GFP-SEU* inflorescences (Fig. 2.3A, B). The ChIP assays revealed that LUG and SEU are associated with the same AP2-binding site (*miR172c-2*) out of

the four AP2 binding sites tested at the *miR172c* promoter and that they both enrich the only AP2-binding site at the *miR172e*. Additionally, they both fail to enrich any of the four sites at the *AGO1*. This remarkably similar specificity of SEU and LUG in their association with promoter elements supports the idea that LUG and SEU may function together as a complex to repress *miR172*.

Transcription co-repressors are recruited to target promoters by their specific association with target-specific DNA-binding transcription factors. The enrichment of specific putative AP2-binding sites at the *miR172c* and *miR172e* by GFP-SEU and GFP-LUG implicates AP2 as a candidate target-specific DNA-binding partner of LUG/SEU. We showed that GFP-SEU was unable to enrich *miR172c-2* and *miR172e* in the absence of the AP2 protein (Fig. 2.3B). Hence, AP2 is required for SEU to associate with *miR172* promoter elements. In addition, AP2 directly and specifically interacts with SEU in yeast and *in planta* (Fig. 2.4), providing further support of AP2 in recruiting SEU to the *miR172* promoters. Since SEU and LUG directly interact (Sridhar et.al., 2006), SEU may in turn recruit LUG to the *miR172* promoters. Previously, *ap2*, *lug*, and *seu* were shown to exhibit synergistic and semi-dominant genetic interactions during floral organ development (Liu and Meyerowitz, 1995; Franks et.al., 2002), revealing functional interactions among these three genes *in vivo*.

Multiple mechanisms underlie A and C antagonism

Integral to the ABCE model is the mutual antagonism between the A and C class genes ,which was proposed based on their genetic interactions and phenotypes (Coen and Meyerowitz, 1991; Bowman et.al., 1991). However the molecular mechanisms underlying the A and C antagonism are not completely understood. The negative regulation of class A activity in the inner two whorls is conferred by at least two mechanisms, the post-transcriptional downregulation of *AP2* by *miR172* (Fig. 2.5a) and the repression of *AP1* transcription by *AG* (Fig. 2.5b) (Gustafson-Brown et.al. 1994; Chen 2004; Scwab et.al., 2005). However, *AG* and *miR172* appear to act independently of each other (Zhao et.al., 2007). The negative regulation of class C genes in outer whorls, on the other hand, depends on the LUG/SEU co-repressors recruited by the AP1/SEP3 (Sridhar et.al., 2006) (Fig. 2. 5c).

Our work reported here revealed a novel mechanism underlying the A-C antagonism. It indicates *AP2*, a miRNA target, is involved in regulating its cognate miRNA, *miR172* (Fig. 2.5d). This may ensure that *miR172* is mainly expressed in the inner two whorls of a floral meristem. *miR172* in turn negatively regulates *AP2* in the inner two whorls by degrading and/or blocking translation of *AP2* mRNA. Thus *AP2* and *miR172* both act in positive feedback loops to promote their own expression in respective A and C domains (Fig. 2.5).

It appears that the miRNA-transcription factor (TF) feedback loop is emerging as a major mechanism in microRNA regulation. Twenty-three TF-miRNAs feedback loops were recently identified in *C. elegans* (Martinez et.al.,

2008). In *Arabidopsis*, two AP2 family members, TOE1 and TOE2, were recently shown to positively regulate *miR172b* during juvenile to adult phase transition (Wu et.al., 2009). It is likely that *miR172* is subject to positive as well as negative regulation by members of the AP2 family. While TOE1/TOE2 may predominantly and positively regulate *miR172* during juvenile to adult phase change and flowering, AP2 may predominantly repress *miR172* transcription to specify proper flower organ identity. This may explain why not all AP2-binding sites in the promoter of *miR172c* are associated with LUG/SEU. It will be interesting to test if similar feedback loops exist for *miR166/miR165* and their target mRNAs and if this is a common theme for all plant transcription factors and the miRNA that regulates them.

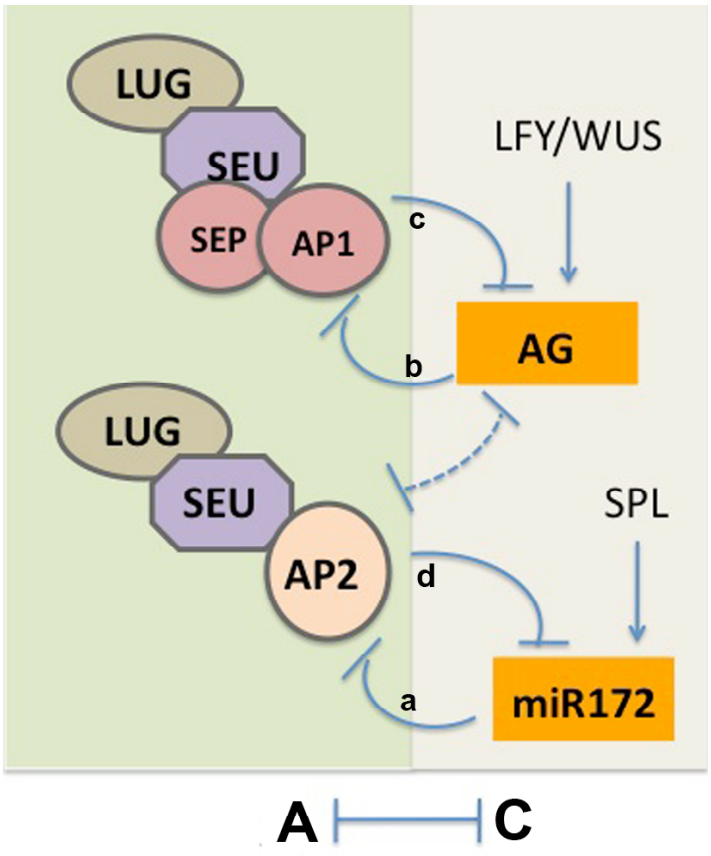

Fig. 2.5: A model on the multiple molecular mechanisms that underlie A and C antagonism. The negative regulation of class A activity in inner two whorls is conferred by the post-transcriptional downregulation of *AP2* by *miR172* (a) and the repression of *AP1* transcription by *AG* (b). The negative regulation of class C gene *AG* in outer floral whorls depends on the LUG/SEU co-repressors recruited by the AP1/SEP MADS domain proteins (c). The repression of *miR172* also requires LUG/SEU mediated by AP2 (d)**.** The dotted bar represents a less well understood regulatory interaction between *AG* and *AP2*. Also indicated are the transcriptional activation of *AG* by LEAFY (LFY) and WUSCHEL (WUS) (52) and the activation of *miR172* by *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL)* (35).

LUG may be an important regulator of several miRNAs

We show that a reduction of *AGO1* in *lug-3* masks more severe developmental defects in the leaves and flowers of *lug-3*. The more severe defects are only manifested when *AGO1* expression level is restored to the wild type level in *lug-3*, revealing previously unknown functions of *LUG* in microRNA

regulation. The severe flower phenotype of *35S::AGO1; lug-3* can be explained only partially by an increase in *miR172* expression, as *35S::AGO1; lug-3* flowers are more severe than *35S::miR172* and *ap2-2* flowers (Chen 2004). For example, the narrower floral organs and leaves of *35S::AGO1; lug-3* could be caused by a mis-expression of *miR165* and *miR166,* which target HD-ZIPIII genes involved in the establishment of leaf and floral organ polarity (Tang et.al., 2003; McConnell et.al., 2001). Small RNA Northern blots reveal a slight elevation of *miR165* and *miR166* in *lug-3* mutants (Fig. 2.1). Future experiments should include examining genome-wide microRNA levels in *lug-3* to identify other microRNA regulated by *LUG*.

LUG regulates the transcription of AGO1

Due to the fundamental importance of *AGO1* in microRNA function and plant development, tremendous efforts have been devoted to dissecting mechanisms that maintain proper *AGO1* homeostasis. These mechanisms include *miR168*-guided cleavage of *AGO1* mRNA, *AGO1*-derived siRNA in *AGO1* silencing, and the stabilization of *miR168* by AGO1 protein (Vaucheret et.al., 2004; Vaucheret and Mallory, 2009; Vaucheret 2009). However, almost nothing is known about the transcriptional regulation of *AGO1* and if such a regulation is important for maintaining *AGO1* levels in plants. Based on similar but widely expressed patterns of *pAGO1::GUS* and *pmiR168::GUS,* it was proposed that the *AGO1-miR168* gene pair is subject to transcriptional co-regulation (Vaucheret et.al., 2006). We show that the *miR168* level remains unchanged in *lug-3*,

indicating that *LUG* regulates *AGO1* independently of *miR168*. Our finding suggests that there are distinct aspects of *AGO1* regulation and that transcriptional regulation plays an important role in establishing or maintaining proper *AGO1* levels thus ensuring normal microRNA function during plant development.

The ChIP experiments (Fig. 2.3A, B) failed to detect any direct association of LUG or SEU to the four AP2-binding sites at the *AGO1* promoter, suggesting that *LUG/SEU* may regulate *AGO1* indirectly. Such an indirect role would be consistent with *LUG/SEU* as repressors, which may repress a negative transcriptional regulator of *AGO1*. However, we could not exclude the possibility that LUG/SEU directly regulates *AGO1* via other transcription factors that bind to regions of *AGO1* promoter not analyzed in our ChIP assays. Nevertheless, our results show that by positively regulating *AGO1* transcription, *LUG* acts as a general regulator of the miRNA pathway.

Transcriptional co-repressors are known to regulate diverse targets and developmental pathways and their target specificity is conferred by the DNA binding partners (Liu 2008). This work reports two new regulatory targets, *miR172* and *AGO1*, of the LUG/SEU co-repressors in addition to the previously reported class C gene *AG* (Liu and Meyerowitz, 1995; Franks et.al., 2002; Sridhar et.al., 2006) and identified AP2 as a pathway-specific DNA-binding partner of LUG/SEU for *miR172* regulation, providing novel insights into co-repressor function in plant development.

2.6 Acknowledgements

We thank R. Steven Conlan for sharing the *lug-3* microarray data, Vidyadhar Karmarkar for initiating the project, Theresa Dinh for sharing unpublished data on the AP2 binding site, Robert Franks for pSEU::GFP-SEU transgenic lines, Beth Krizek for ANT-pGAD424 and AP2 delta-pGAD424 constructs, Detlef Weigel for the AP2 cDNA clone, ABRC stock center for AGO1 cDNA, Joann Conner for constructing *35S::GFP-LUG*, V.V. Sridhar for constructing AP2-pGADT7, and Ranjani Iyer for generating 35S::*GFP-LUG; lug-16* lines. The work is supported by grants from National Science Foundation (IOB0616096 and MCB0744752) to Z. L. Z.L. is partially supported by the University of Maryland Agricultural Experiment Station.

Chapter 3: Intron-Mediated Silencing of the *CAULIFLOWER* Gene in *Arabidopsis*

3.1 Abstract

Most eukaryotic genes contain long non-protein coding intron sequences whose total length within a coding region can sometimes exceed that of the actual coding sequence. Introns can significantly affect gene expression in plants and other eukaryotic organisms in a variety of ways and may be involved in several types of gene regulation, including non-sense- mediated mRNA decay, intron-mediated gene enhancement, transcriptional and post-transcriptional gene silencing. The function of introns remains to be fully understood; however, studies in various organisms have shown that intron-mediated regulation of gene expression is ubiquitous and found throughout a diverse and wide range of organisms (plants, mammals, fungi, nematodes, insects). We provide evidence that the intron of the plant transcription factor *CAULIFLOWER* (CAL) can silence the endogenous *CAL* gene. Specifically, we identified a novel siRNA encoded by the first intron of the *CAL* gene when the intron is over-expressed from the 35S promoter as a transgene. The intron-derived siRNA transcriptionally inhibits *CAL* and alters its expression domain in flowers. Moreover, the intron-derived siRNA appears to silence the *CAL* gene by directing the methylation of the entire first intron of CAL, from which it originates. This silenced phenotype is transmitted through generations, even in the absence of the transgene. Conceivably, our work

on the novel intronic small RNA can be developed further as an effective tool to manipulate desirable traits for research and agricultural applications.

3.2 Introduction

Most eukaryotic genes are interrupted by the presence of long, non-protein coding, and apparently functionless sequences of DNA that are transcribed by RNA polymerase II and subsequently spliced out from the primary messenger RNA transcript. The function of introns remains to be fully understood; however, studies in recent years suggest that they may be involved in several types of gene regulation, both positive and negative. Traditionally, introns have been looked at as deleterious sequences, whose insertion in a bad place may interfere with normal gene expression, and whose transcription consumes vast amounts of metabolic resources and energy. Despite all the obviously negative characteristics of introns, they have been retained within eukaryotic genomes throughout evolution, which suggests that they must play some profound and previously overlooked roles in the cell. One vital role for introns may be their regulation of gene expression.

Studies in various organisms indicate that introns can both positively and negatively alter gene expression. Introns have been shown to regulate eukaryotic gene expression by many different mechanisms, including indirect and direct gene regulation (Maquat 2004; Kertesz et.al. 2006; Jeong et.al. 2006), by enhancers and promoters within introns (Deyholos and Sieburth 2000), by intron-mediated enhancement (IME) (Callis et.al. 1987; Vasil et.al. 1989). In addition, recent discoveries have revealed that small RNAs arising from a variety of intergenic

regions comprise a novel mechanism for post-transcriptional and epigenetic gene regulation (Llave et.al. 2002). Recently, a promoterless construct that contains the second intron of the floral homeotic gene *AGAMOUS (AG)* was shown to give rise to transgenic plants that phenocopy the *ag* mutant phenotype in 80% of the progeny (Zongrang Liu, personal communication). It was found that a small RNA produced from the *AG* second intron can guide the DNA methylation of *AG* locus.

Some introns influence the expression of genes much more than the promoter of a gene. For example, the first intron of *PROFILIN1 (PRF1)*, a gene that encodes an *Arabidopsis* vegetative profilin, is required for the strong constitutive and tissue specific expression of the gene (Jeong et.al., 2006). On the other hand, studies in various plant species including *Arabidopsis* indicate that introns play a pivotal role in NMD, a process that facilitates the degradation of truncated mRNA transcripts that contain a premature stop codon (Maquat et.al., 2004). Additionally, recent advances in small RNA research have reported that there exists a new class of 21bp long noncoding RNAs, called intron-derived microRNAs (Id-miRNAs), transcribed from the introns of genes (Ambros et.al., 2003). Several different types of Id-miRNAs have been identified in human, mouse and *C.elegans* cells; 10 have been mapped in *Arabidopsis* genome and they still remain to be characterized (Rodriguez et.al., 2004; Lin et.al., 2004; Llave et.al., 2002).

Small RNA molecules of about 20-30 nucleotides have been shown to be powerful regulators of gene expression both on the transcriptional and translational level. Plant genomes contain several types of small regulatory RNAs,

including microRNAs and small interfering RNAs (siRNAs). These small RNA molecules mediate gene silencing through initiating post-transcriptional gene silencing (PTGS) or more specifically RNA interference (RNAi), a process that involves small non-coding RNAs associating with nuclease-containing regulatory complexes and then pairing with complementary messenger targets, resulting in mRNA degradation or translational repression (Lee et.al., 1993; Reinhart et.al., 2000; Hammond et.al., 2001). Introduction of dsRNA, various transgene constructs or viruses into plants may trigger PTGS and lead to silencing of endogenous gene targets through mRNA degradation or translation inhibition. Alternatively, siRNAs and more rarely miRNA can mediate gene silencing not only by using the traditional RNAi-silencing machinery, but also through transcriptional gene silencing (TGS) resulting from RNAi-related chromatin remodeling events (Jin et.al., 2004). RNA-directed DNA methylation, an important type of RNAi-like mechanism for gene silencing, has been recognized not only in mammalian cells, but in plant cells as well. It is possible that plant intronic small RNA molecules can also direct epigenetic modifications that repress gene expression. A novel plant-specific protein named RNA polymerase IV (RNAP IV) and RNA-dependent RNA polymerase (RDR2) are required for this process. Initially, RNAP IV was identified as a key component in the biogenesis of more than 90% of plant siRNAs, and only subsequently shown to be involved in siRNA-mediated chromatin condensation (Zhang et.al., 2007; Huettel et.al., 2007). There are several distinct features that characterize plant siRNA-directed DNA methylation. First, methylation occurs primarily at the region of

RNA-DNA sequence similarity, suggesting that RNA-DNA base pairing acts as a substrate for methylation (Pellisier et.al., 1999). Another distinctive feature of plant RNA-directed DNA methylation is that cytosine methylation is the initial epigenetic mark as sequences as short as 30bp can be methylated (Pellisier et.al., 2000). Furthermore, cytosines in all sequence variants become modified (CG, CNG and CNN where N is A, T or C). There are several conserved components of the RNA-directed DNA methylation in plants. First, DNA methylation requires at least two DNA methylatransferases, MET1 and the plant-specific CMT3, which are necessary to maintain continuous CG and CNG methylation during DNA replication (Matzke M., 2005; Chan S.W., 2005). Furthermore, RNA-directed DNA methylation is dependent upon core RNAi proteins that generate and process small RNA's in the RNA-methylation pathway. RNA DEPENDENT RNA POLYMERASE 2 (RDR2), DICER-LIKE 3 (DCL3) and ARGONAUTE 4 (AGO4) play a crucial role in small RNA biogenesis. RDR2 processes various single stranded siRNA precursors into double-stranded RNA precursors, from which the final siRNA molecules are produced. DCL3 catalyzes the enzymatic cleavage of the long dsRNA precursors into 21-24bp duplex siRNAs. The siRNA is finally loaded onto AGO4 which recruits the methylation machinery at the target chromatin site and mediates the DNA methylation (Qi et.al., 2006).

In this study, we have identified and characterized a novel intronic small RNA molecule that antagonizes the function of its host gene. Overexpression of the first intron of the *Arabidopsis* floral specific MADS-box transcription factor *CAULIFLOWER (CAL)* in *ap1* background results in a mutant phenotype that

phenocopies *ap1-1 cal*-1 double mutants (Bowman et.al., 1993). Previous research shows that a mutation in the *CAL* gene locus enhances the *ap1-1* phenotype, whereby in *ap1-1; cal-1* double mutants each meristem that in WT would generate a single flower, consistently behaves like an inflorescence meristem. Thus, each inflorescence meristem gives rise to second, third and fourth order inflorescence meristems until a large number of inflorescences have been produced. Finally, each inflorescence meristem differentiates and gives rise to flowers with increased fertility (Bowman et.al., 1993) In this study, we show that an siRNA located within the first intron of the *CAL* gene inhibits *CAL* activity and affects its expression domain in flowers. Our results indicate that an intron-derived siRNA leads to heterochromatin repression of the *CAL* gene locus. Furthermore, this silenced epigenetic pattern is stably inherited for several generations even in the absence of the transgene *35S::iCAL* that triggers the initial silencing of CAL. Conceivably, intronic small RNA discovered by this study can be used as an effective tool to manipulate gene activity in transgenic plants for research and agricultural purposes.

3.3 Materials and Methods

Plant materials, growth conditions and transformation

Arabidopsis thaliana plants were grown on Metromix soil (Griffin) at 22°C under 16h light/8h dark conditions. *ap1-1, ap1-1 cal-1,* and the transgenic line *35S:: iCAL, ap1-1* are all in the Landsberg *erecta* (Ler) background. *ap1-1; cal-1* seeds were obtained from ABRC stock center (CS6161). For the

construction of *35S:: iCAL, CAL* intron DNA was amplified using intron specific primers 5’ TCCTCTGAATCTTGGTAATTG 3’ and 5’TAGTACCTTCTCCATGCTAC 3’. The amplified *CAL* intron DNA fragment was cloned into pCR8/GW/TOPO using TA cloning kit (Invitrogen, Carlsbad, CA, USA) and then introduced into the pEarleyGate100 plant transformation vector (Earley et al., 2006) using the Gateway® technology (Invitrogen, Carlsbad, CA, USA). The construct was introduced into *Agrobacterium tumefaciens* GV3101 and used to transform *Arabidopsis thaliana ap1-1* (Ler) plants via floral dip. T1 seeds were selected on soil using BASTA. 222 *35S:: iCAL; ap1-1* lines were generated, 91 of which were *ap1-like* and 131 were *cal-like* in phenotype. Detailed analysis was conducted on 11 *35S:: iCAL; ap1-1* transgenic plants with *ap1-like* phenotype and 130 *35S:: iCAL; ap1-1* transgenic lines with *cal-like* phenotype. These lines were used to generate T2 plants for further analyses. Genotyping was performed on genomic DNA from 16 *35S:: iCAL; ap1-1 cal-like* transgenic T2 lines extracted using the DNeasy Plant Mini Kit (Qiagen, MD, USA). PCR conditions were 35 cycles at 94°C for 30s, 55°C for 45s, 72°C for 1 minute using primers specific to the *35S* promoter and the 3’ end of the *CAL* intron. Primer sequences are as follows: forward- 5'GGAAGTTCATTTCATTTGG3'; reverse- 5' CTAGTTAGGGCAAACGAAGG 3'. The *35S:: iCAL-Promoter* construct was generated in the same manner using the following primers to amplify the *CAL* promoter: F 5’ ARGGGAAGGGGTAGGGATTG 3’; R 5’

CAAGATTCAGAGGAGTACTC 3'. None of the transgenic lines (approximately 30 lines) generated showed cal-like phenotype.

After collection, the T1 *35S:: iCAL; ap1-1 ap1-like* and *35S:: iCAL; ap1-1 cal-like* seeds were stored in the dark at ambient temperature in Eppendorf tubes with a hole in the lid but without a desiccant in the box.

For the construction of *iCAL::GUS* and *iCAL::YFP, CAL* intron DNA was amplified using intron specific primers 5' TCCTCTGAATCTTGGTAATTG 3' and 5'TAGTACCTTCTCCATGCTAC 3'. The amplified *CAL* intron DNA fragment was cloned into pCR8/GW/TOPO using TA cloning kit (Invitrogen, Carlsbad, CA, USA) and then introduced into pMDC163 Gateway compatible plant transformation vector (ABRC Stock number CD3-737) and pMTX003 promoterless YFP vector (Reddy lab at UCR), respectively. 5 transgenic *iCAL::YFP* and 15 transgenic *iCAL::GUS* lines were obtained and analyzed.

Northern blot analysis

Total RNA was isolated from inflorescences of T2 *ap1-1, cal-1, 35S::iCAL; ap1-1 (ap1-like)* and *35S::iCAL; ap1-1 (cal-like)* transgenic plants using TRI® Reagent (Sigma, St. Louis, MO, USA). For small RNA blot, a 30 μg aliquot of total RNA was separated by a 15% acrylamide gel. RNA was transferred to Hybond-N nylon membrane (Amersham, Little Chalfont, Buckinghamshire, UK) by electroblotting. RNA was cross-linked onto the membrane using EDC (N-(3-Dimethylamionopropyl)-N'-ethylcarbodiimide

hydrochloride) (Sigma, St. Louis, MO, USA) according to a published protocol (Pall and Hamilton, 2008). *siRNA* was detected using PCR-amplified *CAL* intron (same primers were used as in the initial cloning) as a probe labeled with alpha P^{32}-dCTP (Perkin-Elmer) with Ready-To-Go DNA Labeling Beads accordingly to the manufacturer's instruction (Amersham, Little Chalfont, Buckinghamshire, UK). The small RNA blot was hybridized and washed according a protocol at http://web.wi.mit.edu/bartel/pub/protocols.html. Small RNA blots were also probed with the full-length *CAL* cDNA and 5'UTR sequences. The primers used to generate these probes were as follows: for the *CAL* cDNA probe F 5' ACATTACCATCATTAGAAAA 3' , R 5' CATTGCTCCCCGAAATACAA 3'; for the *CAL 5'UTR* probe 5' F' CCTTCCCCAATACCAAGTTA 3', F 5' GGAAACCTCGGCATCACAAA 3'.

For mRNA blots, the *CAL* cDNA fragment was excised using the *NotI* and *AscI* enzymes from ABRC clone stock # CD3-736 (NSF grant 0418891) and labeled with alpha P^{32}-dCTP (Perkin-Elmer) with Ready-To-Go DNA Labeling Beads accordingly to the manufacturer's instruction (Amersham, Little Chalfont, Buckinghamshire, UK). A 15µg aliquot of total RNA was separated on a 1% agarose gel, transferred onto BrightStar-Plus membrane, hybridized and washed using the Northern Max-Gly kit and accompanying manual (Ambion Inc., Austin, TX, USA).

In situ Hybridization

The *CAL in situ* probe was generated by PCR amplification of cDNA using gene specific primers 5' ATCCAATGTGAGCAGCTGAA 3' and 5' GCTAATACGACTCACTATAGGGTGCTCCCCGAAATACAAG 3'containing a T7 RNA polymerase-binding site (underlined). T7 RNA polymerase (New England Biolabs, Ipswich, MA, USA) was used to transcribe digoxygenin-labeled UTP (Roche GmbH, Mannheim, Germany) probes. Tissue was fixed in 4% paraformaldehyde (Sigma, St Louis, MO, USA) and embedded in Paraplast X-tra (Monoject Scientific, St Louis, MO, USA). 8 μm sections were fixed to Probe-on Plus slides at 42°C (Fisher Scientific, Pittsburgh, PA, USA). Procedures for *in situ* pre-hybridization, hybridization and detection were performed as described previously (Carr and Irish, 1997).

Western Blot

ap1-1, 35S::iCAL; ap1-1 (ap1-like) and *35S::iCAL; ap1-1 (cal-like)* flower tissues were ground to powder in liquid nitrogen, and used for protein isolation and for Western blotting to detect CAL protein. The tissue powder was mixed with 1 volume of Protein Extraction Buffer (150mM NaCl. 20mM Tris-HCl, pH7.5, 10mM $MgCl_2$, 8% glycerol, 0.1% NP-40, 1mM PMSF) and boiled for 5 minutes. The proteins were resolved in 15% SDS/polyacrylamide gel and transferred to a PVDF membrane (Invitrogen, Carlsbad, CA, USA) using a semi-dry transfer apparatus (BioRad, Hercules, CA, USA) for 3 hours at 4 °C. The membranes were then incubated with 5% milk in TBS (137 mM NaCl / 20 mM

Tris HCL, pH7.6) for 1 hour at room temperature and then in anti-CAL antibody (Santa Cruz Biotechnology, 1:125 diluted in TBST (0.05% Tween 20)) overnight at 4°C. After three washes in TBST, the membranes were incubated with the secondary AP-conjugated antibody for 1 hour at room temperature. The membranes were washed again in TBST and signals were detected using Western blue (Promega, Madison, WI, USA).

Bisulfite Sequencing

Genomic DNA was extracted from *ap1-1, 35S::iCAL; ap1-1 (ap1-like)* and *35S::iCAL; ap1-1 (cal-like)* flower tissues using Qiagen Dneasy plant mini Kit. The Bisulfite sequencing was performed using the Zymo Research EZ DNA methylation kit. (Zymo Research, Orange, CA). For the bisulfite conversion 300ng floral genomic DNA was treated with EZ DNA Methylation-Direct™ Kit for each reaction. PCR conditions were as follows: JumpStart REDTaq DNA polymerase (Sigma) 12.5 ul, Primer 1(10uM) 0.4 ul, Primer 2(10uM)0.4 ul, H2O 9.7 ul,DNA 2 ul. Cloning was performed using the Promega pGEM-T easy KitpGEM-T easy vector 0.5 ul. Cloning reaction was setup as follows: insert DNA: 1.5 ul,buffer: 2.5 ul,ligase: 0.5 ul. 15 colonies were sequenced. Methylation analysis was conducted using http://www.gmi.oeaw.ac.at/en/cymate-index/cymate-v2/. Primers used to amplify the *CAL* intron were as follows: MCalU1: CCAAAATTTCCCTTATTRTCTTCTCCCAT; MCalL TTATTATTAAATGGGAAAAAATGAAGAGT; MCalU2 AAAACTCT TCATTTTTTCCCATTTAATA; McalL GTGAGAGTTAGGTGYAATTAGYTGT.

3.4 Results

Overexpression of CAULIFLOWER (CAL) first intron in ap1-1 phenocopies cal-1; ap1-1

To identify candidate *Arabidopsis thaliana* genes regulated by their introns we chose to overexpress the introns of five genes, *FLOWER LOCUS C (FLC), STERILE APETALA (SAP), ACCELLARATED DEATH 6 (ACD6), SHATTERPROOF (SHP) and CAULIFLOWER (CAL) (*Werner et.al., 2005*;*Byzova et.al., 1999; Rate et.al., 1999; Liljegren et.al., 2000; Bowman et.al., 1993) in the respective backgrounds, in order to determine whether a mutant phenotype of each gene would be observed (Table 3.1). The results in Table 3.1 reflect the transgenic plants obtained by overexpressing the five introns in both the forward and reverse orientation. For the *CAULIFLOWER* intron, 119 transgenic plants were generated when overexpressing the intron in the reverse direction, and 103 transgenic plants were obtained when overexpressing the intron in the forward direction. These candidate genes were chosen based on their easy to detect mutant phenotype, as well as relatively large intron size (≥1kb).

Gene Intron	Background	Intron Size	Silencing Observed?	Transgenics Analyzed
CAULIFLOWER	*ap1-1*	*1kb*	**yes**	222
STERILE APETALA	*Ler*	*2.8kb*	no	~ 40
FLOWER LOCUS C	*FRT-sf2*	*3.6kb*	no	~ 300
ACCELERATED DEATH 6	*acd6*	*1.5kb*	n/a	n/a
SHATTERPROOF 1	*shp2*	*1.3kb*	no	~ 120

Table 3.1: Genes analyzed for intron-mediated gene silencing. Numbers of transgenic lines represent combined results of forward and reverse constructs.

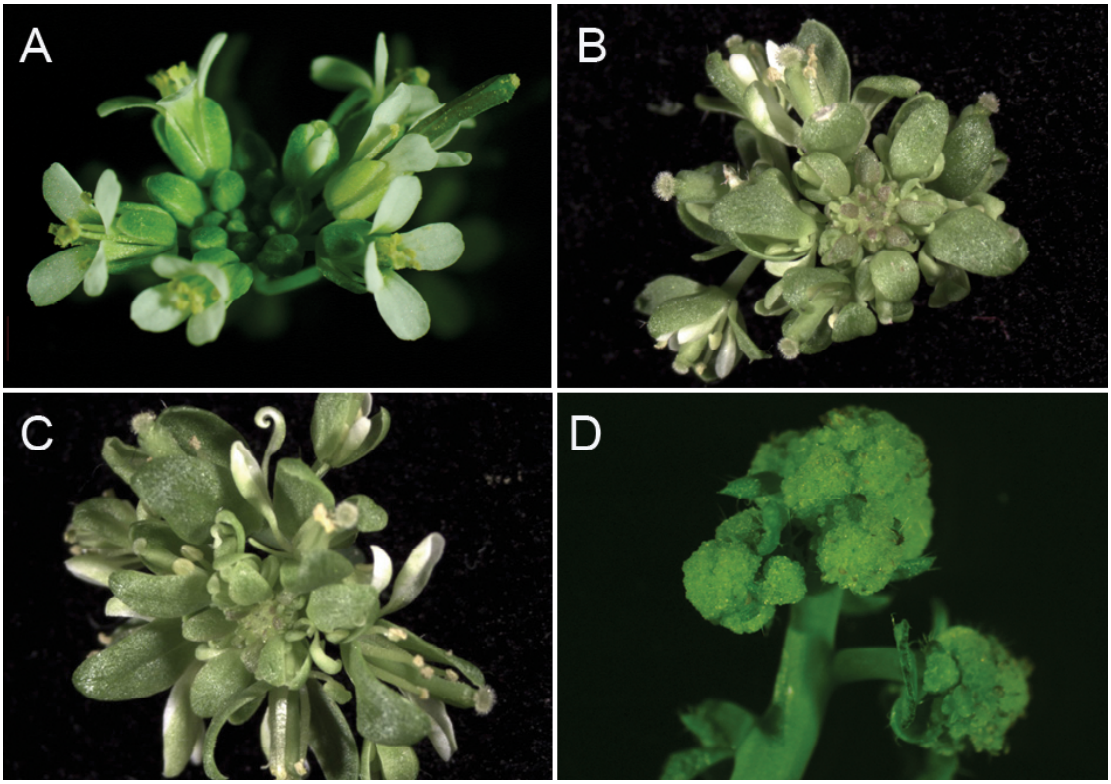

Fig 3.1: Overexpressing the largest *CAULIFLOWER* intron in *ap1-1* mutant phenocopies a *cal-1 ap1-1* flower. A. WT inflorescence. B. *ap1-1* inflorescence. C. Inflorescence of *35S::iCAL;ap1-1* showing an *ap1-like* phenotype. D. Inflorescence of *35S::iCAL;ap1-1* showing a *cal*-like phenotype.

Only overexpression of the *CAL* intron in *ap1-1* background resulted in a mutant phenotype that is identical to *cal-1*; *ap1-1* (Fig.3.1). Overexpression of the *CAL* promoter results in transgenic plants that exhibit an *ap1-1* mutant phenotype suggesting that the silencing phenomenon is specific only to the overexpression of the *CAL* intron. In *ap1* mutants first whorl organs are converted into bract-leaf-like organs, second whorl organs are often absent, and third and fourth whorl organs are normal (Irish and Sussex 1990) (Fig. 3.1B). Furthermore, the *ap1* phenotype is significantly enhanced by a mutation at the *CAL* gene locus (Bowman et.al., 1993), resulting in flowers in which the floral meristem is converted into an inflorescence meristem that gives rise to numerous meristems arranged in a spiral phylotaxy (Ditta et.al., 2004). Introducing *35S::iCAL* into *ap1-1* plants resulted in two distinct types of transgenic plants. 41% of the transgenics resembled *ap1* (*ap1-like*) (Fig. 3.1C) and 59% resembled *ap1-1*; *cal-1* (*cal-like*) (Fig. 3.1D). To conduct further analysis on the two types of transgenic plants, we followed several *ap1-like* T1 lines and *cal-like* T1 lines into the T2 generation for two separate transformation experiments (Table 3. 2; Table 3.3). Initially, both *ap1-like* and *cal-like* T1 lines segregated *ap1-like* and *cal-like* progeny in the T2 generation at a ratio of about 1:1. Surprisingly, seeds from all five *cal-like* T1 lines, planted three weeks later after initial seed collection, segregated close to 100% *cal-like* progeny (Table 3.2; Table 3.3). The segregation ratio was followed for the five T1 *cal-like* lines planted every week for nine weeks, and the close to 100% *cal-like* progeny segregation persisted (Table 3.3). This suggests that there is a time-dependent conversion within the seeds collected

from T1 plants from *ap1-like* flowers into *cal-like* flowers. Therefore, some epigenetic conversion from unsilenced to silenced CAL gene occurs during the first three weeks of seed maturation after collection. In Table 3.2 the first count was conducted with seeds planted immediately after collection from T1 *cal-like* lines, while the second count was performed with the same seeds after approximately 30 days. In Table 3.3, week 1 T2 plants were generated from T1 seeds planted immediately after collection, and weeks 2, 3, 4, 5, 6, 7, 8, 9 were planted at the beginning of each consecutive week. This analysis of the transgenic plants overexpressing the *CAL* intron suggests that there exists an unknown mechanism of time-dependent intron-mediated gene silencing of *CAL*.

Line	First Count	First Count	Second Count	Second Count
	WT	*cauliflower*	WT	*cauliflower*
35s::iCAL (T_1-A)	4	15	2	7
35s::iCAL (T_1-B)	10	14	1	13
35s::iCAL (T_1-C)	8	5	2	15
35s::iCAL (T_1-E)	9	9	2	33
35s::iCAL (T_1-F)	1	5	0	39
35s::iCAL (T_1-G)	9	12	0	25
35s::iCAL (T_1-H)	5	8	4	38
35s::iCAL (T_1-J)	11	17	5	29
35s::CAL (T_1-K)	7	15	1	45
Total	64	100	17	244
% M/WT		61		93

Table 3.2: Transgenic T1 *cal-like* lines exhibit a higher number of mutants after 30 days of seed storage. All transgenic plant lines shown in the table have a *cal-like* phenotype.

Line 1	Week 1		Week 2		Week 3		Week 4		Week 5		Week 6		Week 7		Week 8		Week 9	
	ap1	*cal*	*ap1*	*cal*	*ap1*	*cal*	*ap1*	*cal*	*ap1*	*cal*	*ap1*	*cal*	*ap1*	*cal*	*ap1*	*cal*	*ap1*	*cal*
T_1-A	4	8	3	6	1	5	2	10	0	15	0	36	1	24	0	49	0	29
T_1-B	11	12	12	19	2	23	2	20	4	13	1	19	0	29	0	18	1	37
T_1-C	9	16	22	30	0	13	2	26	1	25	2	41	1	14	1	28	1	17
T_1-D	11	19	33	40	0	25	3	33	3	28	0	27	0	35	1	53	0	32
T_1-E	13	12	16	32	3	33	2	25	0	27	0	33	0	22	1	41	0	36
T_1-1	9	18	10	16	5	3	16	13	9	13	5	12	11	18	8	18	5	13
T_1-2	3	25	7	8	8	6	9	17	6	11	11	13	7	20	16	10	18	16
T_1-3	6	18	33	32	15	6	16	6	10	4	17	9	17	23	9	13	20	12

Table 3.3: Time-dependant conversion of *35S::iCAL; ap1-1* into *cal-like.* Lines T1 A-D exhibit *cal-like* phenotype. Lines T1 1-3 exhibit *ap1-like* phenotype.

Intron-mediated silencing of CAL resulted from a reduction as well as an alteration of expression domain of CAL mRNA

How does the *CAL* intron-mediated silencing of endogenous *CAL* occur? If the *35S* promoter driven *CAL* intron can cause degradation of endogenous *CAL* mRNA precursor (mRNA has no homology to the intron) to cause loss of function of *CAL*, we should observe a reduction of *CAL* mRNA in *cal-like* lines and no reduction of mRNA in *ap1-like* lines. Surprisingly, Northern blot analysis showed no difference in *CAL* mRNA levels in *ap1-1, ap1-like,* and *cal-like* (Fig. 3.2A). However, Western blot analysis showed a slight reduction of the CAL protein in *cal-like* flowers compared to *ap1-1* and *ap1-like* (Fig 3.2B).

One explanation is that the equal amount of *CAL* mRNA may be due to the significant increase in the number of floral meristems in the *cal-like* plants compared to *ap1-1* and *ap1-like* plants. Since the floral meristem is known to express a high level of *CAL* mRNA (Bowman et.al., 1993), it leads to higher

mRNA accumulation in *cal*-like plants which may mask a simultaneous reduction of mRNA in individual floral meristems. Therefore, using Northern blots, we compared the expression level of *CAL* mRNA between *cal-like* flowers and *cal-1; ap1-1* plants. *cal-1* is caused by a missense point mutation resulting in a reduction of *CAL* function (Bowman et.al., 1993). Indeed, a 2-fold decrease of *CAL* mRNA level was detected in *cal-like* flowers compared to *cal-1; ap1-1* (Fig. 3.2C), suggesting that the *CAL* intron can reduce *CAL* mRNA via either posttranscriptional or transcriptional gene silencing.

To elucidate the precise molecular mechanism responsible for the intron-mediated silencing of *CAL,* we tested if the over-expressed intron in *cal-like* plants can produce a small RNA. We performed a small RNA blot of *ap1-1, ap1-like* and *cal-like* inflorescence tissues using the entire *CAL* first intron as a probe (Fig. 3.2D). We detected a small RNA only in RNA extracted from *cal-like* flowers (Fig. 3.2D), suggesting that a small RNA is produced from the *CAL* intron that might be responsible for the *cal-like* phenotype. We did not detect a small RNA when we probed RNA extracted from *ap1-1, ap1-like* and *cal-like* inflorescence tissues using either the *CAL* cDNA or 5' UTR sequences as a probe, suggesting that the small RNA is specifically generated from the *CAL* intron (Fig. 3.2 E, F).

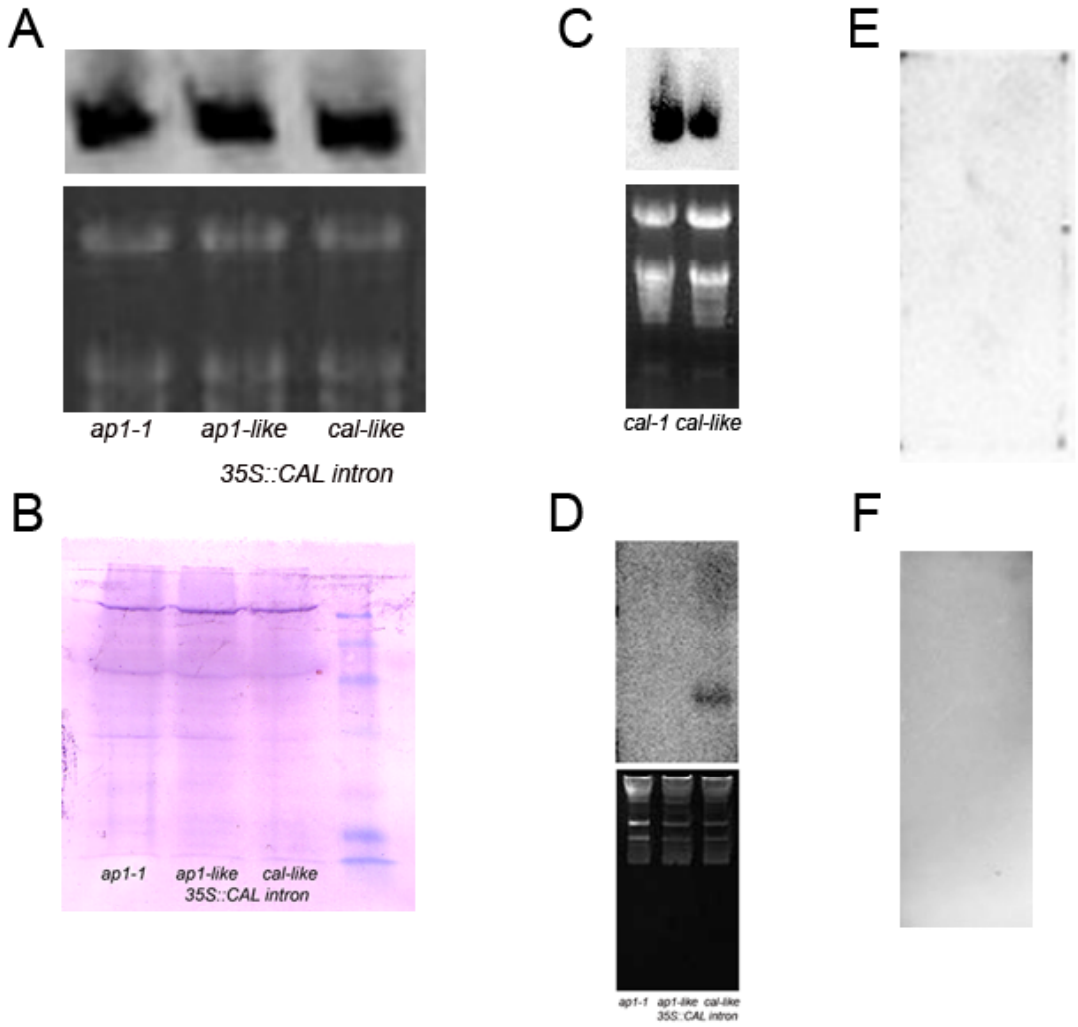

Fig 3.2: Reduction of the *CAL* product in *cal-like* background. A. Northern blot of RNA samples extracted from flowers of *ap1-1*, *ap1-like* and *cal-like* plants shows similar expression of *CAL*. B. Western blot analysis indicates CAL protein levels decrease in *cal*-like flowers. C. Northern blot of RNA samples extracted from *cal* and *cal*-like mutant flowers shows reduced expression of *CAL* in *cal*-like background. D. Small RNA Northern blot probed with the full-length *CAL* intron indicated the presence of an siRNA in *cal*-like flower tissue. E. Small RNA Northern blot probed with the full-length *CAL* cDNA failed to identify a small RNA. F. Small RNA Northern blot probed with the full-length *CAL* 5'UTR failed to identify a small RNA.

To determine if the expression domain of *CAL* is also affected in *cal-like* plants, we conducted *in situ* hybridizations comparing *CAL* mRNA expression domain in *ap1-1, ap1-like*, *cal-1; ap1-1,* and *cal-like* flowers (Fig 3.3). Comparable to published studies on *CAL* expression, in *ap1-1* and *ap1*-like flowers, *CAL* is expressed throughout young floral meristems (Fig 3.3A, B) (Kempin et.al., 1995). The *cal-1* mutation is a missense point mutation that should not affect its mRNA expression domain (Kempin et.al., 1995). Thus, in *cal-1; ap1-1, CAL* mRNA is likely expressed throughout the floral meristems as is in wild type (Fig 3.3C). Thus *cal-1: ap1-1* serves as a better control for *cal-like* plants as both *cal-1; ap1-1* plants and *cal-like* plants similarly have an increased number of floral meristems. We observed that in *cal*-like flowers, there is a reduction in the amount and spatial distribution of *CAL* (Fig 3.3D). Specifically, the *CAL* mRNA expression domain is restricted only to the center of the floral meristems and is greatly reduced. The result indicates that the over-expressed *CAL* intron RNA caused a reduced level and altered spatial domain of endogenous *CAL* mRNA. Thus, the difference in expression level and expression domain of *CAL* mRNA between the *ap1-like* and *cal-like* transgenic lines may be responsible for their different phenotypes. In the *ap1-like* class transgenic lines, *CAL* mRNA is not affected by the *CAL* intron; perhaps these *35S::iCAL* lines expressed a lower level of intronic RNA.

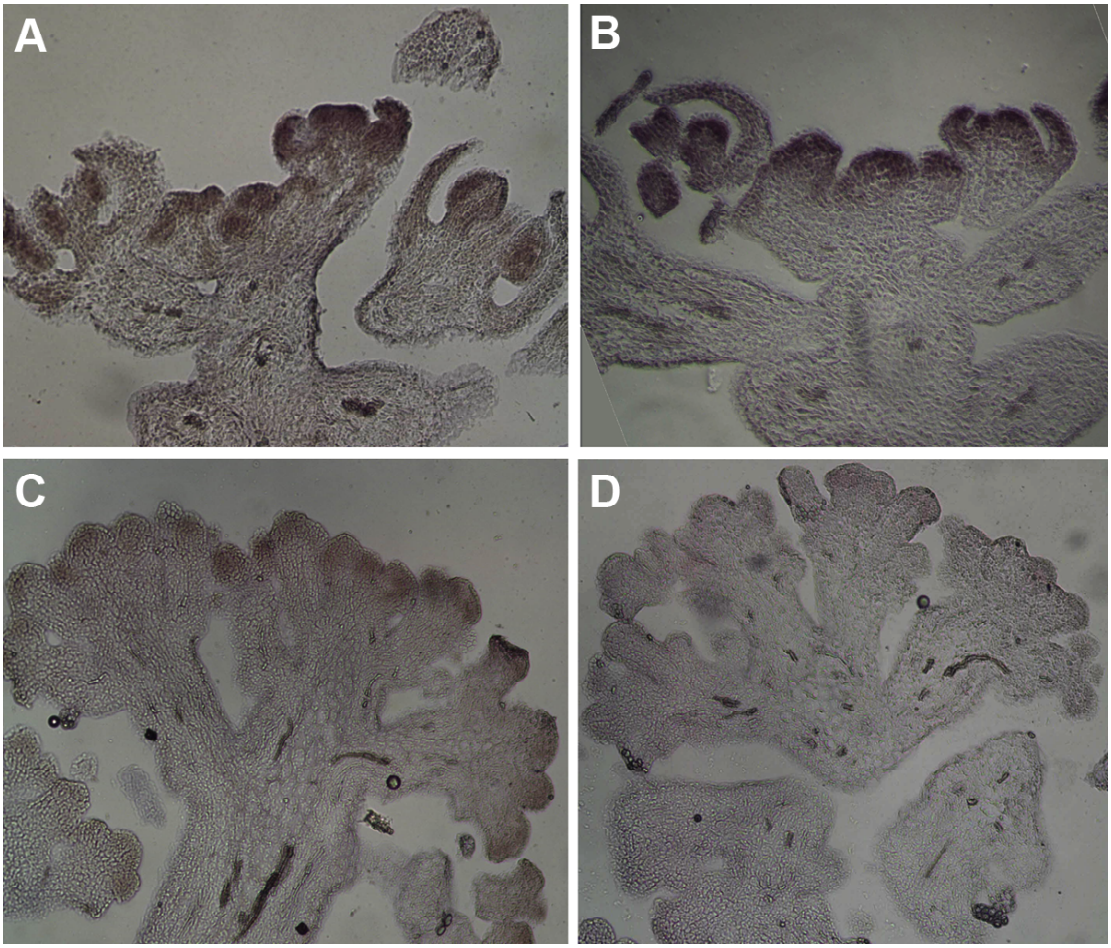

Fig 3.3: *In-situ* hybridization analysis of *CAL* expression. A. *ap1-1* inflorescence. B. *ap1-like* inflorescence. C. *cal-1; ap1-1* inflorescence. D. *cal-like* inflorescence

The cal-like phenotype can be inherited for several generations even in the absence of the 35S::iCAL transgene that initiated the silencing

If the *CAL* intron-derived small RNA silences endogenous *CAL* via modifications of its chromatin, it is likely that some of the silenced lines can inherit the silenced state without inheriting the transgene. This would explain why 100% of the progeny are *cal-like* in the T2 generation of *cal-like* lines (Table 3.2). To test this hypothesis, we genotyped 32 individual T2 *cal-like* plants from five independent T1 lines for the presence of the transgene, a sample of which are

shown in Figure 3.4A. Surprisingly, only 50% of the *cal-like* plants had the transgene, suggesting that the other 50% *cal-like* plants do not have the transgene. Thus, the mutant phenotype could be transmitted in the absence of the *35S::iCAL* transgene and epigenetic mechanisms must be employed.

One possible mechanism of such epigenetic inheritance is that the seeds could inherit the *CAL* intronic siRNA, which is involved in guiding epigenetic silencing of *CAL*. Alternatively, the epigenetic state once initiated can be maintained automatically through maintenance of the established epigenetic marks in the germline during meiosis and throughout all the somatic cells of the progeny *cal-like* plants.

Our results thus far together with the fact that intron sequences cannot directly pair and target mRNA, suggest that transcriptional gene silencing maybe responsible for the *cal-like* phenotype.

To test how vigorous the epigenetic inheritance is for those transgenic lines that have lost their transgene, we analyzed T3 progeny for the presence of the transgene by PCR, from line 7 that possesses transgene and line 9 that possesses no transgene) (Fig 3.4B). While all 24 line 7 and 42 line 9 T3 progeny from both lines had a *cal-like* phenotype, only individuals from line 7 had the transgene (Fig. 3.4B). Currently we are analyzing the T4 generation of line 9.

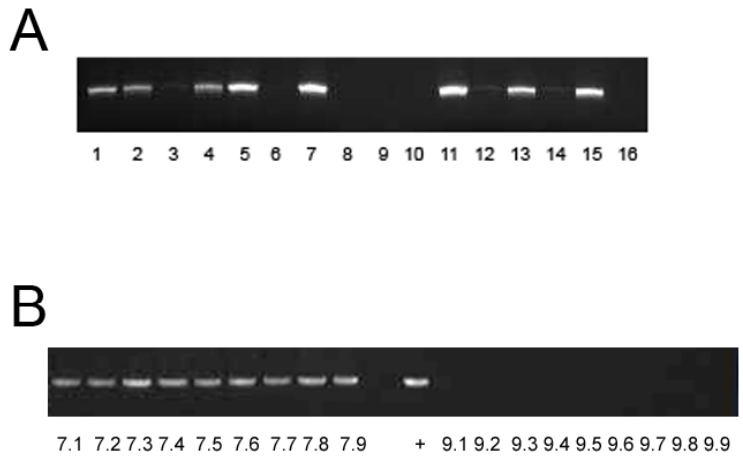

Fig 3.4: Epigenetic inheritance of the *cal-like* phenotype. A. Genotyping of 16 different *cal*-like T_2 plants from a single T1 *cal-like* line. PCR products indicate the presence of the transgene. Absence of PCR products in some T2 lines indicates that the *cal*-like phenotype is not dependent on the presence of a *35S::iCAL* transgene. B. PCR detection of the transgene in T3 generation. Lines 7.1 to 7.9 are progeny from T2 line 7 shown in A. Lines 9.1 to 9.9 are progeny of T2 line 9 shown in A. + is WT used as a control.

CAL-intron DNA is methylated in cal-like plants

siRNAs are known to play key roles in RNA-directed DNA methylation (RdDM) and subsequent heterochromatin formation (Chan et.al. 2004). To determine if the small RNA detected in the *35S::iCAL* flowers was an siRNA that epigenetically silences *CAL*, we collaborated with Dr. Zongrang Liu's lab to conduct bisulfite sequencing of the *CAL* intronic sequences. We analyzed the methylation status of the endogenous *CAL* first intron using genomic DNA extracted from *ap1-1*, *ap1-like* and *cal-like* flowers. We observed a significantly increased percentage of DNA methylation in the *cal-like* sample compared to *ap1-1* and *ap1-like* samples (Fig 3.5B, Table 3.4). Analysis of the whole *CAL* intron reveals the methylation of cytosines at three different types of methylation

sites, designated as CG, CHH and CHG with H being either an A, C or a T, but not a G. From the average values we can conclude that the *CAL* intron is methylated in the *cal-like* plants with CHG methylation being the predominant form (Fig 3.5B). These results suggest that the over-expressed *CAL* intron can trigger epigenetic silencing of *CAL* via DNA methylation, possibly carried out by RNA-directed DNA methylation.

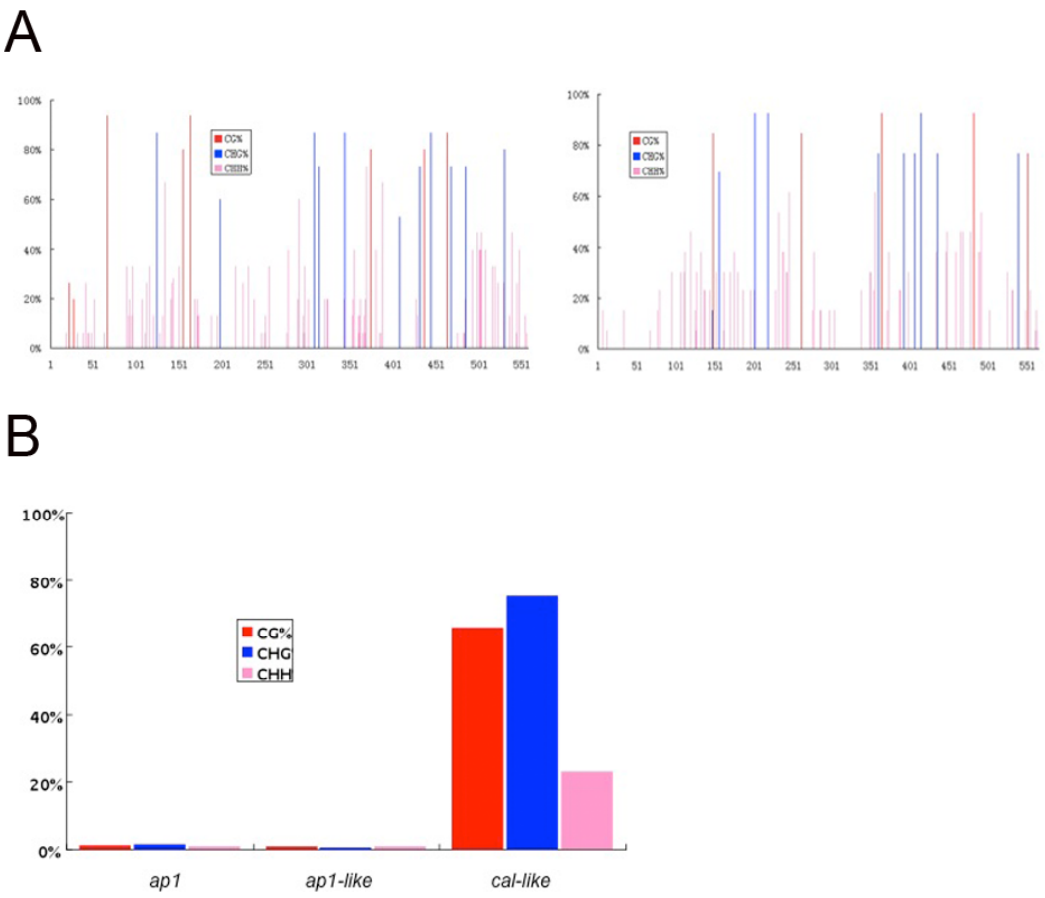

Fig 3.5: Bisulfite sequencing reveals significant DNA methylation in the first intron of *CAL* in *cal-like* plants. A. Percentage methylation at the individual CG, CHG, and CHH in the first intron of *CAL* in cal-like plants. The 1KB intron is separated into two fragments whose methylation is shown in the left and right diagram. B. Average methylation frequency in percentage for CG, CHG, and CHH in the first introns of *ap1-1, ap1-like* and *cal-like* plants.

	Primer		Sequencing ID	Sequencing number	Methylation			Average		
	sense	anti-sense			CG%	CHG%	CHH%	CG%	CHG%	CHH%
ap1	MCalU2	MCalL2	Ym159	12	0.0167	0.0152	0.0077	**0.0111**	**0.0143**	**0.0092**
	MCalU1	MCalL1	Ym135	8	0	0.0125	0.0123			
ap1-like	MCalU2	MCalL2	Ym149	14	0.0071	0.0065	0.0102	**0.0092**	**0.0066**	**0.0096**
	MCalU1	MCalL1	Ym148	15	0.0133	0.0067	0.0085			
cal-like	MCalU2	MCalL2	Ym150	15	0.5600	0.7576	0.2097	**0.6605**	**0.7538**	**0.2312**
	MCalU1	MCalL1	Ym160	13	0.8615	0.7462	0.2741			

Table 3.4. Bisulfite sequencing resuts for the methylation status of *CAULIFLOWER* intron.

A

	ap1-1	cal-like
Line 1	4	8
Line 2	6	6
Line 3	5	6
Line 4	6	8
Line 5	4	8
Line 6	6	5

Phenotypic results of *ap1-1 X 35S::iCAL; ap1-1; cal-like*

B

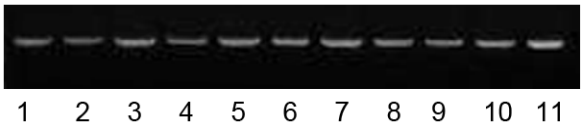

Fig 3. 6: A. Phenotypic results of *ap1-1 x 35S::iCAL; ap1-1 cal-like* crosses. Each line represents the F1 progeny of a different cross. B. Genotyping for the presence of the *35S::iCAL* transgene in F1 progeny of *ap1-1 x 35S::iCAL; ap1-1 cal-like* crosses.

In order to determine whether the *CAL* intron methylation pattern and subsequently the *cal-like* phenotype could be transmitted to the next generation via pollen, we crossed *ap1-1* carpels with the pollen from *cal-like (35S::iCAL; ap1-1)* plants. Progeny of six different crosses were obtained and each line segregated 50% *cal-like* (Fig 3.6A), some of which were genotyped to see if any does not possess the transgene even they showed the *cal-like* phenotype. As shown on Fig. 6B, all the cal-like F1 plants tested possess the transgene, suggesting that their phenotype is due to the presence of the *35S::iCAL* transgene and not due to transmission in the male gametes the epigenetic state of the *cal-like* parent (Fig 3.6B).

3.5 Discussion

Multiple levels of intron-mediated gene regulation

Previous data on intron-mediated gene silencing for *AG* (Zongrang Liu, unpublished) suggests that large introns (>1kb) may have the ability to trigger RNA-directed DNA methylation. To test how prevalent this intron-mediated gene silencing is and what properties an intron should possess in order to trigger silencing of endogenous genes, we overexpressed introns of five different genes and analyzed transgenic plants for mutant phenotypes. These introns are chosen because of their large sizes and possibility of possessing regulatory elements. We show that another flower gene, *CAL,* can be silenced by its largest intron. We were not able to detect the intron-mediated silencing phenomenon for any of the other four genes tested, suggesting that intron-mediated silencing is specific to

certain introns, which possibly possess poorly characterized regulatory elements such as intron-derived small RNAs or cryptic promoters. The *CAL* intron appeared not to contain any promoter elements as when we generated 5 *iCAL::YFP* and 15 *iCAL::GUS* transgenic lines. 5 *iCAL::YFP* and 15 *iCAL::GUS* were analyzed but no signal was observed in flowers. (B. Grigorova and C. Mara, unpublished). However, we are able to detect a small RNA generated from the intron in *cal-like* plants, suggesting that the *CAL* first intron can produce a novel siRNA. The intron-mediated silencing in this study differs from previously observed mechanisms of intron-mediated gene regulation in that the small RNA involved is introduced exogenously and not encoded by the intron of the target gene. Nevertheless, our work indicates that intron-mediated gene silencing may potentially prove a common gene regulatory mechanism as it is the second intron known to have such an effect in addition to the second intron of *AG* (Zongrang Liu, personal communication).

An intron-derived small RNA silences CAL

We identified an intron-encoded siRNA within the first intron of the *CAL* gene that correlated with the ability to silence *CAL* function. Surprisingly, we detected no change of *CAL* RNA levels in *ap1-like* flowers compared to *cal-like* flowers. However, the extra meristematic tissue of the *cal*-like flowers may be masking the reduction of the *CAL* RNA levels in these mutants. Thus, we compared *CAL* expression in *cal*-like vs. *ap1-1 cal-1* flowers and found a reduction in *CAL* expression. Additionally, the in situ hybridization data showed that the *CAL* intronic siRNA mediates an alteration in the spatial domain of *CAL*

within the flower, limiting *CAL* expression to the center of the floral meristem and reducing the expression level. This data elaborates on previously hypothesized RNAi mechanisms suggesting that siRNAs may mediate gene silencing in a temporal and spatial manner. Since the small RNA produced from the intron has no homology to the mRNA of *CAL* the silencing is likely caused by TGS, rather than PTGS and this is supported by the DNA methylation observed in *cal-like* plants but not in *ap1-like* plants.

Time-dependent epigenetic silencing of CAL by an intron-derived siRNA

We propose that an siRNA derived from the first intron of *CAL* can direct the epigenetic silencing of the *CAL* gene. Specifically, the novel siRNA may antagonize the function of its host gene by mediating the methylation of the first intron. Methylation of the intron may result in chromatin condensation of the entire gene locus, which prevents transcription by RNA polymerase II. Furthermore, these epigenetic modifications are stable and inherited in subsequent generations similar to genomic imprinting in mammalian cells. Thus, next generation plants lacking the *35S::iCAL* transgene still exhibit the *cal-like* mutant phenotype and this can be maintained in several generations afterwards. Additionally, methylation occurs in a time-dependent manner whereby seeds that initially yield a wild-type phenotype can be converted to seeds that yield a *cal-like* mutant phenotype. Namely, T1 *cal*-like plants segregate *ap1-like* and *cal-like* in a 1:1 ratio, but after approximately 3 weeks 100% of the plants have a *cal-like* phenotype. However the *ap1-like* plants in T1 are able to segregate *cal-like* plants in T2 but they do not exhibit the time-dependent change in the ratio between *ap1-*

like and *cal-like* T2 progeny. We propose that the *cal-like* and *ap1-like* T1 plants may be due to differences in the dosage of the transgene expression. Specifically, *ap1-like* T1 plants may carry only one copy of the transgene and a low concentration of siRNA transcripts. In contrast, T1 *cal-like* plants may carry several insertions of the transgene and a high concentration of the siRNA transcripts. Thus, in the T2 generation only those progeny of the *ap1-like* parent carrying the transgene (*35S::iCAL*) show a *cal-like* phenotype as they need to consistently synthesize enough siRNA to cause a phenotype. Alternatively, T2 progeny from *cal-like* parents all have *cal*-like phenotypes due to either the presence of the transgene or the siRNA molecules that could be inherited from the germline. Conversion of 100% of the T2 progeny to the *cal-like* phenotype may require one month for the siRNA transcript concentration to reach threshold levels. Our data suggests this time-dependent epigenetic process occurs in the seeds but further analysis is required to understand this phenomenon in greater detail.

Significance and application of intron-derived siRNA

Our study provides a new frontier in the discovery of new gene silencing methods. We show that overexpression of the *CAL* intron leads to the silencing of the *CAL* gene and methylation of the gene locus by an siRNA in *Arabidopsis*. In the same manner, constructs containing promoter-driven introns, containing small RNA molecules, can potentially be exploited in a unique fashion to generate transgenic *Arabidopsis* and crop plants carrying silenced genes of interest without necessarily carrying the transgene. For example, the intron of the *CAL* homologue

in strawberry could be used to create *cal-like* strawberry with increased fertility. These strawberry mutants could have wide implications as an attractive alternative to traditional transgenic crops (GMOs) since they do not require the presence of the transgene for enhancement of specific traits, which can be stably inherited.

Further investigation into similar stably inherited intron-derived siRNAs may provide novel strategies to combat genetic diseases in humans. Specifically, it has been proposed that there might exist a correlation between human disease and intronic small RNAs, as numerous introns containing small RNAs seem to be involved in RNAi-related chromatin silencing mechanisms (Jin et.al., 2004). For instance, fragile X syndrome occurs as a result of an erroneous intronic expansion, resulting in dysregulation of a specific 3'-UTR intronic microRNA that leads to heterochromatin repression of the whole gene locus. Such an event alters the condition-specific and time-specific manner of expression of a intron-derived small RNA encoding gene and results in the genetically inherited mental retardation that characterizes fragile X syndrome. Thorough understanding of the mode of action of the *CAL* intronic siRNA may provided the groundwork for the successful design of siRNA-based drugs for future gene therapies, aimed to cure or alleviate the symptoms of genetic diseases. Moreover, the construction of an artificial intron-derived small RNA system has recently become a successful strategy for knockdown of selected oncogenes and viral genome replication (Lin et.al., 2004). Man-made introns carrying small RNA precursors have already been successfully used in triggering RNAi-like gene silencing in human prostate cancer

cells. Conceivably, intronic small RNAs can be also used as an effective tool to generate not only transgenic mammalian cells, but also transgenic plants for agricultural purposes. Similar constructs can potentially be used to enhance crop disease resistance and yield through the long-term and efficient suppression of specific plant genes.

Acknowledgements

I would like to thank Dr. Zongrang Liu and his graduate student Yang Yangzhou for their contribution towards the bisulfite sequencing experiment and personal communication regarding the intron-mediated regulation of *AGAMOUS*. Also, I am grateful to Dr. Chloe Mara for her collaboration during this project.

Chapter 4: Conclusions

4.1 Summary

Small RNAs, including microRNAs (miRNAs) and small interfering RNAs (siRNAs) play a key role as gene regulators in both higher plants and animals. These evolutionary conserved 21-24 nucleotide long RNA molecules have generated considerable interest in developmental biology since their discovery in the worm *Caenorhabditis elegans* more than a decade ago. Research in various model organisms has provided abundant evidence for their role in many molecular interactions including regulation of gene expression, developmental timing, defense against viruses and stem cell maintenance. The functions of small RNA molecules are still largely unknown. Both microRNAs and siRNAs are negative regulators of gene expression that guide target mRNAs for degradation through the RISC complex. Regulation of gene expression requires precise control and synchronization as well as perception and integration of cellular and environmental signals. In plants, the formation of flowers depends not only upon coordination of the floral homeotic genes but also upon other regulatory factors. In this thesis, I explore the role of microRNAs in flower development. Specifically, I show that the plant co-repressor LUG directly and negatively regulates *miR172*, while indirectly and positively regulates *AGO1*. The repression of *miR172* expression occurs at the promoter level via the tethering of a LUG-SEU-AP2 co-repressor complex to a putative AP2 binding site located within the *miR172* promoter. Thus, my research reveals the existence of a negative feedback loop between *miR172* and AP2, which ensures their correct expression pattern

within the flower domains. In addition, I describe and characterize a novel small interfering RNA encoded within the intron of the meristem- specific gene *CAULIFLOWER*. This intron-derived siRNA mediates the transcriptional and epigenetic silencing of *CAL* in a time-dependent manner whose mechanisms remains largely unknown.

The Liu lab research is focused on understanding the molecular mechanisms of flower development in *Arabidopsis* as well emerging plant model systems such as strawberry. Specifically, this lab has extensively characterized the regulatory role of *LEUNIG,* a transcriptional co-repressor. The ability of plants to produce specific cell types and organs at the right time and place during flower development depends largely on the coordination of the ABC homeotic genes. The domains of their expression are regulated by the integrated expression and function of LEUNIG and *miR172*. Additionally, the lab is also expanding its research goals into characterizing novel modes of gene regulation such as intron-encoded siRNAs and advancing new approaches to the development of transgenic crops.

My work shows the LUG co-repressor complex negatively controls *miR172*, an important regulator of the A-class gene *AP2*. In recent years much research has been focused on understanding the role of miR172 in flower development, but how the miRNA itself is regulated remains largely unknown. The tissue-specific and cell-specific function of miRNAs indicates that miRNA gene expression is precisely regulated in the plant. Conceivably, their abundance in cells can be controlled on multiple levels. Currently available data shows that

miRNA accumulation depends upon the correct function of several components of the siRNA pathway. However, my study shows how *miR172* is regulated on the level of gene expression through an AP2 binding site. Other miRNA promoters have also been shown to contain binding sites for known transcription factors which indicates that various higher order regulatory complexes could be involved in maintaining the correct miRNA gene expression. Because LUG is a transcriptional co-repressor and the *35S::AGO1 (lug-3)* mutant phenotype defects can only be partially explained with an increase in miR172 expression, it is possible that LUG is a component in more than one of those transcription repression complexes.

AGO1 plays a crucial role in the RNAi pathway as the slicer component of the RISC complex. Therefore, understanding how *AGO1* is regulated remains a focus for small RNA research. Currently, there is only one known direct regulator of *AGO1, miR168.* My research shows that LUG indirectly and positively regulates *AGO1.* Because *miR168* levels are unchanged in *lug-3,* there must exist another currently unknown direct inhibitor of AGO1.

Central to the ABC model is the antagonism between A class and C class genes. The molecular mechanisms underlying this mutual inhibition remain elusive. In my thesis, I provide new insights into how the A-C antagonism occurs. It indicates that while *miR172* acts to negatively regulate *AP2* in the inner two whorls, AP2 itself represses *miR172* expression in the outer two whorls of the flower. Thus, *miR172* and AP2 are components of a negative feedback loop, which ensures their correct spatial expression pattern. A focus for future research

could be the discovery of similar negative or positive regulatory feedback loops for other miRNAs and their target genes.

Another aspect of my thesis works is characterizing intron-mediated gene regulation. In this study I show that a floral-specific gene, *CAULIFLOWER,* is negatively regulated by its largest intron. Specifically, I have identified and characterized a novel small RNA encoded within the largest intron *CAULIFLOWER* that silences its expression in a temporal and spatial manner. The intron-derived siRNA can direct not only the transcriptional, but the epigenetic silencing of the gene. Specifically, the novel siRNA antagonizes the function of *CAL* by mediating the methylation of its first intron. Additionally, methylation occurs in a time-dependent manner whereby plants that initially show a wild-type phenotype convert to a mutant phenotype in the course of 10-14 days, suggesting that the siRNA transcripts may be stably inherited through the germline and able to mediate the epigenetic silencing of *CAL* within the dormant seeds. Further analysis needs to be conducted in order to understand this phenomenon in more detail.

4.2 Future directions

My research indicates that LUG directly and negatively regulates *miR172, miR165, miR166* and indirectly and positively regulates *AGO1*. The severe developmental defects of *lug-3* plants with restored WT *AGO1* levels suggest that LUG may regulate more than one miRNA. Future studies may focus on identifying and characterizing additional miRNA targets of LUG. Additionally,

genome-wide microarray experiments could be used to examine the levels of other miRNAs in *lug-3* mutants.

The only known direct regulator of *AGO1* is *miR168*. Since levels of *miR168* remain unchanged in *lug-3* flowers, we hypothesize that there exists another, unknown direct negative regulator of *AGO1*. Although the identity of this regulator remains elusive, our data suggests that LUG acts upstream of it to inhibit its function. Possibly, a microarray screen using an inducible p*LUG::LUG-GR* system could identify additional direct targets of LUG, one of which may be involved in regulating *AGO1*.

My studies of the *CAULIFLOWER* intron provide new insights into intron-mediated gene regulation. Previous research attributes regulatory role to introns, but the mechanisms remain elusive. I identified a novel small interfering RNA encoded within the first intron of the *CAL* gene. The exact location of the siRNA within the intron as well as its sequence still needs to be determined. In order to characterize the siRNA more fully different segments spanning the entire *CAL* intron could be used as probes in a small RNA blot experiment. Additionally, deep sequencing could be a viable option for determining the exact coding sequence of the novel intronic siRNA.

I show that an siRNA derived from the first intron of *CAL* can direct not only the transcriptional, but the epigenetic silencing of the gene. Specifically, the novel siRNA antagonizes the function of its host gene by mediating the methylation of the first intron. Additionally, methylation occurs in a time-

dependent manner whereby plants that initially show a wild-type phenotype convert to an *ap1-1 cal1-1* mutant phenotype. This conversion takes places in the course of approximately 10-14 days. A Northern blot or a qRT-PCR analysis performed with seed RNA collected at different time-points could determine whether an increase in the siRNA levels occurs in the dormant seeds.

My study reveals the regulatory role of a specific intron-derived siRNA in the cellular processes. Intronic small RNA molecules might play a central role in more than one gene-silencing pathways both at the transcriptional and post-transcriptional level. Thus, constructs with promoter-driven introns, containing small RNA molecules, can potentially be exploited to generate transgenic *Arabidopsis* and crop plants carrying silenced genes of interest without necessarily being mutants. For example, the intron of the *CAL* homologue in strawberry could be used to create *cal*-like strawberry mutants with increased fertility. These strawberry mutants could have wide implications as an attractive alternative to traditional transgenic crops (GMOs) since they do not require the presence of the transgene for enhancement of specific traits, which can be stably inherited. Current work in our lab is focused on identifying the *CAULIFLOWER* homolog in strawberry and overexpressing its intron to yield similar results in a crop plant.

References

Ambros, V., Lee R.C., Lavanway A., Williams P.T., Jewell D. (2003). MicroRNAs and other tiny endogenous RNAs in *C.elegans.* Current Biology **10**: 807-818.

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

Aufsatz W., Mette M.F., van der Winden J., Matzke M., Matzke A.J.M. (2002). HDA6, a putative histone deacetylase needed to enhance DNA-methylation induced by double-stranded RNA. EMBO J. **21**: 6832-6841.

Azhakanandam S, Nole-Wilson S, Bao F, & Franks RG. (2008). SEUSS and AINTEGUMENTA mediate patterning and ovule initiation during gynoecium medial domain development. Plant Physiol **146:** 1165-1181.

Bao N., Lye K. and Bathron, K. (2004). MicroRNA Binding Sites in *Arabidopsis* Class III HD-ZIP mRNAs Are Required for Methylation of the Template Chromosome. Dev Cell **7**: 653-662.

Baumberger N & Baulcombe DC. (2005). Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. Proc Natl Acad Sci U S A **102:** 11928-11933.

Bohmert K, *et al.* (1998). AGO1 defines a novel locus of Arabidopsis controlling leaf development. Embo J **17:** 170-180.

Bowman, J.L., Alvarez, J., Weigel, D., Meyerowitz, E.M. and Smyth, D.R. (1993). Control of flower development in *Arabidopsis thaliana* by *APETALA1* and interacting genes. Development **119**, 721-743.

Bowman, J. L., Smyth, D.R., and Meyerowitz, E.M. (1991). Genetic interactions among floral homeotic genes of *Arabidsopsis*. Developmet **112**, 1-20.

Bowman, J. L., Smyth, D.R., and Meyerowitz, E.M. (1989). Genes directing flower development in *Arabidopsis*. The Plant Cell **1**, 37-52.

Busch M.A., Bomblies K., Weigel D. (1999). Activation of a floral homeotic gene in Arabidopsis. Science **285**:585-587.

Byzova MV, Franken J, Aarts MG, de Almeida-Engler J, Engler G, Mariani C, Van Lookeren Campagne MM, Angenent GC. (1999*)*. Arabidopsis STERILE APETALA, a multifunctional gene regulating inflorescence, flower, and ovule development.
Genes Dev **13:** 1002-14.

Callis J., Fromm M., Walbot V. (1987). Introns increase gene expression in cultured maize cells. Genes Dev **1**: 1183-1200.

Cao X., Jacobsen S.E. (2002). Role of the *Arabidopsis* DRM methylatransferases in de novo DNA methylation and gene silencing. Curr Biol **13**:2212-2217.

Carr SM & Irish VF. (1997). Floral homeotic gene expression defines developmental arrest stages in Brassica oleracea L. vars. botrytis and italica. Planta **201**: 179-188.

Chan S.W., Henderson I.R., Jacobsen S.E. (2005). Gardening the genome: DNA methylation in *Arabidopsis thaliana*. Nat Rev Genet **6**:351-360.

Chen, X. (2003). A MicroRNA as a Translational Repressor of *APETALA2* in *Arabidospis* Flower Development. Science **303**, 2022-2025.

Chung, B.Y., Simons C., Firth A.E., Brown C.M., Hellens R.P. (2006). Effect of 5'UTR introns on gene expression in *Arabidopsis thaliana.* BMC Genomics **7**: 120.

Coen, E.S., and Meyerowitz, E.M. (1991). The war of the whorls: genetic interactions controlling flower development. Nature *353, 31-37*.

Conner J & Liu Z. (2000). LEUNIG, a putative transcriptional corepressor that regulates AGAMOUS expression during flower development. *Proc Natl Acad Sci U S A* **97:** 12902-12907

Ditta, G, Pinyopich A., Robles P., Pelaz S., and Yanofsky M.F. (2004). The SEP4 gene of Arabidopsis thaliana functions in floral organ and meristem identity. Curr. Biol. **14**, 1935–1940.

Deyholos M.K., Sieburth L.E. (2000). Separable Whorl-Specific Expression and Negative Regulation by Enhancer Elements within the AGAMOUS Second Intron. The Plant Cell **12**: 1799-1810.

Drews, G.N., Bowman, J.L. and Meyerowitz, E.M. (1991). Negative regulation of the Arabidopsis homeotic gene AGAMOUS by the APETALA2 product. Cell **65**:991-1002.

Early K.W. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. Plant Journal **445:** 616-629.

Franks RG, Wang C, Levin JZ, & Liu Z. (2002). SEUSS, a member of a novel family of plant regulatory proteins, represses floral homeotic gene expression with LEUNIG. *Development* **129:** 253-263.

Gonzalez D, Bowen AJ, Carroll TS, & Conlan RS. (2007). The transcription corepressor LEUNIG interacts with the histone deacetylase HDA19 and mediator components MED14 (SWP) and CDK8 (HEN3) to repress transcription. Mol Cell Biol **27:** 5306-5315.

Gregis V, Sessa A, Dorca-Fornell C, & Kater MM. (2009). The Arabidopsis Floral Meristem Identity Genes AP1, AGL24 and SVP Directly Repress Class B and C Floral Homeotic Genes. Plant J.

Gregis V, Sessa A, Colombo L, & Kater MM. (2006). AGL24, SHORT VEGETATIVE PHASE, and APETALA1 redundantly control AGAMOUS during early stages of flower development in Arabidopsis. *Plant Cell* **18:** 1373-1382.

Gustafson-Brown, C., Savidge, B. and Yanofsky, M. (1994). Regulation of the *Arabidopsis* floral homeotic gene *APETALA1*. Cell 76:131-143.

Hammond SM, Caudy AA, Hannon GJ. (2001) Post-transcriptional Gene Silencing by Double-stranded RNA. Nature Rev Gen **2**: 110-119.

Hollender CaL, Z. (2010). Bimolecular Fluorescence Complementation (BiFC) Assay for Protein-Protein Interaction in Onion Cells Using the Helios Gene Gun. Journal of Visualized Experiments. In press.

Honma, T., and Goto, K. (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature **409**, 525-529.

Huettel B., Kanno T., Daxinger L., Buchner E., Van der Winden J., Matzke A., Matzke M. (2007). RNA-directed DNA methylation mediated by DRD1 and Pol IVb: A versatile pathway for transcriptional gene silencing in plants. EMBO **1769**: 358-374.

Irish, V. and Sussex, I. (1990). Function of the apetala-1 gene during Arabidopsis floral development. The Plant Cell **2**, 741-753.

Ito T, Takahashi N, Shimura Y, & Okada K (1997.) A serine/threonine protein kinase gene isolated by an in vivo binding procedure using the Arabidopsis floral homeotic gene product, AGAMOUS. Plant Cell Physiol **38:** 248-258.

Jack, T. (2001). Relearning our ABCs: new twists on an old model. Trends in Plant Science **6**, 310-316.

Jeong Y.M., Mun J. H., Lee I., Woo J.C., Hong C.B., Kim S.G. (2006). Distinct roles of the first introns on the expression of *Arabidopsis* profiling gene family members. Plant Physiology **140**: 196-209.

Jeong Y.M., Jung E.J., Hwang H., Kim H., Lee S., Kim. S. (2008). Roles of the first intron on the expression of Arabidopsis (*Arabidopsis thaliana*) genes for actin and actin-binding proteins. Plant Science **176**: 58-65.

Jin P., Alisch R.S. and Warren S.T. (2004). RNA and microRNAs in fragile X mental retardation. Nat Cell Biol **6**: 1048-1053.

Jofuku KD, den Boer BG, Van Montagu M, & Okamuro JK. (1994). Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. *Plant Cell* **6:** 1211-1225.

Jung JH, *et al*. (2007). The GIGANTEA-regulated microRNA172 mediates photoperiodic flowering independent of CONSTANS in Arabidopsis. Plant Cell **19:** 2736-2748.

Kempin, S.A., Savidge B., and Yanofsky M.F. (1995). Molecular basis of the cauliflower phenotype in Arabidopsis. Science **267**:522-525.

Kertesz S., Kerenyi Z., Merai Z., Bartos I., Palfy T., Barta E., Silhavy D. (2006). Both introns and long 3'-UTRs operate as *cis*-acting elements to trigger nonsense-mediated decay in plants. Nucl Acids Res **34**: 6147-6157.

Khorova A., Reynolds A. and Jayasena S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. Cell **115:** 209-216.

Kidner CA & Martienssen RA. (2005). The role of ARGONAUTE1 (AGO1) in meristem formation and identity. Dev Biol **280:** 504-517.

Kidner CA & Martienssen RA. (2004). Spatially restricted microRNA directs leaf polarity through ARGONAUTE1. Nature **428:** 81-84.

Kim M.J., Kim H., Shin J.S., Chung C.H., Ohlrogge J.B., Suh M.C. (2006). Seed-specific expression of sesame microsomal oleic acid desaturase is controlled by combinatorial properties between negative *cis*-regulatory elements in the *SeFAD2* promoter and ehancer in the 5'-UTR intron. Mol Genet Genomics **276**: 351-368.

Kooiker M., Airoldi C.A., Losa A., Manzotti P.S., Finzi L., Kater M.M., Colombo L. (2005). BASIC PENTACYSTEINE1, a GA binding protein that induces conformational changes in the regulatory regions of the homeotic *Arabidopsis* gene *SEEDSTICK*. Plant Cell **17**: 722-729.

Krizek BA & Fletcher JC. (2005). Molecular mechanisms of flower development: an armchair guide. Nat Rev Genet **6:** 688-698 (in eng).

Krizek, B. A. and Meyerowitz, E.M. (1996). The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. Development **122**: 11-22.

Krizek BA & Sulli C. (2006). Mapping sequences required for nuclear localization and the transcriptional activation function of the Arabidopsis protein AINTEGUMENTA. Planta **224:** 612-62.

Lee R.C. and Ambros V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. Science **294:** 862-864.

Lee R.C., Feinbaum R.L., Ambros V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell **5**: 843- 854.

Lee, Y., Ahn C., Han J. (2003). The nuclear RNAse III Drosha initiates microRNA processing. Nature **425**: 415-419.

Lee, Y., Kim, M., Han J. (2004). MicroRNA genes are transcribed by RNA polymerase II. EmboJ. **23:** 4051-4060.

Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF. (2000). SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature **404**: 766-70.

Liljegren S.J., Gustafson-Brown C, Pinyopich A., Ditta G.S., Yanofsky M.F. (1999). Interactions among *APETALA1*, *LEAFY* and *TERMINAL FLOWER1* specify meristem fate. Plant Cell 11:1007-1018.

Lin S.L., Chang D., Wu D-Y, Ying S.Y. (2003). A novel RNA splicing-mediated gene silencing mechanism potential for genome evolution. Biochemical and Biophysical Research Communications. **3**: 754-760.

Lin S.L., Chuong C.M., Ying S.Y.(2001). A novel mRNA-cDNA interference phenomenon for silencing bcl-2 expression in human LNCaP cells. Biochemical and Biophysical Research Communications **3**: 639-644.

Lin S.L. and Ying S.Y. (2004). Novel RNAi therapy- Intron-derived microRNA drugs. Drug Design Rev **1**: 247-255.

Liu, Z. and Karmarkar, V. (2008). Croucho/Tup1 family co-repressors in plant development. Trends in Plant Sciences **13**, 137-144.

Liu, Z. and Meyerowitz, E.M. (1995). LEUNIG regulates AGAMOUS expression in Arabidopsis flowers. Development **121**:975-991.

Liu, Z. and Liu, Z. (2008). The second intron of *AGAMOUS* drives carpel- and stamen-specific expression sufficient to induce complete sterility in *Arabidopsis.* Plant Cell Rep **27**: 855-863.

Llave C., Kasschau K.D., Rector M. A., Carrington J. (2002). Endogeneous and silencing-sssociated small RNAs in plants. The Plant Cell **14:** 1605-1619.

Lohmann JU*, et al.* (2001). A molecular link between stem cell regulation and floral patterning in Arabidopsis. Cell **105:** 793-803.

Lu, S. and Cullen, B.R. (2003) Analysis of the stimulatory effect of splicing on mRNA production and utilization in mammalian cells. RNA **9**: 618–630.

Mallory AC, Elmayan T, & Vaucheret H. (2008). MicroRNA maturation and action--the expanding roles of ARGONAUTEs. (Translated from eng) Curr Opin Plant Biol **11:** 560-566.

Mallory AC & Vaucheret H. (2009). ARGONAUTE 1 homeostasis invokes the coordinate action of the microRNA and siRNA pathways. EMBO Rep **10**: 521-526.

Martin A, *et al.* (2009). Graft-transmissible induction of potato tuberization by the microRNA miR172. Development **136:** 2873-2881.

Matzke M.A., Birchler J.A. (2005). RNAi-mediated pathways in the nucleus. Nat Rev Genet **6**:24-35.

Maquat, LE. (2004). Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nature Reviews in Cell Biology **5**: 89-99.

Mette M.F., Aufsatz, J., van der Winden M., Matzke M., Matzke A.. (2000). Transcriptional gene silencing and promoter methylation triggered by double-stranded RNA. EMBO J **19:** 5194-5201.

Mizukami, T. and Ma, H. (1992) Ectopic expression of the floral homeotic gene AGAMOUS in transgenic Arabidopsis plants alters floral organ identity. Cell **71**, 119–131.

Moazed D. (2009). Small RNAs in transcriptional gene silencing and genome defence. Nature **457**: 413-420.

Mosher R., Melnyk C., Kelly K, Dunn R., Studholme D. & Baulcombe D. (2009). Uniparental expression of PolIV-dependent siRNAs in developing endosperm of *Arabidopsis*. Nature **46:** 283-286.

Mukhopadhyay A, Deplancke B, Walhout AJ, & Tissenbaum HA. (2008). Chromatin immunoprecipitation (ChIP) coupled to detection by quantitative real-time PCR to study transcription factor binding to DNA in Caenorhabditis elegans. Nat Protoc **3:** 698-709.

Nam, J., dePamphillis, C.W., Ma, H., and Nei, M. (2003). Antiquity and evolution of the MADS-box gene family controlling flower development in plants. Mol boil Evol **20**, 1435-1447.

Nott, A., Le Hir, H. and Moore, M.J. (2004). Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. Genes Dev. **18**: 210–222.

Okamuro, J.K., Caster, B., Villarroel, R., Van Montagu, M., and Jofuko, K.D. (1997). The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in *Arabidopsis*. PNAS **94**, 7076-7081.

Parcy F., Nilsson O., Busch M.A., Lee I., Weigel, D. (1998). A genetic framework for floral patterning. Nature **395**, 561-566.

Pelissier T., Thalmeir S., Kempe D., Sanger H-L., Wassenegger M. (1999). Heavy *de novo* methylation at symmetrical and non-symmetrical sites is a hallmark of RNA-directed DNA methylation. Nucl Acid Res **27**: 1625-1634.

Pellisier T., Wassenegger M. (2000). A DNA target of 30bp is sufficient for RNA-directed heterochomatin silencing and spreading. RNA **6:** 55-65.

Qi Y.,He X., Wang X.J., Kohany O., Jurka J., Hannon G.J. (2006). Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. Nature **443:** 1008-1012.

Rate D, Cuenca J., Bowman G., Guttman D., and Greenberg J. (1999). The Gain-of-Function **Arabidopsis acd6** Mutant Reveals Novel Regulation and Function of the Salicylic Acid Signaling Pathway in Controlling Cell Death, Defenses, and Cell Growth. The Plant Cell **11:** 1695-1708

Reinhart B.J., Slack F.J., Basson M., Pasquinelli A.E., Bettinger J.C., Rougvie A.E, Horvitz H.R. and Ruvkun G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenohabditis elegans. Nature **403**: 901-906.

Rodriguez A., Griffiths-Jones S., Ashurst J.L., Bradley A. (2004). Identification of mammalian micro-RNA host genes and transcription units. Genome Research **10A**: 1902-1910.

Rose A.B. (2002). Requirements for intron-mediated enhancement of gene expression in *Arabidopsis.* RNA **8**: 1444-1453.

Rose A.B. (2004). The effect of intron-mediated enhancement of gene expression in *Arabidopsis.* Plant Journal **40**: 744-751.

Schwarz D.S., Hutvagner G., Du T., Xu Z., Aronin N., and Zamore,P.D. (2003). Assymetry in the assembly of the RNAi enzyme complex. Cell **115:** 199-208.

Scwab R., Palatnik JF, Riester M., Schommer C, Schmid M. and Weigel D. (2005). Specific effects of microRNAs on the plant transcriptome. Dev Cell **4**: 517-527.

Sijen T., Vijn I., Rebocho A., van Blokland R., Roelofs D., Mol J.N., Kooter J.M. (2001). Transcriptional and posttranscriptional gene silencing are mechanistically related. Curr Biol **11:** 436-440.

Sitaraman J, Bui M, & Liu Z. (2008). LEUNIG_HOMOLOG and LEUNIG perform partially redundant functions during Arabidopsis embryo and floral development. Plant Physiol **147:** 672-681.

Sridhar VV, Surendrarao A, Gonzalez D, Conlan RS, & Liu Z. (2004). Transcriptional repression of target genes by LEUNIG and SEUSS, two interacting regulatory proteins for Arabidopsis flower development. Proc Natl Acad Sci U S A **101:** 11494-11499.

Sridhar, V.V., Surendrarao, A., and Liu, Z. (2006). APETALA1 and SEPALLATA3 interact with SEUSS to mediate transcription repression during flower development. Development **133**, 3159-3166.

Tang G, Reinhart BJ, Bartel DP, & Zamore PD (2003). A biochemical framework for RNA silencing in plants. Genes Dev **17:** 49-63.

Teixeira F.K., Heredia F., Sarazin A., Roudier F., Boccara M., Ciaudo C., Cruaud C., Poulain J., Berdasco M., Fraga M.F., Voinnet O., Wincker P., Esteller M., Colot V. (2009). A role for RNAi in the selective correction of DNA metylation defects. Science **323**: 1600 – 1604

Theissen G & Saedler H (2001) Plant biology. Floral quartets. Nature **409**:469-471.

Vaucheret H. (2008). Plant ARGONAUTES. Trends Plant Sci **13:** 350-358 (in eng).

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

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

Vaucheret H. (2009). AGO1 homeostasis involves differential production of 21-nt and 22-nt miR168 species by MIR168a and MIR168b. PLoS One **7:** e6442.

Wagner D., Sablowski R.W.M., Meyerowitz E.M. (1999). Transcriptional activation of APETALA1 by LEAFY. Science **285**, 582-584.

Walter M*, et al*. (2004). Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. Plant J **40**: 428-438.

Weigel D., Alvarez J., Smyth D.R., Yanofsky M.F., and Meyerowitz, E.M. (1992). *LEAFY* controls floral meristem identity in Arabidopsis. Cell **69**, 843-859.

Weigel D., Meyerowitz E.M. (1993). Activation of floral homeotic genes in Arabidopsis. Science **261**, 1723-1726.

Jonathan D. Werner , Justin O. Borevitz N. Henriette Uhlenhaut, Joseph R. Ecker , Joanne Chory and Detlef Weigel. (2005). FRIGIDA-Independent Variation in Flowering Time of Natural *Arabidopsis thaliana* Accessions. Genetics **170,** 1197-1207.

Wu G, *et al*. (2009). The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. Cell **138:**750-759.

Ying S.Y., Lin S.L. (2004). Intron-derived microRNAs-fine-tuning of gene function. Gene **1**: 25-28.

Ying S.Y. and Lin, S.L. (2005). Current perspectives in intronic micro RNAs (miRNAs). **13**: 5-15.

Zhang X., Henderson I., Lu C., Green P.J., and Jacobsen, S.E. (2007). Role of RNA polymerase IV in plant small RNA metabolism. PNAS **104**: 4536-4541.

Zhao L, Kim Y, Dinh TT, and Chen X. (2007). miR172 regulates stem cell fate and defines the inner boundary of APETALA3 and PISTILLATA expression domain in Arabidopsis floral meristems. Plant Journal **51**, 840-849.