Discovery of LEAFY Transcriptional Complex Components Necessary for Flower Formation in Arabidopsis thaliana

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Nirodhini Srimali Siriwadana, B.Sc.

Graduate Program in Department of Plant Cellular and Molecular Biology

The Ohio State University

2011

Dissertation Committee:

Dr. Rebecca Lamb, Advisor

Dr. David Somers

Dr. Iris Meier

Dr. Patrice Hamel
ABSTRACT

For successful reproduction, angiosperms must form fertile flowers in the appropriate positions and at the appropriate time. The timing of flower formation is especially important for annual plants, such as the model Arabidopsis thaliana used in this study. Floral fate is controlled by a set of genes termed floral meristem identity genes, the most important of which is the plant-specific transcription factor LEAFY (LFY). LFY’s function can be divided into two phases, which are both temporally and genetically separable. In the first phase, LFY controls the expression of genes, such as other floral meristem identity factors, that control phyllotaxy and organ number. In the second phase, LFY is necessary for activation of the floral organ identity genes, which are responsible for differentiating the organs of the flower from each other.

LFY is a transcription factor that does not have transcriptional activation or repression activities on its own and therefore must act in multiprotein complexes. Region- and stage-specific control of LFY target genes is provided by co-factors. However, until very recently little information was available about other factors that are present in such LFY-containing complexes. This work aims to identify components of LFY transcriptional complexes and characterize their function(s) in flower development. We demonstrate that LFY acts as a homodimer, which is mediated by a domain that is conserved across land
plant LFY orthologs, suggesting that dimerization is important for the function of the entire LFY family. We report the isolation of putative LFY-interacting proteins and their preliminary characterization. Finally, we present an in depth analysis of one such LFY-interacting protein, FRIEND OF LEAFY1 (FOL1), which is a novel C2H2 zinc finger transcription factor. FOL1 is expressed in inflorescence meristems, floral meristems, male and female gametophytes and in embryos. Based on gain of function and loss of function analyses, FOL1 appears to function in control of meristem size and gametophyte development. We also discuss FOL1’s physical and genetic interactions with LFY.

The conservation of LFY sequence and functions across angiosperms suggests that most if not all of the LFY-interacting proteins identified in Arabidopsis will also function with LFY orthologs in other plants, including important crops. Therefore this work will provide a basis for the development of new strategies to increase agronomical values such as increased yield by manipulating LFY-interacting proteins in economically important crops.
Dedicated to the most amazing person in my life, my mother (‘amma’).
ACKNOWLEDGMENTS

I could not have done without the help, guidance and critical thinking of my advisor, Dr. Rebecca Lamb. Her positive attitude towards my work and also a friendly atmosphere made my PhD possible. I would like to thank my committee, Drs. David Somers, Iris Meier and Patrice Hamel for their constructive criticism and suggestions, which made my research work, stand strong.

Two past lab members Dr. Sachin Teotia and Andrew Moreo initiated this project in the lab and deserves a special note for their work. A special thanks to former graduate student Dr. Srilakshmi Makkena for moral support, for technical help and being there for me as a friend. I would like to thank a former undergraduate, Matthew Tegowski, who helped me tremendously with the work in Chapter 4 and especially for making me laugh in stressful times. I wish to thank undergraduates, Sterling Field, Irene Gentzel and Alyssa Larue for helping me with my research work. My sincere thanks to Matteo Citarelli for the help received with the alignment figures and with computer issues.

I am specially thankful to my Aronoff and Rightmire “friends in science”, especially to Sowmya Venkatakrishnan my batch mate, not only for help in research work but also for being there for me every step of the way. Sivaramakrishnan Muthuswamy, Mintu Desai,
Joanna Boruc, Thushani Rodrigo-Peiris for encouraging words. I will never forget my dear friends Rosario Barbieri, Elisa Leyva-Guerrero, Vanessa Falcao and Zoe Gokhale with whom I spent most of my time in my first few years.

A special thanks to all the staff members of our department, especially to Rene Hickey, Laurel Shannon and Eduardo Acosta for helping me whole-heartedly whenever I needed help. Also a special thanks to the Biological Sciences green house staff, Joan Leonard and Emily Yoders-Horn for help and support in growing plants.

I thank ABRC for supplying us with seeds and Drs. T. Nakagawa, John Lindbo and Keith Slotkin for providing vectors.

Special thanks to my dear mother and my sister, without them I couldn’t have done all this. To my “friends in life” Sharon, Indi, Mekhala, Chamika, Poorani, Shekar, Harshi and Suresh for their encouraging words, laughs and relaxing times which no doubt gave me strength and motivation, especially in the last few months of my PhD.

Above all to my love, my husband Ravindra Amunugama for helping me in my most stressful times, in every possible way. Most of all, by loving me and by taking care of our little son. And finally to my happiness, my son Rehan Luc Amunugama for making me the happiest mom on earth, you are the joy of my life!
VITA

2003..............................................................B.Sc. Honors, Plant Biology, University of Colombo, Sri Lanka

2003-2004.................................................. Teaching Associate, Plant Biology,
University of Colombo, Sri Lanka

2004 to present ..........................................Graduate Teaching Associate, Department of Plant Cellular and Molecular Biology,
The Ohio State University, U.S.A

FIELD OF STUDY

Major Field: Plant Cellular and Molecular Biology
## TABLE OF CONTENTS

ABSTRACT ................................................................................................................................. ii

ACKNOWLEDGMENTS ................................................................................................................ v

VITA ........................................................................................................................................ vii

FIELD OF STUDY ....................................................................................................................... vii

TABLE OF CONTENTS ............................................................................................................... viii

LIST OF TABLES ........................................................................................................................ xii

LIST OF FIGURES ..................................................................................................................... xiii

CHAPTER 1: The Poetry of Reproduction: Flower Formation in *Arabidopsis thaliana* ................................................................................................................................. 1

1.1 Arabidopsis flower development ......................................................................................... 2
   1.1.1 Events at the transition to flowering ............................................................................. 2
       1.1.1a Converting the VM to IM - the roles of flowering pathways, flowering time
           regulators and integrators .............................................................................................. 4
       1.1.1b FM initiation - the role of the floral meristem identity genes ......................... 10
       1.1.1c Flower organ development - the role of organ identity genes ..................... 14
       1.1.2 Maintaining FM identity ......................................................................................... 16

1.2 LEAFY (LFY) .................................................................................................................... 19
   1.2.1 Evolution of LFY ......................................................................................................... 19
   1.2.2 Role of LFY in flowering ............................................................................................. 20
       1.2.2.1 LFY is involved in precise timing of the floral transition as well as
           maintaining floral fate ................................................................................................. 22
1.2.2.2 LFY controls floral homeotic gene expression and is necessary for floral
organ identity and differentiation .................................................. 24
1.2.3 LFY as a transcription factor .................................................. 24
1.2.4 LFY co-regulators ................................................................. 25

1.3 Research Objectives .................................................................. 28

1.4 Significance of this study .......................................................... 30

REFERENCES ............................................................................... 40

CHAPTER 2: Flower Formation in *Arabidopsis thaliana* is Dependent on
Dimerization of LEAFY ................................................................ 52

2.1 INTRODUCTION .................................................................. 53

2.2 RESULTS ............................................................................. 56
  2.2.1 LFY protein homodimerizes .................................................. 56
  2.2.2 LFY dimerizes through a conserved region in its N-terminus .... 57
  2.2.3 Overexpressing LFY dimerization domain interferes with endogenous LFY
      function ............................................................................. 59
  2.2.4 Mutations within the dimerization domain abrogate dimer formation .... 62
  2.2.5 Dimerization is essential for LFY function ................................ 63

2.3 DISCUSSION ...................................................................... 65
  2.3.1 LFY forms dimers ................................................................. 65
  2.3.2 LFY-LFY association is necessary for function ...................... 68
  2.3.3 Conclusion ......................................................................... 69

2.4 EXPERIMENTAL PROCEDURES .......................................... 70
  2.4.1 Plant material and growth conditions ..................................... 70
  2.4.2 Constructs and cloning .......................................................... 70
  2.4.3 Generation of Arabidopsis transgenic lines and transient infiltration of
      *N. benthamiana* leaves ............................................................... 72
  2.4.4 Yeast two-hybrid interaction assay .......................................... 73
  2.4.5 Bimolecular Fluorescence Complementation Assay ............... 73
  2.4.6 RNA isolation and quantitative Real Time-PCR .................... 74
  2.4.7 Computational analysis .......................................................... 74
  2.4.8 In vitro transcription/translation, invitro binding and electrophoretic mobility
      shift assay ........................................................................... 75
  2.4.9 Co-immunoprecipitation assay .............................................. 76
  2.4.10 GUS staining ..................................................................... 77
  2.4.11 *lfy* mutant complementation experiments ........................ 77
  2.4.12 Phenotypic analysis of transgenic lines .................................. 78
REFERENCES

CHAPTER 3: Identification of LEAFY Complex Components in Arabidopsis thaliana

3.1 INTRODUCTION

3.2 RESULTS

3.2.1 35S::LFY-TAP compliments lfy-6 mutants

3.2.2 Purification and identification of LFY complex components

3.2.3 Putative LFY complex components selected

3.2.4 LFY complex components are localized in the nucleus

3.2.5 CHR9 interacts with LFY

3.2.6 Loss of CHR9 causes seed lethality

3.2.7 LMI1 interacts with LFY

3.2.8 Overexpression of LMI1 causes floral defects

3.3 DISCUSSION

3.3.1 Putative LFY-interacting proteins identified by TAP purification

3.3.2 LMI1 interacts with LFY

3.3.3 Conclusions

3.4 EXPERIMENTAL PROCEDURES

3.4.1 Plant material and growth conditions

3.4.2 Constructs and cloning

3.4.3 TAP tag purification and detection

3.4.4 Mass spectrometry analysis

3.4.5 Yeast two-hybrid interaction assay

3.4.6 Bimolecular Fluorescence Complementation Assay

3.4.7 Computational analysis

3.4.8 Genomic DNA preparation and genotyping

3.4.9 Co-immunoprecipitation assay

3.4.10 lfy-6 mutant complementation experiments

3.4.11 Phenotypic analysis of transgenic lines

REFERENCES

CHAPTER 4: The C2H2 Zinc Finger Transcription Factor FRIEND OF LEAFY1 (FOL1) has Roles in Meristem Identity and Size Control in Arabidopsis thaliana

4.1 INTRODUCTION

4.2 RESULTS

x
4.2.1 *FOL1* has orthologs across the land plants .......................................................... 158
4.2.2 The spatial and temporal expression pattern of *FOL1* marks the sites of cell proliferation ........................................................................................................... 159
4.2.3 Phenotypic analysis of insertion mutant lines exhibits the necessity of *FOL1* for gametophyte development ...................................................................................... 161
4.2.4 Knock down of *FOL1* causes meristem defects ...................................................... 162
4.2.5 Ectopic expression of *FOL1* suggests a meristem function ........................................ 166
4.2.6 *FOL1* interacts with LFY ......................................................................................... 167
4.2.7 *FOL1* interacts genetically with LFY ........................................................................ 169

4.3 DISCUSSION ..................................................................................................................... 171
4.3.1 *FOL1* marks the sites of cell proliferation ............................................................... 171
4.3.2 *FOL1* is necessary for gametophyte development ................................................... 172
4.3.3 *FOL1* controls inflorescence meristem size .............................................................. 173
4.3.4 *FOL1* controls floral meristem identity and size ....................................................... 175
4.3.5 *FOL1* is a LFY interacting partner .......................................................................... 177
4.3.6 Conclusions ............................................................................................................... 178

4.4 EXPERIMENTAL PROCEDURES ................................................................................ 179
4.4.1 Plant material and growth conditions ...................................................................... 179
4.4.2 Constructs and cloning ............................................................................................. 179
4.4.3 Yeast two-hybrid interaction assay ......................................................................... 182
4.4.4 Bimolecular Fluorescence Complementation Assay (BiFC) ..................................... 182
4.4.5 Genomic DNA preparation and genotyping .............................................................. 183
4.4.6 RNA isolation and semi quantitative Real Time-PCR (RT-PCR) .............................. 183
4.4.7 Co-immunoprecipitation assay .............................................................................. 25
4.4.8 GUS staining ........................................................................................................... 184
4.4.9 Phenotypic analysis of transgenic lines .................................................................... 184
4.4.10 Alexander Staining ............................................................................................... 185
4.4.10 Computational Analysis ......................................................................................... 185

REFERENCES ...................................................................................................................... 206

A COMPLETE LIST OF REFERENCES ............................................................................. 210

APPENDIX A: A Complete list of proteins identified from TAP tag-purification ...... 230
LIST OF TABLES

2.1 Trangenic lines examined ........................................................................................................ 99
2.2 Both conserved domains and dimerization are required for LFY function ........ 100
2.3 List of primers used in Chapter 2 .......................................................................................... 101
2.4 Probes used for EMSA ........................................................................................................... 104
3.1 Putative LFY-interacting proteins selected from TAP purification ......................... 143
3.2 List of primers used in Chapter 3 ........................................................................................ 144
4.1 List of primers used in Chapter 4 ........................................................................................ 203
A.1 A complete list of proteins identified from TAP tag-purification ....................... 230
LIST OF FIGURES

1.1. The *Arabidopsis thaliana* reproductive transition ........................................ 31
1.2. Schematic representation of reproductive and floral transitions ...................... 32
1.3. Integration of flowering time pathways to regulate the expression of floral
    pathway integrators: *FT*, *SOC1* and *LFY* .............................................. 33
1.4. Regulation of FM identity ............................................................................. 34
1.5. The ABCE model of floral development ............................................................ 35
1.6. Maintenance of FM identity ............................................................................ 36
1.7. *LFY* is a floral meristem identity gene ......................................................... 37
1.8. Ectopic expression of *LFY* is sufficient to convert meristems into flowers ....... 38
1.9. LFY and its orthologs contain two conserved domains ..................................... 39
2.1. LFY homodimerizes in yeast ........................................................................... 79
2.2. LFY-LFY homodimers form *in planta* ............................................................. 80
2.3. LFY homodimers exist in both *N. benthamiana* leaves and Arabidopsis
    inflorescences ........................................................................................................ 81
2.4. Ectopic expression of *LFY* causes formation of solitary flowers .................. 82
2.5. A schematic representation of LFY deletion fragments used in this study ........ 83
2.6. The N-terminal conserved domain of LFY mediates homodimerization .......... 84
2.7. The entire conserved domain is necessary for LFY dimerization ...................... 85
2.8. The N-terminal conserved domain mediates LFY homodimerization \textit{in vivo}..... 86
2.9. Overexpression of LFY\textsubscript{Δ229-420}-GFP causes a \textit{lfy}-like phenotype.......................... 87
2.10. LFY\textsubscript{Δ229-420} forms a higher order DNA binding complex with LFY............. 88
2.11. The dominant negative phenotype caused by \textit{35S:: LFY\textsubscript{Δ229-420}-GFP} is due to
        reduced LFY target gene expression and correlates with transgene expression... 89
2.12. The LFY dimerization domain is conserved across land plants and may form an
        alpha helical structure ........................................................................................................... 91
2.13. Mutations within the LFY dimerization domain abrogate dimerization ........ 92
2.14. Mutations within the dimerization domain abrogate LFY interaction in yeast ... 93
2.15. Mutations within the dimerization domain abrogate LFY-LFY association \textit{in vivo}
        ........................................................................................................................................ 94
2.16. Both conserved regions of LFY are required for function................................. 95
2.17. Overexpression of LFY\textsubscript{L/A}-GFP causes weak \textit{lfy}-like phenotypes ............. 96
2.18. Dimerization is essential for LFY function \textit{in planta} .............................................. 97
2.19. LFY\textsubscript{L/A} cannot activate \textit{AP3} expression .................................................. 98
3.1. Tandem affinity purification of LFY-containing complexes from Arabidopsis
        inflorescences ........................................................................................................................ 134
3.2. Gene structures of the genes encoding putative-LFY binding proteins............. 135
3.3. Putative LFY-binding proteins are found in inflorescence and in same subcellular
        compartments as LFY ........................................................................................................... 136
3.4. LFY and CHR9 physically associate \textit{in vivo} ................................................................. 137
3.5. Reduced fertility is observed in \textit{chr9-1} plants ............................................................. 138
3.6. LMI1 interactions in Y2H assay .................................................................................... 139

xiv
3.7 LMI1 and LFY physically associate in vivo ........................................ 140
3.8 Overexpression of LMI1 causes floral development defects .................. 142
4.1 Structural features of FOL1 protein and gene ........................................ 186
4.2 Phylogenetic tree of FOL1 and its orthologs from selected land plants .......... 187
4.3 Alignment of FOL1 and its orthologs from selected land plants ................ 188
4.4 Expression pattern of FOL1 ................................................................. 189
4.5 FOL1p::GUS expression in developing flowers and embryos ................. 190
4.6 FOL1 expression is confined to very young seedlings and reproductive tissues 191
4.7 fol1-1 and fol1-2 have gametophyte defects ...................................... 192
4.8 FOL1 amiRNA lines have reduced transcript accumulation .................... 193
4.9 FOL1 knockdown causes meristem defects ........................................ 194
4.10 FOL1 knockdown causes meristem size defects .................................. 195
4.11 35S::FOL1 causes an arrested SAM in the L.er background .................... 196
4.12 Ectopic expression of FOL1 causes reduced stature and terminal flowers in the Col-0 background ......................................................... 197
4.13 FOL1 homodimerizes ................................................................. 198
4.14 FOL1-LFY interaction in planta ......................................................... 199
4.15 FOL1 interacts with the N-terminus of LFY ...................................... 201
4.16 FOL1 and LFY interact genetically ..................................................... 202
CHAPTER 1

The Poetry of Reproduction: Flower Formation in *Arabidopsis thaliana*
“The flower is the poetry of reproduction. It is an example of the eternal seductiveness of life.”

Jean Giraudoux

Reproduction is essential for all organisms. In angiosperms, the unit of reproduction is the flower. Flowering plants undergo a transition from a vegetative phase, during which leaves with associated lateral shoots are produced, to a subsequent reproductive phase, during which flowers are been made (Figure 1.1). Reproductive success depends on initiating flowering at the right time and maintaining reproductive fate until the plant successfully sets seeds. The precise timing of reproduction is especially important in annual plants, which only flower once. The transition to reproduction is influenced not only by internal developmental parameters but also by environmental signals the plant perceives. This introduction will concentrate on reproduction in Arabidopsis thaliana, a model annual angiosperm.

1.1 Arabidopsis flower development

1.1.1 Events at the transition to flowering

Plants, unlike animals, have indeterminate growth, which is mediated by meristems. Meristems are groups of undifferentiated stem cells that give rise to the plant body. At germination, seedlings contain two such meristems, the shoot apical meristem (SAM) and the root apical meristem (RAM). The SAM produces lateral organs at its flanks as well as producing those cells that form the plant stem. During vegetative growth, the SAM
produces lateral leaves with associated axillary meristems. In Arabidopsis, the transition to reproduction can be broken into two phases. During the first phase, the vegetative meristem (VM) becomes an inflorescence meristem (IM; Figure 1.2a). This transformation of the VM into an IM is tightly regulated by both endogenous (hormones, age) and environmental (quality and quantity of light, temperature and day length) pathways. These pathways are integrated by a group of genes known as flowering time genes. The IM will produce several cauline leaves with axillary meristems that recapitulate the development of the primary IM. In addition, Arabidopsis has a rosette growth habit during its vegetative phase, but upon reproductive transition, internode elongation begins, giving rise to the bolt. In the second phase of Arabidopsis reproduction, the lateral organs produced by the IM become flowers. These flowers are not subtended by leaves (bracts); these structures are suppressed. Floral fate of the lateral meristems is due to the action of a number of floral meristem identity genes.

Unlike shoots, flowers are determinate structures that give rise to a set number of floral organs and then cease growth. Floral meristems (FMs) initially have a growth phase during which they increase in size. They then begin to produce floral organs in a whorled pattern, starting at their flanks. These organs are (in order of position from outside to inside) sepals, petals, stamens and carpels (Figure 1.2b). The identities of these organs are controlled by the activity of floral homeotic genes, which are activated by the floral meristem identity genes.
1.1.1a Converting the VM to IM- the roles of flowering pathways, flowering time regulators and integrators

The transition to reproduction and the pathways and signals that control this transition are well studied. Experiments beginning in the 1920s have demonstrated that different plants have varying requirements to trigger flowering (Garner and Allard, 1920). In Arabidopsis, a number of forward genetic screens have identified many genes that are involved in control of flowering time. Subsequent genetic analysis has defined five pathways in Arabidopsis that control this process: photoperiod, vernalization, autonomous, gibberellic acid (GA) and developmental age pathways (Amasino, 2004; Araki, 2001; Bastow and Dean, 2003; Boss et al., 2004; Jack, 2004; Martinez-Zapater, 1994; Mouradov et al., 2002; Simpson and Dean, 2002; Sung and Amasino, 2004). Given that there are multiple pathways that regulate the floral transition and feed into an integrated output (Parcy, 2005), it is important to understand how these pathways interact. A recent review (Srikanth and Schmid, 2011) summarizes all the players of these five pathways in depth. Here, I will only concentrate on the confluence of these pathways on the floral transition through the floral pathway integrators: FLOWERING LOCUS T (FT), SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1) and LEAFY (LFY) as categorized in (Parcy, 2005). Expression of these floral integrators is controlled by multiple signals of the flowering time pathways. The combined input determines expression level, which in turns determines the time of reproduction.
Classical experiments revealed that the leaves are the site of signal perception in the photoperiodic pathway even though the events of floral invocation happen at the shoot apex (Garner and Allard, 1920; Garner and Allard, 1922; Zeevaart, 1976). These early experiments led to the investigation of the component(s) that perceive the signals in the leaves and get transported to the shoot apex to induce flowering, which were termed ‘florigen’. A key component of this mobile flowering signal is the highly conserved gene, \( FT \), which is one of the floral pathway integrators (Abe et al., 2005; Huang et al., 2005; Wigge et al., 2005). Mutations in \( FT \) cause significant delays in flowering and, conversely, overexpression causes precocious flowering, demonstrating that \( FT \) is necessary and sufficient for the reproductive transition (Kardailsky et al., 1999; Kobayashi et al., 1999; Koornneef, 1991). While a large number of genes are involved in triggering reproduction, the decision of when to flower largely depends on the level of \( FT \) expression. The \( FT \) gene encodes a small globular protein with similarities to Raf kinase inhibitory protein (RKIP) and a phosphatidylethanolamine-binding protein (PEBP) (Kardailsky et al., 1999; Kobayashi et al., 1999). This protein is expressed in the vasculature of the leaves and translocates to the SAM (An et al., 2004; Corbesier et al., 2007; Jaeger and Wigge, 2007; Lifschitz et al., 2006; Lin et al., 2007; Mathieu et al., 2007; Takada and Goto, 2003; Wigge, 2011). In the SAM, FT interacts with FLOWERING LOCUS D (FD; Abe et al., 2005; Wigge et al., 2005), a basic leucine zipper transcription factor that is expressed in shoot apex. FT and FD act in a transcriptional complex known to activate the floral meristem identity genes (Abe et al., 2005; Wigge et al., 2005).
SOC1, another floral pathway integrator, encodes a MADS box transcription factor. SOC1 is expressed in the leaves and in the shoot apex. Transcription of this gene undergoes a sharp increase in the shoot apex at the floral transition (Borner et al., 2000; Lee et al., 2000; Samach et al., 2000). Many flowering pathway inputs activate SOC1 (Hepworth et al., 2002; Moon et al., 2003; Searle et al., 2006). SOC1 has been shown to promote activation of the master regulator of flower identity LFY (Lee et al., 2008).

The final floral pathway integrator discussed in this introduction is LFY. LFY is a plant specific transcription factor (Busch et al., 1999) that acts as a meristem identity gene and a regulator of organ identity genes (which will be discussed in detail later in this chapter). It has been shown to regulate expression of other floral meristem identity genes (Busch et al., 1999; Parcy et al., 1998; Saddic et al., 2006; Wagner et al., 1999; William et al., 2004; Winter et al., 2011), which is why it is considered a floral integrator gene as well as a floral meristem identity gene.

Control of floral integrator expression by flowering time pathways

The photoperiodic pathway- Arabidopsis is a facultative long day plant and has much accelerated time to flowering under long days. The critical day length for Arabidopsis is about 16 hours of light/ 8 hours of dark. The cascade of events that measure day length and subsequently signal the initiation of flowering are referred to as the photoperiod pathway. Light is perceived by several classes of photoreceptors (phototropins, cryptochromes and phytochromes) in the leaves (Lariguet and Dunand, 2005; Li and
Yang, 2007; Toth et al., 2001). These photoreceptors regulate the expression of many genes through the circadian clock (Fornara et al., 2009; Imaizumi et al., 2005; Sawa et al., 2007) and a recent review (Srikanth and Schmid, 2011) summarizes the players involved. These players then have both direct and indirect effects on CONSTANS (CO) mRNA expression as well as its protein stability and accumulation (Imaizumi et al., 2005; Sawa et al., 2007; Suarez-Lopez et al., 2001; Valverde et al., 2004 and reviewed in Srikanth and Schmid, 2011). CO encodes a zinc finger transcription factor (Putterill et al., 1995), which is expressed in the vasculature of the leaves (Takada and Goto, 2003). It directly turns on FT expression in the leaves, thereby promoting flowering (Kobayashi et al., 1999; Samach et al., 2000; Takada and Goto, 2003; Wigge et al., 2005). CO is also known to induce expression of LFY (Nilsson et al., 1998a; Samach et al., 2000; Simon et al., 1996) and SOC1 (Lee et al., 2000; Samach et al., 2000), other floral pathway integrators (Figure 1.3).

**Temperature controls time to flower**—FT expression is regulated through temperature in two ways. Many plants need an exposure to a cold period (vernalization) to flower, including some ecotypes of Arabidopsis; in addition, the ambient temperature Arabidopsis grows in has an effect on the timing to flower, with higher temperatures accelerating the reproduction transition. The well-studied gene FLOWERING LOCUS C (FLC) is the key regulator in the vernalization pathway (Michaels and Amasino, 1999). FLC is a MADS box transcription factor that directly binds to regulatory regions of FT and SOC1 and represses their expression, thereby delaying flowering (Figure 1.3;
FLC is epigenetically silenced by vernalization, thus relieving the floral repression (Gendall et al., 2001; Levy et al., 2002). As shown in Figure 1.3, repression of FLC removes the antagonistic effect on FT and SOC1 thereby allowing their expression.

**The GA pathway**- Under short days Arabidopsis will still flower, although it will take much longer. GA is crucial for flowering under these conditions (Blazquez et al., 1998; Wilson et al., 1992). GA regulates the expression of FT (Figure 1.3; (Hisamatsu and King, 2008)). In addition, when GA, which is mobile, enters the SAM it activates SOC1 and LFY (Figure 1.3) via the repression of the DELLA proteins, which are growth-suppressing factors and through MYB transcription factor expression (Blazquez et al., 1998; Moon et al., 2003).

**The autonomous pathway**- The constitutive or autonomous pathway regulates flowering independently of day length. The major players in this pathway encode mostly chromatin remodeling and RNA processing factors (see review (Srikanth and Schmid, 2011)). These genes promote flowering indirectly by repressing FLC expression (Koornneef et al., 1998) in both the leaf and SAM, thereby promoting expression of FT to trigger flowering (Figure 1.3).

**Developmental age** - Compared to above discussed pathways, not much is known about the pathways that perceive developmental age. However, expression of a recently
discovered microRNA, *miR156*, decreases with increasing age of the plant, suggesting that this might be a mechanism by which age is measured (Wang et al., 2009; Wu et al., 2009). The regulation of flowering by *miR156* is thought to involve miRNA-mediated repression of a class of genes encoding transcription factors that activate important flowering genes, the *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL)* family. The *miR156* regulated *SPLs* are expressed in SAM, where it is thought they have an important FT-independent role in promotion of flowering though activation of *SOC1* (Wang et al., 2009). Recent work also revealed that *LFY* is a direct target of *miR156*-regulated SPLs (Figure 1.3; (Yamaguchi et al., 2009)).

**Floral pathway integrators**

Overall, the above discussed flowering pathways converge mainly on three major floral pathway integrators, *FT*, *SOC1* and *LFY* (Figure 1.3; (Blazquez and Weigel, 2000; Kardailsky et al., 1999; Kobayashi et al., 1999; Lee et al., 2000; Lee et al., 2008; Liu et al., 2008; Simpson and Dean, 2002)). As discussed above, before the transition to flowering, *FLC* represses expression of *SOC1* and *FT*, inhibiting reproductive development. Floral promoters act to increase the expression of *FT* either by repressing the expression of *FLC* and thereby activating *FT* and *SOC1* expression or by directly activating the expression of *FT* and *SOC1* to activate floral meristem identity genes (Helliwell et al., 2006; Searle et al., 2006). *FT* might participate in the control of *SOC1* expression, although the evidence is not conclusive (Schmid et al., 2003). Several of the flowering time pathways also directly regulate expression of *LFY*. These integrators are
not only convergence points for flowering time pathways, but also appear to regulate each other’s expression. SOC1 binds to the LFY promoter in vivo to activate LFY expression directly (as shown in Figure 1.3 and Figure 1.4; (Lee et al., 2008; Liu et al., 2008)). A gradual increase in SOC1 expression in the SAM promotes a gradual increase in LFY in primordia until a threshold is reached, leading to FM identity. Mutant and overexpression studies of FT support the finding that FT upregulates LFY as well (Kardailsky et al., 1999; Schmid et al., 2003).

1.1.1b FM initiation-the role of the floral meristem identity genes

Once an IM is formed, it will begin generating FMs on its flanks after formation of 2-5 cauline leaves. Floral meristem identity genes are required to specify the lateral meristems as flowers. To date, the known meristem identity proteins in Arabidopsis are LFY (Blazquez et al., 1997; Nilsson et al., 1998a; Weigel et al., 1992), the related MADS box transcription factors APETALA1 (AP1; (Bowman et al., 1993; Mandel et al., 1992) (Ferrandiz et al., 2000b; Irish and Sussex, 1990)), CAULIFLOWER (CAL; (Ferrandiz et al., 2000b; Kempin et al., 1995)) and FRUITFUL (FUL; (Ferrandiz et al., 2000b)) and the class 1 HD-Zip transcription factor LATE MERISTEM IDENTITY1 and its paralogs (Saddic et al., 2006). All of the genes encoding these proteins are expressed in FMs (Hempel et al., 1997; Mandel et al., 1992; Saddic et al., 2006). In Arabidopsis, LFY and AP1 are the two most important floral meristem identity regulators (Bowman et al., 1993; Ferrandiz et al., 2000a; Huala and Sussex, 1992; Mandel and Yanofsky, 1995; Weigel et al., 1992). In addition to these positive promoters of floral identity, there is a negative
regulator of floral fate, *TERMINAL FLOWER1 (TFL1)*, which is expressed in the IM and encodes a member of the CETS (CENTORADIALIS (CEN), TFL1, FT) family, related to phosphatidylethenalamine binding proteins and FT (Liljegren et al., 1999; Ratcliffe et al., 1999). TFL1 prevents the IM from expressing floral meristem identity genes and becoming a flower, therefore maintaining its indeterminate nature.

LFY was first recognized for its function in flower meristem development. *lfy* mutants are slightly late flowering, produce extra cauline leaves and have abnormal floral-like structures in which there is homeotic transformation of floral organs to leaf-like structures (see Figure 1.7). LFY directly activates expression of the other floral meristem identity genes, *AP1, CAL, FUL* and *LMII* (Figure 1.4; (Saddic et al., 2006; William et al., 2004; Winter et al., 2011)). Combining a mutation in *LFY* with *AP1* and *CAL* mutations produces plants that have shoots in the place of flowers (Ferrandiz et al., 2000a). LFY will be discussed in more depth below.

The MADS box transcription factor-encoding gene *AP1* is the other major floral meristem identity regulator. It is expressed throughout the very young FM before becoming confined to the outer two whorls at stage 3 of floral development (Mandel et al., 1992). *AP1* expression is directly activated by LFY as well as by FT (Figure 1.4; (Abe et al., 2005; Wagner et al., 1999; Wigge et al., 2005)). *AP1* is involved in regulating genes promoting floral organ formation and repressing flowering time genes to maintain the floral fate of the meristem (Hill et al., 1998; Liu et al., 2007; Ng and Yanofsky, 2001;
In ap1 mutants, extra cauline leaves are made before the formation of flowers. ap1 flowers have leaf-like sepals and no petals, although stamen and carpel development are normal. In addition, ectopic flowers form in the axils of the leaf-like sepals, a phenotype that has been interpreted as a floral meristem identity defect (Bowman et al., 1993; Irish and Sussex, 1990). Loss of AP1 function is not as severe as loss of its orthologs in other groups of angiosperms due to the presence of the CAL gene. CAL is a paralog of AP1 found in Brassicas that is partially redundant with AP1 (Kempin et al., 1995). ap1-1; cal-1 mutants show complete transformation of flowers into IMs, although loss of CAL alone has no phenotype (Bowman et al., 1993). However, eventually the IMs of ap1-1; call-1 plants will form differentiated flowers that resemble the flowers of ap1 single mutants. This is due to the activity of FUL, the closest gene to AP1/CAL in the Arabidopsis genome (Ferrandiz et al., 2000b). Loss of all three of these genes leads to a severe meristem identity defect (Ferrandiz et al., 2000b). AP1, CAL, FUL appear to be potential activators of LFY (Figure 1.4; (Ferrandiz et al., 2000b)).

In addition to the genes discussed above, LMI1 and its paralogs are important for floral meristem identity (Saddic et al., 2006). An lmi1 mutant enhances the meristem defects of the weak lfy-10 allele, hence the name LATE MERISTEM IDENTITY1 (Saddic et al., 2006). lmi1 mutants on their own have very subtle or no meristem defects. LMI1 acts upstream of CAL and together with LFY regulates CAL expression directly (Figure 1.4; (Saddic et al., 2006)).
In Arabidopsis, the IM remains indeterminate and does not form a terminal flower. This is due to the activity of TFL1 in the IM. AP1 and LFY repress TFL1 expression in floral meristems, suppressing IM fate (Figure 1.4; (Liljegren et al., 1999; Parcy et al., 2002)). TFL1 in turn suppresses the expression of AP1 and LFY in the IM (Ratcliffe et al., 1998). The balance between these genes is what regulates shoot architecture (Bradley et al., 1997). In fact, during the domestification of soybean (Glycine max), mutant alleles of the GmTFL1 gene were selected because they conferred a determinate growth habit, an agronomically important trait (Tian et al., 2010).

**Flower development stages**

Morphological changes that take place during flower development have been well studied in Arabidopsis. Flower development can be divided into 12 stages (Smyth et al., 1990). The initial setting aside of FM founder cells from the IM is considered stage 0 (Long and Barton, 2000). At stage 1, the FM emerges as an outgrowth on the flanks of the IM, with each new FM forming at an angle of 130°-150° to the one before. At stage 2, the FM develops further into a ball like structure, which is now morphologically distinguishable from the IM. At stage 3, the primordia of the sepals arise and thereafter the rest of the floral organs appear in order. At stage 5, the petal and stamen primordia become visible. Stage 7 begins when primordial of the growing stamen becomes stalked at the base creating the filament. Stage 13 is when the buds open and anthesis takes place (Smyth et al., 1990).
1.1.1c Flower organ development - the role of organ identity genes.

Once the meristem has transitioned to a floral fate, it gives rise to a determinate floral structure, consisting of four whorls of organs: sepals, petals, stamens and carpels in a whorled phyllotaxy. The identity of the floral organs depends on the activity of floral homeotic genes, which can be divided into four classes, A, B, C and E (reviewed in (Krizek and Fletcher, 2005)). The A, B and C genes act in a combinatorial manner to specify each organ type. The relatively recently identified SEPALLATA genes (SEP1-4) function redundantly with other homeotic genes in specifying floral organs (class E; Figure 1.5; (Ditta et al., 2004; Pelaz et al., 2000)). Class A genes function alone to specify sepal identity in the outermost whorl. Class A and B genes together specify petals in the second whorl. Class B and C genes specify stamens in the third whorl and class C alone functions to specify carpal identity in the inner most whorl (Figure 1.5). A and C genes also negatively regulate each other (Bowman et al., 1991b; Coen and Meyerowitz, 1991; Weigel and Meyerowitz, 1994). All the floral homeotic genes encode MADS box transcription factors (Irish, 2010; Sablowski, 2010) with the exception of one APETALA2 (AP2), which encodes a founding member of a plant specific group of transcription factors, the AP2/EREBP family (Riechmann and Meyerowitz, 1998).

AP1 and AP2 are class A genes. As mentioned above, AP1 is also a floral meristem identity gene and is initially expressed throughout the floral meristem (Mandel et al., 1992). However, due to the negative regulation by the C gene product AGAMOUS (AG), AP1 expression is restricted to whorls 1 and 2 at stage 3 where it functions in sepal and
petal identity (Bowman et al., 1991a; Gustafson-Brown et al., 1994; Mandel et al., 1992). 

*AP2* is expressed throughout the floral meristem, but its mRNA gets restricted to whorl 1 and 2 through the action of a microRNA, *miR172*, which negatively regulates its accumulation in the 3rd and 4th whorls (Aukerman and Sakai, 2003; Chen, 2004).

In Arabidopsis, two B class genes, *APETALA3 (AP3)* and *PISTILLATA (PI)*, function to specify petal and stamen identity. *AP3* and *PI* and are expressed in whorls 2 and 3 (Bowman et al., 1989; Goto and Meyerowitz, 1994; Jack et al., 1992). Their gene products bind to DNA as an obligate heterodimer (Riechmann et al., 1996).

The C class gene *AG* ’s expression is restricted to whorls 3 and 4. As mentioned above, A and C gene functions negatively regulate each other. *AP1* together with two transcriptional repressors, *LEUNIG (LUG)* and *SEUSS (SEU)*, binds to regulatory sequences of *AG*, preventing *AG* expression in the outer two whorls (Sridhar et al., 2004; Sridhar et al., 2006).

The *SEP* genes (class E) are expressed throughout the floral meristem and are necessary to differentiate the floral organs from leaves. They encode four redundant MADS box transcription factors. A triple mutant of class E genes (*sep1; sep2; sep3*) has all floral organs replaced by sepals (Pelaz et al., 2000) while the quadruple mutant (*sep1; sep2; sep3; sep4*) has leaves in the place of floral organs (Ditta et al., 2004). These findings
suggest that the functions of class E genes are to pre-establish the floral context for the other organ identity genes to function in.

1.1.2 Maintaining FM identity

Once the FMs are made, it is important to maintain their identity so floral reversion is prevented. Floral reversion is the reinitiation of shoot growth from a partially formed flower and was first described in 1880 (Buckhout, 1880). In Arabidopsis, floral reversion has been observed in floral meristem identity mutants including *lfy* and *ap1* under the noninductive short day photoperiod (Bowman et al., 1993). This phenotype is partially due to the inappropriate expression of three flowering time genes, *AGAMOUS-LIKE 24 (AGL24), SHORT VEGETATIVE PHASE (SVP) and SOCI* (Liu et al., 2007; Yu et al., 2004). Loss of function mutants in these three genes individually or in combination reduces reversion defects seen in *ap1* mutants (Liu et al., 2007). It has been shown that AP1 binds directly to promoters of these genes and represses their expression (Figure 1.6a; (Gregis et al., 2008; Liu et al., 2007)). Overexpression of *AGL24* promotes the transformation of FMs into IMs, overexpression of *SOCI* further enhances this and overexpression of *SVP* converts the FMs into vegetative shoots (Liu et al., 2007; Yu et al., 2004). In contrast to AP1, LFY might not have a direct effect on these flowering time genes (Gregis et al., 2008; Yu et al., 2004). It is possible LFY maintains FMs through its regulation of *API* (Figure 1.6a; (William et al., 2004)). Clearly, an important function of floral meristem identity genes is to repress, directly or indirectly, expression of flowering time genes to maintain floral fate.
The class E organ identity genes also contribute in maintaining the FM identity by repressing flowering time gene activity. The study of the sep1; sep2; sep3; sep4 quadruple mutant has shown that in this background secondary meristems arise, similar to what is seen in ap1 mutants. This indicates that SEPs not only act as floral organ identity genes but also have a role in FM specification (Ditta et al., 2004). Chromatin immunoprecipitation has shown that SEP3 directly binds to the promoters of both AGL24 and SVP (Figure 1.6a; (Gregis et al., 2008)). sep mutants have AGL24 and SVP expressed in the FM beyond their normal time of expression. This data suggests that SEPs act as repressors of AGL24 and SVP to maintain FM identity (Figure 1.6a).

While flowering time genes must be turned off to ensure proper FM development, the plant needs to prevent early differentiation of the floral organs by controlling the expression of floral homeotic genes as well. Therefore, flowering time genes cooperate with floral meristem identity genes to repress floral homeotic gene expression until the proper time in FM development. The class B and C genes are repressed before stage 3 of floral development by the repression of SEP3 expression (necessary for their activation) by the direct activity of AGL24, SVP and SOC1 (Figure 1.6a; (Liu et al., 2009b)). In soc1; agl24;svp triple mutants, SEP3 together with LFY synergistically activates Class B and C genes early, resulting in floral defects (Liu et al., 2009b). Overall, maintaining AGL24, SVP and SOC1 at appropriate levels at appropriate times is crucial to FM. Overexpression of these genes will cause floral reversion where as lower expression will cause precocious differentiation of floral organs.
Another group of regulators repressing floral organ identity gene expression are the chromatin regulators. Epigenetic control of organ identity genes is an area of study in which there is growing interest. Flower development is particularly sensitive to epigenetic control, possibly due to the strict regulation required. The Polycomb Group (PcG) proteins inhibit transcription by maintaining a repressed chromatin state of inactive genes through the formation of multiprotein complexes that catalyze the tri-methylation of lysine 27 of histone H3 (H3K27me3) on target genes (Chanvivattana et al., 2004; Farrona et al., 2008; Goodrich et al., 1997; Katz et al., 2004). (Liu et al., 2009a) gives a list of such proteins that inhibit the transcription of floral homeotic genes. The first plant polycomb protein indentified, CURLY LEAF (CLF), was found because of its function in repressing AG and AP3 expression in leaves (Goodrich et al., 1997) and AG in the inflorescence stem and in the first two whorls of the flower (Goodrich et al., 1997; Schubert et al., 2006). EMRYONIC FLOWER 2 (EMF2) and FERTILIZATION INDEPENDENT ENDOSPERM (FIE), which are also PcG proteins, repress AP3 and AG during vegetative growth (Chen et al., 1997; Kinoshita et al., 2001; Yoshida et al., 2001). EMBRYONIC FLOWER 1 (EMF1) together with EMF2 represses AG expression during vegetative development (Calonje et al., 2008). TERMINAL FLOWER 2 (TFL2), also known as LIKE-HETEROCHROMATIN PROTEIN 1 (LHP1), is expressed in proliferating cells of meristems, including in the FM, and binds to chromatin marked with H3K27me3 (Zhang et al., 2007). TFL2 directly associates with regulatory regions of AP3, PI, AG and SEP3 and represses their expression during vegetative growth (Figure
Clearly, epigenetic control of floral homeotic genes is essential for proper development.

Flowers are determinate structures. Therefore, they must stop organ production and growth after formation of the two carpels of the 4th whorl. The C class gene \( AG \) is necessary for this. WUSCHEL (WUS) is a homeobox transcription factor that is necessary to maintain all shoot meristems in an indeterminate state (Lenhard et al., 2001; Lohmann et al., 2001). AG represses \( WUS \) expression in the FM, terminating meristematic identity of the floral meristem (Figure 1.6a; (Lenhard et al., 2001; Lohmann et al., 2001)).

1.2 LEAFY (LFY)

LFY is a plant-specific transcription factor that does not belong to any known transcription factor family (Busch et al., 1999). Unlike many other transcription factors that have evolved by gene duplication to form multigene families (Riechmann and Ratcliffe, 2000), \( LFY \) is present as a single copy in most of the angiosperms. This makes LFY unique among transcription factors in plants.

1.2.1 Evolution of LFY

The \( LFY \) gene is conserved throughout land plant species, from bryophytes (the moss \( Physcomitrella patens \)) to flowering plants (Maizel et al., 2005). In moss, the \( LFY \)
orthologs \textit{PpLFY1} and \textit{PpLFY2} regulate the first division of the zygote (Tanahashi et al., 2005). In gymnosperms and angiosperms, \textit{LFY} is associated with reproductive structure formation (cones and flowers, respectively). Studies in gymnosperms identified two \textit{LFY}-like genes: \textit{LFY} and \textit{NEEDLY (NLY)} (Albert et al., 2002); at the base of the angiosperms, the \textit{NLY}-like gene disappeared and only the \textit{LFY}-like gene persisted (Albert et al., 2002). Studies done using 144 different species across the land plants demonstrated that the ability to complement the Arabidopsis \textit{lfy} mutant decreases as the evolutionary difference from Arabidopsis increases and that this is due to changes in the DNA binding specificity of the more distantly related proteins (Maizel et al., 2005).

Although control of floral meristem identity is the core function of \textit{LFY} genes in angiosperms, in some species additional roles have been acquired. Some of these functions include involvement in SAM development in tobacco (Ahearn et al., 2001), compound leaf development in legumes (Hofer et al., 1997) and tomato (Molinero-Rosales et al., 1999) and panicle branching in rice (Kyozuka et al., 1998a). In this Chapter, I will introduce only LFY’s functions in flower formation, concentrating on the model angiosperm \textit{Arabidopsis thaliana}.

\textbf{1.2.2 Role of LFY in flowering}

\textit{LFY} plays two main roles in flowering, which are both temporal and genetically separable (Parcy et al., 1998). Firstly, \textit{LFY} acts as a meristem identity regulator and activates other important floral meristem identity regulators. During this phase of
activity, LFY regulates phyllotaxy in the flower as well as meristem size and organ number. Secondly, LFY is necessary for activation of floral organ identity genes. LFY also directly or indirectly helps to maintain FM identity.

Loss of LFY function causes meristems that would normally make flowers to instead produce cauline leaves and associated lateral shoots (Figure 1.7a, b; (Weigel et al., 1992)). Although flowering time is slightly delayed (Blazquez et al., 1997), LFY function is apparently not essential for the transition to reproduction, since this take place normally. Eventually lfy plants will form structures that have both shoot-like and flower-like characteristics and consist of many leaf-like organs and abnormal carpels in a partially spiral phyllotaxy (Figure 1.7f; (Weigel et al., 1992)). This is due to the fact that, in Arabidopsis, later-arising flowers have only a partial requirement for LFY because the AP1 meristem identity gene, which shares some functions with LFY, can become activated independently (Bowman et al., 1993; Huala and Sussex, 1992; Wigge et al., 2005). lfy mutant flower-like structures are subtended by bracts that normally are suppressed in Arabidopsis (Figure 1.7d; (Huala and Sussex, 1992; Schultz and Haughn, 1991; Weigel et al., 1992)). Conversely, overexpression of LFY leads to early flowering and causes all axillary meristems to produce solitary flowers and termination of the apical meristem in a flower (Figure 1.8; (Weigel and Nilsson, 1995)). Therefore, LFY is both necessary and sufficient for floral identity.
1.2.2.1 LFY is involved in precise timing of the floral transition as well as maintaining floral fate.

LFY expression is first detectable in leaf primordia at a very low level and increases until a certain threshold is reached; once the threshold is reached, the primordia are specified as flowers. In other words, the level of LFY in the plant is the trigger to produce flowers (Blazquez et al., 1997; Hempel et al., 1997). Thus, when the number of copies of LFY is altered, timing of flower formation is changed (Blazquez et al., 1997). The level of LFY reflects the quantity and the quality of different flowering signals the plant perceives (Blazquez et al., 1998; Blazquez and Weigel, 2000; Hempel et al., 1997; Lee et al., 2008; Nilsson et al., 1998a). Previous studies done on flowering time mutants show that in late flowering mutants, LFY expression is delayed, while in early flowering mutants, its expression is accelerated (Nilsson et al., 1998a). Since LFY has been shown to be downstream of all the known pathways that control flowering time (Aukerman et al., 1999; Blazquez et al., 1998; Nilsson et al., 1998b; Weigel et al., 2000) and expression of the other key meristem identity gene in Arabidopsis, AP1, is observed only after the floral transition has been initiated (Hempel et al., 1997; Mandel et al., 1992; Simon et al., 1996) and LFY is a direct activator of AP1 transcription (Wagner et al., 1999) as well as at least two other meristem identity genes, CAL (William et al., 2004) and LMI1 (Saddic et al., 2006; Winter et al., 2011), LFY is the key player in the transition (Figure 1.4). This is supported by data from other angiosperms, where the orthologs of LFY are the predominant floral meristem identity genes and the AP1 orthologs have a much more minor role (for example, tomato FALSIFLORA (Molina-Rosales et al., 1999) and
Once the transition to flower formation has occurred, it is important to maintain this status for successful reproduction, especially in annual plants like Arabidopsis. Studies done in Arabidopsis shows that this is archived by the mutual feedback regulation of the meristem identity genes. LFY activates *AP1* and *CAL* by binding their regulatory sequences (Wagner et al., 1999; William et al., 2004), while in turn *AP1* and *CAL* further promote *LFY* expression (Figure 1.4). Another requirement to maintain floral identity is repression of the negative meristem identity regulator *TFL1*. This is achieved by direct repression of *TFL1* in the floral meristems by LFY and AP1 and repression of *LFY* and *AP1* expression in the IM by TFL1 (Figure 1.4; (Bowman et al., 1993; Bradley et al., 1997; Liljegren et al., 1999; Ratcliffe et al., 1998; Shannon and Meekswagner, 1993; Weigel et al., 1992)). This balance between TFL1 and the floral meristem identity genes allow the plants to transit to flowering in a timely manner and to maintain a stable switch to flowering, which is important for the survival of the plant species. As discussed in the previous section, AP1 plays an important role in the precise timing of early events in the establishment of the floral meristem by repressing *AGL24, SVP* and *SOC1* (Figure 1.6a; (Gregis et al., 2008; Liu et al., 2007)). Therefore, a complex network exists between floral meristem identity genes, flowering time genes and inflorescence meristem identity genes to produce flowers in the correct time and place.
1.2.2.2 LFY controls floral homeotic gene expression and is necessary for floral organ identity and differentiation.

After initiating the meristem identity switch, LFY has a second role in flower development through activation of the floral organ identity genes, particularly class B genes (Figure 1.6a; (Weigel and Meyerowitz, 1994)). LFY directly activates the expression of the B class genes AP3 and PI (Lamb et al., 2002; Winter et al., 2011). AP1 also regulates these genes (Ng and Yanofsky, 2001). LFY directly activates the class C gene AG (Busch et al., 1999). However, LFY is not absolutely required for AG expression as lfy mutants have detectable amounts of AG expression (Weigel and Meyerowitz, 1993). LFY also directly activates the SEP genes (Winter et al., 2011). As discussed previously, LFY is also necessary for full activation of AP1 expression (Wagner et al., 1999). Therefore, LFY is required for proper expression of most of the known A, B, C and E genes, with the exception of AP2.

1.2.3 LFY as a transcription factor

LFY is a plant-specific transcription factor that directly binds to its target genes through a DNA binding domain with similarity to helix-turn-helix motifs (Busch et al., 1999; Hames et al., 2008). LFY has two highly conserved domains, a N-terminal domain had been hypothesized to be involved in transcriptional activation and a highly conserved, well-studied C-terminal DNA binding domain (Figure 1.9; (Coen et al., 1990; Maizel et al., 2005)). Outside of these regions there is very little similarity. Genome wide binding and expression studies have showed that LFY can act as both a transcriptional activator
and repressor (Wagner et al., 1999; William et al., 2004; Winter et al., 2011). However, LFY does not seem to have either activation or repression activity on its own (R.S. Lamb, unpublished data; (Busch et al., 1999)). LFY regulates gene expression by recognizing pseudopalindromic sequence elements (CCANTGT/G) in the promoters of its target genes (Busch et al., 1999; Lamb et al., 2002; Lamb and Irish, 2003; Lohmann et al., 2001; Parcy et al., 1998). The crystal structure of the LFY C-terminus bound to DNA, shows that the DNA binding domain comprises of seven helices and two short beta-strands connected by short loops (Hames et al., 2008). Helices α2 and α3 occupy the major groove and mediate most of the DNA contacts. This has more accurately defined the binding sequence as T/ANNNCCANTGG/TNNNNT/A (with the center of the pseudopalindrome underlined; (Hames et al., 2008)). This motif has the previously defined consensus as the core. The three putative LFY binding sites described in the API promoter (Benlloch et al., 2011) also conform well to this new definition. Several recent papers have also found this expanded consensus sequence for LFY binding sites in flowering-related genes (Moyroud et al., 2011; Winter et al., 2011).

1.2.4 LFY co-regulators

LFY and known LFY co-factors act together to regulate region/domain- and stage-specific floral homeotic gene expression. As discussed before, LFY is expressed during vegetative phase and increases gradually till floral transition and in young flower floral buds LFY is uniformly expressed (Blazquez et al., 1997). How does uniformly expressed LFY lead to whorl-specific expression of organ identity genes? The results of the studies
done with *LFY-VP16* (Parcy et al., 1998) suggested the involvement of LFY co-factors in activating organ identity genes. The current evidence shows that LFY is responsible for domain specific and stage specific activation of ABCE genes though acting together with co-factors.

UNUSUAL FLORAL ORGANS (UFO) was shown to physically interact with LFY both *in vitro* and *in vivo*. UFO is expressed in a domain similar to that of the B class genes *AP3* and *PI*. Even though LFY is uniformly expressed in young buds, because of the expression pattern of *UFO*, LFY activates B class genes only in the second and third whorls (Levin and Meyerowitz, 1995; Wilkinson and Haughn, 1995). Experimentally it was shown that LFY and UFO both were needed to activate *AP3::GUS* activity (Honma and Goto, 2000; Parcy et al., 1998). UFO is an F-box protein (Ingram et al., 1995; Samach et al., 1999). F-box proteins form part of SCF ubiquitin ligase complexes that polyubiquitinate proteins, targeting them for destruction via the 26S proteasome (Ni et al., 2004; Sullivan et al., 2003; Wang et al., 2003). (Samach et al., 1999) proposed that UFO might be degrading negative regulators of *AP3* expression. However, the current model is that UFO is involved in modifying LFY in order to enhance its transcriptional activity, although no evidence has been found for ubiquitnation of LFY (Chae et al., 2008). Similar activities for F-box proteins has been reported previously in yeast and mammals (Muratani and Tansey, 2003).
*WUSCHEL* is one of the key genes involved in maintaining stem cell populations of the meristems. In stage 6 flowers the termination of WUS expression is what causes floral meristem termination (Lenhard et al., 2001; Lohmann et al., 2001). In addition, *WUS* also functions as a co-factor of LFY in regulating expression of *AG* in the inner whorls. LFY and WUS bind very closely together on *AG* cis-elements but do not physically interact with each other (Busch et al., 1999; Hong et al., 2003; Lenhard et al., 2001; Lohmann et al., 2001). Both LFY and WUS have to bind to activate *AG* (Lohmann et al., 2001). *WUS* is expressed only in the center of the meristem (Mayer et al., 1998), giving spatial specificity to the activation of *AG*.

PERIANTHIA (PAN) is a bZIP protein that also activates *AG* (Maier et al., 2009; Das et al., 2009). LFY and PAN bind to binding sites in the *AG* second intron to promote *AG* expression and a *lfy* mutation enhances the floral defects of *pan* mutations, suggests PAN most likely acts as a co-factor of LFY (Maier et al., 2009; Das et al., 2009).

SEP3 was recently shown to be a LFY co-factor in addition to its function as an E-class gene (Castillejo et al., 2005; Liu et al., 2009b; Kaufmann et al., 2009). LFY and SEP3 work together to activate ABCE genes (Castillejo et al., 2005). LFY and SEP3 were shown to physically interact using *in vitro* GST-immunoprecipitation assay (Liu et al., 2009c).
Another co-factor of LFY is encoded by *SPLAYED (SYD)*. SYD is a member of the Snf2p ATPase family of chromatin remodeling factors. It was identified though a genetic screen done to isolate enhancers of weak *lfy* mutant phenotypes (Wagner and Meyerowitz, 2002). SYD acts as an LFY-dependent repressor of meristem identity switch although it is not known if they function in the same physical complexes.

A very recent publication (Winter et al., 2011) identified potential LFY cofactor motifs from a *de novo* bioinformatic motif analysis. Two such motifs, GA-repeat hexamers and octamers, were proposed to help recruit stage specific LFY co-factors. This type of repeats is often found in Polycomb Responsive Elements (PREs). It is already known that polycomb factors such as CLF regulate LFY targets, suggesting that these proteins also act as LFY co-factors.

Although effort has been made in identifying LFY co-factors, only two proteins (UFO and SEP3) have been shown to physically interact with LFY, making it likely that other factors in LFY transcriptional complexes remain to be identified.

**1.3 Research Objectives**

The research work carried out over the years on LFY has identified the roles LFY performs as a master regulator in flowering. However, until very recently little information was available about other factors that are present in LFY transcriptional
complexes. The research work described in this thesis will focus on understanding the protein composition and function of LFY-containing transcriptional complexes.

Chapter 2- Like many other transcription factors, LFY activates its target genes by recognizing a psedopalindromic region in the promoters of these genes (Benlloch et al., 2011; Busch et al., 1999; Hong et al., 2003; Lamb et al., 2002; Lohmann et al., 2001; Moyroud et al., 2011; Parcy et al., 1998; Winter et al., 2011). Binding to a palindromic region increases binding affinity for DNA as well as increases the specificity by having doubling the length of the binding site. We show that LFY homodimerizes and this homodimerization is important for its function in planta. The dimerization is mediated by the conserved N-terminal domain, which is a likely reason that this domain is conserved across land plants.

Chapter 3- As described, LFY acts as a transcription factor at early and late stages of flower development. LFY does not have strong transcriptional activity by itself (R.S. Lamb, unpublished data; (Busch et al., 1999)). LFY is expressed during vegetative phase and increases gradually till floral transition and in young flower stages 0-7 (Blazquez et al., 1997). How does uniformly expressed LFY lead to whorl-specific expression of organ identity genes? There is strong evidence that shows LFY acts with other regulatory factors in performing its functions in different stages and in different floral domains (Chae et al., 2008; Lee et al., 1997; Lenhard et al., 2001; Lohmann et al., 2001; Parcy et al., 1998; Wagner and Meyerowitz, 2002). In Chapter 3, we report our biochemical
isolation of putative LFY-interacting proteins in Arabidopsis thaliana. We provide preliminary characterization of four of these proteins.

Chapter 4- From our biochemical isolation of LFY-containing complexes, we identified a novel C2H2 zinc finger transcription factor that we named FRIEND OF LFY1 (FOL1). We present our characterization of this protein and the gene that encodes it. FOL1 is expressed in inflorescence meristems, floral meristems, male and female gametophytes, embryos and young seedlings. Based on gain of function and loss of function analysis, FOL1 appears to function in control of meristem size and gametophyte development. We also discuss FOL1’s physical and genetic interactions with LFY.

1.4 Significance of this study

The conservation of LFY across plant species suggests that most if not all of the LFY-interacting proteins identified in Arabidopsis will also function with LFY orthologs in other plants. LFY expression promotes early flowering in a number of commercially important crops, including rice (Oryza sativa) and poplar (Populus trichocarpa) (Kyozuka et al., 1998a; Kyozuka et al., 1998b; Rottmann et al., 2000). Therefore unraveling the transcriptional complexes in which LFY performs its functions will provide a basis for the development of new strategies to increase agronomical values such as increased yield by manipulating LFY-interacting proteins in economically important crops.
Figure 1.1. The *Arabidopsis thaliana* reproductive transition. During the vegetative phase, rosette leaves are produced without internode elongation. Upon transition to reproduction, internode elongation is activated to form the bolt. Arrowheads indicate indeterminate growth.
Figure 1.2. Schematic representation of reproductive and floral transitions. (a) Conversion of a vegetative meristem (VM) into inflorescence meristem (IM) and the formation of floral meristem (FM) at the flanks of the IM. L- lateral leaf primordia. (b) Micrograph of a wild type flower showing floral organs: sepals, petals, stamens and carpels.
Figure 1.3. Integration of flowering time pathways to regulate the expression of the floral pathway integrators FT, SOC1 and LFY. The photoperiod pathway regulates CO, which activates FT, SOC1 and LFY expression. The vernalization pathway represses FLC, which is a repressor of FT and SOC1 expression, thereby promoting flowering. The autonomous pathway also represses FLC to regulate the expression of FT and SOC1. The GA pathway acts directly on LFY and SOC1 to promote flowering. The developmental age pathway acts through miRNA regulated SPL genes to activate SOC1 and LFY.
Figure 1.4. Regulation of FM identity (Figure modified from (Liu et al., 2009a)). Flowering pathway integrators $FT$, $SOC1$ and $LFY$ activate meristem identity genes ($LFY$, $AP1$, $FUL$, $CAL$ and $LMI1$). $TFL1$, a repressor of floral meristem identity is regulated through $LFY$ and $AP1$. Blue arrows indicate promoting, red linkers indicates repressive effects. Asterisks indicate direct transcription regulation.
Figure 1.5. The ABCE model of floral development. Diagrams of flowers showing sepals, petals, stamens, and carpel in whorl 1-4, respectively. (a) Original ABC model. (b) ABCE model. Underneath the flower is the model aligned with the respective organ. The colors used in the ABC model for different classes and the colors marking the organs are color matched. Barred lines indicate mutual repression between A and C class genes.
Figure 1.6. Maintenance of FM identity (a) Schematic diagram representing the players in FM identity maintenance. Blue arrows indicate promoting activity, red lines indicates repressive effects. (b) Schematic representation of flower development stages and the players involved in FM identity maintenance. At each stage the active genes are given below. Stronger expression is indicated with larger, bold font. IM-inflorescence meristem. FM-floral meristem. FM$_i$- incipient/ immature FM.
Figure 1.7. *LFY* is a floral meristem identity gene. Micrographs of Arabidopsis plants and flowers. (a) L. *er* plant. (b) *lfy*-6 plant. Numbers indicate cauline leaves with subtending branches along the primary stem. (c) Top of L. *er* inflorescence. (d) Top of *lfy*-6 inflorescence. (e) WT flower. (f) *lfy*-6 mutant flower. Note partially spiral phyllotaxy in (f). White arrows indicate bracts, red arrows indicate partially fused carpel-like organs, yellow arrowhead indicates leaf-like sepals.
Figure 1.8. Ectopic expression of LFY is sufficient to convert meristems into flowers. Micrographs of 35S::LFY plants. Black arrows indicate solitary flowers formed by axillary meristems, red arrows indicate terminal flowers.
Figure 1.9. LFY and its orthologs contain two conserved domains. Amino acid alignment of selected LFY orthologs from across land plants. The conserved N-terminal region is marked with a red bar. The conserved C-terminal DNA binding domain is marked with a blue bar. At, Arabidopsis thaliana; Pt, Populus trichocarpa; Vv, Vitis vinifera; FALSIFLORA. Solanum lycopersicon LFY ortholog; FLO, Antirrhinum majus LFY ortholog; RFL, Oryza sativa LFY ortholog; Nymod, Nymphaea odorata; PRFLL, Pinus radiata LFY ortholog; Wel, Welwitschia mirabilis; Cr, Ceratopteris thalictroides; Sel, Selaginella moellendorfii; Pp, Physcomitrella patens.
REFERENCES


Garner, W.W., and H.A. Allard. 1920. Effect of the relative length of day and night on flowering and fruiting of plants.


Sridhar, V.V., A. Surendrarao, D. Gonzalez, R.S. Conlan, and Z. Liu. 2004. Transcriptional repression of target genes by LEUNIG and SEUSS, two

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


CHAPTER 2

Flower Formation in *Arabidopsis thaliana* is Dependent on Dimerization of LEAFY
2.1 INTRODUCTION

Coordinated control of eukaryotic gene expression is mediated by the action of transcription factors (TFs). Many studies in different systems have revealed that TFs have modular structures, consisting of DNA-binding domains as well as transcription activation domains, dimerization domains (homo- or hetero-dimer interfaces) and/or nuclear localization domains (Frankel and Kim, 1991; Lee, 1992). TFs are classified into families by their DNA-binding domains and potential for dimerization. For example, the bHLH TF family is characterized by a helix-loop-helix (HLH) region that mediates dimerization and an alpha-helical basic region that binds DNA ((Amoutzias et al., 2004; Amoutzias et al., 2008) and references therein). bZIP TFs have a basic region that binds DNA and a coiled-coil leucine zipper (LZ) domain that directs dimerization (Amoutzias et al., 2008; Vinson et al., 2006).

In order to function, many TFs must form homodimers and/or heterodimers with a family member (Marianayagam et al., 2004). In such transcription factors, dimerization is the key for both DNA binding affinity and specificity and biological activity. In general, TFs carry their DNA binding and dimerization activities in highly conserved domains. Usually, the DNA binding region has the most conservation while the dimerization domain has less similarity (Amoutzias et al., 2004; Amoutzias et al., 2008; Simionato et al., 2007). Controlling nuclear localization of the TF (Glass, 1994; Jiang et al., 1997),
improving stability of TFs and generating new protein binding interfaces (Marianayagam et al., 2004) are some of the ways dimerization can facilitate transcriptional activity. Dimerization in eukaryotes is known to regulate many processes including the cell cycle, reproduction, development, cellular homeostasis, signaling, metabolism and cell death (Amoutzias et al., 2008; Desvergne et al., 2006; Levy and Darnell, 2002; Wietek and O'Neill, 2007).

*Arabidopsis thaliana* LFY is a unique, plant-specific transcription factor that cannot be grouped with any well-studied TF families. LFY and its orthologs from other land plants contain a DNA binding domain in their C-termini. A crystal structure of the C-terminus of Arabidopsis LFY bound to DNA revealed structural similarities with the DNA binding domains of helix-turn-helix (HTH) proteins such as homeodomain TFs (Hames et al., 2008). However, the structure and function of the rest of the LFY protein is not known. Alignment of LFY and its orthologs across land plants reveals the presence of two conserved regions (Figure 1.9), the highly conserved DNA binding domain in the C-terminus and another conserved region in the N-terminus that has been hypothesized to contribute to transcriptional activation (Maizel et al., 2005).

Like many other transcription factors, LFY activates its target genes by recognizing a pseudopalindromic region in the promoters of these genes with the consensus sequence CCANTGT/G (Busch et al., 1999; Hill et al., 1998; Hong et al., 2003; Lamb et al., 2002; Lohmann et al., 2001; Parcy et al., 1998). Binding to a palindromic region not only increases the binding affinity for DNA but also, by having double the length of the
binding site, increases specificity. We hypothesized that LFY forms homodimers that bind DNA and that this dimerization is important for its function in planta. In this study, we show that LFY homodimerizes. We demonstrate that the conserved N-terminal region of LFY mediates this interaction. Most importantly, studies done in Arabidopsis thaliana demonstrates that dimerization of LFY is important for its function as a transcription factor.
2.2 RESULTS

2.2.1 LFY protein homodimerizes

It is known that LFY binding sites in the promoters of its direct target genes consist of pseudopalindromic sequence elements (Busch et al., 1999; Lamb et al., 2002; Lohmann et al., 2001; Parcy et al., 1998), which led us to investigate LFY protein dimerization. We used yeast two-hybrid assays (Y2H), Bimolecular Fluorescence Complementation (BiFC) and co-immunoprecipitation (co-IP) assays to investigate this association. LFY interacted with itself strongly in yeast (Figure 2.1). To confirm LFY homodimerization in planta, we carried out BiFC assays in transiently transformed Nicotiana benthamiana leaves. LFY was able to bring the two halves of YFP together, indicating dimerization. YFP fluorescence could be seen in the nucleus and in punctate points in the cytoplasm (Figure 2.2a). The fluorescence in the cytoplasm may represent staining at plasmadesmata, since LFY has been shown to traffic through these structures (Wu et al., 2003). The background fluorescence signal from the negative controls used is negligible (Figure 2.2b).

In order to confirm and extend the yeast and BiFC results, co-IP assays were performed. Proteins extracted from N. benthamiana leaves transiently co-transformed with 35S::LFY-GFP and 35S::LFY-MYC or 35S::CBF3-GFP (C-REPEAT BINDING FACTOR 3 (Gilmour et al., 2000), a nuclear-localized transcription factor) and
35S::LFY-MYC were used to perform co-IP studies. Anti-GFP antibody was used to immunoprecipitate LFY-GFP or CBF3-GFP (Figure 2.3a). As expected based on the yeast two-hybrid and BiFC results, only LFY-GFP was able to co-IP LFY-MYC, supporting the presence of higher order LFY complexes. Immunoprecipitation of LFY-MYC using an anti-MYC antibody also brought down LFY-GFP but not CBF3-GFP (data not shown). Further co-IP experiments were done using stable transgenic Arabidopsis plants co-expressing 35S::LFY-GFP and 35S::LFY-TAP or 35S::GFP and 35S::LFY-TAP. Overexpression of LFY in Arabidopsis causes precocious flowering and converts the primary shoot meristem into a terminal flower and axillary meristems into solitary flowers (Weigel and Nilsson, 1995), which was also visible in lines we used (Figure 2.4). In addition, the LFY fusion protein lines used had been crossed into the lfy-6 null mutant background and shown to complement the mutation (shown in other sections-Figure 2.16 and Figure 3.1). IPs using anti-GFP antibody was done on proteins extracted from inflorescences. The LFY-TAP protein was co-IPed with LFY-GFP, but not with GFP alone (Figure 2.3b). Taken together, these results support the existence of higher order LFY complexes in plants.

2.2.2 LFY dimerizes through a conserved region in its N-terminus

LFY orthologs are found across the land plants. Earlier work has identified two conserved regions within LFY-like proteins, one at the N-termini and one at the C-termini of the proteins (Figure 1.9; (Maizel et al., 2005)). The C-terminal conserved region has been shown to be the DNA binding domain. The N-terminal conserved region has some function in strengthening DNA binding, but is not essential for this process. Its function
is currently unknown. We hypothesized that this region might mediate homodimerization of LFY and its orthologs.

To further investigate the domain(s) involved in LFY dimerization, a series of LFY deletion variants was made, as shown in Figure 2.5. Initially, we used the Y2H system to identify the domain(s) mediating homodimerization. Analysis by both growth on selection plates (Figure 2.6a) and quantitative assay (Figure 2.6b) was done. Those constructs that include the conserved N-terminal domain (LFY_{A229-420}, LFY_{A121-420} and LFY_{A}) exhibit strong homodimerization activity, while no construct that does not contain amino acids 46-121 dimerizes. In fact, as shown in Figure 2.6b, the LFY_{A229-420} protein interacts more strongly than full length LFY. This suggests that the full length protein may be present in a conformation that hinders dimerization. We determined the minimal N-terminal fragment required for the dimerization by developing further deletions of the N-terminal half of LFY. LFY_{A}, which includes only amino acids 46-121, still homodimerizes (Figure 2.6). We deleted a further 20 amino acids from the C-terminus of LFY_{A} (Figure 2.7a) to determine if residues at the end of the conserved domain are necessary for dimerization. This construct has lost the ability to form dimers (Figure 2.7b), indicating that the entire conserved domain is essential for LFY-LFY interactions.

In order to validate these findings in planta, we performed co-IP studies using transient transfections in *N. benthamiana*. 35S:: LFY_{A229-420-GFP} or 35S:: LFY_{A1-229-GFP} were co-infiltrated with 35S::LFY-MYC. Both deletion constructs, like full length LFY, localize in
the nucleus (Figure 2.8c). However, only LFY\textsubscript{\textDelta229-420}-GFP was able interact with LFY\textsubscript{\textDelta229-420}-MYC and LFY-MYC (Figure 2.8a, b), supporting the conclusion that the N-terminal half of LFY contains the dimerization domain.

2.2.3 Overexpressing the LFY dimerization domain interferes with endogenous LFY function.

If dimerization is important for LFY function, interfering with this should reduce function. Therefore, we overexpressed the N-terminal half of LFY containing the dimerization domain using the 35S promoter in a wild type background to examine the effect on endogenous LFY function. These plants displayed lfy-like phenotypes suggesting that the function of endogenous LFY is reduced. 20 independent transgenic lines were examined, of which 14 (70%) displaced such phenotypes (Table 2.1). The transgenic lines had extra branches subtended by cauline leaves (Figure 2.9a) and their flowers consisted largely of leaf-like sepals, petals and stamens were absent and the carpels partially unfused (Figure. 2.9c). Most of the lines recovered were sterile and could not be propagated past the T\textsubscript{0} generation. These sterile lines were classified as having severe phenotypes (Table 2.1; Figure 2.11f, j, m). We also recovered several moderate and mild lines (Table 2.1; Fig. 2.11g, k, n). The LFY\textsubscript{\textDelta229-420} protein can interact with full length protein to form a LFY-LFY\textsubscript{\textDelta229-420} complex (Figures 2.8a and 2.6). These results suggest that LFY\textsubscript{\textDelta229-420} acts dominant negatively, titrating wild type LFY away from functional transcriptional complexes.
At least three possible mechanisms could explain the reason for the dominant negative effect observed with LFY\textsubscript{A229-420}. First, the transgenes could be silencing the endogenous LFY gene. Given the number of transgenic lines examined and the high percentage that have an lf\textsubscript{y}-like phenotype, this is unlikely. Second, the heterodimer or higher order complex that is formed when LFY\textsubscript{A229-420} is expressed might be nonfunctional or less functional. This could be due to lack of dimerization affecting DNA binding and/or stability or ability to activate transcription. Thirdly, overexpression of LFY\textsubscript{A229-420} might titrate some essential binding partner(s) of LFY. The second and third possibilities are not mutually exclusive. The third possibility is difficult to test, as there has been no report of a protein that binds to the N-terminus of LFY to date. To test the second model by which LFY\textsubscript{A229-420} might interfere with endogenous LFY function, we performed Electrophoretic Mobility Shift Assay (EMSA) and quantitative real time-PCR (qRT-PCR).

In order to examine DNA binding, we used EMSA assays with a DNA probe containing a previously characterized LFY binding site from the \textit{APETELA1 (API)} promoter (Busch et al., 1999; Hames et al., 2008; Parcy et al., 1998). In our hands, LFY forms two complexes on DNA (Figure 2.10b). The lower complex we assume to be an LFY-LFY homodimer. The upper band is a LFY oligomer of unknown stoichiometry. Although LFY\textsubscript{A229-420} does not bind to DNA alone (Figure 2.10a), when increasing amounts of it were added to the LFY binding reaction, a higher order complex was formed and eventually essentially replaced the LFY-LFY lower complex (Figure 2.10a). The fact that this complex runs higher than the upper band present with LFY alone suggests either that
this is an oligomer of unknown composition rather than a LFY- LFYΔ229-420 dimer or that the confirmation of the LFY- LFYΔ229-420 dimer is retarding its progress through the gel. This demonstrates that LFYΔ229-420 does not interfere with LFY DNA binding in vitro and furthermore than its dominant negative activity in planta is unlikely to be due to an inability to bind target promoters. In order to determine if the LFY- LFYΔ229-420 complex is less stable or has weaker DNA binding than the LFY-LFY complex, we performed competition EMSAs with increasing amounts of cold probe (Figure 2.10b right half). Surprisingly, the LFY- LFYΔ229-420 complex was less sensitive to competition, suggesting that this complex may bind DNA more strongly than LFY-LFY and/or that it is more stable.

The LFYΔ229-420-containing complex can bind DNA at least as strongly as the LFY complex, suggesting that its ability to interfere with LFY function in vivo is not due to lack of ability to bind to target promoters when bound to endogenous LFY. Therefore, we examined the expression of endogenous LFY and two LFY target genes, AP1 and AP3, as well as the transgene in planta using qRT-PCR. For all lines examined, which included some exhibiting severe (plant lines 3 and 21) and moderate or mild (plant lines 26 and 1-2) lfy-like phenotypes, the transgene was robustly expressed (Fig. 2.11a). Expression of endogenous LFY was reduced in almost all lines examined (Fig. 2.11b). This can be explained by the feedback loops on LFY expression by downstream targets, particularly AP1 (Liljegren et al., 1999). Transgenic lines with severe lfy-like phenotypes had less expression of AP3 than the plants with intermediate or mild phenotypes (Figure 2.11d) and lower expression of AP1 as well (Figure 2.11c). In other words the severity of the
phenotype can be explained by lowered LFY target gene expression. Taken together, the EMSA and qRT-PCR data supports the hypothesis that the dominant negative phenotype seen when LFY_{A229-420} is overexpressed is likely caused by the shorter protein interacting with endogenous LFY, forming a complex less able to activate transcription of target genes and outcompeting transcriptionally competent LFY complexes.

2.2.4 Mutations within the dimerization domain abrogate dimer formation

Our work has demonstrated that, in Arabidopsis, LFY amino acids 46-120 are essential for dimerization. This region is highly similar across land plants (Figure 2.12). Within this region there are five conserved leucine residues (positions 69, 77, 84, 91 and 99 in Arabidopsis LFY), which appear in a X7X8X7X8 pattern. This pattern has some similarity to the spacing of leucines in leucine zippers, suggesting these residues may be essential for protein-protein contact. Therefore we mutated these leucines to alanines (within the context of the full length LFY protein) singly, in pairs and in total and tested the effect on dimerization using yeast two-hybrid assays. In addition, two non-leucine but still conserved residues, arginine 59 (R_{59}) and methionine 79 (M_{79}), were also mutated to alanine (Figure 2.12). When single leucine residues were substituted with alanine, the strength of the dimerization interactions was significantly reduced (Figure 2.13a). When either two leucine residues (Figure 2.13b) or all five residues (LFY_{L/A}) were mutated (Figure 2.14a and b), no interaction above background could be detected. Interestingly, as shown in Figure 2.14a, LFY_{L/A} could not interact with wild type LFY. These results support the hypothesis that the leucines are important for LFY dimerization. Mutation of R_{59} and M_{79} also abolished dimerization in Y2H assay (Figure 2.14c), suggesting that the
structure of the dimerization domain is very sensitive to changes in sequence. To confirm the yeast two-hybrid results, co-IP experiments done with LFY<sub>L/A</sub> in transiently infiltrated <i>N. benthamiana</i> leaves. LFY<sub>L/A</sub>-GFP, although localized to the nucleus (Fig. 2.15b) was unable to immunoprecipitate LFY<sub>L/A</sub>-MYC (Figure 2.15a), supporting the hypothesis that the leucine residues are important for dimerization.

Although the DNA binding domain of LFY has been crystallized and its structure determined (Hames et al., 2008), the same has not been done for either the dimerization domain or full length LFY. The sequence of the dimerization domain (amino acids 46-121) was analyzed using the Phyre structural prediction program (Kelley and Sternberg, 2009) and the I-TASSER protein modeling program (Roy et al., 2010; Zhang, 2007). The results suggest that this region consists of four alpha helices (Figure 2.12 a and b), which would create a large protein-protein interaction interface.

### 2.2.5 Dimerization is essential for LFY function

We have shown that LFY forms homodimers or oligomers <i>in vivo</i> and this is mediated by a conserved region in the N-terminus of the protein. To test the functional significance of this domain and dimerization <i>in vivo</i>, we generated transgenic plants expressing various LFY deletions and missense mutations in the background of the null <i>lfy</i>-6 allele and analyzed the ability of the transgenes to complement (Table 2.2). Full length LFY-GFP driven by the 35S promoter is able to partially rescue the <i>lfy</i>-6 mutant (Figure 2.16a). However, deletion of either the N-terminal dimerization domain or the C-terminal DNA binding domain produces a non-functional protein unable to complement (Figure 2.16a),
although the proteins accumulate (Figure 2.16b). As expected, both conserved regions of LFY are essential for function.

In order to study the functional consequence of LFY-LFY interactions \textit{in vivo}, $35S::LFY_{L/A}$-GFP transgenic lines were generated. Surprisingly, in the L. er background most of the transformants (Table 2.1) showed weak $lfy$-like mutant phenotypes, similar to those shown by the partial loss of function $lfy$-5 allele (Figure 2.17; (Maizel et al., 2005; Weigel et al., 1992)). This suggests that this variant may interfere weakly with the endogenous LFY protein, perhaps by titrating an interacting protein. However, the $35S::LFY_{L/A}$-GFP transgene could not complement the $lfy$-6 allele (Figure 2.18; Table 2.2). This supports the hypothesis that dimerization is essential for LFY function \textit{in planta}.

In order to study the $AP3$ target gene expression driven by $35S::LFY_{L/A}$-GFP, we used Arabidopsis $AP3p::GUS$; $lfy$-6 lines. $AP3p::GUS$; $lfy$-6 lines will show residual $AP3$ promoter driven expression at the base of the second and third whorl of the floral organs (Lamb et al., 2002), as seen in Figure 2.19b. Introduction of $35S::LFY$-GFP into this background induces expression of $AP3::GUS$ (Fig. 2.19a); however, introduction of $35S::LFY_{L/A}$-GFP does not (Fig. 2.19b). This further supports that dimerization is essential for LFY function \textit{in planta}.
2.3 DISCUSSION

Proper flower formation is essential for plant reproduction and has been the focus of much interest since Goethe (Goethe, 1790). Floral meristem identity is controlled by several transcription factors, the most important of which is LFY, a plant specific factor. LFY orthologs are found across the land plants (Maizel et al., 2005) and two conserved regions characterize these proteins, a C-terminal DNA binding domain and a N-terminal region of unknown function. Although primary sequence did not reveal any similarity to known transcription factor families, recent structural work has shown that the DNA binding of Arabidopsis thaliana LFY has similarity to helix-turn-helix (HTH) proteins (Hames et al., 2008). The HTH domain is widely used, especially in prokaryotes where it may be the most widely used DNA-binding motif (Harrison and Aggarwal, 1990; Huffman and Brennan, 2002). In this study we analyze the function of the N-terminal conserved domain and find it mediates dimerization of LFY and that this is essential for LFY function in planta.

2.3.1 LFY forms dimers

Previous work had shown that the C-terminal LFY DNA binding domain exists as a monomer in solution and that, when bound to DNA, although two DNA binding domains are found at the LFY pseudopalindromic binding sequence, only a few amino acids of these are in contact (Hames et al., 2008). Therefore, it was hypothesized that LFY exists...
as a monomer (Hames et al., 2008). However, we have found that full length LFY exists as a dimer in yeast and in planta, and this dimerization is not dependent on DNA (Figures 2.1, 2.2 and 2.3). We hypothesized that the N-terminal conserved domain of LFY mediates this interaction. This is fact the case (Figures 2.6 and 2.7). Interestingly, in the yeast two-hybrid system, the strength of the self-interaction between LFY_{Δ229-420} is stronger than that of full length LFY (Figure 2.6b). This suggests that the confirmation of LFY is such that, in the absence of DNA and/or cofactor binding, the DNA binding domain partially inhibits dimerization.

LFY is a plant specific factor. Its dimerization domain is conserved among all the LFY orthologs, but does not have extensive primary sequence similarity to any known domain. Our deletion analysis demonstrates that the entire conserved N-terminal region is necessary for LFY-LFY interaction, as deletions from either the N or C-terminus eliminate binding (Figure 2.8). In addition, although residues outside of 46-121 are not necessary for dimerization, they do contribute to interaction strength, as constructs that remove residues between 121 and 220 do not interact as strongly in yeast (Figure 2.6). Examination of amino acids 46-121 reveals that they contain conserved leucines spaced in a similar manner to leucine zippers (Figure 2.12), suggesting that this region may form a similar structure. Mutation of these leucines within the dimerization domain eliminated LFY-LFY association (Figures 2.13, 2.14 and 2.15), supporting the importance of these residues. However, mutation of two non-leucine conserved residues, R_{59} and M_{79}, also eliminated self-interaction (Figure 2.14c), suggesting that all or most of the conserved residues are important, at least in yeast. Modeling of the LFY dimerization domain using
two different modeling programs, Phyre (Kelley and Sternberg, 2009) and I-TASSER (Roy et al., 2010), suggest that it forms four alpha helices that would provide a large interaction face (Figure 2.12). The accuracy of these predictions remains to be tested.

Overexpression in the wild type background of a LFY construct consisting of only the N-terminal half of the protein (LFY$_{\Delta 229-420}$) caused $lfy$-like phenotypes (Figures 2.9), suggesting that this protein can interfere with endogenous LFY function. This protein can interact with full-length LFY protein (Figures 2.6 and 2.7), supporting the idea that the complexes containing LFY$_{\Delta 229-420}$-LFY dimers are made and further that they must be nonfunctional. DNA binding studies revealed that the N-terminal half of LFY does not interfere with binding; this is not surprising, as it has been previously demonstrated that monomers of the LFY DNA binding domain can associate with DNA (Hames et al., 2008). In fact, if anything, mixing full length LFY and LFY$_{\Delta 229-420}$-GFP creates a complex that binds DNA more strongly than LFY alone (Figure 2.10). Examination of LFY target gene expression in $35S::LFY_{\Delta 229-420}$ transgenic plants revealed that expression of $AP1$ and $AP3$ is reduced (Figure 2.11). This suggests that the LFY- LFY$_{\Delta 229-420}$-GFP heterodimer may bind to LFY binding sites and prevent functional LFY-LFY complexes from activating transcription. The heterodimer is likely to be nonfunctional due to a paucity of other factors that need the C-terminal region of LFY to be recruited to the complex and/or due to the presence of only one LFY DNA binding domain. A similar situation has been seen with fusion proteins associated with acute lymphoblastoid leukemia, which consist of the DNA binding domain of PAX5 (without its activation domain) fused to the C-terminus of C20orf112, a protein of unknown function. These
fusion proteins multimerize at PAX5 binding sites and remain bound very stably, preventing wild type PAX5 binding and target gene regulation (Kawamata et al., 2011). We hypothesize that a similar situation is seen when LFY_{Δ229-420-GFP} is overexpressed; this deletion stabilizes DNA binding through an unknown mechanism and prevents functional LFY transcriptional complexes from binding to activate transcription.

### 2.3.2 LFY-LFY association is necessary for function

We have demonstrated that LFY can associate with itself. This association is necessary in planta. Point mutations that eliminate the ability of LFY to dimerize in yeast (LFY_{L/A}) are unable to complement the ify-6 null mutation (Figure 2.18). This lack of complementation is not due to an inability of the protein to accumulate or properly localize to the nucleus (Figure 2.18d). The role of the N-terminal domain of LFY in dimerization is likely to at least partially explain the evolutionary conservation of this region of the protein across land plants. However, the N-terminus may have other functions in addition to mediating LFY-LFY interactions. For example, there may be other factors that form part of LFY-containing transcriptional complexes that bind to the N-terminal region. In fact, we have identified such a factor (FOL1; see Chapter 4). In addition, other LFY cofactors may need to interact with a dimer of LFY and may not interact with LFY monomers. The fact that the only positive result from a yeast two-hybrid screen with LFY as a bait was LFY itself may support this (A. Moreo and R.S. Lamb, unpublished data). Regardless of any other functions mediated by this conserved region, dimerization of LFY is essential and the domain is necessary for this function.
Further work will be necessary to extend the function of this region beyond LFY-LFY association.

2.3.3 Conclusion

LFY can associate with itself and this association is necessary in planta. In vitro and in vivo assays identified N-terminal conserved region as the domain involved in dimerization. This domain is conserved among all LFY orthologs in land plants, but does not have extensive primary sequence similarity to any known domain. The minimal region competent to dimerize comprises amino acids 46-121. Structural predictions suggest this region consists of alpha helical structures, which are capable of making an interaction surface. Overexpression of the N-terminus of LFY interferes with endogenous LFY. EMSAs suggest that complexes containing this truncated protein bind DNA more stably than wild type complexes, but are unable to activate target gene expression. Mutants of the residues abolishing dimerization in planta are also unable to complement the *lfy-6* null mutation, supporting the idea that LFY-LFY association is necessary for LFY’s function in Arabidopsis.
2.4 EXPERIMENTAL PROCEDURES

2.4.1 Plant material and growth conditions

All Arabidopsis plants used in this study were in the L. er ecotype. The lfy-6 and lfy-5 alleles have been previously described (Weigel et al., 1992). Seeds were vernalized for 3-5 days before growth in soil (Fafard-2 Soil Mix-Superior growers supply Inc. http://www.superiorgrowers.com). Seeds from transgenic lines were sterilized with 70% ethanol and 10% (v.v) sodium hypochlorite and were plated on Murashige and Skoog (MS- Phytotechnology Laboratories, http://www.phytotechlab.com) medium containing sucrose (10g/L), MES (0.5g/L) and agar (6g/L) with kanamycin (50µg/ml) for selection. The plates were incubated at 4°C in the dark for 3 days before placing in a CU-36L plant growth chamber (Percival Scientific, http://percival-scientific.com) under long day conditions (16 hours light at 80µmol m⁻² s⁻¹ irradiance). Resistant seedlings were then transferred to soil. All plants were grown at 22°C under long day conditions with 50% relative humidity in controlled growth chambers (Enconair Ecological Chambers, http://biochambers.com). Nicotiana benthamiana seeds were also grown in above described soil and growth conditions.

2.4.2 Constructs and cloning

The LFY coding region was cloned into pENTR/D-TOPO Gateway entry vector (Invitrogen, http://www.invitrogen.com) following the manufacturer’s protocol and this
was subsequently used as a template to generate deletion variants and mutations used in this study. The primers used are listed in Table 2.3. PCR was done using Platinum Pfx DNA Polymerase (Invitrogen). The PCR products introduced to *pENTR/D-TOPO* vector were sequenced at the Plant-Microbial Genomics Facility (PMGF) at The Ohio State University. Individual residues were mutated using QuickChange Site-Directed Mutagenesis Kit and QuickChange MultiSite Directed Mutagenesis Kit (Agilent, [http://www.genomic.agilent.com](http://www.genomic.agilent.com)) according to the instructions provided. Plasmids carrying the mutations were sequenced at the PMGF. *LFY* deletion and mutation variants in *pENTR/D-TOPO* vector were then recombined with binary destination vectors *pGWB5* and *pGWB6* with a GFP tag and *pGWB20* and *pGWB21* with a MYC tag at the C- and N-termini, respectively (binary destination vectors were a gift from T. Nakagawa, Shimane University, Matsue, Japan); these were then introduced into *Argobacterium tumefaciens* strain GV3101.

For yeast two-hybrid assays, the relevant *pENTR/D* clones were recombined with *pDEST22* and *pDEST32* destination vectors containing the GAL4 Activation Domain (AD) and DNA Binding Domain (BD), respectively (ProQuest Two-Hybrid system, Invitrogen). These plasmids were transformed into yeast strain PJ694A (James, 2001) following the protocol in the Clontech Yeast Protocol Handbook.

*In vitro* protein production vectors were produced by creating PCR products of full length *LFY* and *LFY*<sub>329-420</sub> with EcoR1 and XhoI flanking sites (primers in Table 2.3) and cloning them into the pTNT vector (Promega, [http://www.promega.com](http://www.promega.com)) using restriction
digests and ligation. The restriction enzymes were obtained from New England BioLabs (NEB; http://www.neb.com). Digestions were done at 37°C overnight. The cut DNA was subsequently treated with alkaline phosphatase (CIP; NEB) and ligated using T4 DNA Ligase Kit (Invitrogen) overnight at room temperature.

BiFC vectors p2YN (amino acids 1-158), p2YC (amino acid 159-238) and pYN-I and pYC-I carrying either fragment of split YFP in C- and N- terminal positions, respectively, were used (vectors from Dr. John A. Lindbo, The Ohio State University, Wooster, OH). For cloning the gene into these vectors, restriction sites and regions overlapping with BiFC vectors were added to the LFY coding region through PCR (primers are provided in Table 2.3). Cloning into the vectors was done using In-Fusion cloning (Clonetech, http://www.clontech.com).

**2.4.3 Generation of Arabidopsis transgenic lines and transient infiltration of Nicotiana benthamiana leaves**

*A. tumefaciens* strain GV3101 carrying insertion plasmids were selected on LB plates containing 50 μg/ml kanamycin, 50 μg/ml gentamycin and 25 μg/ml rifampicin. Arabidopsis wild type plants were transformed by the floral dip method (Clough and Bent, 1998). Transformants were selected on MS plates with 50 μg/ml kanamycin. Transgenic Arabidopsis plants expressing LFY-GFP and LFY-TAP or GFP alone and LFY-TAP under the 35S promoter was generated by performing crosses between individual lines. *N. benthamiana* plant leaves were co-infiltrated transiently with *Agrobacterium* as described in (Voinnet et al., 2003).
2.4.4 Yeast two-hybrid interaction assay

Yeast competent cell preparation and transformations were performed as in (Dohmen et al., 1991). \textit{pDEST22-LFY} deletion and mutant variants were co-transformed with \textit{pDEST32-LFY} variants studied. As negative controls, empty vectors were co-transformed with above vectors carrying \textit{LFY} variants. The positive control used in this experiment was a well-studied strong interaction between Krev1 and RalGDS that was included with the ProQuest two-hybrid system. Transformants were selected on double dropout media (SD-leucine-tryptophan). Interactions were selected on triple dropout selection plates (SD-leucine-tryptophan-histidine). Three independent colonies from each transformation were selected to perform quantitative β-galactosidase liquid assays as described (Lamb et al., 2002).

2.4.5 Bimolecular Fluorescence Complementation Assay (BiFC)

BiFC was performed in transiently transformed \textit{N. benthamiana} leaves. \textit{Agrobacterium} with BiFC vectors carrying full-length LFY along with split YFP were co-infiltrated using the protocol published in (Ohad et al., 2007). Empty vectors with one half of the YFP together with the vector carrying LFY gene fused to the other half of the YFP was used as negative controls as well as the two empty vectors. The presence of YFP was observed and photographed three days after infiltrations using a confocal laser scanning microscope (Nikon D-ECLIPSE C1 90i and Camera- Nikon C1-SHS, http://www.nikoninstruments.com) using the GFP filter.
2.4.6 RNA isolation and quantitative Real Time-PCR (qRT-PCR)

Total RNA was extracted from inflorescence apices using RNeasy Plant Mini Kit following the manufacture’s protocol (Qiagen, http://www.qiagen.com). Total cDNA was synthesized from 1μg of RNA using the BluePrint 1st Strand cDNA Synthesis Kit with oligo (dT) primers (TaKaRa, http://www.takara-bio.us). qRT-PCR was carried out using four independent transgenic lines carrying 35S::LFYΔ229-420-GFP that had varying phenotypic severity as well as lfy-6 mutant and wild type plants. qRT-PCR was done with TaqMan Gene Expression Assays (Applied Biosystems, http://www.appliedbiosystems.com) with FAM probes made for LFY (At02270390_m1), AP3 (At02188782_g1), AP1 (At02226242_g1), GFP (At1L179), ACT2 (At02329915_s1) and RAD23 (At02163241_g1). The LFY probe used only detected endogenous LFY. The GFP probe was used to quantify the transgene. TaqMan probes for ACT2 and RAD23 were used as reference genes. Three Biological replicates of qRT-PCR were performed in technical triplicates. qRT-PCR was performed in a 96- well plate with 3μl cDNA (1:3 diluted) and 5μl master mix (TaqMan Gene Expression Master Mix- Applied Biosystems) in a total reaction volume of 10μl. The reactions were done at the PMGF in a BioRad CFX96 (http://www.bio-rad.com).

2.4.7 Computational analysis

Coiled-coil prediction was carried out using Multicoil software available on-line ((Wolf et al., 1997); http://groups.csail.mit.edu/cb/multicoil/cgi-bin/multicoil.cgi). The putative secondary structure of the LFYΔΔ was examined using the Protein homology/analogy
recognition engine (Phyre; Kelley and Sternberg, 2009). The sequence alignment of LFY\textsubscript{ΔΔ} was done using the MUSCLE3.8.31 multiple alignment tool, using default settings (Edgar, 2004). The alignment picture was generated using TEXshade (Beitz, 2000) and Jalview2.0 (Waterhouse et al., 2009). 3D structure predictions were done by submitting LFY sequence to I-TASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER; (Roy et al., 2010; Zhang, 2007)). PyMOL software (www.pymol.org) was used to mark the putative helices of the dimerization region in red.

2.4.8 In vitro transcription/translation, in vitro binding and EMSAs

Full length LFY and LFY\textsubscript{Δ229-420} polypeptides were synthesized using the TNT T7 Quick Coupled Transcription/Translation System (Promega) as specified by the manufacturer. The LFY\textsubscript{Δ229-420} protein was concentrated using Amicon Ultra 0.5 ml Centrifugal Filters (Millipore, www.millipore.com) using manufacturer’s guidelines. Synthesized proteins were run on a 10% SDS-PAGE gel, blotted onto nitrocellulose membrane (Bio-Rad) and probed with polyclonal anti-LFY antibody (aV-16, 1:300, Santa Cruz Biotechnology, http://www.scbt.com) and with anti-goat IgG-horseradish peroxidase (HRP; Sigma-Aldrich, http://www.sigmaaldrich.com) used at 1:10,000. Protein amounts were quantified on film using a densitometer (Molecular Dynamics 375) and ImageQuant Software (GE, http://www.gelifesciences.com).

To carryout EMSAs, 5’-Cy5 labeled single stranded oligonucleotide probes from Sigma-Aldrich (listed in Table 2.4) were annealed to non-fluorescent complimentary oligonucleotides in annealing buffer (10mM Tris pH7.5, 150mM NaCl and 1mM EDTA)
according to (Lamb et al., 2002). Protein and probe binding reactions were carried out in binding buffer according to (Lamb et al., 2002), at room temperature for 20 min and run on 6% polyacrylamide gels. Gels were scanned using Typhoon 9400 Scanner (GE, [http://www.gelifesciences.com](http://www.gelifesciences.com)). Cy-5 scanning mode (633nm wavelength) was used and the image was saved and analyzed using ImageQuant TL software.

### 2.4.9 Co-immunoprecipitation assays

Arabidopsis inflorescence apices and infiltrated leaves of *N. benthamiana* were collected and ground in liquid nitrogen into fine powder. Cold buffer containing 50mM Tris-HCl, pH7.5, 150mM NaCl, 0.5% Nonidet P-40, 1mM EDTA, 3mM DTT, 1mM phenylmethylsulphonyl fluoride and protease inhibitor cocktail (1:100, Sigma-Aldrich) was used to extract proteins followed by 14,000rpm centrifugation. Immunoprecipitations were done using polyclonal anti-GFP antibody (abcam-290, Abcam, [http://www.abcam.com](http://www.abcam.com)) bound overnight to Protein A Agarose beads (Invitrogen). Protein extracts were incubated with the above preparation for 1.5 hours. All experiments were done at 4°C. After washing four times with the above cold buffer, immunoprecipitates were resuspended in 50µl of 2X SDS-PAGE loading buffer. Proteins were separated on a 10% SDS-PAGE gel (Sambrook et al., 1989), transferred to nitrocellulose membranes and probed with alkaline phosphatase (AP)-conjugated anti-GFP antibody (1:1500, Rockland, [http://www.rockland-inc.com](http://www.rockland-inc.com)) or monoclonal anti-cMYC antibody (1:500, Sigma-Aldrich). Anti-mouse IgG-peroxidase (1:10,000, Sigma-Aldrich) was used as the secondary antibody. SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, [http://www.piercenet.com](http://www.piercenet.com)) and Immune-Star AP Substrate
Pack (BioRad) were used to detect AP- and peroxidase-conjugated antibodies, respectively. GeneMate Blue Lite Autorad Film (Bioexpress, http://www.bioexpress.com) was used to detect signal and developed on a KONICA MINOLTA SRX-101A film developer.

2.4.10 GUS staining

AP3p::GUS lines (Hill et al., 1998) were crossed with lfy-6/lfy-6 plants. F1 seed was selected for the transgene with kanamycin and allowed to self. F2 AP3p::GUS; lfy6-6/lfy-6 lines were isolated and analyzed. These confirmed lines were then crossed with 35S::LFY-GFP lines and 35S::LFY_L/A-GFP lines. β-glucoronidase (GUS) enzymatic activity was detected as described (Hill et al., 1998). GUS stained inflorescences were observed under a Nikon SMZ745T (Nikon, http://www.nikoninstruments.com) dissecting microscope and photographed using OmniVID Camera (LW Scientific, http://www.lwscientific.com).

2.4.11 lfy-6 mutant complementation experiments

35S::LFY-TAP/+, 35S::LFY-GFP/+, 35S::LFY_A1-129-GFP/+, 35S::LFY_A229-420-GFP/+ and 35S::LFY_L/A-GFP/+ lines were crossed to lfy-6 mutants. F2 generation plants were selected for presence of the transgene and genotyped. dCAPS genotyping using the primers listed in Table 2.1 and restriction digest with BstAPI (NEB) was used to screen plants. In the F2 generation, phenotypic analysis was done to check for complementation of lfy-6 null mutants. Lines were analyzed to look for floral defects and number of branches formed before flower formation. Presence of petals and stamens, presence of
fused carpels and number of sepals in the outer most whorl of the flower were some of the characters analyzed. We compared phenotypes of these lines with *lfy*-6 null mutants and wild type grown in the same growth area at the same time.

2.4.12 Phenotypic analysis of transgenic lines

Phenotypic analysis of the over expression transgenic lines were done using homozygous lines except for 35S::*LFY*Δ29-420-GFP lines, which could not be propagated. Solitary flower formation was the phenotype scored in *LFY* overexpression plants. 35S::*LFY*Δ29-420-GFP and 35S::*LFY*L/A-GFP plants were analyzed mainly for number of branches and flower phenotypes. Wild type plants typically make 2-3 branches before flower formation, whereas *lfy*-6 mutant makes more branches (Weigel et al., 1992). *lfy* flower-like structures also are often subtended by bracts (Weigel et al., 1992), a feature also scored in transgenic lines. Other features scored were the replacement of sepals with leaf-like organs with an increased number of trichomes and branched trichomes and the absence or reduction in number of petals and stamens in flower-like structures. In addition, the amount of carpel fusion in the pistil was observed.
Figure 2.1. LFY homodimerizes in yeast. Vector combinations used to co-transform yeast cells are given as prey/bait. Three independent transformants were analyzed for each combination. The interaction between Krev1 and RaLGDS was used as a positive control. Empty prey and bait vectors in above shown combinations were used as negative controls. Error bars indicate standard deviations.
Figure 2.2. LFY-LFY homodimers form in planta. Micrographs of *N. benthamiana* leaves (bright field, epifluorescence and merged images). (a) LFY tagged at either the C-terminus (LFY-2YC/LFY-2YN) or the N-terminus (YC1-LFY/YN1-LFY) form homodimers. (b) Negative controls consisting of two empty vectors or one empty vector with one LFY fusion protein did not show any fluorescence. Arrowheads indicate nuclei and arrows indicate punctuate staining in the cytoplasm. Scale bars indicate 50μm.
Figure 2.3. LFY homodimers exist in both *N. benthamiana* leaves and Arabidopsis inflorescences. (a) Transiently transfected *N. benthamiana* leaves transformed with 35S::LFY-MYC and 35S::LFY-GFP or 35S::LFY-MYC and 35S::CBF3-GFP were used for GFP immunoprecipitations. (b) Transgenic Arabidopsis lines carrying 35S::LFY-GFP and 35S::LFY-TAP or 35S::LFY-TAP and 35S::GFP were used for GFP immunoprecipitations. Asterisks mark non-specific bands.
Figure 2.4. Ectopic expression of LFY causes formation of solitary flowers. (a) 35S::LFY-GFP; 35S::LFY-TAP plant. (b) 35S::LFY-GFP plant. (c) 35S::LFY-TAP plant. Arrowheads indicate solitary flowers while red arrow indicates terminal flower.
Figure 2.5. A schematic representation of LFY deletion fragments used in this study. The red boxes represent the conserved N-terminal region (amino acids 46-121). The cyan boxes represent the conserved DNA binding domain (amino acids 229-388). Δ represents the amino acid range deleted. ΔΔ represents deletion of amino acids 1-46 and 121-420.
Figure 2.6. The N-terminal conserved domain of LFY mediates homodimerization. (a) LFY dimerization can support growth on -Leu-Trp-His media. 1: AD-LFY construct/BD-LFY construct; 2: AD/BD-LFY construct; 3: AD-LFY construct/BD. (b) Quantitative yeast two-hybrid assays of LFY deletion fragments. Vector combinations used to co-transform yeast cells are given as prey (AD)/bait (BD). Three independent transformants were analyzed for each combination. Positive control: LFY homodimerization. Negative controls: empty vector combinations. Error bars indicate standard deviation.
Figure 2.7. The entire conserved domain is necessary for LFY dimerization. (a) Schematic representation of constructs used, as in Figure 2.5. (b) Yeast two-hybrid assays done as in Figure 2.6. Error bars indicate standard deviation.
Figure 2.8. N-terminal conserved domain mediates LFY homodimerization in vivo.
Transiently transfected *N. benthamiana* leaves with (a) 35S::LFYΔ229-420-GFP and (b) 35S::LFYΔ1-229-GFP along with 35S::LFY-MYC were subjected to Co-IP with anti-GFP antibody and the pull downed proteins were detected using anti-GFP and anti-MYC antibodies. LFY homodimerization was used as a positive control. (c) Nuclear localization of LFYΔ229-420-GFP and LFYΔ1-229-GFP in *N. benthamiana* leaves. The scale bars indicate 50μm.
Figure 2.9. Overexpression of 35S::LFY<sub>Δ229-420</sub>-GFP causes a lfy-like phenotype. Light micrographs of Arabidopsis plants. (a) Adult plants as indicated. (b-d) Close ups of the tops of inflorescences of lfy-6 (b) 35S::LFY<sub>Δ229-420</sub>-GFP (c) and L. er. (wild type) (d). Insets are individual flowers. Arrows point to leaf-like sepals, while arrowheads indicate exposed ovules. (e) Western blot showing LFY<sub>Δ229-420</sub>-GFP expression in a representative transgenic line (1) and in wild type (2). Asterisks indicate non-specific bands.
Figure 2.10. LFY_{A29-420} forms a higher order DNA binding complex with LFY. (a) Increasing amounts of LFY_{A29-420} to LFY results in switch to a higher order complex but addition of a nonspecific protein Oct2A does not. (b) Probe competition assays. Red arrowheads indicate a higher order complex while black asterisk indicate a presumed LFY-LFY dimer.
Figure 2.11. The dominant negative phenotype caused by $35S::LFY_{1229-420}^{\Delta}$GFP is due to reduced LFY target gene expression and correlates with transgene expression.
Figure 2.11: continued

(a-d) Gene expression studies done with wild type, lfy-6 and four independent 35S::LFYΔ229-420-GFP transgenic lines (line 3, 21, 26 and 1-2). Line 3 and 21 show severe phenotypes. Line 26 and 1-2 show milder phenotypes. Expression was normalized to that of both ACTIN2 and RAD23. (a) Transgene expression. (b) LFY expression. (c) AP1 expression. (d) AP3 expression. (e-l) Micrographs of Arabidopsis plants. (e and i) lfy-6. (f and j) 35S::LFYΔ229-420-GFP plants with severe phenotype (line 21). (g and k) 35S::LFYΔ229-420-GFP plants with milder phenotype (line 1-2). (h and l) L. er (wild type). (m and n) Close-up of flowers of 35S::LFYΔ229-420-GFP plants with severe phenotype (line 21; m) and with milder phenotype (line 1-2; n).
Figure 2.12. The LFY dimerization domain is conserved across land plants and may form an alpha helical structure. (a) Alignment of the LFY dimerization domain from LFY orthologs from select land plants. Amino acids mutated in this study are indicated. The predicted secondary structure of Arabidopsis LFY is shown below the alignment. The predictions were done using Phyre (Kelley and Sternberg, 2009) and MultiCoil (Wolf et al., 1997). At, Arabidopsis thaliana; Pt, Populus trichocarpa; Vv, Vitis vinifera; FALSIFLORA. Solanum lycopersicon LFY ortholog; FLO, Antirrhinum majus LFY ortholog; RFL, Oryza sativa LFY ortholog; Nymod, Nymphaea odorata; PRFLL, Pinus radiata LFY ortholog; Wel, Welwitschia mirabilis; Cr, Ceratopteris thalictroides; Sel, Selaginella moellendorfii; Pp, Physcomitrella patens. (b) The predicted 3D structure of LFY protein modeled using I-TASSER (Roy et al., 2010; Zhang, 2007). Indicated in red are the helices predicted to form from the LFY dimerization domain (amino acid residues 46-121).
Figure 2.13. Mutations within the LFY dimerization domain abrogate dimerization. (a and b) Quantitative Y2H assays. Vector combinations used to co-transform yeast cells are given as prey/bait. Three independent colonies were used. Positive control: homodimerization of LFY. Negative controls: empty prey and bait vectors in above shown combinations. Error bars indicate standard deviation.
Figure 2.14. Mutations within the dimerization domain abrogate LFY interaction in yeast. (a, b) Y2H assays on -Leu-Trp-His selection media. (a) LFY_{UA} does not associate with itself or wild type LFY. (b) Quantitative Y2H liquid assay as described in Figure 2.6. Error bars indicate standard deviation. (c) Dimerization is lost when arginine 59 or methionine 79 are replaced by alanines. 1: AD-LFY construct/BD-LFY construct. 2: AD/BD-LFY construct. 3: AD-LFY construct/BD.
Figure 2.15. Mutations within the dimerization domain abrogate LFY-LFY association in vivo. (a) IPs done from transiently transformed *N. benthamiana* leaves as indicated. (b) LFY\textsubscript{LA}-GFP localizes to the nucleus in the *N. benthamiana* leaf.
Figure 2.16. Both conserved regions of LFY are required for function. (a) Micrographs of adult plants. Insets are individual flowers. A: Wild type; B: 35S::LFY-GFP; lfy-6; C: 35S::LFY\(_{\Delta1-229}\)-GFP; lfy-6; D: 35S::LFY\(_{\Delta229-420}\)-GFP; lfy-6; E: lfy-6. (b) Western blots with anti-GFP antibody showing the expression of the GFP-tagged proteins from representative transgenic lines. 1: Wild type; 2: lfy-6; 3: 35S::LFY\(_{\Delta229-420}\}-GFP; lfy-6; 4: 35S::LFY\(_{\Delta1-229}\}-GFP; lfy-6; 5: 35S::LFY-GFP; lfy-6. Asterisks indicate a non-specific band recognized by the anti-GFP antibody.
Figure 2.17. Overexpression of $LFY_{LA}$-GFP causes weak $lfy$-like phenotypes. (a) Micrographs of adult plants as indicated. (b) Close up of $35S::LFY_{LA}$-GFP inflorescence. (c) Close up of a $35S::LFY_{LA}$-GFP flower. Arrowhead indicates an unfused carpel. (d) Western blot probed with anti-GFP antibody showing expression of $LFY_{LA}$ protein from a representative transgenic line. 1: $35S::LFY_{LA}$-GFP. 2: Wild type. Red asterisk indicates $LFY_{LA}$ while the black asterisk indicates a non-specific bind recognized by the anti-GFP antibody.
Figure 2.18. Dimerization is essential for LFY function in planta. Micrographs of Arabidopsis plants. (a) Adult plants as indicated. 35S::LFY_{L/A}-GFP did not complement the lfy-6 mutant phenotype. (b and c). Close up of the inflorescence (b) and a flower (c) of a 35S::LFY_{L/A}-GFP plant. Red arrowhead indicates exposed ovules in the carpelloid structure. White arrowhead indicates bracts. Arrow indicates extra trichomes on the sepals. (d) Micrograph of GFP expression in transgenic plants, which did not complement.
**Figure 2.19. LFY<sub>L/A</sub> cannot activate AP3 expression.** Micrographs of GUS stained inflorescences with indicated genotypes. Insets show micrographs of GFP expression in transgenic plants of the same lines.
Table 2.1. Transgenic lines examined.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Ecotype</th>
<th>Background</th>
<th>Number of lines examined</th>
<th>Number of lines with phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Severe</td>
</tr>
<tr>
<td>35::LFYΔ229-420-GFP</td>
<td>L. er</td>
<td>Wild type</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>35::LFYΔL/GFP</td>
<td>L. er</td>
<td>Wild type</td>
<td>16</td>
<td>5</td>
</tr>
</tbody>
</table>

Initial (T₀) transgenic lines were analyzed due to sterility. See text and Figure 2.11 for definition of phenotypic severity.
Table 2.2. Both conserved domains and dimerization are required for LFY function.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Ecotype</th>
<th>Background</th>
<th>Complemented</th>
</tr>
</thead>
<tbody>
<tr>
<td>35S::LFY-GFP</td>
<td>L. er</td>
<td>lfy-6</td>
<td>Yes (75%)</td>
</tr>
<tr>
<td>35::LFYΔ1-229-GFP</td>
<td>L. er</td>
<td>lfy-6</td>
<td>No</td>
</tr>
<tr>
<td>35::LFYΔ229-420-GFP</td>
<td>L. er</td>
<td>lfy-6</td>
<td>No</td>
</tr>
<tr>
<td>35::LFYΔL/F-GFP</td>
<td>L. er</td>
<td>lfy-6</td>
<td>No</td>
</tr>
</tbody>
</table>

At least 3 independent lines were analyzed in these studies in the F₂ generation.
Table 2.3. List of primers used in Chapter 2.

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Primer name</th>
<th>Primer Sequence (5'-3')</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>LFY</em> BiFC</td>
<td>LFYp2YCF</td>
<td>TACGAACGATAGTTAATTAACATGGATCCTGAAGGTTTCACGAG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>LFYp2YCR</td>
<td>CACCTCCTCCACTAGTGAACGCAAGTCGTGCAGGCGCCGC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>LFYp2YNF</td>
<td>TACGAACGATAGTTAATTAACATGGATCCTGAAGGTTTCACGAG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>LFYp2YNR</td>
<td>CACCTCCTCCACTAGTGAACGCAAGTCGTGCAGGCGCCGC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>LFYpYC-1F</td>
<td>CTGGAGGAGGCAGGCAGACATGGATCCTGAAGGTTTCACGAG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>LFYpYC-1R</td>
<td>ATTTGCGGAATCTAGAATAGAAACGCAAGTCGTGCAGGCGCCGC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>LFYpYN-1F</td>
<td>CTGGAGGAGGCAGGCAGACATGGATCCTGAAGGTTTCACGAG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>LFYpYN-1R</td>
<td>ATTTGCGGAATCTAGAATAGAAACGCAAGTCGTGCAGGCGCCGC</td>
<td>This Study</td>
</tr>
<tr>
<td><em>LFY</em> variants</td>
<td>LFYGWF</td>
<td>CACCATGGATCCTGAAGGTTTCAC</td>
<td>This Study</td>
</tr>
<tr>
<td>cloning</td>
<td>LFYGWR-S</td>
<td>CTAGAAAAACGCAAGTCGTGCAG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>LFYGWR</td>
<td>GAAACGCAAGTCGTGCAG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>LFY229F</td>
<td>CACCATGGAGAGACAGAGGAGAGCATCCG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>LFY121F</td>
<td>CACCATGGAGGAGAAGGAGATCTCTAGAAGC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>LFY121R-S</td>
<td>CACTCTTCTCTGGCAATCGTCCTGCATGAC</td>
<td>This Study</td>
</tr>
</tbody>
</table>

Continued
Table 2.3: continued

<table>
<thead>
<tr>
<th>LFY Site Directed Mutagenesis</th>
<th>L69A</th>
<th>GGCGAAGATAGCGGAGGCAGGTTTTACGGCGAGC</th>
<th>This Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>L69A_antisense</td>
<td>GCTCGCCGTAAACCTGCTCCGCTATCTCGCC</td>
<td>This Study</td>
<td></td>
</tr>
<tr>
<td>L76A</td>
<td>GTTTTACCGCGAGACGCGCTGGGTATGAAGGACG</td>
<td>This Study</td>
<td></td>
</tr>
<tr>
<td>L76A_antisense</td>
<td>CGTCTTCTATCCACAGCCGCTGCGCCGTAAAAC</td>
<td>This Study</td>
<td></td>
</tr>
<tr>
<td>L84A</td>
<td>GGGTATGAAGGAGGAGGAGGCTGAAGAGATGATGATGATG</td>
<td>This Study</td>
<td></td>
</tr>
<tr>
<td>L84A_antisense</td>
<td>ACTATTCATCATCTCTCTACCATAGTTACGCTCCCTCTCCCATACCC</td>
<td>This Study</td>
<td></td>
</tr>
<tr>
<td>L91A</td>
<td>GCTGGAAGAGATGATGATGATGATGATAGTTACCTCTCATATCTTTCCGTGGGAG</td>
<td>This Study</td>
<td></td>
</tr>
</tbody>
</table>

Continued
Table 2.3: continued

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>L91A_antisense</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CTCACAACGAAAGATGAGGACTATTCATCATCTCTTTCAAGC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>L99A</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TCATATCTTTTCAGGATGGCTTTCTTTGTCACGC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>L99A_antisense</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GTGCCGTTCCACCAAAAGAAGCTCACAACCACCCCAAAGATATGA</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>M79A</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GAGCAGCCTTTGTGGAAGCGAAGCACGAGAGGCTT</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>M79A_antisense</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AAGCTCCTCGCTCTCCGC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R59A</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CGTGCGTCAGGATAGCTTTTCTACACGGGCACG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R59A_antisense</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GCCGCCGTGTAGAAAGCTATACCAGCGAGGACCG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>103</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>pTNT-LFYF</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TATACTCGAGATGGATGCTGAGTTTC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>pTNT-LFYR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TATAATGCTTTTATAGAAACGCAAGTGTCAG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>pTNT-229R</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TATAGAATCTTTATTATGCTCCCAAACACAC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>T7EEV</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AAGGCTAGAGTACCTAAATACGA</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>T7 TERM</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GCCCCAAAGGGGTTATTGCTAG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>dCAPS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>lfy-6F</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AAGCAGCCTGCTGCGGTGTCAGCTGTGTT</td>
<td><a href="http://www.weigelworld.org/resources/mutants/caps">http://www.weigelworld.org/resources/mutants/caps</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>lfy-6R</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TATGGATCTGGAAGTTTCTACG</td>
<td></td>
</tr>
</tbody>
</table>
### Table 2.4: Probes used for EMSA

<table>
<thead>
<tr>
<th>Probe Name</th>
<th>Sequence (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP1WT'</td>
<td>'TTGGGGAAGGACCAGTGGTCCGTACAATGT</td>
</tr>
<tr>
<td>AP1WT</td>
<td>ACATTGTACGGACCACCTGGTCCTTCCCCAA</td>
</tr>
</tbody>
</table>
REFERENCES


CHAPTER 3

Identification of LEAFY Complex Components in *Arabidopsis thaliana*. 
3.1 INTRODUCTION

Transcriptional regulation is one of the fundamental processes in biology governing the development of an organism in every aspect, including morphology, function and behavior and thus the survival of the organism. In eukaryotes, transcriptional regulation of gene expression is a complex process, partly due to the highly condensed packing of DNA into chromatin. DNA is tightly wrapped around an octameric core of histone proteins, arranged in repeating units called nucleosomes (Kornberg, 1974; Kornberg and Thomas, 1974; Luger et al., 1997; Thomas and Kornberg, 1975). Depending on the degree of compaction, DNA may be accessible and able to be transcribed or be present in heterochromatin that maintains the genes in an inactive state (Kornberg, 1974). To coordinate development, genes are turned on and off in a pre-programmed fashion giving proper spatial and temporal gene expression patterns and specificity. This is mainly controlled by DNA binding transcription factors (TFs) that recruit other regulatory proteins in a combinatorial manner, including chromatin remodeling factors to help increase or decrease accessibility of the locus to the RNA polymerase II complex, in order to regulate expression in different regulatory contexts (Britten and Davidson, 1969). This is especially true for TFs that play roles during development.

Many different regulatory proteins combine to form multi-protein transcriptional complexes or enhanceosomes in eukaryotic systems (Wolffè and Drew, 1989) reviewed
in (Merika and Thanos, 2001)). The particular activity of the complex depends on the components of the complex. TFs can be present in vivo in several varyingly composed complexes; by interacting with different components TFs can broaden their spectrum of recognized DNA target sequences (Garvie and Wolberger, 2001; Klemm et al., 1998). In addition, interactions with different regulatory proteins allow a TF to regulate different genes at different development stages and act as either an activator or a repressor of gene expression. Many examples of such complexes are described and well studied in animal systems (Merika and Thanos, 2001; Vanden Berghe et al., 2006). In plants, even though many TFs have been identified, much less information is available about the complexes in which they act. However, the recent advances in the use of interaction proteomics, tandem affinity purification methods, and computational analysis have accelerated discoveries in this field (Van Leene et al., 2011; Van Leene et al., 2007; Van Leene et al., 2008).

As discussed in Chapter 1, LFY is a novel plant-specific TF. Currently limited information is available on how it acts to regulate transcription of its direct targets. LFY acts as a TF both in early and late stages of flower development. At early stages LFY acts to specify floral meristem identity (Liljegren et al., 1999; Weigel et al., 1992). Later, LFY plays a role in regulating organ identity homeotic genes (Busch et al., 1999; Lamb et al., 2002; Wagner et al., 1999; Wagner et al., 2004; Weigel and Meyerowitz, 1993). LFY’s involvement in early and late stages can be genetically separated (Parcy et al., 1998). LFY does not have strong transcriptional activation activity by itself (R.S. Lamb, unpublished data; (Busch et al., 1999)) and strong evidence exists that shows that LFY
acts with other regulatory factors in performing its functions (Chae et al., 2008; Lee et al., 1997; Lenhard et al., 2001; Lohmann et al., 2001; Parcy et al., 1998; Wagner and Meyerowitz, 2002). Considering the ubiquitous early floral meristem expression of LFY (Blazquez et al., 1997; Parcy et al., 1998) and the fact that LFY activates different classes of homeotic genes in different domains/whorls of the flower, it is obvious LFY performs its functions in different complexes recruiting different proteins to activate different genes.

To date, some information is available on LFY co-regulators/interacting proteins during floral organ identity specification. As mentioned in Chapter 1, LFY acts synergistically with the homeodomain protein WUSCHEL (WUS) to activate AGAMOUS (AG), a C class gene responsible for carpel development and floral determinacy (Lenhard et al., 2001; Lohmann et al., 2001). The AG regulatory sequence carries adjacent LFY and WUS binding sites suggesting they could be binding simultaneously or in a common complex (Hong et al., 2003; Lohmann et al., 2001), although this has not been demonstrated. Since WUS is expressed only in the center the meristem, AG is only transcribed in the inner two whorls of the flower (Lenhard et al., 2001; Lohmann et al., 2001). Another TF that has a binding site near that of LFY and is known to genetically interact with LFY is the direct LFY target LATE MERISTEM IDENTITY1 (LMI1), which encodes a Class I homeodomain-leucine zipper (HD-Zip) TF. LMI1 has been shown to act together with LFY to activate CAL expression (Saddic et al., 2006). The LFY binding site and LMI1 binding site in CAL promoter are around 100 bp apart, suggesting a
possible interaction between these proteins LFY and LMI1 (Saddic et al., 2006), although it has not been demonstrated.

SEP3 was recently shown to be a LFY co-factor in addition to its function as an E-class gene (Castillejo et al., 2005; Liu et al., 2009b) (Kaufmann et al., 2009). LFY and SEP3 work together to activate ABCE genes (Castillejo et al., 2005). LFY and SEP3 were shown to physically interact using in vitro GST-immunoprecipitation assay (Liu et al., 2009a)

LFY directly activates APETALA3 (AP3), a B class gene responsible for petal and stamen development (Lamb et al., 2002). In lfy mutants, petal and stamen identity is severely affected (Weigel and Meyerowitz, 1993). AP3 expression is confined to whorls two and three and in a small group of cells at the base of the first whorl (Jack et al., 1992). LFY has been shown to require UNUSUAL FLORAL ORGANS (UFO) to activate AP3 transcription (Ingram et al., 1995; Lee et al., 1997; Levin and Meyerowitz, 1995; Wilkinson and Haughn, 1995). LFY and UFO physically interact at the AP3 promoter (Chae et al., 2008) and UFO overexpression phenotypes require functional LFY, suggesting that UFO is a co-factor of LFY (Lee et al., 1997; Samach et al., 1999).

In addition to these other TFs, LFY has also been shown to need chromatin remodeling factors for function (Wagner and Meyerowitz, 2002). TFs need to bind to the core promoter of the gene and recruit other transcriptional machinery into a compacted chromatin structures. Chromatin remodeling factors will loosen the chromatin using ATP
hydrolysis to allow transcriptional regulators to bind; other factors will modify histones to influence chromatin state (Carlson and Laurent, 1994). Examples of such modifications are phosphorylation, methylation, acetylation, glycosylation, ADP-ribosylation, sumoylation and ubiquitination. Recent work in plants has identified a number of chromatin factors that are recruited to transcriptionally active sites by TFs (for example, (Hernandez et al., 2007; Yu et al., 2008)). SPLAYED (SYD) is a member of the SWI/SNF ATPase family. It has been shown to be necessary for LFY to activate class B gene expression, although no physical interaction between these proteins has been seen (Wagner and Meyerowitz, 2002). That the state of chromatin at LFY active sites is important for their regulation is also supported by the fact that mutations in the CURLY LEAF (CLF) gene, which encodes a polycomb group protein involved in repression of gene expression, cause misexpression of several of these targets, including AP3 (Goodrich et al., 1997).

As discussed above, only two proteins, UFO and SEP3, have been shown to bind to LFY physically and be in a transcriptional complex with it. Identification of components of LFY transcriptional complex(s) will provide us more insight into the function of LFY and transcriptional regulation in plants. Preliminary FPLC gel filtration experiments demonstrated that LFY migrates in a higher order complex or complexes, suggesting that other LFY-interacting proteins remain to be identified (R.S. Lamb, unpublished data). To begin to identify these complexes, we have taken a biochemical approach to identify LFY-interacting partners in Arabidopsis thaliana. Tandem affinity purification (TAP) combined with mass spectrometry (MS)-based protein identification has been a powerful
tool in genome-wide screens for protein complexes in plants (Brown et al., 2006; Rohila et al., 2004; Rohila et al., 2009; Van Leene et al., 2007). Therefore, we used this approach to elute possible LFY transcriptional complex components and identified them using LC/MS/MS analysis. We identified four possible LFY-interacting proteins: At1g03750/CHROMATIN REMODELING FACTOR9 (CHR9), At5g22610, At3g48430/RELATIVE OF EARLY FLOWERING6 (REF6) and At4g06634 and describe preliminary characterization of the first three in this chapter. In addition, we have tested the possible interaction of LMI1 and LFY, based on genetic evidence and demonstrate that these proteins interact together.
3.2 RESULTS

3.2.1 35S::LFY-TAP complements ify-6 mutants

In order to perform a biochemical isolation of a protein complex, you need a very good antibody. The anti-LFY antibodies available to us at the beginning of this study, including one generated in our laboratory, did not have a high enough specificity for use in such a purification. Therefore, we decided to take a tandem affinity purification (TAP) approach to isolation LFY-containing complexes and generated a construct containing the LFY coding sequence cloned in frame to a TAP tag. TAP tag purification has been used successfully to identify interacting proteins in many systems (Doolittle et al., 2009; Rohila et al., 2004; Van Leene et al., 2007), especially because of the capability of purifying proteins using two sequential steps, to eliminate non-specific binding and enrich for targets. LFY is expressed at low levels. We were concerned that using the LFY native promoter would not enable us to isolate enough protein to proceed with. Therefore, we generated LFY-TAP overexpression lines using the 35S constitutive promoter (35S::LFY-TAP). These plants display the LFY gain of function phenotype as described in Chapter 2 (Figure 2.4c). The LFY-TAP line used was crossed into the ify-6 null background and shown to partially complement the mutant phenotype (Figure 3.1a), in a similar manner to untagged LFY driven under the 35S promoter.
3.2.2 Purification and identification of LFY complex components

The TAP procedure relies on two sequential purifications using two different affinity steps. The tandem affinity purification (TAP) tag we used is composed of a calmodulin binding domain (CBD) and two IgG binding domains from *Staphylococcus aureus* Protein A, separated by a TEV protease cleavage site (Figure 3.1b; (Rigaut et al., 1999)). This allows an initial purification using IgG beads followed by release of product by proteolysis and a second purification using calmodulin, thus enriching for product.

Since LFY is expressed in young floral buds, we purified proteins from inflorescence tissues containing floral buds that were still closed, which includes the inflorescence meristem as well as young floral buds. Initially, small-scale TAP tag purifications were done to standardize the protocol, to minimize the loss of LFY-TAP proteins as well as to minimize false positive bindings. As seen in Figure 3.1c, this procedure allowed an enrichment of the LFY-TAP product. The shift in the size of the band is due to the LFY-TAP protein being cleaved through the TEV cleavage site, eliminating the IgG binding domain. Staining of all the proteins present in the gel revealed that other proteins were coming down with LFY-TAP (Figure 3.1d). Two prominent bands corresponding to the IgG heavy and light chains, at 50 and 25 kDa respectively are seen (Figure 3.1d). In addition, several other bands were observed. The sample was then sent to Proteomic Research Services (Ann Arbor, MI), which performed the protein isolation and mass spectrometry (MS) after in gel digestion with trypsin. LC/MS/MS was done on the samples to identify the proteins present. Seven bands (band A-E and L-H) excised from
the gel were analyzed (Figure 3.1d). The raw data sent from Proteomic Research Services is listed in Appendix A.

From the initial list of putative LFY-interacting proteins, we used proteomic (quality and the quantity of MS data; number of unique peptides, number of occurrences, sum ion score) and bioinformatics tools (expressions pattern, subcellular localization, biological activity) to prioritize those proteins most likely to be bona fide LFY-binding proteins (Appendix A). First, we used a published list of contaminating proteins found in 46 different TAP purifications in Arabidopsis (Rohila et al., 2006) and eliminated those on our list that were contained on that list (Appendix A - highlighted in yellow). Examples of these proteins include HSP-70-related proteins, RUBISCO and ATP synthase. Since the TAP tag used in our experiment also contains a calmodulin-binding module (Figure 3.1b), we also eliminated all calmodulins and related proteins from our list (Appendix A – highlighted in green). Once these assumed containments were eliminated, we took into account several properties to further prioritize the proteins. Only proteins with multiple peptide hits were considered. We further considered predicted subcellular localization of the proteins and predicted biological functions. From these considerations, we came up with a short list of putative LFY-interacting proteins (Table 3.1). Encouragingly, the LFY-TAP protein itself was identified as a top pick. In addition as expected, transcription factors, chromatin remodeling factors, and other regulatory proteins were identified. It is quite possible that we have only identified some of the interacting partners of LFY. We collected floral buds at different stages of development, which contain different LFY-containing complexes, diluting any one such complex. In addition, due to the size
increase that occurs as flower development proceeds, older developmental stages are biased for in any such collection. In addition, only relatively highly expressed and stable proteins are likely to be isolated. Despite these limitations, we have identified LFY-interacting proteins (see below).

3.2.3 Putative LFY complex components selected

*At1g03750 (CHR9- CHROMATIN REMODELING 9)*- This poorly characterized gene is a member of SWI/SNF ATPase chromatin remodeling factor family. This family is large in Arabidopsis and includes the previously described SYD protein, known to be necessary for some of LFY’s function (Wagner and Meyerowitz, 2002). This family is characterized by the SNF2 domain and helicase domains (Figure 3.2a) found in proteins involved in transcriptional regulation, DNA-repair and other nuclear functions (Eisen et al., 1995). Examples of some of the other SWI2/SNF2 members that play a role in development include PICKLE (also known as GYNOS) that controls differentiation of carpels and regulates the transition from embryonic to vegetative state in Arabidopsis (Eshed et al., 1999; Ogas et al., 1999). PIE1 controls flowering time in Arabidopsis (Noh and Amasino, 2003), BRAHAMA (BRM) controls flowering and shoot development (Farrona et al., 2004) and CHR11 controls female gametophyte development (Huanca-Mamani et al., 2005). SYD is relatively distantly related to CHR9, but in the same family as BRM (Bezhani et al., 2007). The expression profile of *CHR9* in different developmental tissues, according to the AtGenExpress atlas ((Schmid et al., 2005); ([http://www.weigelworld.org/resources/microarray/AtGenExpress/](http://www.weigelworld.org/resources/microarray/AtGenExpress/)) includes senescing leaves, floral organs and seeds, with the highest expression in pollen. The Arabidopsis
Electronic Fluorescent Pictograph Browser (eFP Browser- (Winter, 2007); http://www.bbc.botany.utoronto.ca/efp/) also shows CHR9 expression in floral organs, the highest being in the pollen and seed. Our preliminary RT-PCR data indicates expression in the inflorescence (Figure 3.3a).

*At5g22610*- This uncharacterized protein belongs to the F-box (FBX) protein family, characterized by the F-box and functioning in SCF E3 ubiquitin ligase complexes (Cardozo and Pagano, 2004). In *A. thaliana* alone 700 FBX proteins have been reported (Gagne et al., 2002; Risseeuw et al., 2003; Xu et al., 2009). In addition to the F-box, At5g22610 contains both leucine-rich-repeats (LRRs), often involved in protein-protein interactions, and a plant specific domain of unknown function, the F-box associated domain (FBD; Figure 3.2b; (Gagne et al., 2002)). The previously characterized LFY-interacting protein UFO is also an F-box protein; however, it is not closely related to At5g22610 (Gagne et al., 2002). At5g22610 is a member of a small subfamily of F-box proteins that includes five other F-box proteins, none of which are characterized (Gagne et al., 2002). *At5g22610* is not represented on the Affymetrix ATH1 microarray; therefore, public expression databases were not helpful. Our preliminary RT-PCR data indicates expression in the inflorescence but not in seedling and roots (Figure 3.3b).

*At3g48430 (REF6- Relative of Early Flowering 6)*- REF6 is a JmjC and N domain containing histone demethylase involved in epigenetic control of gene expression (Noh et al., 2004b). There are 15 potentially active Jumonji (Jmj) domain histone demethylases in *Arabidopsis* (Lu et al., 2008). REF6 contains four C2H2 zinc finger domains in addition
to the Jmj domains (Figure 3.2c). REF6 has been shown to demethylate H3K27me3/2 when overexpressed (Lu et al., 2011). Histone H3 lysine 27 trimethylation (H3K27me3) imparts gene repression (Lu et al., 2011). By demethylating this residue, REF6 helps control flowering time by activating FT and SOC1, leading to accelerated flowering (Noh et al., 2004b). EARLY FLOWERING 6 (ELF6), a paralog of REF6, also functions in flowering time regulation but in an opposite manner (Noh et al., 2004a). Recently it was shown that REF6 interacts with BES1 in brassinosteroid (BR) signaling and biosynthesis. BES1 is a TF and recruits ELF6 and REF6 to regulate target gene expression (Yu et al., 2008). BRs are known to regulate flowering time as well, linking the known functions of this protein (Domagalska et al., 2007).

\textit{At4g03364-} At4g03364 is a novel uncharacterized C2H2-type zinc finger TF. The characterization of this protein will be described in detail in Chapter 4.

\textbf{3.2.4 LFY complex components are localized in the nucleus}

As discussed above, the genes encoding the putative LFY-interacting proteins \textit{CHR9}, \textit{At5g22610} and \textit{REF6} are expressed in floral tissues, overlapping the expression pattern of \textit{LFY}. In order for these proteins to interact with LFY, they must be present in the same subcellular compartment as well. We created GFP fusion proteins of the selected genes driven by the \textit{35S} promoter and performed transient transfection into tobacco leaves to examine subcellular distribution. As discussed previously, LFY-GFP is found in the nucleus and also in cytoplasmic punctate structures that might represent plasmodesmata (Figure 3.3c). Both CHR9 (Figure 3.3d) and At5g22610 (Figure 3.3e) are found in both
the nucleus and cytoplasm, overlapping the localization of LFY. This would allow interaction between these proteins and LFY. REF6 has been shown to localize to the nucleus previously (Yu et al., 2008).

3.2.5 CHR9 interacts with LFY

We isolated the putative LFY-interacting proteins through biochemical purification. In order to confirm physical interactions between these proteins and LFY, co-IP assays with proteins expressed in *N. benthamiana* leaves were performed. When 35S::LFY-GFP and 35S::CHR9-MYC were co-transformed, LFY-GFP was able to co-IP CHR9-MYC (Figure 3.4). Therefore, CHR9 is a true LFY-interacting protein.

3.2.6 Loss of CHR9 causes seed lethality

In order to study the function of *CHR9*, loss of function mutants are necessary. Three independent SALK insertion lines (Alonso et al., 2003), *SALK_064383* (chr9-1; Figure 3.5d), *SALK_064382* (chr9-2) and *SALK_087472* (chr9-3) were screened. All alleles produced homozygous plants, which appeared relatively normal. Transcript levels have not been assessed for these lines. In *chr9-1* homozygous plants, the first 5-10 siliques on every branch are abnormally small compared with wild type. These siliques contain misshapen and shrunken seeds (Figure 3.5a, c). Later formed siliques develop to a normal size with fewer shrunken seeds (Figure 3.5b). This phenotype is similar to that of *syd* and *brm* single mutants and the *syd; brm* double mutant (Bezhani et al., 2007). The lack of full expressivity may be explained if proper seed development depends on SWI/SNF chromatin-remodeling factors and/or that *chr9-1* is only a partial loss of function allele.
Insertion lines were also identified for the other loci encoding putative LFY-interacting proteins. Three insertion lines, *SALK_117573*, *SALK_123361* and *SALK_029402*, were available for the F-box encoding locus *At5g22610*. The first two produced homozygous plants. One insertion line, *SALK_133806*, was studied for REF6, which also produced homozygous plants. These lines have not been further characterized.

3.2.7 *LMI1* interacts with LFY

As mentioned in the Chapter 1, *LMI1* encodes a class I homeodomain ZIP TF and is a direct target of LFY. It then cooperates with LFY in the transcriptional activation of *CAL*. We hypothesized that LMI1 might interact with LFY. To test this, we first did yeast two-hybrid assays. LMI is able to homodimerize (Figure 3.6a) as has been reported for other class I HD-ZIP proteins (Johannesson et al., 2001; Meijer et al., 2000; Sessa et al., 1993). In addition, LMI1 can interact with LFY in yeast (Figure 3.6b).

In order to confirm the interaction between LMI1 and LFY *in vivo*, we performed co-IP studies using transient transfections in *N. benthamiana*. When 35S::*LMI1-GFP* was co-infiltrated with 35S::*LFY-MYC*, LMI-GFP was able to IP LFY-MYC with LMI-GFP (Figure 3.7a). This results supports an interaction between LFY and LMI1 *in vivo*. In order to confirm the co-IP results, we performed BiFC experiments in tobacco leaves. YFP fluorescence is seen in the nucleus when the *LFY* and *LMI1* constructs are present together (Figure 3.7b), confirming the physical interaction between these two proteins.
LMI1 localizes to the nucleus (Figure 3.7d), overlapping the subcellular localization of LFY. According to (Saddic et al., 2006), LMI1 shows expression in the bract primordial during bolting, and strongly expressed in the incipient flower primordial after bolting, overlapping with LFY's expression pattern. Thereafter, the expression proceeds to the young floral primordia to stage 1 and 2 flowers, where LFY is still expressed. Later in floral development LMI1 is expressed in sepals and stamens at stage 7 and in petals in at stage 9.

3.2.8 Overexpression of LMI1 causes floral defects

Loss of function mutations in LMI1 have been described (Saddic et al., 2006). We generated overexpression lines to examine gain of function phenotypes, which show severe floral defects (8/10 independent lines). Abnormalities were seen in all floral organs (Figure 3.8a-g). The inflorescence structure is different from wild type, as there is less internode elongation between flowers, resulting in a clump of flowers at the top of the inflorescence (Figure 3.8a-c). This might reflect production of more flowers or more rapid production of flowers in addition to or instead of problems with internode elongation. Sepals and petals of 35S::LMI1 flowers are serrated and vary in number and size (Figure 3.8d-f). In most of the flowers, petals are entirely absent (Figure 3.8g). Stamens are underdeveloped (Figure 3.8d-g). Carpels are stunted and protrude from the undeveloped sepals, petals and stamens (Figure 3.8a-g). This suggests that floral development is sensitive to the dose of LMI1.
3.3 DISCUSSION

The floral meristem identity gene LFY encodes a plant specific transcription factor. Previously, very little information has been available about what other proteins may function together with LFY to regulate gene expression. Here we describe the identification and initial characterization of CHR9, REF6, At5g22610 and LMI1, putative components of LFY transcriptional complexes.

3.3.1 Putative LFY-interacting proteins identified by TAP purification

We used TAP purification to isolate LFY-containing protein complexes. This method has been previously used successfully to isolate protein complexes in Arabidopsis (Rohila et al., 2004; Rubio et al., 2005). From this purification, we identified several putative LFY-interacting proteins, which are discussed below.

CHR9 is a member of the helicase/SWI/SNF chromatin remodeling factor family (Shaked et al., 2006). This family uses the energy of ATP hydrolysis to remodel chromatin structures, allowing transcription factors access to DNA and also functions in DNA repair. At least 39 proteins in the Arabidopsis genome encode members of the SWI/SNF family (Verbsky and Richards, 2001). CHR9 has been classified as being more related to SWI/SNF family members with roles in DNA repair; however, RNAi studies revealed that this is not its likely function (Shaked et al., 2006). Another member of the
SWI/SNF family, SYD, has previously been shown to work in cooperation with LFY (Wagner and Meyerowitz, 2002). SYD acts as a repressor of LFY activity before floral transition under non-inductive conditions and acts as a co-activator in induction of B and C class floral homeotic genes. The closest related protein to SYD in the Arabidopsis genome, BRM, also acts to activate floral homeotic gene expression (Hurtado et al., 2006).

Our expression studies demonstrate that CHR9 is expressed in floral organs, pollen and seeds. We identified 3 putative insertional mutants in the CHR9 locus; one of these, chr9-1, contains an insertion close to the 5′ end of the gene (Figure 3.5d). chr9-1 homozygous plants had aborted ovules or undeveloped seeds in their fruit, suggesting that its function may be necessary in ovule or seed development or both. This phenotype resembles the phenotype of syd; brm double mutants (Bezhani et al., 2007). It is possible that seed development depends on many SWI/SNF chromatin remodeling factors, which explains the incomplete expressivity observed. The interaction of CHR9 with LFY was validated by co-IP, suggesting that it is present in an LFY-containing transcriptional complex. As described for other SWI/SNF factors in plants (Farrona et al., 2004; Hurtado et al., 2006; Kwon et al., 2005; Sarnowski et al., 2005; Su et al., 2006; Wagner and Meyerowitz, 2002), CHR9 could be acting as either a co-activator or a repressor in transcriptional complexes. It is possible that CHR9 is acting as such in LFY-containing complex by ensuring accessibility to DNA through chromatin remodeling. However, the function of CHR9 in LFY-dependent gene expression remains to be elucidated.
We also identified another factor that acts at the chromatin level, REF6, which is known to function in flowering time regulation (Noh et al., 2004b). REF6 is a histone H3 lysine 27 demethylase that specifically demethylates H3K27me3 and H3K27me2. H3K27me3 methylation is mediated by polycomb group proteins (reviewed in (Ito and Sun, 2009)), promoting gene repression. REF6 acts to reverse this mark, relieving repression and allowing gene expression (Lu et al., 2011). Ectopic expression of REF6 allows expression of several direct targets of LFY, including AP1, AP3 and AG, in seedlings where they are not normally expressed (Lu et al., 2011). This is similar to the loss of function phenotype of CLF, a polycomb group histone methyltransferase, known to repress expression of these genes (Katz et al., 2004). This suggests that REF6 is acting with LFY to activate its target genes. Although we have not confirmed the physical interaction between LFY and REF6, there is other unpublished evidence that REF6 is interacting with LFY (personal communication-Dr. Doris Wagner). Further work needs to be done to explore the relationship between REF6 and LFY.

UFO has been shown to be physically associated with LFY (Chae et al., 2008) and be necessary to activate B class genes (Ingram et al., 1995; Lee et al., 1997; Levin and Meyerowitz, 1995; Wilkinson and Haughn, 1995). UFO belongs to the F-box family of E3 ubiquitin ligases. We identified another F-box protein in our TAP purification, At5g22610. At5g22510 contains LRR repeats and a FBD domain in addition to the F-box and belongs to a subfamily containing at least three other proteins (At5g22730, At5g44490, At5g22720; (Gagne et al., 2002)). None of these proteins have been
characterized. *At5g22610* is expressed in floral tissues (Figure 3.3b). No further work has been done on this protein to date.

### 3.3.2 LMI1 interacts with LFY.

*LMI1* was originally identified as a target gene of LFY and encodes a transcription factor that cooperates with LFY in the transcriptional activation of *CAL* (Saddic et al., 2006). In the *CAL* promoter the identified LMI1 and LFY binding sites are only 100 bp apart (Saddic et al., 2006). We hypothesized that LMI1 and LFY might physically interact to regulate transcription. These two proteins can interact directly in yeast 2-hybrid analysis (Figure 3.6b) and *in planta* by co-IPs and BiFC (Figure 3.7). *LMI1* is strongly expressed in bract primordia during bolting and in incipient flower primordia, overlapping with the expression of both *LFY* and *CAL*. It is expressed throughout stage 1 and 2 flowers but is expressed only in sepals and stamens at stage 7 and in petals at stage 9 (Saddic et al., 2006). Our data suggests that LMI1 may form a subunit of LFY-containing transcriptional complexes that act early in floral development,

### 3.3.3 Conclusions

We have used biochemical approaches to identify components of LFY-containing complexes. As might be expected, we recovered other transcription factors as well as chromatin remodeling and modifying enzymes. Interestingly, of the four genes encoding proteins identified by TAP purification (*CHR9, REF6, At5g22610* and *FOL1* (see Chapter 4)), only *CHR9* is represented on the ATH1 GeneChip. Therefore, these genes would be not be identified by expression data mining approaches. In addition, mutations
in two of the genes, \textit{chr9-1} and \textit{fol1} (see Chapter 4) appear to be either gametophytic or embryonic lethal. This means that mutations in these genes would be unlikely to be isolated in traditional genetic screens for factors acting in floral development. \textit{At5g22610} is a member of a multigene family; it may be redundant with other members of its family, making isolation based on mutant phenotype less likely. Therefore, using the TAP approach to identify LFY-interactors has proven to be successful in identifying novel actors in floral development.
3.4 EXPERIMENTAL PROCEDURES

3.4.1 Plant material and growth conditions

The SALK lines (Alonso et al., 2003) used in this study were obtained from the Arabidopsis Biological Resource Center (ABRC; http://www.arabidopsis.org/abrc) at The Ohio State University, Columbus, OH. *lfy*-6; *35S::LFY-TAP* lines were created by crossing *lfy*-6 mutant lines with homozygous *35S::LFY-TAP* lines. Arabidopsis plants used in this study were in Landsberg erecta (L. er) background and all SALK lines used were in Columbia (Col-0) background. The *lfy*-6 allele was previously described (Weigel et al., 1992).

Plant growth conditions for Arabidopsis and *N. benthamiana* are as described in Chapter 2. Arabidopsis transformants were selected on selection plates as described in Chapter 2.

3.4.2 Constructs and cloning

Cloning of *CHR9, At5g22610, REF6 and LMI1* coding regions into pENTR/D-TOPO Gateway entry vector was done as described in Chapter 2. The primers used are listed in Table 3.2. The PCR products were sequenced at the PMGF and were then recombined into binary destination vectors *pGWB5* (GFP tag at C-terminus) and into *pGWB20* (MYC tag at C-terminus). The entry vector carrying the *LFY* coding region was recombined into *pGWB29* (TAP tag C-terminus; a gift from T. Nakagawa, Shimane University,
Matsue, Japan). These were then introduced into *Argobacterium tumefaciens* strain GV3101 for Arabidopsis plant stable transformation and *N. benthamiana* transient co-infiltrations. The 35S::empty-GFP vector in *Agrobacterium* was obtained from Dr. Iris Meier, The Ohio State University, Columbus, OH.

BiFC vectors *p2YN* (YFP amino acids 1-158), *p2YC* (YFP amino acid 159-238) and *pYN*-1 and *pYC*-1 carrying either fragment of split YFP in C- and N- termini respectively were used as in Chapter 2. The procedure described in Chapter 2 for cloning *LMII* into these vectors was used using the primers provided in Table 3.2. Cloning into the vectors and transformation in to *Agrobacterium* was done as in Chapter 2. Previously cloned *LFY* in BiFC vectors were used for this experiment (Chapter 2).

For yeast two-hybrid assays, the *pENTR/D* clones with *LMII* coding region were recombined with *pDEST22* and *pDEST32* destination vectors with GAL4 Activation Domain (AD) and DNA Binding Domain (BD), respectively (ProQuest Two-Hybrid system, Invitrogen). Transformation into yeast strain PJ694A and Y2H assay was done as in Chapter 2.

### 3.4.3 TAP tag purification and detection

Inflorescence tissue containing primarily unopened buds was collected. Twenty-five grams (fresh weight) was used for this study. Total proteins were extracted using 1:1(v/v) of extraction buffer as in (Zhao et al., 2008) with the following modifications: 0.1% NP-40 instead of 0.25% dodecyl maltoside. The TAP tag purification using Protein A
Agarose beads (Invitrogen) and Calmodulin Affinity Resin (Stratagene) was performed as described in (Zhao et al., 2008). Western blots to detect the enriched LFY-TAP protein was done as described in Chapter 2 using Peroxidase Anti-Peroxidase (anti-PAP) (Sigma, P-1291) antibody to detect protein A. Proteins were separated using SDS-PAGE gels and silver stained with Silver Stain Plus (Bio-Rad).

3.4.4 Mass spectrometry analysis

Mass spectrometry analysis was performed by Proteomic Research Services, Inc. (Ann Arbor, MI, http://www.prsproteomics.com). In-gel digestion was performed on a ProGest workstation, incubating at 37°C for four hours with trypsin. The resulting peptide mix were analyzed by nano LC/MS/MS on a Micromass Q-Tof 2 and/or a Finnigan LTQ. 15μl of hydrolysate were processed on a 75μm C18 column at a flow-rate of 200nL/min. MS/MS data were searched using MASCOT (http://www.matrixscience.com) according to specified parameters. A detailed protocol is provided in http://www.prsproteomics.com).

3.4.5 Yeast two-hybrid interaction assay

Yeast competent cells preparation and transformations, and interaction studies were performed as described in Chapter 2. pDEST22-LMI1 were co-transformed with pDEST32-LMI1 and pDEST22-LMI1 was co-transformed with pDEST32-LFY and vice versa. As negative controls, empty vectors were co-transformed with above vectors carrying LMI1 or LFY.
3.4.6 Bimolecular Fluorescence Complementation Assay (BiFC)

BiFC was performed in transiently transformed \textit{N. benthamiana} leaves and observations were done as described in Chapter 2. \textit{Agrobacterium} with BiFC vectors carrying \textit{LFY} and \textit{LMII} were co-infiltated. Empty vectors with one half of the YFP together with the vector carrying \textit{LMII} gene fused to the other half of the YFP was used as negative controls as well as the two empty vectors.

3.4.7 Computational analysis

On-line tools were utilized to study the available and predicted information on \textit{CHR9}, \textit{At5g22610} and \textit{REF6}. Gene information was gathered from The Arabidopsis Information Resource (TAIR) database (http://www.arabidopsis.org). Protein domains were identified using National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov) and Universal Protein Resource (UniProt) databases (http://www.uniprot.org).

3.4.8 Genomic DNA preparation and genotyping

Genomic DNA was extracted from leaves as described in (Teotia and Lamb, 2009). This was used as template in PCR reactions, to verify the presence of the T-DNA insertion in plants from SALK collection. LBb1 and RP primers were used to genotype the insertion and LP and RP was used to amplify the wild type locus (for primers refer to Table 3.2). The sequence for above primers was obtained from SALK website (http://signal.salk.edu/). PCR was done using Choice-Taq Blue DNA Polymerase

3.4.9 Co-immunoprecipitation assay

Co-IP assays with transiently transformed *N. benthamiana* leaves were done following the protocol in Chapter 2.

3.4.10 *lfy*-6 mutant complementation experiments

35S::LFY-TAP lines were crossed to *lfy*-6 mutants. F2 generation plants were genotyped. dCAPS primers (Neff et al., 1998) were used to screen for *lfy*-6; 35S::LFY-TAP lines (listed in Table 3.2). In the F2 generation phenotypic analysis was done to check for complementation of *lfy*-6 null mutants. Complimented lines were analyzed to look for floral defects and number of branches formed before flower formation. The presence of petals and stamens, fused carpels and four sepals in the outer most whorl of the flower were some of the characters analyzed. We compared phenotypes of these lines with *lfy*-6 null mutants and wild type grown in the same growth area at the same time.

3.4.11 Phenotypic analysis of transgenic lines

Homozygous overexpression lines of *LMI1* were screened for floral defects. Phenotypic analysis was attempted on *chr9-1*. Plants were observed at all stages, looking mostly for floral development defects. Siliques were dissected to observe the misshapen and shrunken seeds.
Figure 3.1. Tandem affinity purification of LFY-containing complexes from Arabidopsis inflorescences. (a) Micrographs of flowers of the indicated genotypes. (b) Amino acid sequence of the TAP tag fused in frame to LFY with important motifs indicated. (c) Western blot showing results of TAP purification. Anti-PAP antibody was used to develop Western. Size of the product is shown on the right of the blot. (d) Representative silver stained gel of final product of the large-scale TAP purification. Size (in kiloDaltons) is indicated on the right. The letters refer to the areas of the gel that were excised and processed further.
**Figure 3.2. Gene structures of the genes encoding putative LFY-binding proteins.** Predicted protein domains are shown below the exons (boxes) that encode them and final protein product size in amino acids is shown at right. (a) *CHR9*. (b) *At5g22610*. (c) *REF6*. 
Figure 3.3. Putative LFY-binding proteins are found in inflorescences and in the same subcellular compartments as LFY. (a and b) RT-PCR showing expression in inflorescence. (a) *CHR9*. (b) *At5g22610*. ACTIN was used as the reference gene. IN-inflorescence SE-seedling RT-root (c-e) Subcellular localization in *N. benthaminana* leaves (c) LFY-GFP. (d) CHR9-GFP. (e) At5g22610-GFP. Arrowheads indicate the nuclei and the arrows indicate the cytoplasm.
Figure 3.4. LFY and CHR9 physically associate in vivo. Western blots showing the result of co-immunoprecipitation from transiently transfected N. benthamiana leaves. LFY-GFP was able to bring down CHR9-MYC and LFY-MYC (red arrowheads). Co-IP done with At4g06634 (white arrowhead) will be discussed in Chapter 4.
Figure 3.5. Reduced fertility is observed in chr9-1 plants. (a-c) Micrographs of siliques. (a) Early arising siliques from wild type (upper silique) and chr9-1 (lower silique) plants. The silique from chr9-1 has misshapen and shrunken seeds indicated by an arrowhead. Wild type has fully developed seeds and few unfertilized ovules, marked by an arrow. (b) A later arising chr9-1 silique with several misshapen and shrunken seeds (arrowheads) and unfertilized ovules (arrow) and normal developed seeds (asterisks) (c) Mature, fully desiccated chr9-1 showing shrunken seeds (arrowhead). (d) Schematic representation of CHR9 locus indicating insertion site of T-DNA in chr9-1 allele. Filled boxes represent exons and lines introns. Inverted triangle represents the T-DNA insertion position, in the first exon.
Figure 3.6. LMI1 interactions in Y2H assays. (a) LMI1 homodimerizes. (b) LMI1 interacts with LFY. Yeast transformed with constructs marked above the figure. Selected on SD-Leu-Trp-His dropout selection media.
Figure 3.7. LMI1 and LFY physically associate in vivo. (a) Western blots showing the results of co-immunoprecipitation from transiently transfected N. benthamiana leaves. Only LMI-GFP was able to bring down LFY-MYC (red arrowhead) but not GFP alone.

Continued
Figure 3.7: continued

Asterisk mark non-specific bands. (b-d) Confocal micrographs of transfected *N. benthamiana* leaves. Constructs used are indicated at the right or at the bottom of the panel. (b) LFY and LMI1 interact, creating fluorescence in the nucleus. Inset is magnification of the nucleus. (c) No fluorescence is detected when empty vectors are used. (d) LMI1-GFP is localized to the nucleus.
**Figure 3.8. Overexpression of LMII causes floral development defects.** (a-g) Micrographs of 35S::LMII-GFP transgenic plants. (a-c) Inflorescences. (d-g) Flowers. (g) Sepals were removed to reveal absence of petals and poorly developed stamens and carpels. Yellow arrowheads indicate sepals, red arrowheads indicate petals and arrow indicate stamens. Protruding carpels marked with asterisks.
Table 3.1. Putative LFY-interacting proteins selected from TAP-purification.

<table>
<thead>
<tr>
<th>Category</th>
<th>AGI ID</th>
<th>Gene Name</th>
<th>Gene Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transcription Factors</td>
<td>At5g61850</td>
<td>LFY</td>
<td></td>
</tr>
<tr>
<td></td>
<td>At4g06634</td>
<td>Unknown</td>
<td>C2H2-type Zinc finger</td>
</tr>
<tr>
<td>Chromatin Remodeling</td>
<td>At3g48430</td>
<td>REF6</td>
<td>Histone demethylase</td>
</tr>
<tr>
<td></td>
<td>At1g03750</td>
<td>CHR9</td>
<td>SWI/SNF chromatin remodeling</td>
</tr>
<tr>
<td>F-box</td>
<td>At5g22610</td>
<td>Unknown</td>
<td>F-box/LRR</td>
</tr>
</tbody>
</table>
Table 3.2. List of primers used in Chapter 3.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Primer name</th>
<th>Primer Sequence (5’-3’)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMI1 BiFC cloning</td>
<td>LMI1p2YCF</td>
<td>TACGAACGATAGTTAATTAACCATGGAGTGGTGCAACAACGAGCAAC</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>LMI1p2YCR</td>
<td>CACCTCCTCCACTAGTGGGGTATGACGGCCAGCCAGGAG</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>LMI1pYC-1F</td>
<td>CTGGAGGAGGCAGCGCCTGGAAGTGTCACCAACAGAGCAAC</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>LMI1pYC-1R</td>
<td>ATTTGCGGACTCTAGATCGGTATGACGGCCAGCCAAGGAG</td>
<td>This study</td>
</tr>
<tr>
<td>CHR9 coding region cloning</td>
<td>CHR9GWF</td>
<td>CACCATGTCGTTGATACACACTTTCAAGAGA</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>CHR9GWR</td>
<td>TTACTTGACTCTTCTAAAGTC</td>
<td>This study</td>
</tr>
<tr>
<td>REF6 coding region cloning</td>
<td>pDNRREF6F</td>
<td>GGGGACAAGGTGGTCAAAAAAACAGGCTCCATGCGGTTCCAGGAGGTC</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>pDNRREF6R</td>
<td>GGGGACCACTTTGACTCAAGAAAGCTGGGTCCCTTTGTCTGTCG</td>
<td>This study</td>
</tr>
<tr>
<td>At5g22610 coding region cloning</td>
<td>FBGWF</td>
<td>CACCATGGAGGAAAAATATAGAGAAAAAGAATATGCG</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>FBGWR</td>
<td>ACTCAATGGCACCATAATAGGATTCAGCAAC</td>
<td>This study</td>
</tr>
<tr>
<td>LMI1 coding region cloning</td>
<td>LMI1F</td>
<td>CACCATGGAGTGTCACCAACAGGAG</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>LMI1R</td>
<td>TCAGGGGTATGACGGCCAGCC</td>
<td>This study</td>
</tr>
</tbody>
</table>

Continued
Table 3.2: continued

<table>
<thead>
<tr>
<th></th>
<th>dCAPS</th>
<th>lfy-6F</th>
<th><a href="http://www.weigelworld.org/resources/mutations/caps">http://www.weigelworld.org/resources/mutations/caps</a></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AAG CAG CCG TCT GCG GTG TCA GCA GCT GTT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTG TCA ATT TCC CAG CAA GAC AC</td>
<td></td>
</tr>
<tr>
<td>SALK line</td>
<td>S133806LP</td>
<td>GCCTTTGTTACTCTGATGAACATCTCTC</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td>genotyping</td>
<td>S133806RP</td>
<td>AAGCCATGGAACACATCTTTG</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>S087472LP</td>
<td>GATAGGTTTCCTTTCCATTTTCGGG</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>S087472RP</td>
<td>CGGAATATCTGCAAGCAAGAG</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>S117573LP</td>
<td>GAAGAGAATCCATGCTTGCCAC</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>S117573RP</td>
<td>ATTTGGAAGTTGGTGGCAGTGC</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>S064382LP</td>
<td>CGAGAGAATTCTCCAAAACCG</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>S064382RP</td>
<td>TCCAGAATTTTGAGAAGAACCATC</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>S064383LP</td>
<td>CACAGAAAGAGGTGTCTGTGTGTC</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>S064383RP</td>
<td>TCGGAATCTCTCAAATTGATG</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>S029402LP</td>
<td>CATTGGCTCTCTCTGTTGCAATTG</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>S029402RP</td>
<td>ACCCTTGAATATTCTTTATGGGG</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>S123361LP</td>
<td>TTGTGATTGATGGCCCTAGAC</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>S123361RP</td>
<td>GCTAGTCTCGGGTCATAGCACC</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>LBb1</td>
<td>GCGTGGACCCTTGCTGCAACT</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
</tbody>
</table>
REFERENCES


CHAPTER 4

The C2H2 Zinc Finger Transcription Factor FRIEND OF LEAFY1 (FOL1) has Roles in Meristem Identity and Size Control in *Arabidopsis thaliana*
4.1 INTRODUCTION

As discussed in the previous chapter, we used Tandem Affinity Purification (TAP) followed by mass spectrometry to identify individual components of LFY complexes. One protein identified was a novel C2H2-type zinc finger protein, which we have named FRIEND OF LEAFY1 (FOL1).

Zinc finger proteins (ZFPs) play a critical role in many cellular functions, including DNA binding, RNA binding and protein-protein interactions. In general, zinc finger proteins can be classified based on the number and the order of the cysteine (C) and histidine (H) residues that bind zinc ions. Classes include C2H2, C2C2, C2HC, C2C2C2C2 and C2HCC2C2 (Klug and Schwabe, 1995; Mackay and Crossley, 1998; Sanchez-Garcia and Rabbitts, 1994). Among these, the C2H2-type ZFPs are one of the best-studied classes (Laity et al., 2001). These proteins act as transcription factors with the zinc fingers binding DNA (Pabo et al., 2001). C2H2 proteins constitute one of the largest families of transcription factors in plants (Englbrecht et al., 2004).

The C2H2-type motif was first identified in the transcription factor TFIIIA in studies of *Xenopus laevis* oocytes (Miller et al., 1985). Some C2H2-type ZFPs are still referred to as TFIIIA-type proteins or are sometimes referred to as Kruppel-type proteins, based on
the *Drosophila* Kruppel protein (Miller et al., 1985; Ollo and Maniatis, 1987). C2H2-type ZFPs display a wide range of functions through DNA or RNA binding, and through involvement in protein-protein interactions (Grishin, 2001). C2H2-type ZFPs contain one of the best-characterized DNA binding motifs found in eukaryotes. This 28-30 amino acid motif contains two conserved C and two conserved H residues binding to one zinc atom tetrahedrally (Pabo et al., 2001). This zinc-dependent structure is required for the DNA binding interaction. In the absence of the zinc ion or if the zinc fingers lose their ability to fold through residue mutations, DNA binding is abolished (Frankel et al., 1987; Lee et al., 1989; Miller et al., 1985; Pavletich and Pabo, 1991). Each finger forms two beta-strands and one alpha helix (Wolfe et al., 2000) and has an amino acid sequence motif of X2-Cys-X2, 4-Cys-X12-His-X3, 4, 5-His (Pabo et al., 2001). The C2H2 zinc finger motif recognizes the DNA sequence by binding in the major groove through its alpha helix (Wolfe et al., 2000). Transcription factors (TFs) typically contain several zinc fingers (each with the beta/beta/alpha structure), facilitating multiple contacts along the DNA, increasing binding specificity.

In *Arabidopsis thaliana*, an *in silico* analysis identified 176 C2H2 ZFPs, constituting one of the most abundant families of putative TFs in this organism (Englbrecht et al., 2004). Only a minority of these is conserved in other eukaryotes while 81% are plant-specific (Englbrecht et al., 2004). Two structural features distinguish most of the plant C2H2 ZFPs. In plants, the zinc fingers are parted by long spacers that vary in length and sequence (Sakamoto et al., 2004; Takatsuji, 1999), whereas in yeast and mammals the
zinc fingers are clustered and separated by short spacers (6-8 amino acids) known as H-C linkers (Klug and Schwabe, 1995). In addition, most of C2H2-type ZFPs in plants have an invariant QALGHH motif within the finger, which is lacking in other organisms (Takatsuji, 1999). It has been shown that this motif is critical for DNA binding activity in those proteins that contain it and mutating any of these residues reduces or completely abolishes DNA binding ability (Kubo et al., 1998). For example, the Arabidopsis SUPERMAN protein necessary for the proper spatial development of reproductive floral tissues, which contains one C2H2 zinc finger, losses its function when the second glycine of the QALGHH motif was mutated to aspartic acid (Sakai et al., 1995). Englbrecht et al., 2004 further subdivided the C2H2 type ZFPs of Arabidopsis into three different sets (A, B and C), each divided into subsets (C1, C2 and C3), which in turn were further divided into different families and subclasses. Set A consists of ZFPs with tandem ZF arrays containing up to five ZFs. Set B consists of tandem ZF arrays with more than five ZFs. Set C contains a single ZF or several dispersed ZFs. This set can be further classified depending on the spacing between the two invariant zinc coordinating H residues by 3 (C1), four (C2) or five (C3) amino acid residues. A complete list of all Arabidopsis proteins categorized into these classes is given in (Englbrecht et al., 2004). Many ZFPs in plants have already been characterized and shown to be involved in many processes such as the regulation of floral organogenesis, gametogenesis, leaf initiation, lateral shoot initiation and stress response (Takatsuji, 1999).

FOL1 is a novel unknown C2H2-type ZFP encoded by locus At4g06634, and has five
zinc finger motifs, out of which four are placed in a tandem array and one separated from this array (Figure 4.1). All five zinc fingers contain the amino acid sequence motifs typically found in C2H2-type ZFPs. FOL1 falls into set A according to (Englbrecht et al., 2004) and has at least two putative orthologs in maize, TRANSCRIPTION REPRESSOR MAIZE 1 (TRM1) and YIN-YANG1 (YY1; (Xu et al., 2001)), which are members of the YY1-family. The founding member of this family is an animal protein, YY1, which has been shown to act as either a transcriptional activator or repressor depending on context. TRM1 was shown to bind to YY1-like DNA binding sites and to two other sites with no sequence similarity to the YY1 site. Like TRM1, FOL1’s four tandem zinc fingers are of the Kruppel-type. Members belonging to Kruppel/ TFIIIA group contain finger structures with (Tyr/Phe)-X-Cys-X2, 4-Cys-X3-Phe-X5-Leu-X2-His-X3, 5-His conserved sequence (Berg, 1990). According to (Englbrecht et al., 2004), TRM1 and FOL1 have tandem fingers with rare spacing and unusually short linkers, features that are conserved in the putative orthologs from beyond plants. FOL1’s zinc finger regions do not contain the QALGGH motif characteristic of most plant C2H2-type ZFPs.

In addition to its five zinc fingers, FOL1 also contains a putative SFPI domain, which is a transcriptional repressor domain regulating the G2/M transition, and a glutamic acid rich region near its C-terminus that might function as a transcriptional activation domain (Figure 4.1; domains as annotated at the National Center for Biotechnology Information - NCBI). In addition, a phosphoserine at residue 284 is predicted.
The *FOL1* gene produces at least three splice variants identified (Figure 4.1c), with the first variant the most abundant (TAIR). The gene consists of six exons (Figure 4.1c). The only other *Arabidopsis thaliana* protein similar to FOL1 is At5g04240, a C2H2-type zinc finger family protein/jumonji family protein. This match is only due to the conserved zinc finger motifs and is unlikely to reflect similar function. Therefore, *FOL1* is a unique gene in the Arabidopsis genome, facilitating its study.

This chapter presents our work characterizing the unique and important *FOL1* gene in *Arabidopsis thaliana*. *FOL1* is expressed in the inflorescence and floral meristems, in developing floral organs, male and female gametophytes and developing embryos and young seedlings. It is not expressed in vegetative tissues. Loss and gain of function studies suggest that *FOL1* functions in meristem size control and gametophyte development. In addition, *FOL1* interacts with *LFY*, both physically and genetically, and has both LFY-dependent and independent functions. We demonstrate that *FOL1* is an essential gene in Arabidopsis and is necessary for proper reproductive meristem development.
4.2 RESULTS

4.2.1 FOL1 has orthologs across the land plants.

*FOL1* is a single copy gene in Arabidopsis; however, it has orthologs across the land plants. As shown in Figure 4.2, the FOL1 family of proteins is found from the moss *Physcomitrella patens* into angiosperms. In these orthologs, the four tandem zinc fingers and the isolated 5th zinc finger are highly conserved (Figure 4.3). The putative SPF repression domain detected in FOL1 overlaps with the 5th zinc finger and is also conserved in all the identified plant FOL1 orthologs. However, the Glu-rich region found at the C-terminus is only conserved in the angiosperm orthologs, and not in the moss or *Selaginella* proteins (Figure 4.3). Currently, no orthologs have been identified in the vascular and seed plants between *Selaginella* and angiosperms; therefore it is unclear when the FOL1 clade acquired this sequence. Although there is some similarity between FOL1 and the human YY1 transcription factor, this similarity is confined to the four tandem zinc fingers (which are found at the C-terminus of YY1 (Shi et al., 1991)). It is unclear if FOL1-like proteins are true orthologs of YY1. However, the conservation of this FOL1-like proteins in all land plants is consistent with a role in LFY ortholog function, since LFY-like proteins are also found in all land plants (Maizel et al., 2005).
4.2.2 The spatial and temporal expression pattern of FOL1 marks sites of cell proliferation

FOL1 is an uncharacterized gene; therefore, there is no available experimental data. In addition, FOL1 is not represented on the Affymetrix ATH1 chip, the most common microarray used generate the publically available expression data for Arabidopsis thaliana. Expression pattern information for this gene thus had to be generated de novo. RT-PCR data indicates FOL1 is expressed in inflorescence and siliques (Figure 4.4a); however, this does not provide spatial expression data. We generated a construct consisting of a 2.6kb region 5’ upstream from the transcription start site of FOL1 fused to the coding region of the beta-glucoronidase (GUS) encoding gene (FOL1p::GUS). Ten independent initial transformation events were analyzed. Staining was observed only in the reproductive tissues, in young inflorescences and siliques (Figure 4.4b, d), with reduced expression in mature flowers and siliques. Within the flower, expression of FOL1p::GUS is present in developing petals (Figure 4.5a inset), stamens and carpels (Figure 4.5a). This data is consistent with the RT-PCR data. Intense staining was present in the male and female gametophytes (Figure 4.5a-c) and in the very young developing embryo inside the young seed (Figure 4.5d).

In the next generation (T1), we examined FOL1p::GUS expression throughout the lifecycle of Arabidopsis. For this, a time course experiment was performed from germination to 27-day-old plants. Staining was done every 2 days starting 3 days after the break from vernalization. As Figure 4.6a clearly shows, we observed FOL1p::GUS
expression in 3-day-old germinating seedlings. By day 5-7 the expression has diminished, with some staining, perhaps an artifact due to the long half life of the GUS product, observed in the vasculature in the leaves (Figure 4.6b, c). No FOL1p::GUS expression was observed during the rest of the vegetative phase of the plant (Figure 4.6d-f). This appears to contradict the RT-PCR data presented earlier; however, 10-day-old seedlings had been the source of the cDNA used in that experiment. This explains why no FOL1 expression was seen in those seedlings using RT-PCR. At day 15, FOL1p::GUS expression was observed in the shoot meristem area (Figure 4.6g), which by now has transformed into an inflorescence meristem, and also in the incipient floral meristems, suggesting that FOL1 is either involved in the transition from VM identity to IM identity or its expression is activated shortly after IM initiation. At day 25 (Figure 4.6l) and 27 (Figure 4.6m) FOL1 expression was observed at the base of the inflorescence axis where the axillary buds are forming branches. This further supports the observation that FOL1 is expressed at the initiation or just after the inflorescence meristem is formed.

Overall, the expression pattern of FOL1::GUS suggests that FOL1 is expressed at sites of cell proliferation, including male and female gametophytes, the embryo and in the inflorescence and young floral meristems.
4.2.3 Phenotypic analysis of insertion mutant lines exhibits the necessity of FOL1 for gametophyte development

In order to determine the function of FOL1, it is important to identify loss of function mutants. Two T-DNA insertional mutants in this locus were identified in the publically available mutant collections (SALK_040806C and SALK_041031; http://www.signalsalk.edu). FOL1 has 6 exons; in both the lines the insertion was predicted to be near the end of the fifth exon, which we confirmed through sequencing (Figure 4.7a).

The SALK_040806C line, hereafter referred to as foll-1, is homozygous for the insertion in FOL1. RT-PCR data revealed that a truncated FOL1 transcript is made in these plants (Figure 4.7b). This transcript is predicted to form a protein of about 21kD, which is half the size of the full-length protein. We predict that foll-1 would be a hypomorphic allele, but not a null allele. Although foll-1 appears to be phenotypically normal, examination of the siliques formed revealed the presence of a large number of aborted or unfertilized ovules (Figure 4.7c). Furthermore, Alexander staining of the anthers of the stamen revealed inviable pollen compared to wild type (Figure 4.7d). This preliminary data suggest that FOL1 may be necessary for gametophyte development, consistent with its expression during this process.

Unlike foll-1, homozygous plants were not observed for the SALK_041031 line, hereafter referred to as foll-2. This could be due to gametophytic or embryonic lethality of the
allele. Alexander staining (Figure 4.7d) shows a lot of inviable pollen in anthers of $foll$-2/+ plants. In addition, preliminary germination counts done with this line revealed a germination rate of only 36.2% compared of that of wild type, 75.29%. These results suggest that $FOL1$ is necessary for gametophytic and/or embryo development. In order to determine whether the insertion in $FOL1$ is responsible for these phenotypes or if another linked mutation causes the lethality, a rescue construct carrying $FOL1$ gene under its native promoter ($FOL1p::gFOL1$) was introduced into these plants heterozygous for the insertion. If this construct compliments the phenotype observed, this would support the hypothesis that lack of functional $FOL1$ causes lethality in Arabidopsis and that it is an essential gene.

4.2.4 Knock down of $FOL1$ causes meristem defects

The inconsistent results seen with the insertional alleles of $FOL1$, and possible gametophytic defects and embryo lethality involved, makes it difficult to study the functionality of the $FOL1$ gene. Therefore, we took a different approach to knockdown $FOL1$ gene expression through artificial microRNAs (amiRNAs; (Ossowski et al., 2008; Schwab et al., 2006).

Three different amiRNA constructs (C1, C2 and C3) were used that would target three different regions of the $FOL1$ mRNA (see experimental procedure and Figure 4.8a). Two promoters were used to drive expression of the amiRNAs: the constitutive 35S promoter
(Benfey and Chua, 1990; Benfey et al., 1990) and an estradiol inducible promoter (Curtis and Grossniklaus, 2003).

The number of transformants recovered carrying amiRNA constructs driven by the 35S constitutive promoter was unusually low in number, especially with the C1 and C2 constructs. With C1, we were able to recover 13 independent transformants and with C2 only 3 transformants, compared to C3 and inducible constructs for which we were able to recover 32 or more transformants from comparable amounts of seeds plated on selection. Plants used to transform were grown in controlled conditions in growth chambers at the same time and transformation replicates were performed with C1 and C2, eliminating the possibility of defects in the parent plants as an explanation for the low recovery. It is possible this transformation efficiency is due to the lethality issue previously observed with $fol1$-1 and $fol1$-2. To date, only 35S lines have been analyzed.

The inflorescences of second generation ($T_2$) of transgenic plants carrying 35S::amiRNA constructs were used to analyze the expression levels of $FOL1$ in these lines. As shown in Figure 4.8, four lines had significant knockdown of expression from C1 (amiRC1b-d) and C2 (amiRC2c) constructs compared to wild type. These knockdown lines had altered development compared to wild type. The phenotypes we observed could be largely categorized into inflorescence meristem development defects, floral meristem identity defects and flower development defects.
We first examined inflorescence development, since *FOL1* is strongly expressed in the inflorescence meristem. All amiRNA knockdown plants observed had extra branches subtended by cauline leaves before producing flowers. Typically, a wild type plant will have 2-3 branches before lateral meristems become floral. *FOL1* knock down plants had as many as 5-7 branches before making flowers. In addition, many of these plants had branches coming from the stem where there were no visible cauline leaves (Figure 4.9b-e), which is never observed in wild type Arabidopsis. Branches arise from axillary meristems formed in association with leaves and not independently. *FOL1* knockdown plant inflorescences also occasionally formed cauline leaves with no visible branch (Figure 4.9c). These defects suggest a defect in lateral meristem identity and development. Some FOL1 knockdown plants also appeared to have floral meristem identity problems, since several times the inflorescence produced a leaf with a branch, then a flower and then another leaf with a branch, suggesting that the lateral meristems produced from the flank had mixed identity (Figure 4.9e). Taken together, these defects suggest that *FOL1* functions in floral meristem identity.

In addition to these defects in lateral meristem identity, knockdown of *FOL1* also caused changes in the inflorescence meristem itself. Some of the *35S::amiRNA* plants had fasciated stems (Figure 4.10a, b), often thought to be caused by changes in the size of the meristem. Those plants also had a short inflorescence stem (Figure 4.10a, b). Some of these plants had a fused stem that appeared as 2-3 stems forming a helix that spiraled around each other (Figure 4.9c). There were phyllotaxy problems observed for all *FOL1*
knockdown plants (Figure 4.9 and Figure 4.10a, b). On a wild type Arabidopsis inflorescence, flowers arise in a regular spiral pattern with approximately equal internode lengths between them upon maturity (Figure 4.9a). However, on FOL1 knockdown plants, flowers were not placed in such a manner, resulting in an irregular pattern of flowers/siliques with reduced internode elongation that was more variable in length between nodes (Figure 4.9b-e and 4.10a, b). This was also seen for the branches of many of the 35S::amiRNA lines (Figure 4.9b, c). Sometimes, more than one lateral meristem appeared to form at a single node, resulting in as many as 7 branches growing from a single node (Figure 4.9b, d). Similar defects have been observed in Arabidopsis mutants with enlarged inflorescence meristems, suggesting FOL1 is necessary for regulation of the size of meristems.

In addition to floral meristem identity defects, FOL1 knock down plants had abnormal flowers. Some flowers had extra petals and sepals (Figure 4.10c). Similar to some of the inflorescence defects described above, extra floral organs can arise from a larger than normal floral meristem. In addition, the sepals of 35S::amiRNA plants had many more trichomes present than seen on wild type sepals (Figure 4.10c). This can be an indication that the sepals have a partial leaf-like identity.
4.2.5 **Ectopic expression of FOL1 suggests a meristem function.**

*FOL1* overexpression/ectopic expression lines (*FOL1OX*) were generated as another approach to characterize this novel gene (*35S::FOL1*). We transformed both Colombia-0 (Col-0) and *L. erecta* (*L.er*) ecotypes.

The most striking phenotype observed in the *L. er* background was an arrested SAM (Figure 4.11). These plants only developed a couple of abnormal leaves before the SAM produced pin-like organs and then arrested. Based on severity, we grouped the lines into classes. Class I is the most severe, with an arrested SAM that lead to lethality without any further development (Figure 4.11a, b). The SAM of these plants appears to differentiate. Fifty percent of the transformants from independent transformation events belonged to Class I (4/8 independent transformants). Class II plants had an arrested SAM but eventually axillary buds grew out and the plants developed relatively normally (Figure 4.11c-e). Class III plants had no observed abnormalities. Quantitative RT-PCR was not done to relate the expression levels to the severity of the phenotype. These phenotypes suggest that *FOL1*, when ectopically expressed in the SAM, can compromise the function of this meristem.

*35S::FOL1* expressed in Col-0 showed two phenotypes we did not observe in the *L. er* background. Some of the transgenic lines displayed short stature (2/6 lines from independent transformation events). The dwarf plants observed did not grow pass 36mm in height (Figure 4.12a, b). The extremely small stature suggests defects in cell
proliferation in the plant. In addition, cell elongation may be less than wild type, although this was not examined. This meristematic deficiency is also shown in the flowers of these plants, which were infertile and missing floral organs, suggesting a smaller meristem. Several other independent 35S::FOL1 lines appeared normal during vegetative development. However, these lines had floral defects including incorrect numbers of organs as well as organ identity defects. As shown in Figure 4.12c and e, some branches ended in a terminal flower or even several incomplete flowers. These flowers had fewer organs than wild type and showed homeotic transformation of the outer whorl sepals into carpelloid structures as well as more than one carpel (Figure 4.12c-e).

Ectopic overexpression of FOL1 caused different phenotypes in the two different backgrounds. This suggests that unique modifiers of FOL1 function may exist in the two backgrounds. It is interesting to note, however, that the phenotypes observed in both cases suggests that FOL1 functions in meristem size and identity.

4.2.6 FOL1 interacts with LFY
FOL1 was isolated from the biochemical analysis we performed (discussed in Chapter 3) to identify interacting partners of LFY. This interaction was not observed with Y2H assay, possibly due to an indirect interaction of FOL1 with LFY or due to the requirement of LFY and/or FOL1 homodimerization for the interaction, since we know that LFY homodimerizes (Chapter 2) and that FOL1 dimerizes in yeast and in BiFC analysis done in N. benthamiana leaves (Figure 4.13). To confirm a physical interaction between FOL1
and LFY *in planta*, we used BiFC and co-IP assays from transiently transformed *N. benthamiana* leaves. Both of these assays demonstrated that these two proteins interact *in vivo* (Figures 4.14 and 4.15). In the BiFC assay, YFP fluorescence could be seen in the nucleus in distinct speckles/subnuclear foci when the FOL1 and LFY fusion proteins were co-expressed (Figure 4.14a). This is an interesting observation as LFY-GFP and FOL1-GFP alone (Figure 4.14c) do not localize in to speckles as distinct as those formed when the two proteins are together. Although the identity and significance of these speckles/foci remains to be determined, the specificity of the interaction could indicate an important process FOL1 and LFY perform together. LFY-GFP alone is diffuse in the nucleoplasm and also forms cytoplasmic puncta (Figure 4.14c), which are likely to represent plasmadesmata, as LFY as been shown to traffic through these structures (Wu et al., 2003). FOL1-GFP is found only in the nucleus where there is some speckling, but not as dramatic as when it is co-expressed with LFY (Figure 4.14c). To confirm that the relocalization of LFY within the nucleus is caused by an association with FOL1 and not an artifact of the BiFC system, we co-expressed LFY-GFP with FOL1-MYC. In these cells, LFY is found at the plasmadesmata and in nuclear speckles, consistent with the BiFC results (Figure 4.14d).

To further confirm the interaction between FOL1 and LFY, co-IPs were performed. Proteins extracted from *N. benthamiana* leaves co-transformed with 35S::LFY-MYC and 35S::FOL1-GFP or 35S::CBF3-GFP were used. An anti-GFP antibody was used to immunoprecipitate FOL1-GFP or CBF3-GFP (Figure 4.15a). LFY-MYC was able to co-
IP with FOL1-GFP but not with CBF3-GFP. This further confirms the interaction of FOL1 with LFY. To delineate the region of LFY that interacts with FOL1, LFY deletion variants were used in co-IP studies. $35S::LFY_{\Delta 229-420}$-MYC or $35S::LFY_{\Delta 1-229}$-MYC were co-infiltrated with $35S::FOL1$-GFP. As shown in figure 4.15b, only LFY$_{\Delta 229-420}$-MYC was pulled down with FOL1-GFP (Figure 4.15b). This suggests that LFY interaction with FOL1 is through the N-terminal half of LFY, which includes the conserved LFY dimerization domain.

4.2.7 FOL1 interacts genetically with LFY

We demonstrated the physical interaction of LFY and FOL1 in vivo using BiFC and co-IP studies. In addition, knockdown of FOL1 causes floral meristem identity defects, similar to some aspects of the lfy phenotype. We therefore took a genetic approach to determine if the physical interactions between FOL1 and LFY are important for their functions in planta. For this, we crossed plants with altered levels of LFY and FOL1. Two LFY alleles were used, the null allele lfy-6 and a weak hypomorphic allele lfy-5. These were crossed with $35S::FOL1$-GFP lines. Crosses done with lfy-6 are yet to be analyzed. In plants heterozygous for both lfy-5 and $35S::FOL1$-GFP we observed a number of phenotypes not seen in either lfy-5 mutant plants or $35S::FOL$-GFP plants. One such phenotype was the solitary flower formation in the axils of cauline leaves instead of the expected branch (Figure 4.16a, c, d). This indicates that a determinate floral meristem developed where an indeterminate inflorescence meristem should be. A similar phenotype is seen in $35S::LFY$ plants (Weigel and Nilsson, 1995). Observation of a
phenotype characteristic of higher *LFY* expression in a reduced *LFY* background was not an expected result. In addition to the axillary meristem defect, we also observed suppression of internode elongation in the inflorescence compared with wild type and *lfy*-5 (Figure 4.16a), which resulted in short, stunted plants. The transheterozygotes we observed suggests that the two genes may act in a common genetic pathway, and this interaction is probably dependent on the ratio of FOL1/LFY.
4.3 DISCUSSION

FRIEND OF LFY1 (FOL1) is a novel C2H2-type zinc finger protein, which we identified through a biochemical assay done to discover components of LFY transcriptional complexes. This gene is not represented on the Affymetrix ATH1 chip, and there is no other experimental data available. In this study we describe the characterization of FOL1, a new player in flower development, in addition to understanding the role it performs together with LFY in a transcriptional complex in planta.

4.3.1 FOL1 marks the sites of cell proliferation

In order to understand the function of a gene, it is important to understand its expression pattern. Expression pattern information for this gene had to be generated de novo. Results using FOL1p::GUS transgenic lines indicate that FOL1 is expressed in the developing embryo within the young seed (Figure 4.5d). Expression continued until 3 days after germination and was not observed during vegetative development (Figure 4.6). FOL1 is expressed again at day 15 in the SAM, which has now transitioned to inflorescence identity, and also in incipient floral meristems (Figure 4.5a), suggesting that FOL1 is either involved in transition from VM to IM identity or activated shortly after IM initiation. FOL1 is also expressed in developing axillary buds (which make secondary inflorescence branches), supporting the interpretation that it is involved in IM fate or
activated just after the IM is formed (Figure 4.6l,m). In the flower, expression was observed in developing petals, male and female gametophytes in young flower buds (Figure 4.5) and in siliques (Figure 4.4d) and the expression diminishes as the flower matures. By and large, our RT-PCR and \textit{FOL1p::GUS} expression studies show that \textit{FOL1} expression is prominent in sites of cell proliferation. This expression pattern is consistent with functional studies we have done, as \textit{FOL1} has been involved in meristem development and gametophyte development.

\textit{FOL1} was isolated as an interacting partner of LFY. The expression of \textit{LFY} is found at low levels in leaves (see Chapter 1) and then confined to early flower meristems where it is ubiquitously expressed (Weigel et al., 1992). We have shown \textit{FOL1} expression in floral meristems (Figure 4.4b and 4.5a). In other words the expression pattern of \textit{FOL1} overlaps with that of \textit{LFY} consistent with the interaction observed. The broader expression pattern of \textit{FOL1} suggests it has LFY-independent functions in addition to any roles with LFY.

\textbf{4.3.2 \textit{FOL1} is necessary for gametophyte development}

In order to examine gene function, we obtained two putative loss of function \textit{FOL1} alleles from publically available T-DNA insertion collections. However, the study of \textit{fol1-1} and \textit{fol1-2} was difficult due to a low germination rate of homozygous \textit{fol1-1} seeds and due to possible gametophytic or embryonic lethality of \textit{fol1-2}. Alexander staining of anthers of \textit{fol1-2/+} plants revealed a high number of inviable pollen compared to wild
type (Figure 4.7d). fol1-2/+ siliques also contained a high percentage of aborted ovules or unfertilized ovules (Figure 4.7c). In fol1-2, homozygous mutant plants were never recovered. These observations suggest that FOL1 is necessary for gametophytic and/or embryo development. This correlates with the FOL1 expression pattern we observed in gametophytes in anthers and carpels and in young developing siliques (Figure 4.5).

Due to the difficulty we encountered working with the putative insertional alleles of FOL1, we generated amiRNA lines to knock down FOL1 expression to examine function of this gene. Preliminary analysis of these plants and 35S::FOL1 ectopic expression plants suggests that FOL1 functions in controlling meristem proliferation.

4.3.3 FOL1 controls inflorescence meristem size

FOL1 amiRNA lines had a number of developmental defects. FOL1 is expressed in inflorescence meristems (Figures 4.4b and Figure 4.5a), suggesting it may function there. A prominent defect observed in FOL1 knockdown plants was stem fasciation, where the stem resembled 2-3 stems fused and spiraled around each other (Figure 4.9 and 4.10). A similar phenotype has been described in mutants, such as clv1, where the inflorescence meristem is enlarged (Leyser and Furner, 1992). Knockdown plants also displayed irregular phyllotaxy along the bolt and reduced internode elongation, also associated with increases in meristem size (Leyser and Furner, 1992). The axillary IMs formed in association with cauline leaves sometimes bifurcated, producing two branches, which can happen when the meristem size is large. Taken together, this suggests that FOL1
functions to restrict the size of IMs. Since LFY is not expressed in IMs, this would be a LFY-independent function of FOL1.

FOL1 involvement in meristem size control is further supported by the phenotype of FOL1 overexpression transgenic plants. These 35S::FOL1 lines drive higher expression of the gene and also ectopic expression, for example in vegetative meristems. The most striking phenotype observed in 35S::FOL1 lines in the L.er background was an arrested SAM (Figure 4.11). 50% of the independent transgenic lines isolated produced a couple of leaves or pin-like structures and arrested. This suggests that ectopic FOL1 in VMs caused them to stop proliferating and producing lateral organs. In the Col-0 background we observed different abnormalities in 35S::FOL1 lines, but they are still in line with FOL1 being involved in regulation of meristem proliferation. Two of the independent lines were severely dwarfed (Figure 4.12a,b), suggesting cell proliferation and/or elongation defects. These lines were sterile and could not be propagated. The other lines were normal in height but had inflorescence and floral defects. Some of the inflorescence branches ended with a terminal flower or with several incomplete flowers, suggesting that the IM had transitioned to a floral meristem. Sepals of these terminal flowers were transformed into carpelloid structures. Often these terminal flowers did not maintain this floral fate, but rather made floral organs, then sent out a branch, and then continued making floral organs (Figure 4.12c-e), suggesting that the IM was not completely transformed to a flower.
4.3.4 FOL1 controls floral meristem identity and size

The phenotypes of 35S::FOL1 lines in the Col-0 background suggested that FOL1 may function in floral meristem identity. This is also supported by loss of function studies. An increase in the number of secondary branches subtended by cauline leaves before producing flowers has been previously observed in lfy mutants (Weigel et al., 1992), ufo mutants (Levin and Meyerowitz, 1995) and in bop1; bop2 mutants (Xu et al., 2010), all of which are known to be involved in floral meristem identity. We also observed this phenotype in amiRNA plants, suggesting that FOL1 may function in floral meristem identity. In addition, many FOL1 knock down plants made branches that were not subtended by a cauline leaf (Figure 4.9 b-e) or a cauline leaf without an axillary branch (Figure 4.9c). The absence of a cauline leaf from the base of shoots preceding the first flower has also been observed in mutants with floral meristem identity defects (Xu et al., 2010). Further indication that FOL1 functions in floral meristem identity is the fact that in some knockdown lines we observed a branch then a flower and then another branch (Figure 4.9e), suggesting that the switch to making flowers was not complete. Together these phenotypes indicate that FOL1 function is necessary for floral meristem identity. Since LFY also functions in floral meristem identity, this maybe an LFY-dependent function of FOL1. FOL1 amiRNA lines also have floral defects. These flowers contain extra sepals and petals (Figure 4.10c). This phenotype has been associated with mutations in genes that regulate floral meristem size (Fletcher et al., 1999; Leyser and Furner, 1992) and suggests that FOL1 knock down plants have larger flower meristems and that FOL1 functions in floral meristem size control.
How *FOL1* functions in meristem size control is still not clear. We can speculate that it may be involved in regulating *WUS*, a homeodomain transcription factor important for meristem functions. Strong *wus* mutants die at the seedling stage without a SAM (Laux et al., 1996). In the SAM and IM, WUS functions to maintain the stem cells in an undifferentiated state and control of its expression is essential to regulate meristem size (Mayer et al., 1998). Too little *WUS* will lead to small meristems while too much leads to large meristems. Expression of WUS is controlled by a feedback loop with the CLV pathway. WUS activates expression of the *CLV* genes (Brand et al., 2000; Fletcher et al., 1999; Schoof et al., 2000). CLV1 and CLV2 form a receptor kinase complex at the plasma membrane, which recognizes the CLV3 peptide ligand. The CLV3 signal transduction pathway represses *WUS* expression in a feedback loop, regulating meristem size (reviewed in (Carles and Fletcher, 2003)). We can speculate that FOL1 maybe inhibiting the function of WUS or is responsive to signals from CLV in order to regulate size. Further work is needed to resolve this.

In the flower, *WUS* expression needs to be turned off at stage 5/6 in order for determinate growth (Lenhard et al., 2001; Lohmann et al., 2001). During floral development, WUS and LFY together activate *AG* expression by binding to adjacent sites on the *AG* regulatory region (Busch et al., 1999; Lenhard et al., 2001; Lohmann et al., 2001). AG is required for stamen and carpel development and determinate growth of the floral meristem (Bowman et al., 1991; Bowman et al., 1989; Yanofsky et al., 1990). In a negative feedback loop, AG directly represses *WUS* expression, causing the floral
meristem to terminate (Lenhard et al., 2001; Lohmann et al., 2001). Therefore, in the flower WUS activates its own negative regulator. This suggests a possible role for FOL1 with LFY in the floral meristem in regulating AG and/or WUS expression. However, this remains to be explored.

4.3.5 FOL1 is a LFY interacting partner

Biochemical isolation of LFY containing complexes identified FOL1 as a putative component of such complexes. We confirmed this physical interaction between FOL1 and LFY using BiFC (Figure 4.14) and co-IP (Figure 4.15) in planta. In the BiFC assay, the relocalization of LFY from a general nuclear localization into distinct loci in the nucleus, perhaps marking sites of transcription, when associating with FOL1 was an intriguing observation. Interestingly, FOL1 interacts with the N-terminal region of LFY (Figure 4.15), in the area necessary for dimerization (see Chapter 2). To our knowledge, FOL1 is the first protein identified to interact with the N-terminus of LFY.

The physical interaction between FOL1 and LFY, the overlap in their expression patterns and the commonalities between FOL1 knockdown phenotypes and lfy phenotypes suggests that these proteins may regulate transcription together during floral development. This interaction is also supported by a genetic interaction between the two genes. 35S::FOL1/+; lfy-5/+ plants displayed meristem identity issues in which an axillary IM was occasionally replaced with a floral meristem (Figure 4.16). Shoot conversion to a solitary flower is observed in LFY overexpression lines (Figures 1.8 and
2.4). It is unclear how the meristems of 35S::FOL1/+; lfy-5/+ plants are acquiring this unexpected identity change in a reduced LFY background. The phenotypes we observed are not seen either in lfy-5, lfy-5/+ or in FOL1 overexpression lines. These transheterozygotes suggests the two genes may act in a common genetic pathway, and their interaction probably depends on the ratio of FOL1/LFY.

4.3.6 Conclusions

Our biochemical isolation work identified a novel C2H2-type zinc finger protein as an interacting partner of LFY, FRIEND OF LFY1 (FOL1), whose transcript is not represented on the ATH1 microarray. fol1 mutations appear to be either gametophytic or embryonic lethal. Therefore, these genes would be not identified by expression data mining approaches or by traditional genetic screens for factors acting in floral meristem identity. Our analyses demonstrate that FOL1 has roles in gametophytic and embryonic development, in inflorescence and flower meristem size control and floral meristem identity in Arabidopsis thaliana. Its role in floral meristem identity is likely to be at least partially LFY-dependent, as supported by the genetic interaction between the two genes. However, FOL1 also functions independently of LFY, particularly in the gametophyte, embryo and inflorescence meristem.
4.4 EXPERIMENTAL PROCEDURES

4.4.1 Plant material and growth conditions

The *fol1-1* (*SALK_040806C*) and *fol1-2* (*SALK_041031*) lines used in this study were obtained from the ABRC and are T-DNA insertion lines from the SALK collection (Alonso et al., 2003). *lfy-6/+; 35S::FOL1-GFP* and *lfy-5/+; 35S::FOL1-GFP* lines were created by crossing mutant lines with homozygous *35S::FOL1-GFP* lines. Arabidopsis plants used in this study were in the L.<i>er</i> and Col-0 ecotypes. The *lfy-6* and *lfy-5* alleles were previously described (Weigel et al., 1992).

Plant growth conditions are as described in Chapter 2.

4.4.2 Constructs and cloning

The *FOL1* coding region was cloned into *pENTR/D-TOPO* Gateway entry vector (Invitrogen) as described in Chapter 2. The primers used are listed in Table 4.1. The PCR products introduced to *pENTR/D-TOPO* vector were sequenced at the PMGF. *FOL1* in *pENTR/D-TOPO* vector was then recombined into binary destination vectors *pGWB5* (introducing a C-terminal GFP tag) and into *pGWB20* (introducing a C-terminal MYC tag) (binary destination vectors were a gift from T. Nakagawa, Shimane University,
Matsue, Japan). Constructs containing LFY coding sequence and the deletions were the same used in Chapter 2.

The promoter region of FOL1, a 2615 bp region 5’ of the transcription start site, was amplified from genomic DNA using Phusion HF DNA Polymerase (NEB). The primers used to clone the promoter region are given in Table 4.1. This region was cloned into pDONR221 Gateway entry vector (Invitrogen) using the manufacturer’s protocol. After sequencing at the PMGF, this was recombined in the pGWB3 binary destination vector to drive expression of the GUS gene.

The FOL1 genomic region, including the 5’ and 3’ UTRs, was amplified using genomic DNA as template and the primers listed in Table 4.1, using Phusion HF DNA polymerase as above. The FOL1 promoter and the genomic region were inserted into pDONR221P1-P5r and pDONR221P5-P2, respectively (Invitrogen) and recombined together into the pGWB4 binary destination vector (to introduce a C-terminal GFP tag) using the MultiSite Gateway Pro Kit (Invitrogen) following the manufacturer’s instructions. In order to clone the FOL1 ‘rescue construct’ carrying the promoter and the genomic region, the pGWB1 binary destination vector (no tag) was used instead.

The amiRNA constructs targeted against FOL1 were designed using the WMD3 tool (http://wmd.weigelworld.org). Three sequences with high confident levels and one mismatch were selected. These sequences only targeted FOL1 and not other locus in the
The constructs chosen were: C1-TTAAGGGCATGCGTACGGCCG, C2-TAAGGGCATGCGTATGCCCGA and C3-TTTAAGGATGGGTCGCCTCTA (see Figure 4.7a for the target location). The plasmid pRS300, which contains a MIR319a precursor sequence, was obtained from Dr. Keith Slotkin, The Ohio State University. PCR to obtain the amiRNA fragments were done using Pfu DNA Polymerase (Stratagene). For every construct, four fragments were generated; the primers used are listed in Table 4.1. These fragments were then fused into one precursor using PCR with the primers amiRNA modA and modB as in Table 4.1. These fused constructs were then cloned into the pENTRD/TOPO (Invitrogen) vector according to the manufacturer’s instructions. These were sequenced at the PMGF and recombined using manufacturer’s guidelines into pGWB2 to drive expression under the 35S constitutive promoter or into pMDC7 to drive expression under an estradiol inducible element. All the above cloned plasmids were then introduced into Argobacterium tumefaciens strain GV3101 and was used to transform Arabidopsis wild type plants as described in Chapter 2.

For yeast two-hybrid assays, the FOL1-pENTR/D clone was recombined with pDEST22 and pDEST32 destination vectors with the GAL4 Activation Domain (AD) and DNA Binding Domain (BD), respectively (ProQuest Two-Hybrid System, Invitrogen). These plasmids were transformed into yeast strain PJ694A (James, 2001) following the protocol published in the Clontech Yeast Protocol Handbook.
For BiFC, vectors p2YN (YFP amino acids 1-158), p2YC (YFP amino acids 159-238) and pYN-1 and pYC-1 carrying either fragment of split YFP in C- and N- termini respectively were used (vectors were from Dr. John A. Lindbo, The Ohio State University, Wooster, OH). The procedure described in Chapter 2 for cloning LFY into these vectors was used to clone FOL1 using the primers provided in Table 4.1. These fusion constructs were transformed into Agrobacterium for N. benthamiana infiltrations. Generation of Arabidopsis transgenic lines and transient infiltration of N. benthamiana leaves were done as described in Chapter 2.

4.4.3 Yeast two-hybrid interaction assay

Yeast competent cell preparation, transformations and interaction studies were performed as described in Chapter 2. pDEST22-FOL1 was co-transformed with pDEST32-FOL1. As negative controls, empty vectors were co-transformed with above vectors carrying FOL1.

4.4.4 Bimolecular Fluorescence Complementation Assay (BiFC)

BiFC was performed in transiently transformed N. benthamiana leaves and observations were done as described in Chapter 2. Agrobacterium with BiFC vectors carrying full-length LFY and FOL1 along with split YFP were co-infiltrated. Empty vectors with one half of the YFP together with the vector carrying FOL1 gene fused to the other half of the YFP was used as negative controls as well as the two empty vectors.
4.4.5 Genomic DNA preparation and genotyping
Genomic DNA was extracted from leaves as described in (Teotia and Lamb, 2009). This was used as template in PCR reactions, to verify the presence of the T-DNA insertion in plants from the SALK collection. LBb1 and RP primers were used to genotype the insertion and LP and RP was used to amplify the wild type locus (for primers refer Table 4.1). The sequence for above primers was obtained from SALK website (Alonso et al., 2003). PCR was done using Choice-Taq Blue DNA Polymerase (Denville Scientific Inc) using a BioRAD iCycler version 4.006.

4.4.6 RNA isolation and semi quantitative Real Time-PCR (RT-PCR)
Total RNA extraction and cDNA synthesis from inflorescence apices of amiRNA transgenic plants were done using the protocol described in Chapter 2. RT-PCR on amiRNA lines was performed using primers listed in Table 4.1. The ACTIN2 gene was used as the reference gene in RT-PCRs performed. PCR was done using Choice-Taq Blue DNA Polymerase using the BioRAD iCycler version 4.006.

4.4.7 Co-immunoprecipitation assay
Co-IP assays with infiltrated leaves of N. benthamiana were done using the protocol described in detail in Chapter 2.
4.4.8 GUS staining

Rosette leaves, stems, cauline leaves, whole inflorescences and siliques were collected from initial (T₀) FOL1::GUS transgenic lines. In the next generation (T₁), every 2 days starting on day 3 after the break from vernalization, the seedlings or the plants were collected and stained. β-glucuronidase (GUS) enzymatic activity was detected as described in (Hill et al., 1998). GUS stained inflorescences were observed under a Nikon SMZ745T (Nikon) dissecting microscope and photographed using an OmniVID Camera (LW Scientific).

4.4.9 Phenotypic analysis of transgenic lines

amiRNA knockdown lines were scored for plant height, phyllotaxy defects, number of cauline leaves before the first flower, number of petals on flowers, presence of cauline leaves with no branches and vice versa, and number of axillary meristems.

Overexpression lines in the Col-0 ecotype were screened in the first generation. Floral development defects, plant height, and occurrence of terminal flowers were scored. Overexpression lines in the L. er ecotype were scored in first generation for growth arrest. Homozygous lines in this ecotype background were also analyzed for phenotypes in the next generation.

For the loss of function studies, phenotypic analysis was done on homozygous fol1-1 (SALK_040806C) and heterozygous fol1-2 (SALK_041031) plants. Plants were observed
at all stages, looking mostly for floral development or identity defects, branching defects, plant height, aborted ovules and pollen viability. All seed germination rate experiments were conducted on MS media by counting the number of seeds plated, and then how many germinated.

*lfy5/+; 3S::FOL1-GFP/+* plants were scored for solitary flower formation instead of a branch at the axils of the cauline leaf as well as the height and internode length.

### 4.4.10 Alexander Staining

Pollen viability of wild type and SALK lines were examined using the method described in (Alexander, 1969).

### 4.4.11 Computational Analysis

The sequence alignment was done as described in Chapter 2. The multiple alignment was subjected to a maximum-likelihood analysis as described in (Citarelli et al., 2010) to create the phylogenetic tree.
Figure 4.1. Structural features of FOL1 protein and gene. (a) Amino acid sequence of the longest isoform of FOL1. The five zinc finger motifs are in red (residues 79-103, 108-132, 138-162, 168-193 and 230-255), the putative SFP1 domain is underlined, and the glutamic acid rich region is in blue (residue 341-386). The circled residue is the predicted phosphoserine residue at 284. (b) Schematic diagram of FOL1 protein showing five zinc fingers (red boxes), four tandemly spaced and one isolated and glutamic acid rich region (blue box). (c) Splice variants of FOL1. At4g06634.1 is considered the primary isoform. The filled boxes denote exons, the lines denote introns and the unfilled boxes are the untranslated regions (UTRs). Arrow above points to 5’ to 3’ direction.
Figure 4.2. Phylogenetic tree of FOL1 and its orthologs from selected land plants. Graphical representation of the maximum-likelihood phylogenetic tree of FOL1 orthologs. This tree is based on multiple alignments shown in Figure 4.3. Branch supports values are indicated at the nodes as computed in PhyML using an aLRT non-parametric Shimodaira-Hasegawa-like procedure and a midpoint rooting method.
Figure 4.3. Alignment of FOL1 and its orthologs from selected land plants. The zinc finger (ZnF) motifs are marked with a red bar and the glutamic acid rich region is marked with a blue bar. The putative SPF1 domain marked with a black bar. At, *Arabidopsis thaliana*; Al, *Arabidopsis lyrata*; Pt, *Populus trichocarpa*; Rc, *Ricinus communis*; Gm, *Glycine max*; Mt, *Medicago truncatula*; Vv, *Vitis vinifera*; Br, *Brassica rapa*; Nt, *Nicotiana tabacum*; Osj, *Oryza sativa Japonica*; Hv, *Hordeum vulgare*; Zm, *Zea mays*; Sb, *Sorghum bicolor*; Sm, *Selaginella moellendorffii*; Pp, *Physcomitrella patens*. 
Figure 4.4. Expression pattern of FOL1. (a) RT-PCR showing expression of FOL1 only in whole inflorescence and siliques. ACTIN was used as the reference gene. IN-inflorescence RL-rosette leaves SE-seedling SI-silique RT-root. (b-k) GUS expression; (b) Whole inflorescence showing expression in developing areas where cell division is still occurring. (d) Developing siliques showing expression in ovules at the ends of the siliques. (b, d, f, h, j) FOL1::GUS lines. (c, e, g, i, k) wild type. (b, c-inflorescence; d, e-silique; f, g-cauline leaves; h, i-stem; j, k-rosette leaves).
Figure 4.5. *FOL1p::GUS* expression in developing flowers and embryos. (a) Inflorescence with early (red arrowheads) and later stage flowers showing expression in petals (inset-asterisk), anthers (black arrow) and developing carpel (black arrowhead). (b) At later stages, *FOL1* expression is confined to developing gametophytes. Insets: male gametophyte/pollen (left) and female gametophyte (right). (c) Expression within the ovule continues into embryo development (red arrow). (d) Developing embryo within an immature seed coat with *FOL1p::GUS* expression.
Figure 4.6. *FOL1* expression is confined to very young seedlings and reproductive tissues. (a) *FOL1p::GUS* expression is seen at the root tip, hypocotyl and cotyledon of a 3-day-old seedling. (b) 5-day-old seedling. (c) 7-day-old seedling. (d) 9-day-old seedling. (e) 11-day-old plant. (f) 13-day-old plant. (g) By 15 days, *FOL1* expression is seen in the apical meristem (circle). (h) 17-day-old plant *FOL1* expression seen in inflorescence and floral meristems. (i) 19-day-old plant. (j) 21-day-old plant with expression in developing anthers and carpels. (k) 23-day-old plant. (l) 25-day-old plant with expression in co-inflorescence meristems. (m) 27-day-old plant *FOL1* expression seen at the axillary buds (circle).
Figure 4.7. *fol1-1* and *fol1-2* have gametophyte defects. (a) Diagrammatic representation of the *FOL1* gene, inverted triangle indicating the insertion site for *fol1-1* and *fol1-2*. The filled boxes represent exons and the lines introns. Primer pairs used to amplify transcripts are marked below the diagram. (d) RT-PCR showing only the transcript before the insertion is made. (c) *fol1-1* siliques (right) have a large number of aborted ovules (arrowheads) compared to wild type (left). (d) Alexander staining of genotypes as indicated. Insets show viable pollen (stained pink) and inviable pollen (stained blue). Asterisks mark inviable pollen in *fol1-1*.
Figure 4.8. *FOL1* amiRNA lines have reduced transcript accumulation. (a) Schematic representation depicting the regions amiRNA constructs are targeted to. Primers used in RT-PCR are shown below the diagram. (b) RT-PCR of amiRNA transgenic lines (C1 and C2) and wild type. Asterisks marks the bands of knockdown lines selected for further analysis. *ACTIN* was used as the reference gene.
Figure 4.9. *FOL1* knockdown causes meristem defects. (a) Wild type plant. (b) *FOL1* knockdown line amiRC1b. (c, d and e) *FOL1* knockdown line amiRC1b, amiRC1c. Phyllotaxy problems can be seen in all plants. A red arrowhead marks a branch without a subtending cauline leaf. Yellow arrowheads mark cauline leaves without a branch. White arrowhead marks the unusual alternation of a branch, a flower and a branch. White arrows marks multiple branches at a single node.
Figure 4.10. *FOL1* knockdown causes meristem size defects. (a and b) Knockdown lines showing stem fasciation and short inflorescence stem. (c) Flower from a knockdown plant (right) next to a wild type flower (left) showing more petals and a plump phenotype. Asterisks marks the petals. White color arrow marks the fasciated stem. Red arrow points to extra trichomes on the sepals.
Figure 4.11. 35S::FOL1 causes an arrested SAM in the Ler background. (a and b) Class I severe phenotype, showing arrested SAM and plants completely arrest growth. (c, d, e) Class II phenotype with arrested SAM but the axillary buds grow out and plant grow successfully. Arrows point to arrested SAMs. Red arrowheads point to the leaf-like abnormal organs. White arrowheads points to axillary buds growing out in class II plants.
Figure 4.12. Ectopic expression of FOL1 causes reduced stature and terminal flowers in the Col-0 background. (a and b) Dwarfed plants. (c, d and e) Terminal flower phenotype in the apices of inflorescence. Arrow points to the dwarf plant in (a). Numbers indicate pistils of the terminal flowers. Red arrowheads points to carpelloid organs/sepal-like structures. White arrowheads point to exposed ovules on these carpelloid organs.
Figure 4.13. FOL1-FOL1 homodimerizes. (a) Y2H interaction assay showing homodimerization on selection plate. Vector combinations used to co-transform yeast cells are given as bait/prey. (b) Micrographs of BiFC analysis done in *N. benthamiana* leaves, showing interaction with C-terminal tagged YFP. The negative controls with the two empty vectors, or one empty with one FOL1 fused protein did not show fluorescence. The scale bars indicate 50 μm.
Figure 4.14. FOL1-LFY interaction in planta. Micrographs of *N. benthamiana* leaves: bright field, epifluorescence and merged images, from left to right.

Continued
Figure 4.14: continued

(a) FOL1 and LFY tagged at either the C-terminus or the N-terminus interact. Fluorescence is seen in distinct speckles in the nucleus (red arrowheads), which is not seen with either protein alone. (b) Negative controls consisting of two empty vectors or one empty vector with one FOL1 fusion protein did not show any fluorescence. (c) FOL1-GFP and LFY-GFP showing nuclear localization. FOL1-GFP shows speckled nature in the nucleus but not distinct as when FOL1 interacts with LFY. LFY-GFP marks the nucleoplasm and also shows the punctate points in the cytoplasm (white arrowheads). (d) LFY-GFP forms speckles in the nucleus when FOL1-MYC is present with LFY-GFP. The scale bars indicate 50μm.
Figure 4.15. FOL1 interacts with N-terminus of LFY. (a and b) Western blots of co-immunoprecipitations done using transiently transformed *N. benthamiana* leaves. (a) FOL1-GFP was able to bring down LFY-MYC (red arrowhead) but CBF3-GFP could not. (b) FOL1-GFP was able to pull down LFY$_{\Delta229-420}$-MYC (red arrowhead) but not LFY$_{\Delta1-229}$-MYC showing FOL1 interaction with N-terminal domain of LFY.
Figure 4.16. FOL1 and LFY interact genetically. (a) Compared to wild type the \textit{lfy-5/+; 35S::FOL1/+} plants have reduced internode elongation. (b) \textit{lfy-5} plant. (c, d) Close ups of solitary flower in axil of cauline leaf (circles).
<table>
<thead>
<tr>
<th>Experiment</th>
<th>Primer name</th>
<th>Primer Sequence (5’-3’)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>FOL1</em> BiFC cloning</td>
<td>ZnFp2YCF</td>
<td>TACGAACGATAGTTAATTAACATGGATCATCAAAATTATCAATAC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>ZnFp2YCR</td>
<td>CACCTCCTCCACTAGTATCTTCATACTCGGTCTCTTCATC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>ZnFp2YNF</td>
<td>TACGAACGATAGTTAATTAACATGGATCATCAAAATTATCAATAC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>ZnFp2YNR</td>
<td>CACCTCCTCCACTAGTATCTTCATACTCGGTCTCTTCATC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>ZnFpYCF-1F</td>
<td>CTGGAGGAGGCGGGCGCATGGATCATCAAAATTATCAATAC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>ZnFpYCF-1R</td>
<td>ATTTGCGGACTCTAGACTAATCTTCATACTCGGTCTCTTC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>ZnFpYN-1F</td>
<td>CTGGAGGAGGCGGGCGCATGGATCATCAAAATTATCAATAC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>ZnFpYN-1R</td>
<td>ATTTGCGGACTCTAGACTAATCTTCATACTCGGTCTCTTC</td>
<td>This Study</td>
</tr>
<tr>
<td><em>FOL1</em> coding region cloning</td>
<td>ZnFGWF</td>
<td>CACCATGGATCATCAAAATTATCAATACC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>ZnFGWR</td>
<td>ATCTTCATACTCGGTCTCTTCATCGTGTC</td>
<td>This Study</td>
</tr>
<tr>
<td><em>FOL1</em> genomic region cloning</td>
<td>ZnFGenomicF</td>
<td>GGGGACAACTTTGTATACAAAAAGTTGCCAACCGGGACTCCATTATCAAAGTAT</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>ZnFGenomicR</td>
<td>GGGGACAACTTTGTACAAAAAGCTGGTAATCTTCATACTCGGTCTTTC</td>
<td>This Study</td>
</tr>
<tr>
<td><em>FOL1</em> promoter region cloning</td>
<td>ZnFPromoterF</td>
<td>GGGGACAAGTTTTGTACAAAAAGCAGGCTCACCAGCTCCAGATGTTGACAG</td>
<td>This Study</td>
</tr>
</tbody>
</table>

Continued
Table 4.1: continued

<table>
<thead>
<tr>
<th>ZnFPromoterR</th>
<th></th>
<th>GGGGACAAACTTTTTGTATACAAAGTTGGGTTAAAGGTGGTGCCTTTTG</th>
<th>This Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOLI amiRNA cloning</td>
<td>amiRNAC1-1</td>
<td>GATTAAGGGCATGCAGCCTTGCTCGCTCTCTTTTGTATCC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAC1-2</td>
<td>GACCGGCCCTACGCGCTGCCCTTTATTCACAGGTGTGTATAG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAC1-3</td>
<td>GACGCCCTACGCGCTGCCCTTTATTCACAGGTGTGTATAG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAC1-4</td>
<td>GAATAAGGGGATGCGTGCTGCTCACTATATATATATCC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAC2-1</td>
<td>GATAGGGCATGCGCTTGCCCGATGCTCTCTTTTGTATCC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAC2-2</td>
<td>GATCGGCCATACGCGCTGCCCTTATTCACAGGTGTGTATAG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAC2-3</td>
<td>GATCGGCCATACGCGCTGCCCTTATTCACAGGTGTGTATAG</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAC2-4</td>
<td>GAATAAGGGGATGCGTGCTGCTCACTATATATATATCC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAC3-1</td>
<td>GATTAAAGGATGGGCTCGCTCTCTCTCTTTTGTATCC</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAC3-2</td>
<td>GATAGGGGACCCATCCTTTAAATCAGGAGATCAATGA</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAC3-3</td>
<td>GATTAAGGGGACCCATCCTTTAAATCAGGAGATCAATGA</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAC3-4</td>
<td>GATTAAGGGGACCCATCCTTTAAATCAGGAGATCAATGA</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAModA</td>
<td>CACCCGATAAGGCTTAGATCAGATACCT</td>
<td>This Study</td>
</tr>
<tr>
<td></td>
<td>amiRNAModB</td>
<td>CGGGCCGCTCTAGAAGCTTAGGGT</td>
<td>This Study</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>amiRNA RT-PCR</th>
<th>ActinF</th>
<th>GGCGATGAAGCTCAATCCAAA</th>
<th>This Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ActinR</td>
<td>GGTCACGACCAGCAAGATCAA</td>
<td>This Study</td>
</tr>
<tr>
<td>RTZnF1502F</td>
<td>AATGGTGGTGAGAGACTCCCCAAATATAC</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
<td></td>
</tr>
<tr>
<td>RTZnF2050R</td>
<td>CTAACTCTTACTCGGTTCTTCTCATCGTCG</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
<td></td>
</tr>
<tr>
<td>SALK line genotyping</td>
<td>LBb1</td>
<td>GCGTGGACCGCTTGCTGCAACT</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>SALK_041031LP</td>
<td>TTGTGTGTGGTATTACCAAGG</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>SALK_041031RP</td>
<td>GAACCGATGCAAAGTTCTGTC</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>SALK_040806CLP</td>
<td>AATGGGTCACAGGTGAATTG</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
<tr>
<td></td>
<td>SALK_040806CRP</td>
<td>GAACCGATGCAAAGTTCTGTC</td>
<td><a href="http://signal.salk.edu/">http://signal.salk.edu/</a></td>
</tr>
</tbody>
</table>
REFERENCES


Teotia, S., S. Muthuswamy, and R.S. Lamb. 2010. RADICAL-INDUCED CELL DEATH1 and SIMILAR TO RCD ONE1 and the Stress-Induced Morphogenetic Response. Plant Signal Behav. 5.


Aukerman, M.J., I. Lee, D. Weigel, and R.M. Amasino. 1999. The Arabidopsis flowering-time gene LUMINDEPENDENS is expressed primarily in regions of


Garner, W.W., and H.A. Allard. 1920. Effect of the relative length of day and night on flowering and fruiting of plants.


Winter, C.M., R.S. Austin, S. Blanvillain-Baufume, M.A. Reback, M. Monniaux, M.F. Wu, Y. Sang, A. Yamaguchi, N. Yamaguchi, J.E. Parker, F. Parcy, S.T. Jensen,


**APPENDIX A**: A complete list of proteins identified from TAP tag-purification

**Table A.1.** Proteins are listed based on the score; top most is the highest score. Highlighted in red are the proteins selected to investigate. Highlighted in green are proteins likely to have bound to calmodulin binding peptide moiety of the TAP tag. Highlighted in yellow are some of the commonly found cellular proteins found in TAP-tagged preparations, which are taken as common contaminants, unless supported by independent data.

<table>
<thead>
<tr>
<th>Band</th>
<th>NCBI ID and protein description</th>
<th>Abb. Name</th>
<th>AGI ID</th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A B C D E</td>
<td>gi</td>
<td>20563223 leafy [Arabidopsis thaliana]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>gi</td>
<td>397482 heat shock protein 70 cognate [Arabidopsis thaliana]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>gi</td>
<td>3962377 heat shock protein 70 [Arabidopsis thaliana]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>gi</td>
<td>15241847 ATP binding [Arabidopsis thaliana]</td>
<td>HSC70-2</td>
<td>AT5G02490</td>
</tr>
<tr>
<td>D</td>
<td>gi</td>
<td>1944432 ribulosebisphosphate carboxylase [Arabidopsis thaliana]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>gi</td>
<td>18402186 ABF4 (ABRE BINDING FACTOR 4); DNA binding / transcription factor/ transcriptional activator [Arabidopsis thaliana]</td>
<td>ABF2</td>
<td>AT3G19290</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>25365080 hypothetical protein A/Ag25180 - Arabidopsis</td>
<td>ARR12</td>
<td>At2g25180</td>
</tr>
<tr>
<td>E</td>
<td>gi</td>
<td>15229784 CAM7 (CALMODULIN 7); calcium ion binding [Arabidopsis thaliana]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>6714410 unknown protein [Arabidopsis thaliana]</td>
<td>DCL2</td>
<td>AT3G03300</td>
</tr>
<tr>
<td>E</td>
<td>gi</td>
<td>2866274 Hypothetical protein [Arabidopsis thaliana]</td>
<td>LHR5</td>
<td>AT1G09250</td>
</tr>
<tr>
<td>B</td>
<td>gi</td>
<td>19759469 unnamed protein product [Arabidopsis thaliana]</td>
<td>MDH9.2</td>
<td>AT5G42330</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>9502367 T12C24.2 [Arabidopsis thaliana]</td>
<td>FSO11.2</td>
<td>AT5G42740</td>
</tr>
<tr>
<td>B</td>
<td>gi</td>
<td>18400540 unknown protein [Arabidopsis thaliana]</td>
<td>At2g44240</td>
<td>105</td>
</tr>
<tr>
<td>C</td>
<td>gi</td>
<td>15226197 ATP binding / kinase/ protein serine/threonine kinase [Arabidopsis thaliana]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>15238990 XRN2 (EXORIBONUCLEASE 2); 5'-3' exonuclease/ exonuclease/ nuclear acid binding</td>
<td>XRN2</td>
<td>AT5G42540</td>
</tr>
<tr>
<td>B</td>
<td>gi</td>
<td>3297824 bZIP transcription factor-like protein [Arabidopsis thaliana]</td>
<td>ABF3</td>
<td>At4g3400</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>2577957 AB13 protein [Arabidopsis thaliana]</td>
<td>RAV1</td>
<td>AT5G3260</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>15231799 transcription factor [Arabidopsis thaliana]</td>
<td>MAG2.16</td>
<td>AT5G14180</td>
</tr>
<tr>
<td>H</td>
<td>gi</td>
<td>30680251 nucleic acid binding / zinc ion binding [Arabidopsis thaliana]</td>
<td>4d806634</td>
<td>99</td>
</tr>
<tr>
<td>N</td>
<td>gi</td>
<td>79586698 REF6 (RELATIVE OF EARLY FLOWERING 6); nucleic acid binding / transcription factor/ zinc ion binding</td>
<td>REF6</td>
<td>AT5G48430</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>62320602 hypothetical protein [Arabidopsis thaliana]</td>
<td>T1G16.80</td>
<td>AT5G27750</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>3036812 ATM-like protein [Arabidopsis thaliana]</td>
<td>NA</td>
<td>AT4G36080</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>L</th>
<th>gi</th>
<th>9759469 unnamed protein product [Arabidopsis thaliana]</th>
<th>MDH9.2</th>
<th>AT5G42330</th>
<th>94</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>gi</td>
<td>15238380 ATP binding / nucleoside-triphosphatase/ nucleotide binding [Arabidopsis thaliana]</td>
<td>T15G18_1</td>
<td>AT4G09360</td>
<td>93</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>15239811 unknown protein [Arabidopsis thaliana]</td>
<td>MNL12_5</td>
<td>AT5G42230</td>
<td>87</td>
</tr>
<tr>
<td>H</td>
<td>gi</td>
<td>15239835 unknown protein [Arabidopsis thaliana]</td>
<td>At5g12010</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>gi</td>
<td>15230484 unknown protein [Arabidopsis thaliana]</td>
<td>At3g44690</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>gi</td>
<td>18396508 hydrolase/ protein serine/threonine phosphatase [Arabidopsis thaliana]</td>
<td>82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>9989051 Similar to DNAJ proteins [Arabidopsis thaliana]</td>
<td>F19K19.3</td>
<td>T15G18_1</td>
<td>82</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>4678213 hypothetical protein [Arabidopsis thaliana]</td>
<td>AT2G39440</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>gi</td>
<td>15239835 unknown protein [Arabidopsis thaliana]</td>
<td>At5g12010</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>gi</td>
<td>12322632 unknown protein, 5' partial [Arabidopsis thaliana]</td>
<td>At1g42470</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>25411423 probable retroelement pol polyprotein [imported] - Arabidopsis thaliana</td>
<td>F23M2_11</td>
<td>At2g11950</td>
<td>76</td>
</tr>
<tr>
<td>H</td>
<td>gi</td>
<td>15219364 unknown protein [Arabidopsis thaliana]</td>
<td>Atg16800</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>gi</td>
<td>10178235 unnamed protein product [Arabidopsis thaliana]</td>
<td>At5g22610</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>gi</td>
<td>18394319 DNA binding / transcripti</td>
<td>74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>gi</td>
<td>2244755 hypothetical protein [Arabidopsis thaliana]</td>
<td>DL3035W</td>
<td>T7H20.20</td>
<td>73</td>
</tr>
<tr>
<td>H</td>
<td>gi</td>
<td>15241656 unknown protein [Arabidopsis thaliana]</td>
<td>At1g31290</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>gi</td>
<td>15241656 unknown protein [Arabidopsis thaliana]</td>
<td>T7H20.20</td>
<td>AT5G01970</td>
<td>72</td>
</tr>
<tr>
<td>E</td>
<td>gi</td>
<td>15246454 unknown protein [Arabidopsis thaliana]</td>
<td>F16D14.11</td>
<td>At2G31270</td>
<td>72</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>30680811 RNA binding / RNA methyltransferase [Arabidopsis thaliana]</td>
<td>F6F22.10</td>
<td>AT2G19870</td>
<td>71</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>9279636 unnamed protein product [Arabidopsis thaliana]</td>
<td>MOJ10.20</td>
<td>AT3G27120</td>
<td>70</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>12323390 hypothetical protein; 73-483-63403 [Arabidopsis thaliana]</td>
<td>F6F22.10</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>gi</td>
<td>15233105 DNA binding / transcription coactivator [Arabidopsis thaliana]</td>
<td>At1g77600</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>gi</td>
<td>15231105 DNA binding / transcription coactivator [Arabidopsis thaliana]</td>
<td>At1g77600</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>2660674 hypothetical protein [Arabidopsis thaliana]</td>
<td>REV1</td>
<td>At1g77600</td>
<td>64</td>
</tr>
<tr>
<td>B</td>
<td>gi</td>
<td>30693433 unknown protein [Arabidopsis thaliana]</td>
<td>At1g38065</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>gi</td>
<td>4467359 Phosphatidylcholine:phosphatidylcholine 4-kinase [Arabidopsis thaliana]</td>
<td>At1g38065</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>gi</td>
<td>15219109 HSP70B, ATP binding [Arabidopsis thaliana]</td>
<td>AT1G60390</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>gi</td>
<td>52354231 hypothetical protein AT1G75150 [Arabidopsis thaliana]</td>
<td>T15G18_1</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>15237334 unknown protein [Arabidopsis thaliana]</td>
<td>At1g53030</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>3834306 EST gb</td>
<td>R65024 comes from this gene. [Arabidopsis thaliana]</td>
<td>SNC1-like,SNW</td>
<td>AT1G77180</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>40846341 SNC1-like protein [Arabidopsis thaliana]</td>
<td>MIE1.9</td>
<td>AT3G14590</td>
<td>64</td>
</tr>
<tr>
<td>B</td>
<td>gi</td>
<td>12323390 hypothetical protein; 73-483-63403 [Arabidopsis thaliana]</td>
<td>MIE1.9</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>gi</td>
<td>3763925 putative calthrin binding protein (epsin) [Arabidopsis thaliana]</td>
<td>At1g53030</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>18406032 DNA binding / transcription coactivator [Arabidopsis thaliana]</td>
<td>MBF1A</td>
<td>At2g42680</td>
<td>63</td>
</tr>
<tr>
<td>H</td>
<td>gi</td>
<td>15227892 DNA binding [Arabidopsis thaliana]</td>
<td>At2g42150</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>gi</td>
<td>15233358 DNA binding [Arabidopsis thaliana]</td>
<td>At4g11400</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>gi</td>
<td>5042343 Hypothetical protein [Arabidopsis thaliana]</td>
<td>At1g65010</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>gi</td>
<td>15235468 unknown protein [Arabidopsis thaliana]</td>
<td>At1g65010</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>gi</td>
<td>4803928 putative XAP-5 protein (Homo sapiens) [Arabidopsis thaliana]</td>
<td>At1g65010</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>gi</td>
<td>15240096 unknown protein [Arabidopsis thaliana]</td>
<td>At1g65010</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>20198135 putative chromosome associated protein</td>
<td>At1g65010</td>
<td>57</td>
<td></td>
</tr>
</tbody>
</table>

Continued
Table A.1: continued

| L | gi|17065024 Unknown protein [Arabidopsis thaliana] | 55 |
| L | gi|15219789 lupeol synthase [Arabidopsis thaliana] | 54 |
| H | gi|18409633 ubiquitin conjugating enzyme/ ubiquitin-like activating enzyme [Arabidopsis thaliana] | 53 |
| L | gi|14916970 ATP synthase subunit alpha, mitochondrial | 53 |
| C | gi|15225561 unknown protein [Arabidopsis thaliana] | 52 |
| L | gi|5541663 putative helicase, fragment [Arabidopsis thaliana] | 52 |
| C | gi|9280321 unnamed protein product [Arabidopsis thaliana] | 51 |
| H | gi|12325155 unknown protein; 58745-68005 [Arabidopsis thaliana] | 50 |
| D | gi|15236573 transporter [Arabidopsis thaliana] | 50 |
| L | gi|4056495 putative kinesin heavy chain [Arabidopsis thaliana] | 50 |
| A | gi|4432796 putative VSF-1-like b-ZIP transcription factor [Arabidopsis thaliana] | 49 |
| C | gi|15219625 unknown protein [Arabidopsis thaliana] | 49 |
| D | gi|7288013 putative protein [Arabidopsis thaliana] | 49 |
| E | gi|8569097 Contains strong similarity to rapamycin associated protein FRAP2 from Homo sapiens gb|U88966 and co | 48 |
| L | gi|4415919 hypothetical protein [Arabidopsis thaliana] | 47 |
| L | gi|15218016 unknown protein [Arabidopsis thaliana] | 47 |
| L | gi|8843818 unnamed protein product [Arabidopsis thaliana] | 47 |
| C | 51 |
| D | gi|15241928 AtSerat1.1 (SERINE ACETYLTTRANSFERASE 52); serine O-acetyltransferase [Arabidopsis thaliana] | 46 |
| C | gi|4704662 uracil phosphoribosyltransferase 1 [Arabidop | 46 |
| L | gi|30690461 ATP binding / ATP-dependent helicase/ helicase/ nucleic acid binding [Arabidopsis thaliana] | 46 |
| B | gi|9759303 protein phosphatase 2C-like [Arabidopsis thaliana] | 45 |
| D | gi|3335344 Similar to serpin gene homolog gb|Z49890 from A. thaliana, [Arabidopsis thaliana] | 45 |
| E | gi|2245104 retrotransposon like protein [Arabidopsis thaliana] | 45 |
| E | gi|15219846 unknown protein [Arabidopsis thaliana] | 45 |
| L | gi|4204276 Hypothetical protein [Arabidopsis thaliana] | 44 |
| D | gi|8885606 retroelement pol polyprotein-like [Arabidop | F16F17.4 | AT5G38040 | 43 |
| B | gi|12321345 jasmonate inducible protein, putative [Arabidopsis thaliana] | 42 |
| D | gi|5733879 ESTs gb|N97074, gb|T13943 and gb|R89965 come from this gene. [Arabidopsis thaliana] | 42 |
| C | gi|8920636 Strong similarity to wall-associated kinase 1 from Arabidopsis thaliana gb|AJ009696 and contains Eu gb|F16F4.8 | F16F17.4 | AT1G21240 | 41 |
| D | gi|50911557 T10024.26 [Arabidopsis thaliana] | 41 |
| L | gi|22327062 ATP binding / kinase/ protein kinase/ protein serine/threonine kinase/ protein-tyrosine kinase [Ara | 41 |
| H | gi|10121909 T79N2.4 [Arabidopsis thaliana] | 41 |
| H | gi|15231810 PIP2A; water channel [Arabidopsis thaliana] | 40 |
| H | gi|4309763 putative retroelement pol polyprotein [Arabidopsis thaliana] | 40 |
| A | gi|8570446 Contains similarity to a kinesin homolog from Arabidopsis thaliana gb|T06733 and contains a Kinesin | 40 |
| A | gi|9294330 unnamed protein product [Arabidopsis thaliana] | 40 |
| L | gi|3047096 similar to eukaryotic protein kinase domains (Pfam: pkinase.hmm, score: 189.74) [Arabidopsis thaliana] | 40 |
| L | gi|15912233 At2g41680/T32G6.20 [Arabidopsis thaliana] | 40 |
| L | gi|15227902 ARAC1; GTP binding [Arabidopsis thaliana] | 40 |
| D | gi|15238905 protein binding / zinc ion binding [Arabidopsis thaliana] | 39 |

Continued
Table A.1: continued

<table>
<thead>
<tr>
<th>Column</th>
<th>Accession</th>
<th>Description</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>gi</td>
<td>8778301</td>
<td>F14J16.5 [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>2827718</td>
<td>retrotransposon - like protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>B</td>
<td>gi</td>
<td>42564994</td>
<td>unknown protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>30678047</td>
<td>aspartic-type endopeptidase/ pepsin A [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>7576220</td>
<td>putative protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>9294677</td>
<td>unnamed protein product [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>D</td>
<td>gi</td>
<td>6562000</td>
<td>hypothetical protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>42569280</td>
<td>unknown protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>18402131</td>
<td>unknown protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>15236141</td>
<td>unknown protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>15223349</td>
<td>unknown protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>4185129</td>
<td>At14a, putative [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>15233088</td>
<td>unknown protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>D</td>
<td>gi</td>
<td>42569161</td>
<td>unknown protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>D</td>
<td>gi</td>
<td>18414574</td>
<td>unknown protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>15081709</td>
<td>Atg52410/F19K6_14 [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>15225116</td>
<td>unknown protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>4678944</td>
<td>putative protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>A</td>
<td>gi</td>
<td>9758971</td>
<td>ABC transporter, ATP-binding protein-like [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>8778646</td>
<td>F5O11.15 [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>12230770</td>
<td>Probable N(2),N(2)-dimethylguanosine</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>15220095</td>
<td>DNA binding [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>H</td>
<td>gi</td>
<td>4587572</td>
<td>Similar to gb</td>
</tr>
<tr>
<td>L</td>
<td>gi</td>
<td>9187883</td>
<td>mitochondrial half-ABC transporter [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>B</td>
<td>gi</td>
<td>9759216</td>
<td>subtilisin-like serine protease [Arabidopsis thaliana]</td>
</tr>
</tbody>
</table>