FUNCTIONAL CHARACTERIZATION OF THE \textit{FT}/\textit{TFL1} GENE FAMILY IN PHOTOPERIODIC FLOWERING IN ARABIDOPSIS AND SOYBEAN

BY

DANIEL PAUL WICKLAND

THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Crop Sciences in the Graduate College of the University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Master’s Committee:
Assistant Professor Yoshie Hanzawa
Professor Lila Vodkin
Professor Brian Diers
ABSTRACT

The flowering transition represents one of the most important developmental milestones in the plant life cycle. Several interconnected pathways contribute to flowering induction. By coordinating flowering time with environmental input, these pathways help to optimize plant adaptation and reproductive success. The photoperiodic flowering pathway relies on day length in order to ensure the coincidence of plant reproductive development with favorable seasonal timing. The projects described here aimed to characterize the functions of the $FT/TFL1$ gene family in photoperiodic flowering control in Arabidopsis and soybean.

Chapter One reviews plant flowering regulation. It opens by discussing the developmental context of flowering and the importance of photoperiodic flowering control to plant reproduction and agriculture. The chapter then details the mechanisms of the photoperiodic flowering pathway in the model species Arabidopsis thaliana. Next, the photoperiodic flowering pathways in rice, soybean, sugar beet and poplar are covered. The remainder of the chapter surveys five other flowering pathways: the vernalization, ambient temperature, gibberellin and autonomous flowering pathways.

Chapter Two reports efforts to characterize the roles of the $FT/TFL1$ gene family in photoperiodic flowering in Arabidopsis. Using an RNA-sequencing approach, we analyzed mRNA expression data to identify genes potentially regulated by TFL1. Known targets of TFL1 as well as potentially further downstream targets were found. Using a transgenic approach based on two induction systems, we aimed to identify the immediate targets of FT and TFL1. The results of this approach suggest that FT controls downstream genes in an indirect manner.

Chapter Three describes research to characterize the functions of soybean FT homologs, particularly in relation to soybean development and evolution. We focused on one such homolog and found that mRNA expression data, gene structure data and functional data suggest that this gene stimulates flowering in wild soybean but is nonfunctional in domesticated soybean.
ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Dr. Yoshie Hanzawa, for offering invaluable guidance during my master’s program. I would also like to acknowledge the technical instruction and support I received from my fellow lab members: Eric Sedivy, Faqiang Wu, Waseem Haider and Minglei Wang. In particular, I would like to recognize the contributions of Faqiang Wu, who cloned $FT_{2c}$ and other soybean $FT$ homologs, and Waseem Haider, who supplied much of the code used for the RNA-sequencing analysis. Finally, I would like to express deepest gratitude to my family and to Gelacio Villalobos for providing the support necessary for me to succeed in graduate school and beyond.
TABLE OF CONTENTS

CHAPTER 1: REVIEW OF FLOWERING REGULATION

1.1 The photoperiodic flowering pathway ...................................................... 1
1.2 Other flowering pathways ........................................................................... 25
1.3 Conclusion ...................................................................................................... 31
1.4 Literature cited ............................................................................................. 32
1.5 Figures ........................................................................................................... 3

CHAPTER 2: ROLES OF FT AND TFL1 IN ARABIDOPSIS
PHOTOPERIODIC FLOWERING CONTROL .......................................................... 59

2.1 Background ..................................................................................................... 59
2.2 Materials and methods .................................................................................. 62
2.3 Results ............................................................................................................ 82
2.4 Discussion ....................................................................................................... 91
2.5 Literature cited .............................................................................................. 94
2.6 Figures .......................................................................................................... 98
2.7 Tables ............................................................................................................ 106

CHAPTER 3: ROLES OF FT AND TFL1 IN SOYBEAN
PHOTOPERIODIC FLOWERING CONTROL ...................................................... 113

3.1 Background ................................................................................................... 113
3.2 Materials and methods ................................................................................ 116
3.3 Results .......................................................................................................... 121
3.4 Discussion ..................................................................................................... 123
3.5 Literature cited .............................................................................................. 124
3.6 Figures .......................................................................................................... 129
3.7 Tables .......................................................................................................... 132
CHAPTER 1
REVIEW OF FLOWERING REGULATION

The transition to flowering represents one of the most important developmental milestones in the plant life cycle. Flowering signifies a plant’s transition from the juvenile vegetative stage of development to the adult reproductive stage (Pidkowich et al., 1999; Benlloch et al., 2007). While in the vegetative phase, the plant’s shoot apical meristem, a structure composed of undifferentiated dividing cells, produces the vegetative primordia that generate leaves and shoots (Benlloch et al., 2007). All aboveground parts of the plant originate from the shoot apical meristem (Benlloch et al., 2007). Eventually, the shoot apical meristem enters the reproductive phase by transforming from a vegetative meristem into an inflorescence meristem (Benlloch et al., 2007; Sablowski, 2007). This juvenile-to-adult transition results in a reproductively competent plant able to respond to both external and internal flowering-inductive signals (Mutasa-Göttgens and Hedden, 2009). Once the plant reaches the reproductive stage, its transition to flowering is defined by the emergence of flower-producing floral meristems from the flanks of the inflorescence meristem (Benlloch et al., 2007; Sablowski, 2007). A number of interconnected pathways contribute to the initiation of flowering: the photoperiodic, vernalization, ambient temperature, gibberellin and autonomous flowering pathways (Simpson and Dean, 2002; Blázquez et al., 2003; Wigge, 2011; Andrés and Coupland, 2012). These pathways enable regulated, seasonal flowering and coordinate flowering with environmental input, which helps to optimize plant adaptation and reproductive success (Wigge, 2011; Andrés and Coupland, 2012). This review chapter focuses primarily on the photoperiodic pathway but surveys the other pathways as well.

1.1 THE PHOTOPERIODIC FLOWERING PATHWAY

In the photoperiodic pathway, a network of transcription factors responds to environmental input to trigger flowering (Figure 1.1) (Andrés and Coupland, 2012).
Photoperiod refers to a plant’s seasonally varying length of daily light exposure (Andrés and Coupland, 2012). Plant photoperiodic responses to the earth’s 24-hour rotational cycle ensure the coincidence of plant reproductive development with proper seasonal timing, leading to optimal seed production, reproductive fitness and yield (Valverde, 2011; Andrés and Coupland, 2012). Regulation of flowering time facilitates the regional adaptation of plants and influences growing seasons (Kim et al., 2012b). This pathway helps to prevent prematurely early and excessively late flowering, which can detrimentally affect plant success. For instance, premature flowering could expose flowers to lingering frigid temperatures and frost from winter, while delayed flowering could prevent plants from concluding their reproductive cycles before winter begins (Inouye, 2008; Beaubien and Hamann, 2011; Anderson et al., 2012).

The external coincidence model accounts for the integration of developmental timing with photoperiodic input (Imaizumi and Kay, 2006). According to this model, light’s resetting of the plant circadian clock produces an endogenous signal that is expressed in a cyclical pattern daily and that initiates responses to photoperiod (Imaizumi and Kay, 2006). “Coincidence” in this model refers to the daily light peak arriving simultaneously with peak expression of the endogenous signal, and this coincidence triggers plant responses to photoperiod (Imaizumi and Kay, 2006). The endogenous signal’s peak expression in the late afternoon coincides with a period of high light intensity, thereby explaining the long-day property of certain plants (Imaizumi and Kay, 2006). Plants belong to one of three groups based on their response to photoperiod (Andrés and Coupland, 2012). Long-day plants flower when the day length exceeds a species-specific threshold length, typically in the spring (Koornneef et al., 1991; Andrés and Coupland, 2012). Short-day plants, on the other hand, flower when the day length falls below a species-specific threshold length, typically in the autumn (Andrés and Coupland, 2012). Unresponsive to photoperiodic inputs, day-neutral plants flower regardless of day length (Andrés and Coupland, 2012).

1.1.1 AGRICULTURAL IMPACTS OF PHOTOPERIODIC FLOWERING

Flowering time is recognized as a key agronomic trait and correlates strongly with crop yields (Koornneef et al., 2004; Itoh et al., 2010; Wilczek et al., 2010). A primary motivation behind studying photoperiodic flowering is to generate knowledge applicable
to agriculture. Through domestication, humans have impacted the regional adaptation of plants, and fine-tuning photoperiodic response offers one way to acclimate plants to different areas (Kim et al., 2012b). Elucidating the genetic regulation of environmentally responsive flowering therefore supports the production of adaptive plant varieties well matched to diverse environments (Kim et al., 2012b). The introduction of earlier flowering varieties could allow crop growth in locations with abbreviated seasons (Roux et al., 2006; Kim et al., 2012a). The introduction of later flowering varieties, with lengthened vegetative phases to produce more biomass for seed production, could increase crop yield in locations with longer seasons (Roux et al., 2006; Kim et al., 2012a; Kim et al., 2012b). Orchestrating flowering times could also facilitate the crossing of genotypes that ordinarily would bloom at different times (Kim et al., 2012a).

Better understanding photoperiodic flowering could also combat two pressing agricultural dilemmas faced by society. A widening gap exists between crop production and accelerating crop demand spurred by global population growth (Ray et al., 2013). Worldwide demand for crop output is expected to double by the year 2050 (Tilman et al., 2011; Ray et al., 2013). Projected estimates for crop production, however, fall far short of this figure. For instance, at current rates, soybean yield will increase by just 65% by 2050 (Ray et al., 2013). The gap between anticipated crop demand and production underscores the need for expertise in plant breeding to improve crop yields. Research into plant environmental response could promote increased and more robust yield or expanded areas of cultivation, thereby helping to bolster global food security (Watanabe et al., 2012).

Knowledge of flowering regulation could also help to alleviate effects of climate change on crop production. Temperature interacts with photoperiodic response and affects plants throughout development (Yan and Wallace, 1998; Craufurd and Wheeler, 2009). Rising global temperatures have already triggered precocious flowering in a variety of plants (Anderson et al., 2012). In fact, several studies have recently shown accelerated flowering time in numerous species due to warming temperatures over time spans of just 50-200 years (Hu et al., 2005; Menzel et al., 2006; Anderson et al., 2012; Ellwood et al., 2013). One meta-analysis of observational studies and warming experiments predicts that flowering will occur 5-6 days earlier in the spring per °C rise in temperature (Wolkovich et al., 2012). Such precocious flowering raises the possibility
of plants—particularly those from populations with little genetic variation—no longer adapting rapidly enough to climate change in the future and becoming extinct (Etterson and Shaw, 2001; Craufurd and Wheeler, 2009; Anderson et al., 2012). Rising temperatures and early flowering abbreviate many crops’ developmental stages, potentially reducing yield as well (Craufurd and Wheeler, 2009). Strategies based on photoperiodic flowering offer a promising method to produce crops robust against climate change. Growing new varieties with region-appropriate photoperiods assists in adapting crops to changing climates (Ludlow and Muchow, 1990; Richards, 2006; Craufurd and Wheeler, 2009). This practice has already allowed breeders and farmers to expand the geographic range and increase the yield of numerous crops (Craufurd and Wheeler, 2009; Lawn and James, 2011).

1.1.2 PHOTOPERIODIC FLOWERING IN ARABIDOPSIS

*Arabidopsis thaliana* (Arabidopsis), a facultative long-day plant, normally flowers under long photoperiods but eventually flowers under non-inductive (i.e. short) photoperiods as well (Amasino, 2010). The plant’s 6-week life cycle, small genome, self-compatibility, and widespread geographic prevalence make it especially amenable to genetic analysis and contribute to its status as a model system (Alonso-Blanco et al., 2009; Weigel, 2012). Genes impacting flowering time generally display conservation across species, and a great number of flowering-related genes in other species, including crop species, have been discovered by identifying orthologs of Arabidopsis flowering-time genes (Wilczek et al., 2010; Watanabe et al., 2012).

**The Circadian Clock**

The photoperiodic pathway in Arabidopsis includes two primary components: the circadian clock and a related day-length measurement mechanism involving the transcription factor CONSTANS (CO), the endogenous signal mentioned earlier (Imaizumi and Kay, 2006). The transcription and translation of particular genes, regulated by a negative feedback loop, constitutes the circadian clock (Yanovsky and Kay, 2003; Izawa, 2007a). The circadian clock maintains particular phases during which it detects input from photoreceptors (Izawa, 2007a). Based on this input, it measures
day length and coordinates responses that initiate or inhibit flowering; the circadian clock directs the expression of genes involved in flowering (Izawa, 2007a).

A number of both morning- and evening-expressed transcription factors comprise the negative feedback loop of the clock’s core oscillator (Imaizumi and Kay, 2006). The evening-expressed transcription factors TIMING OF CAB EXPRESSION1 (TOC1), LUX ARRHYTHMO (LUX) and EARLY FLOWERING 4 (ELF4) induce CIRCadian CLOCK ASSOCIATED 1 (CCA1) and LATE ELONGATED HYPOCOTYL (LHY), morning-expressed factors that reach their highest levels near dawn (Wang and Tobin, 1998; Strayer, 2000; Doyle et al., 2002; Hazen et al., 2005; Onai and Ishiura, 2005; Imaizumi and Kay, 2006). The morning factors CCA1 and LHY, in turn, repress the evening factors TOC1 and LUX in the morning by binding to the promoters of the corresponding genes (Alabadi et al., 2001; Hazen et al., 2005; Imaizumi and Kay, 2006). These transcription factors’ different patterns of expression and interactions throughout the day generate the circadian rhythm (Wang and Tobin, 1998; Yanovsky and Kay, 2003; Imaizumi and Kay, 2006).

**CONSTANS (CO)**

Circadian clock–related processes that control CONSTANS (CO) generation and regulation allow a plant to measure day length (Hayama et al., 2004; Imaizumi and Kay, 2006). The transcription factor CO is a key component of the gene network leading to activation of FLOWERING LOCUS T (FT), a transcriptional regulator that promotes flowering by stimulating expression of meristem identity genes (Putterill et al., 1995; Andrés and Coupland, 2012). Under the influence of mRNA-regulating mechanisms, CO peaks under long photoperiods at the end of the day, when it reaches levels sufficient to trigger FT expression (Suárez-López et al., 2001; Imaizumi et al., 2003; Jang et al., 2008; Andrés and Coupland, 2012). *co* loss-of-function mutants exhibit delayed flowering under both long-day and short-day conditions, indicating an inability to perceive photoperiod without functional CO (Rédei, 1962; Valverde, 2011).

CO regulation takes place at both the transcriptional and post-transcriptional levels (Andrés and Coupland, 2012). The circadian clock and light regulate CO at the transcriptional level (Andrés and Coupland, 2012). The circadian clock controls the expression of genes encoding two proteins that promote CO transcription: GIGANTEA
GI) and FLAVIN-BINDING, KELCH REPEAT, AND F-BOX 1 (FKF1) (Fowler et al., 1999; Nelson et al., 2000; Covington et al., 2001; Mizoguchi et al., 2005). Ten to fourteen hours after dawn (i.e. in the late afternoon) under long-day (i.e. flowering inductive) conditions, GI and FKF1 interact to form a complex that stabilizes FKF1 (Sawa et al., 2007; Fornara et al., 2009; Andrés and Coupland, 2012). Stabilized FKF1 then initiates degradation of the CO transcriptional repressors Cycling DOF Factors (CDFs) (Imaizumi et al., 2003; Izawa, 2007a; Sawa et al., 2007; Fornara et al., 2009; Amasino, 2010; Andrés and Coupland, 2012). CDFs suppress CO transcription by binding to its promoter (Imaizumi et al., 2005; Amasino, 2010). Due to the degradation of CDFs by stabilized FKF1, CO mRNA peaks 12-16 hours after dawn (Andrés and Coupland, 2012). Under short (i.e. non-inductive) photoperiods, CDFs continue suppressing CO transcription at the end of the day because GI and FKF1 reach peak expression at different times, so no appreciable protein interaction occurs (Sawa et al., 2007; Andrés and Coupland, 2012).

Two other circadian clock-controlled proteins repress CO transcription: EARLY FLOWERING 3 (ELF3) and RED AND FAR-RED INSENSITIVE 2 (RFI2) (Kim et al., 2005; Chen and Ni, 2006; Imaizumi and Kay, 2006). ELF3, which inhibits light signaling, indirectly inhibits CO expression (McWatters et al., 2000; Imaizumi and Kay, 2006). Because elf3 loss-of-function mutants express CO, FKF1 and GI at high levels, ELF3 may repress CO by reducing FKF1 and GI expression (Suárez-López et al., 2001; Kim et al., 2005; Imaizumi and Kay, 2006). ELF3 also impacts signaling of the photoreceptor phyochrome B (phyB) and therefore may affect CO post-transcriptionally as well (Kim et al., 2005; Imaizumi and Kay, 2006). Another protein, RFI2, also suppresses CO transcription, and rfi2 loss-of-function mutants flower early and show upregulated CO and FT under both long- and short-day conditions (Chen and Ni, 2006; Imaizumi and Kay, 2006). Like ELF3, RFI2 influences phyB signaling, resulting in delayed flowering and degraded CO (Lin, 2000; Valverde et al., 2004; Chen and Ni, 2006).

A number of mechanisms regulate CO post-transcriptionally, and these mechanisms account for the protein’s activity during certain phases of light (Valverde, 2011; Andrés and Coupland, 2012). These processes promote peak CO levels at the end of the day and limit its abundance at other times. CO protein becomes most stable in the
late afternoon under long day-conditions (Valverde et al., 2004; Imaizumi and Kay, 2006). Under short-day conditions, however, it remains less stable throughout the day (Valverde et al., 2004). The photoreceptors phyA and phyB and cryptochromes cry1 and cry2 direct this light-driven modulation of CO protein abundance and stability (Valverde et al., 2004; Imaizumi and Kay, 2006). Under long photoperiods, phyB facilitates CO degradation in the morning, while phyA and the cryptochromes stabilize CO at dusk in a process involving far-red and blue light detection (Valverde et al., 2004; Chen and Ni, 2006; Imaizumi and Kay, 2006; Izawa, 2007a). Short photoperiods, on the other hand, result in cessation of cry2 and phyA signaling and therefore less stable CO (Izawa, 2007a). In another mechanism of post-transcriptional CO regulation, SUPPRESSOR OF PHYA-105 (SPA1) and its homologs SPA3 and SPA4 decrease CO stability and operate in a pathway that negatively regulates phyA (Valverde et al., 2004; Imaizumi and Kay, 2006; Laubinger et al., 2006). Finally, CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1) forms a complex with SPA1 that degrades CO in the dark via ubiquitination (Jang et al., 2008; Liu et al., 2008b; Amasino, 2010; Zuo et al., 2011; Andrés and Coupland, 2012). During the day, however, light activates the blue-light photoreceptor cry3, which stabilizes CO by binding to the COP1/SPA1 complex and thus impairing its catalytic activity (Zuo et al., 2011; Andrés and Coupland, 2012).

FLOWERING LOCUS T (FT)

CO activates FLOWERING LOCUS T (FT), a transcriptional regulator and the main factor impacting the time at which plants flower (Wigge, 2011; Andrés and Coupland, 2012). FT overexpression accelerates flowering, while the ft loss-of-function mutation retards flowering (Kardailsky et al., 1999; Kobayashi, 1999). Specific expression of FT either in leaf phloem companion cells or in the shoot apical meristem leads to flowering even without CO present (An et al., 2004).

Prior to its identification as FT, a floral stimulus (florigen) found in leaves of induced plants was shown to cause flowering in non-induced plants (Lang et al., 1977; Zeevaart, 1982; Imaizumi and Kay, 2006). Grafting experiments demonstrated that non-induced plants can flower from just one induced leaf transplanted onto them, and similar grafts can also induce flowering between different species (Zeevaart, 1976; Imaizumi and Kay, 2006; Wigge, 2011). In Xanthium strumarium, a single leaf placed
in flowering-inductive conditions spurs a flowering response throughout the plant (Zeevaart, 1976; Imaizumi and Kay, 2006). Transport of florigen occurs because a connection must exist between the light-sensing organs—the leaves—and the flowering organ in the shoot (Imaizumi and Kay, 2006). FT protein moves from the leaves through the phloem until it reaches the shoot apical meristem (Takada and Goto, 2003; Imaizumi and Kay, 2006; Corbesier et al., 2007; Lin et al., 2007; Wigge, 2011). FT-INTERACTING PROTEIN 1 (FTIP1) facilitates FT movement from phloem companion cells to sieve elements (Andrés and Coupland, 2012; Liu et al., 2012). After reaching the meristem, FT participates in transcriptional complexes to influence downstream gene expression (Andrés and Coupland, 2012). Its binding to the transcription factor FLOWERING LOCUS D (FD) forms a complex that regulates meristem identity genes, resulting in promotion of flowering (Abe et al., 2005; Hanano and Goto, 2011; Andrés and Coupland, 2012).

FT expression occurs only at dusk under long-day conditions, when CO binds to the FT promoter to initiate transcription (Kobayashi, 1999; Izawa, 2007a; Tiwari et al., 2010; Andrés and Coupland, 2012). CO induces FT in the leaves; CO expressed specifically in leaf phloem companion cells, but not in the shoot apical meristem, leads to early flowering in both wild-type and co loss-of-function mutant Arabidopsis (An et al., 2004; Ayre and Turgeon, 2004; Amasino, 2010). In addition to its regulation by CO, FT undergoes regulation by a number of other factors, many of which target its promoter region (Adrian et al., 2010; Andrés and Coupland, 2012). For instance, under high temperatures, PHYTOCHROME INTERACTING FACTOR4 (PIF4) binds to FT’s promoter to initiate transcription (Andrés and Coupland, 2012; Kumar et al., 2012). FT transcription is repressed when LIKE HETEROCHROMATIN PROTEIN1 (LHP1) (also called TERMINAL FLOWER 2 (TFL2)) binds to FT’s promoter and modifies histones (Adrian et al., 2010). Mutants for LHP1 demonstrate increased FT expression that promotes early flowering independent of day length (Kotake et al., 2003; Adrian et al., 2010). Finally, the circadian clock-associated proteins GI and CDF1 also regulate FT at the promoter level (Sawa and Kay, 2011; Andrés and Coupland, 2012; Song et al., 2012).
FLOWERING LOCUS D (FD)

The basic leucine zipper transcription factor FD interacts with FT in the shoot apical meristem to modulate transcription of genes involved in flowering induction (Abe et al., 2005; Hanano and Goto, 2011; Andrés and Coupland, 2012). The FT/FD complex promotes expression of SQUAMOSA BINDING LIKE (SPL) transcription factors in the meristem, which results in transcription of meristem identity genes such as LEAFY (LFY), APETALA1 (AP1) AND FRUITFULL (FUL) (Wang et al., 2009; Yamaguchi et al., 2009; Andrés and Coupland, 2012; Jung et al., 2012). When transformed with an FT overexpression construct, fd loss-of-function mutants fail to show the early-flowering phenotype of FT-overexpressing plants, highlighting the FT-FD partnership’s critical role in FT function (Abe et al., 2005; Wigge et al., 2005).

Meristem identity genes downstream of the photoperiodic pathway

The transition to flowering forms the basis of reproductive development and relies on a number of meristem identity genes (Blázquez et al., 2006). These genes encourage cell differentiation in the shoot apical meristem that leads to the irreversible process (in Arabidopsis) of transition to a reproductive inflorescence meristem (Amasino, 2010). After becoming a reproductive inflorescence meristem, the shoot apical meristem generates floral meristems; the actions of meristem identity genes reprogram the primordia to produce reproductive structures instead of vegetative ones (Blázquez et al., 2006; Amasino, 2010). Meristem identity genes integrate internal and external inputs and exhibit reciprocal regulatory interactions (Blázquez et al., 2006). The developmental transition to an inflorescence meristem is associated with rising SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1), SHORT VEGETATIVE PHASE (SVP) and AGAMOUS Like 24 (AGL24) expression (Mutasa-Göttgens and Hedden, 2009). Generation of a floral meristem produced from the inflorescence meristem entails the activity of LEAFY (LFY) and APETALA1 (AP1), which in turn inhibit inflorescence meristem genes (such as SOC1, SVP and AGL24) (Liu et al., 2007; Souer et al., 2008; Mutasa-Göttgens and Hedden, 2009). Therefore, AP1 and LFY expression occurs only in the lateral floral meristems, not in the shoot apical meristem (Liu et al., 2007; Souer et al., 2008; Mutasa-Göttgens and Hedden, 2009). Inhibition of
inflorescence meristem genes permits the floral meristems’ commitment to flowering, while the inflorescence remains indeterminate (Mutasa-Göttgens and Hedden, 2009).

**LEAFY (LFY)**

Considered the central agent of floral induction with an essential role in conferring floral meristem identity, LFY specifies a transcription factor that operates in the early stages of floral development (Maizel et al., 2005; Blázquez et al., 2006; Lee et al., 2008; Andrés and Coupland, 2012). LFY integrates input from multiple pathways to initiate flowering (Liljegren et al., 1999; Weigel and Blazquez, 2000; Parcy, 2005; Benlloch et al., 2007). It helps to maintain meristem identity and activates downstream genes by binding to their regulatory regions; its targets include the flowering-related genes AP1, APETALA 3 (AP3) and AGAMOUS (AG) (Parcy et al., 1998; Busch et al., 1999; Weigel and Blazquez, 2000; Lamb et al., 2002). LFY levels rise with advancing maturity, and LFY levels in the shoot apical meristem influence timing of the flowering transition (Blázquez et al., 1997; Benlloch et al., 2007). High expression of LFY in leaf primordia correlates both with impending floral transition and with sensitivity to flowering-inductive signals (Blázquez et al., 1997; Blázquez et al., 2006).

Mutant and overexpression analysis provides additional insight into LFY function. *lfy* loss-of-function mutants present abnormal floral development, with most flowers partially converted into secondary inflorescence shoot-like structures (Schultz and Haughn, 1991; Huala and Sussex, 1992; Weigel et al., 1992; Blázquez et al., 1997; Blázquez et al., 2006; Benlloch et al., 2007). These mutants undergo a late transition from vegetative to reproductive development, finally flowering even without LFY present due to the eventual expression of flowering-inductive downstream genes (Blázquez et al., 2006; Benlloch et al., 2007). Constitutively expressed LFY, on the other hand, causes early flowering; secondary meristems prematurely transition to floral meristems (Weigel and Nilsson, 1995). Overexpression of LFY also prompts the premature appearance of AP1 in young leaves (Parcy et al., 1998).

A number of processes contribute to LFY regulation. The photoperiodic, gibberellin and autonomous flowering pathways upregulate LFY (Blázquez et al., 2006). In the photoperiodic pathway, LFY activates both AP1 and the related floral inducer CAULIFLOWER (CAL), which in turn reciprocally regulate LFY (Wagner, 1999; William
et al., 2004; Blázquez et al., 2006). The development-related gibberellin hormones act on the *LFY* promoter to stimulate flowering (Blázquez et al., 1997). The gibberellin-deficient *ga* mutant exhibits very low *LFY* expression, failure to flower under short-day conditions, and late flowering under long-day conditions (Blazquez et al., 1998; Blázquez et al., 2006). The introduction of constitutively expressed *LFY* rescues this mutant phenotype (Blazquez et al., 1998; Gocal et al., 2001; Blázquez et al., 2006). Finally, SOC1 promotes *LFY* expression to induce flowering as well (Lee et al., 2008).

**APETALA1 (AP1)**

*AP1* promotes floral meristem identity in conjunction with *LFY*, participates in floral organ formation, and ensures that no reversion to vegetative development occurs (Mandel et al., 1992; Bowman et al., 1993; Blázquez et al., 2006). *AP1* is expressed in young floral meristems following the appearance of *LFY* and functions as an inducer of the flowering transition (Mandel et al., 1992; Parcy et al., 1998; Liljegren et al., 1999; Benlloch et al., 2007). Restricted to young flower primordia, it shows no expression in the inflorescence meristem (Mandel et al., 1992; Mandel and Yanofsky, 1995a). Mutant and overexpression data support *AP1*’s role in promoting floral meristem identity. *ap1* loss-of-function mutants display indeterminate growth and abnormalities associated with floral organ identity, including a partial conversion of flowers into inflorescence shoots (Bowman et al., 1993; Blázquez et al., 2006; Benlloch et al., 2007). Constitutive *AP1* expression causes early flowering, conversion of lateral meristems to flowers, and acquisition of floral identity in the shoot apical meristem (Mandel and Yanofsky, 1995b; Benlloch et al., 2007).

In addition to its induction by *LFY*, *AP1* undergoes regulation by FT and FD (Blázquez et al., 2006). Mutant analysis provides evidence for FT and FD’s involvement in *AP1* modulation. In contrast to the *AP1*-expressing *lfy* mutant, the *ft/lfy* double mutant shows no *AP1* expression, which indicates that FT plays a role in *AP1* induction (Ruiz-García et al., 1997; Blázquez et al., 2006). Moreover, *ft* single mutants show great delays in *AP1* expression in response to flowering-inductive inputs (Schmid et al., 2003; Blázquez et al., 2006). In concert with FT, FD promotes *AP1* transcription, and *fd* mutants show late *AP1* expression (Abe et al., 2005; Wigge et al., 2005; Blázquez et al., 2006).
CAULIFLOWER (CAL)

A paralog of AP1, the MADS-box gene CAL mirrors AP1 in sequence, expression and function (Kempin et al., 1995; Blázquez et al., 2006; Benlloch et al., 2007; Izawa, 2007a). cal loss-of-function mutants show no differences from wild-type flowering; the AP1 present in these mutants can compensate for CAL’s absence (Kempin et al., 1995). In cal/ap1 double mutants, however, reversion of floral meristems into inflorescence meristems takes place, which underscores the important role of both genes in the transition to a floral meristem (Bowman et al., 1993; Blázquez et al., 2006).

FRUITFULL (FUL)

The AP1 paralog FRUITFULL (FUL), also called AGAMOUS–LIKE 8 (AGL8), supports AP1 and CAL function (Blázquez et al., 2006; Izawa, 2007a). In contrast to AP1, it is highly expressed in the inflorescence meristem, but not in floral meristems, following the floral transition (Mandel and Yanofsky, 1995a; Gu et al., 1998; Blázquez et al., 2006). AP1 controls this partitioning because FUL is expressed equally in both the inflorescence meristem and floral meristems of ap1 loss-of-function mutants (Mandel and Yanofsky, 1995a). Therefore, because of its high and specific expression in the inflorescence meristem at the shoot apex, FUL may act in the maintenance of inflorescence meristem identity or in flowering induction (Mandel and Yanofsky, 1995a).

TERMINAL FLOWER 1 (TFL1)

The flowering repressor TERMINAL FLOWER1 (TFL1) performs two main functions (Blázquez et al., 2006; Benlloch et al., 2007). First, it modulates developmental phase length (Bradley, 1997; Ratcliffe et al., 1998; Blázquez et al., 2006). Second, it controls the determinacy status of the inflorescence shoot (Bradley, 1997). In terms of its first function, TFL1 suppresses the transition to the reproductive phase by influencing the shoot apical meristem developmental transitions; the tfl1 loss-of-function mutation shortens the vegetative and reproductive phases and therefore accelerates flowering, while TFL1 overexpression elongates developmental stages and therefore delays flowering (Ratcliffe et al., 1998; Blázquez et al., 2006). TFL1’s second function is to promote inflorescence meristem identity, in contrast to LFY and AP1,
which promote floral meristem identity (Liljegren et al., 1999; Benlloch et al., 2007). In \textit{tfl1} mutants, the inflorescence shoot becomes a floral meristem that produces a terminal flower; the inflorescence switches from indeterminate to determinate growth (Shannon and Meeks-Wagner, 1991; Alvarez et al., 1992; Bradley, 1997; Blázquez et al., 2006; Benlloch et al., 2007). Because the inflorescence shoot of \textit{tfl1} mutants acquires floral identity, \textit{LFY} and \textit{AP1} are expressed in this structure (Ratcliffe et al., 1998).

\textit{TFL1} operates at the molecular level by preventing meristem identity genes (such as \textit{LFY} and \textit{AP1}) from acting at the center of the shoot apex (Ratcliffe et al., 1999). It both inhibits these floral meristem genes’ expression and prevents the meristem from responding to them (Ratcliffe et al., 1999). \textit{tfl1} mutants ectopically express \textit{LFY} and \textit{AP1}, while plants overexpressing \textit{TFL1} show delayed expression of the two genes (Weigel et al., 1992; Bowman et al., 1993; Bradley, 1997; Ratcliffe et al., 1998; Blázquez et al., 2006). Additionally, \textit{TFL1}-overexpressing plants that also overexpress either \textit{LFY} or \textit{AP1} show disrupted floral development, with primary and axillary meristems acquiring shoot identity instead of developing into flowers, which lends more support to \textit{TFL1}’s inhibitory function (Ratcliffe et al., 1999). \textit{TFL1} also disrupts \textit{LFY}’s upregulation of \textit{AP1} (Blázquez et al., 2006). \textit{TFL1} may suppress \textit{LFY}, \textit{AP1} and other meristem identity genes by partnering with \textit{FD}, with which it weakly interacts (Abe et al., 2005; Hanano and Goto, 2011).

Just as it regulates the expression of floral meristem genes, \textit{TFL1} itself is regulated by floral meristem genes. \textit{LFY}, \textit{AP1} and \textit{CAL} block expression of \textit{TFL1} in floral meristems, particularly at the apex periphery (Ratcliffe et al., 1999; Blázquez et al., 2006). In the presence of ectopically expressed \textit{LFY}, no \textit{TFL1} is expressed in the shoot apex (Ratcliffe et al., 1999). Similarly, ectopic expression of \textit{AP1} reduces \textit{TFL1} expression, while the \textit{ap1 cal} double mutant expresses \textit{TFL1} at high levels (Ratcliffe et al., 1999; Blázquez et al., 2006). As transcription factors, \textit{LFY}, \textit{AP1} and \textit{CAL} may inhibit \textit{TFL1} by binding to its promoter (Ratcliffe et al., 1999; Blázquez et al., 2006). Ultimately, the differing temporal and spatial expression patterns of these genes determines meristem identity.
Evolution and Natural Variation

Natural variation within a species stems from spontaneous, naturally occurring mutations preserved by selection (Alonso-Blanco et al., 2009). Regulatory genes, such as ones encoding transcription factors and signal transduction machinery, contribute significantly to flowering-time variation (Alonso-Blanco et al., 2009). Probably a result of flowering time and climate varying together, the geographic distribution of Arabidopsis is categorized into altitudinal and latitudinal clines (Caicedo et al., 2004; Lempe et al., 2005; Weigel, 2012). This distribution suggests that flowering time is under selection (Weigel, 2012). Gene duplication, a process responsible for generating some of the genetic variation upon which selection acts, occurred in 70-80% of all angiosperms (Moore and Purugganan, 2005; Liu et al., 2008a). Often, one copy of a duplicated gene retained the original function, while a second, mutated copy lost the original function and either acquired a new function or became non-functional (i.e. a pseudogene) (Liu et al., 2008a). Almost all angiosperms, including crops, analyzed so far demonstrate polyploidy, which implies at least one universal genome duplication common to all plants (Soltis et al., 2008). Evidence suggests that flowering-related genes underwent a duplication close to the beginning of angiosperm evolution (Soltis et al., 2008). For instance, at least two homologs of the flowering-related genes AP3 and PISTILLATA (PI) exist in all angiosperms (Soltis et al., 2008).

A number of genes involved in flowering are highly conserved throughout the plant kingdom (Andrés and Coupland, 2012). Some flowering genes retain functions equivalent to those found in Arabidopsis, other genes perform different functions despite their homology to Arabidopsis flowering genes, and still other non-Arabidopsis flowering genes have no counterparts in Arabidopsis (Andrés and Coupland, 2012). The CO-FT system is present in all plants examined so far, although its function varies among species (Valverde, 2011). For instance, CO induces FT under long-day conditions and thus promotes flowering in Arabidopsis, but the rice CO ortholog represses the rice FT ortholog under long-day conditions and thus inhibits flowering (Valverde, 2011; Andrés and Coupland, 2012).

Arabidopsis contains a number of CO homologs called CO-like (COL) proteins. COL1 and COL2, for example, share about 67% amino acid identity with CO and, like CO, undergo circadian-clock regulation, but their overexpression has little impact on
flowering (Ledger et al., 2001; Valverde, 2011). The first CO-like gene originated in algae, and this ancestral gene spawned the COL family of genes found throughout the plant kingdom (Zobell et al., 2005; Chia et al., 2008). Typically under circadian control, these genes function in processes involving light (Valverde, 2011). CrCO, the circadian-regulated CO ortholog from the algae Chlamydomonas reinhardtii and an ancestral form of both CO and its rice ortholog, stimulates flowering when expressed in wild-type Arabidopsis (Serrano et al., 2009; Valverde, 2011). Upregulation of FT in Arabidopsis overexpressing CrCO is highly similar to upregulation of FT in Arabidopsis overexpressing CO itself (Serrano et al., 2009). Under the constitutive 35S promoter and phloem-specific SUC2 promoter, CrCO complements the Arabidopsis co mutant, causing early flowering (Serrano et al., 2009). Under the same promoters, the CrCO transgene also elicits early flowering in wild-type Arabidopsis (Serrano et al., 2009). These results indicate conservation between CO and CrCO, particularly in their biochemistry (Serrano et al., 2009; Valverde, 2011). Although they share only 27% sequence identity, their effects on phenotype and FT levels imply that CO and CrCO interact with the same complexes (Serrano et al., 2009; Valverde, 2011).

FT-like genes are especially conserved among many species and share the trait of transcription triggered by species-specific inductive conditions (Andrés and Coupland, 2012). Despite their conservation in flowering-inductive function, these genes are subject to different species-specific regulatory mechanisms that control their expression, which explains variation in photoperiodic response among plant species (Andrés and Coupland, 2012; Pin and Nilsson, 2012). FT-like and TFL1-like genes arose with the emergence of flowering plants (Hedman et al., 2009; Karlgren et al., 2011; Pin and Nilsson, 2012). In terms of molecular evolution, evidence points to the flowering repressor MOTHER OF FT AND TFL1 (MFT) as the ancestral form of FT and TFL1 (Hedman et al., 2009; Andrés and Coupland, 2012). Like FT and TFL1, MFT is a member of the phosphatidyl ethanolamine-binding protein (PEBP) family (Hedman et al., 2009). It appears that during evolution, an MFT-like gene duplication spawned two clades, MFT-like genes and FT/TFL1-like genes, which led to further diversification of function, duplication and eventual divergence into separate FT-like and TFL1-like clades (Hedman et al., 2009; Karlgren et al., 2011; Pin and Nilsson, 2012). FT and TFL1 share ~60% amino acid identity, and swapping one key amino acid between FT and TFL1
converts FT into a flowering inhibitor and TFL1 into a flowering inducer, which suggests conserved biochemical action (Hanzawa et al., 2005; Hanano and Goto, 2011).

1.1.3 PHOTOPERIODIC FLOWERING IN RICE

In contrast to Arabidopsis, rice (*Oryza sativa*) flowers in response to short days, although it eventually flowers under long days as well (Tamaki et al., 2007; Tsuji et al., 2011; Andrés and Coupland, 2012). Considered a model short-day organism as well as a staple crop, rice contains a number of genes orthologous to ones found in Arabidopsis (Izawa, 2007a; Tsuji et al., 2011; Andrés and Coupland, 2012). As in Arabidopsis, the photoperiodic pathway in rice interacts with the circadian clock; changes in day length influence flowering induction (Itoh et al., 2010). Rice, like other short-day plants, produces seeds that germinate in the spring, develops and flowers in the summer under short days, and sets seeds in the autumn (Izawa, 2007a).

**HEADING DATE 1 (HD1)**

HEADING DATE 1 (HD1), the circadian clock-regulated rice ortholog of CO, regulates the flowering transition (Hayama et al., 2003; Izawa, 2007a; Wigge, 2011; Andrés and Coupland, 2012). Unlike CO, however, this protein performs dual roles. First, it promotes flowering under short-day conditions by initiating transcription of the flowering inducer and *FT* ortholog **HEADING DATE 3A (HD3A)** (Figure 1.2) (Hayama et al., 2003; Doi et al., 2004; Komiya et al., 2009; Andrés and Coupland, 2012). In this sense, it behaves under short days as CO behaves under long days (Tsuji et al., 2011). Under short-day conditions, wild-type rice expresses peak **HD1** at night and **hd1** loss-of-function mutants express **HD3A** at lower levels than does wild type (Izawa, 2007a; Tsuji et al., 2011). Unlike CO under short days—which is expressed and subsequently degraded at night—**HD1** protein remains intact at night under short days (Jang et al., 2008; Liu et al., 2008b; Tsuji et al., 2011). Under long-day conditions, **HD1** performs its second function: repressing flowering by suppressing **HD3A** transcription (Figure 1.3) (Hayama et al., 2003; Doi et al., 2004; Komiya et al., 2009). Arabidopsis CO, on the other hand, only promotes flowering and does so under long days (Yanovsky and Kay, 2003; Izawa, 2007b).
Rice FT Orthologs: **HEADING DATE 3A (HD3A) and RICE FLOWERING LOCUS T (RFT1)**

Rice carries two FT orthologs that play critical roles in flowering; when neither is expressed, flowering never occurs (Komiya et al., 2009). Therefore, unlike in Arabidopsis, flowering in rice requires active florigen (Tsuji et al., 2011). Induced by HD1, HD3A promotes flowering under short days, analogous to CO-induced FT promoting flowering under long days in Arabidopsis (Kojima et al., 2002; Izawa, 2007a; Andrés and Coupland, 2012). Flowing under short photoperiods correlates with HD3A expression patterns, and early-flowering rice varieties share a high level of HD3A expression (Takahashi et al., 2009; Tsuji et al., 2011; Andrés and Coupland, 2012). Under short-day conditions, HD3A expression begins in the hours approaching dawn, peaks at or around dawn, and remains at high levels until dusk (Kojima et al., 2002; Izawa, 2007a). Arabidopsis FT expression, on the other hand, occurs only at dusk under long-day conditions (Izawa, 2007a). HD3A expression is suppressed under long-day conditions due to HD1’s second, repressive function (Hayama et al., 2003).

Rice contains a second FT ortholog, **RICE FLOWERING LOCUS T1 (RFT1)**, which induces flowering under long days (Figure 1.3) (Komiya et al., 2009; Pin and Nilsson, 2012). This second florigen encodes a B-type response regulator with no Arabidopsis counterpart (Doi et al., 2004; Tsuji et al., 2011). Like HD3A, RFT1 peaks at dawn under its inductive photoperiod (Komiya et al., 2009). Both RFT1 and HD3A are generated in the leaf under long and short days (respectively) and then move to the meristem (Tamaki et al., 2007; Komiya et al., 2008). After arriving in the shoot apical meristem, both RFT1 and HD3A independently form complexes with 14-3-3 proteins that interact with OsFD1, rice’s FD counterpart (Taoka et al., 2011; Pin and Nilsson, 2012). Once in the nucleus, the RFT1/14-3-3/FD complex or HD3A/14-3-3/FD complex promotes transcription of the AP1 ortholog OsMADS15 (Taoka et al., 2011; Pin and Nilsson, 2012).

**EARLY HEADING DATE 1 (EHD1) and Grain number, plant height, heading date** **(GHD7)**

The B-type response regulator EARLY HEADING DATE 1 (EHD1), which lacks an Arabidopsis ortholog, incites both HD3A and RFT1 expression and flowering under
short-day conditions (Figure 1.2) (Doi et al., 2004). As it requires no HD1 to function, EHD1 is an HD1-independent member of the signaling cascade that induces flowering (Doi et al., 2004). The coincidence of light with circadian oscillations in EHD1 expression initiates HD3A transcription, thereby contributing to flowering promotion under short photoperiods (Doi et al., 2004; Itoh et al., 2010; Osugi et al., 2011). Additionally, the circadian clock component OsGI (Oryza sativa GIGANTEA) upregulates EHD1 expression in a mechanism thought to involve blue-light signaling at the start of the light period that coincides with particular circadian phases (Itoh et al., 2010; Tsuji et al., 2011). During long photoperiods, the protein GHD7 contributes to flowering repression by inhibiting transcription of EHD1 and, consequently, HD3A (Xue et al., 2008; Itoh et al., 2010; Osugi et al., 2011; Tsuji et al., 2011).

Photoperiodic Flowering and Domestication

Rice is a staple crop that feeds billions of people throughout the world (Khush, 1997). As a domesticated plant, rice has undergone artificial selection by humans (Izawa, 2007b). Humans selected varieties adapted to particular geographic ranges, which greatly influenced the distribution of rice flowering-time genes and their alleles (Wu et al., 2013). Cultivated rice enjoys a much broader growing range than its progenitor, Oryza rufipogon (Khush, 1997; Wu et al., 2013). Oryza rufipogon, also called wild rice, grows mainly at tropical latitudes in Indochina and is not grown above 28 degrees N latitude, its northern limit (Izawa, 2007b). Cultivated rice, on the other hand, boasts a northern limit of 45 degrees N (Izawa, 2007b). Domestication, via human artificial selection, influenced this northward spread (Izawa, 2007b).

Rice adaptation to particular latitudes depends in part on flowering-time genes. Cultivars in the most northern regions (which feature long photoperiods) exhibit reduced photosensitivity and flower earlier, which permits harvest before the sterility-inducing cold of winter (Roux et al., 2006; Izawa, 2007b; Komiya et al., 2009). De-repression of HD3A in some rice plants above 40 degrees N contributes to this early flowering (Izawa, 2007b). Additionally, RFT1’s function as a florigen under long-day conditions plays a role in adaptation to northern latitudes (Komiya et al., 2009). For plants in temperate latitudes above 31 but below 40 degrees N latitude, the presence of EHD1 contributes to a strong photoperiodic response and flowering under long days
(Izawa, 2007b). Rice located at latitudes below 31 degrees flowers in response to short days due to the *HD1* and *EHD1* pathways (Izawa, 2007b). Finally, tropical regions feature shorter photoperiods and longer warm seasons (Izawa, 2007b). Rice cultivars containing loss-of-function *HD1* and *EHD1* alleles in such regions show delayed flowering due to their insensitivity to photoperiod, thus allowing a longer vegetative period to support increased seed production (Izawa, 2007b).

1.1.4 PHOTOPERIODIC FLOWERING IN SOYBEAN

Domesticated soybean (*Glycine max*) is the foremost plant source of proteins and oils in the world (Wong et al., 2013). A facultative short-day plant, soybean possesses a polyploid genome and orthologs for the majority of Arabidopsis flowering genes, many of which exist in multiple copies in this legume (Cannon and Shoemaker, 2012; Jung et al., 2012; Watanabe et al., 2012).

**E loci**

Nine loci have been implicated in soybean maturity and flowering time: the *E* loci (maturity loci) 1-9 (Bernard, 1971; Buzzell, 1971; McBlain and Bernard, 1987; Ray et al., 1995; Bonato and Vello, 1999; Cober and Voldeng, 2001a; Cober et al., 2010; Watanabe et al., 2011; Watanabe et al., 2012). In most cases, dominant alleles at these loci retard flowering, while recessive alleles stimulate flowering under both short-day and long-day conditions (Upadhyay et al., 1994; Cober and Voldeng, 2001b; Thakare et al., 2010; Watanabe et al., 2012). In addition to affecting flowering time, these loci impact yield, morphology and stress tolerance (Watanabe et al., 2012). The genes underlying some of these loci have been characterized (Kim et al., 2012a; Watanabe et al., 2012).

Of all the *E* loci, the *E1* locus impacts flowering most strongly (Bernard, 1971; Upadhyay et al., 1994; Stewart et al., 2003; Thakare et al., 2010; Watanabe et al., 2012). Under long days, it promotes flowering if recessive (*e1*) and delays flowering if dominant (*E1*) (Thakare et al., 2010). The gene encoded by this locus remains unknown (Watanabe et al., 2011). The *E2* locus encodes *GmGIa*, a gene orthologous to *GI* in Arabidopsis (Bernard, 1971; Watanabe et al., 2011). Plants homozygous for the recessive allele, *e2*, display early flowering (Watanabe et al., 2011). *E3* encodes a soybean phyA ortholog, *GmphyA3* (Watanabe et al., 2009). Dominant *E3* confers later maturity and
sensitivity to fluorescent light, while recessive $e_3$ confers earlier maturity and insensitivity to fluorescent light (Buzzell, 1971). The $E_4$ locus encodes another soybean phyA ortholog, $GmphyA2$ (Liu et al., 2008a; Watanabe et al., 2012). In Arabidopsis, phyA senses red and far-red light and functions in numerous developmental processes (Casal et al., 1997; Watanabe et al., 2012). In soybean, $GmphyA2$ (encoded by $E_4$) is more responsive to a low red:far-red ratio, while $GmphyA3$ (encoded by $E_3$) is more responsive to a high red:far-red ratio (Watanabe et al., 2009). The observation that $E_3$ and $E_4$ show different responses under long-day conditions depending on the red:far-red ratio indicates their possible roles in phyA’s induction of flowering in soybean (Cober et al., 1996; Cober et al., 2001; Watanabe et al., 2012).

**FT Homologs**

Soybean contains eleven $FT$ homologs implicated in flowering and grouped into five sets of gene pairs (Kong et al., 2010; Watanabe et al., 2012). Two of these homologs, $GmFT2a$ and $GmFT5a$, play a prominent role in flowering initiation. Their expression fluctuates in a circadian fashion and reaches highest levels under flowering-inductive short days but remains low under non-inductive long days, indicating a response to environment consistent with photoperiod-regulated flowering (Kong et al., 2010; Thakare et al., 2010). Expression of both $GmFT2a$ and $GmFT5a$ peaks four hours after dawn (Kong et al., 2010; Watanabe et al., 2012). In transgenic soybean, ectopically expressed $GmFT2a$ and $GmFT5a$ accelerate flowering under long-day conditions and elevate the expression of meristem identity genes such as $GmAP1s$ ($a$, $b$, $c$), $GmLFY2$, $GmSOC1$ and $GmSOC2$ (Nan et al., 2014). Both $GmFT2a$ and $GmFT5a$ interact with $GmFDL19$, a soybean FD ortholog (Nan et al., 2014).

**Meristem identity gene orthologs**

Soybean carries orthologs for a number of the meristem identity genes present in Arabidopsis. Expression of the soybean $AP1$ orthologs $GmAP1a$, $GmAP1b$ and $GmAP1c$ in the shoot apex and in flowers during reproductive growth suggests their conserved role in controlling floral meristem identity (Chi et al., 2011; Nan et al., 2014). Furthermore, overexpression of $GmAP1$ in tobacco accelerates flowering (Chi et al., 2011). Additionally, soybean possesses two $LFY$ orthologs, $LFY1$ and $LFY2$ (Meng et al.,
GmLFY1 is expressed in pods and seeds but not in the shoot apex, which may indicate the gene’s role in seed development (Meng et al., 2007; Nan et al., 2014). GmLFY2, on the other hand, is expressed in the shoot apex in response to flowering-inductive signals (Nan et al., 2014). Finally, soybean contains two SOC1 orthologs, GmSOC1 (also called GmGAL1 or GmSOC1a) and GmSOC1-like (also called GmSOC1b) (Zhong et al., 2012; Na et al., 2013; Nan et al., 2014). Expression of both GmSOC1 and GmSOC1-like oscillates in a circadian pattern, GmSOC1 overexpression promotes flowering in Arabidopsis, and GmSOC1-like overexpression promotes flowering in Lotus corniculatus (Zhong et al., 2012; Na et al., 2013). These results suggest that both GmSOC1 and GmSOC1-like function in flowering induction.

Other orthologs

Soybean possesses several other genes with counterparts in Arabidopsis. Dt1 is the soybean ortholog of Arabidopsis TFL1 (Liu et al., 2010; Tian et al., 2010; Watanabe et al., 2012). When expressed in the early-flowering Arabidopsis tfl1 mutant, Dt1 reestabishes the wild-type phenotype (Tian et al., 2010). The dominant form of the gene is associated with indeterminate growth, while the recessive allele is associated with determinate growth (Bernard, 1971; Cober and Morrison, 2010; Tian et al., 2010). Determinate soybean generally culminates in a shorter height and matures earlier than indeterminate soybean, and the geographic distribution of dominant and recessive Dt1 alleles may function in adaptation to particular photoperiodic environments (Bernard, 1972; Curtis et al., 2000; Tian et al., 2010). Additionally, several orthologs of the Arabidopsis circadian clock-associated genes have been identified in soybean (Liu et al., 2009; Kim et al., 2012a). Expressed in a circadian fashion, GmLCL2 (a CCAI ortholog) and GmTOC1 show conserved function between soybean and Arabidopsis (Liu et al., 2009). Soybean also contains orthologs of the clock-associated genes CDF1 and FKF1 (Kim et al., 2012a).

Photoperiodic Flowering and Domestication

Derived from its wild relative Glycine soja, soybean was domesticated 6,000-9,000 years ago (Guo et al., 2010; Kim et al., 2010). Glycine max underwent a genetic bottleneck during this domestication process, reducing the plant’s genetic diversity by
50% compared to wild soybean and depleting 81% of its rare alleles (Hyten et al., 2006; Guo et al., 2010; Kim et al., 2010; Tian et al., 2010; Kim et al., 2012a). Soybean inhabits an array of latitudes, but individual varieties’ differing sensitivities to photoperiod confine them to strict latitudinal boundaries (Cober and Morrison, 2010; Wong et al., 2013). This large diversity in preferred latitude stems from variation in flowering-associated genes and Quantitative Trait Loci (QTL) (Watanabe et al., 2012; Wong et al., 2013).

The longer photoperiods characteristic of high latitudes are suited to soybean varieties less sensitive or insensitive to photoperiod because they flower relatively early under such conditions (Upadhyay et al., 1994). Beneficial at high latitudes, this flowering in the presence of long days allows reproductive development to occur before cold temperatures arrive; plants less responsive to photoperiod avoid the relatively early winter temperatures and instead flower early in the summer when days are long (Kong et al., 2010; Liu et al., 2010; Wilczek et al., 2010; Kim et al., 2012b). Varieties with greater photoperiod sensitivity are better adapted to regions near the tropics because flowering occurs later during summer when days are short, thereby ensuring a longer period for soybean to accumulate sufficient biomass to support bountiful seed production (Roberts et al., 1996; Cooper, 2003). In the United States, soybean cultivars comprise 10 maturity groups based on preferred latitudinal zones of adaptation (Zhang and State, 2007). The main factor accounting for this grouping is variation in photoperiodic response among the cultivars (Zhang and State, 2007).

A number of strategies have improved soybean adaptation to particular photoperiods and, consequently, have augmented yield. Domestication of soybean played a key role in selecting for plants sensitive to particular regional photoperiods (Kim et al., 2012b). More recently, breeders have exploited natural variation in photoperiod sensitivity to produce higher-yielding cultivars (Upadhyay et al., 1994). Traditional breeding strategies have identified and introduced varieties well suited to certain geographic areas. For instance, the first soybean cultivars brought from the southern United States to Australia generally flowered prematurely in response to the shorter days of this tropical region (Lawn and James, 2011). The shortened vegetative period resulted in reduced plant size and biomass and therefore reduced yield potential (Lawn and James, 2011). Subsequent efforts to identify and introduce cultivars
flowering later in response to shorter days successfully improved soybean yield throughout Australia (Lawn and James, 2011).

Another tactic to improve soybean adaptability and yield involved the long-juvenile trait, which confers an extended vegetative phase at lower latitudes (Carpentieri-Pípolo et al., 2002). Controlled by one or two recessive genes, this trait allows soybean to flower late under short days, which normally induce flowering (Ray et al., 1995; Carpentieri-Pípolo et al., 2002; Kong et al., 2010). By delaying flowering under short days, the long-juvenile trait circumvents the yield-impairing problem of short photoperiods abbreviating the vegetative period (Sinclair and Hinson, 1992; Kong et al., 2010; Lawn and James, 2011). This trait therefore permits growth over more latitudes, particularly lower latitudes, and supports more flexible sowing dates, greater seed production and higher yield (Herbert and Litchfield, 1982; Sinclair and Hinson, 1992; Carpentieri-Pípolo et al., 2002; Kantolic and Slaf er, 2007; Kim et al., 2012b). The simple genetic control of the long-juvenile trait lent itself to a straightforward breeding process (Tomkins and Shipe, 1996). Breeders have exploited this trait to extend the geographic range of soybean. Prior to the 1970s, soybean grew only at latitudes above 22 degrees because early flowering and reduced yield occurred at lower latitudes (Carpentieri-Pípolo et al., 2002). In the 1970s, however, breeders began growing soybean harboring the long-juvenile trait (Carpentieri-Pípolo et al., 2002). The introduction of this gene expanded the crop’s geographic range to less than 15 degrees latitude (Carpentieri-Pípolo et al., 2002). Further identification of flowering-related genes and QTL using molecular genetics and genomics approaches could promote increased or more robust yield in particular geographic regions by fine-tuning soybean’s environmentally responsive developmental stages (Watanabe et al., 2012).

1.1.5 PHOTOPERIODIC FLOWERING IN SUGAR BEET

Cultivated sugar beet (Beta vulgaris ssp. vulgaris) is a biennial long-day crop with a vernalization requirement (Pin et al., 2010; Pin and Nilsson, 2012). Vernalization occurs when several weeks of cold exposure in winter promote flowering in spring (Michaels and Amasino, 1999; Johanson et al., 2000; Werner et al., 2005; Andrés and Coupland, 2012). Cultivated sugar beet must undergo vernalization and attain a size
large enough for vernalization responsiveness, which allows this crop to avoid the small yields associated with early flowering (Pin et al., 2010; Andrés and Coupland, 2012).

The sugar beet genome contains two *FT* orthologs regulated by the circadian clock: *BvFT1* and *BvFT2* (Pin et al., 2010; Andrés and Coupland, 2012). An antagonistic pair, they share high sequence similarity but function in an opposite manner (Pin et al., 2010; Pin and Nilsson, 2012). *BvFT2* promotes flowering, while *BvFT1* represses flowering (Pin et al., 2010; Pin and Nilsson, 2012). Long-day conditions upregulate *BvFT2* expression, and short-day conditions upregulate *BvFT1* expression (Pin et al., 2010; Pin and Nilsson, 2012). Cold temperatures repress *BvFT1* expression (Pin et al., 2010; Pin and Nilsson, 2012).

During initial growth, cultivated sugar beet produces high levels of *BvFT1* and low levels of *BvFT2* (Pin et al., 2010; Pin and Nilsson, 2012). However, the vernalization- and long day-triggered induction of flowering reduces *BvFT1* to a low level (Pin et al., 2010; Pin and Nilsson, 2012). According to one model, *BvFT1* inhibits flowering by repressing *BvFT2* expression before the onset of vernalization (Pin et al., 2010). Vernalization then reduces *BvFT1* expression and upregulates *BvFT2* expression, which then triggers flowering (Pin et al., 2010; Andrés and Coupland, 2012). Flowering requires exposure to long-day conditions after vernalization, and plants deficient in *BvFT2* never flower under long days (Pin et al., 2010; Pin and Nilsson, 2012). The protein BOLTING TIME CONTROL 1 (*BvBTC1*) regulates *BvFT1* and *BvFT2* expression (Pin et al., 2012). In annual sugar beet accessions, *BvBTC1* represses *BvFT1* expression and promotes *BvFT2* expression, which eliminates the vernalization requirement for flowering (Andrés and Coupland, 2012; Pin et al., 2012). In cultivated biennial varieties, however, recessive *BvBTC1* remains inactive until vernalization occurs, after which the protein accumulates enough to repress *BvFT1* and activate *BvFT2* (Andrés and Coupland, 2012; Pin et al., 2012).

1.1.6 PHOTOPERIODIC FLOWERING IN POPLAR

Poplar (*Populus* spp) is a model perennial species (Andrés and Coupland, 2012). As opposed to annual species, which flower once and then die, perennial species undergo repeated cycles of vegetative and reproductive growth (Hsu et al., 2011; Andrés and Coupland, 2012). This recurrent flowering allows them to survive for multiple years.
(Hsu et al., 2011; Andrés and Coupland, 2012). The poplar shoot apical meristem alternates between vegetative and reproductive growth phases each year in a season-dependent manner (Hsu et al., 2011). Poplar’s vegetative cycle occurs during the summer, when warmer temperatures and long days abound (Hsu et al., 2011). During the autumn, short days and reduced temperatures initiate bud set (Bohlenius, 2006; Wigge, 2011). Setting buds in autumn ensures the survival of buds amidst the cold and frost of winter, and buds then open the following spring (Bohlenius, 2006; Wigge, 2011).

Poplar relies on two FT orthologs to coordinate its recurrent, seasonal flowering cycle (Hsu et al., 2011). Low temperatures of winter induce expression of the FT homolog PtFT1, which initiates reproductive growth during the spring (Hsu et al., 2011; Andrés and Coupland, 2012). High temperatures, on the other hand, repress PtFT1 (Hsu et al., 2011). Upregulation of the second FT homolog, PtFT2, occurs under high temperatures and long days during the spring and summer (Hsu et al., 2011; Andrés and Coupland, 2012). This upregulation supports vegetative growth and inhibition of bud set during the autumn (Hsu et al., 2011). Low temperatures suppress PtFT2 expression, and short photoperiods reduce both PtFT1 and PtFT2 expression (Hsu et al., 2011). In response to light under long photoperiods, both PtFT1 and PtFT2 mRNA rises with increasing poplar CO (PtCO2), which parallels the CO-FT interaction in Arabidopsis and which may signify functional conservation (Bohlenius, 2006; Andrés and Coupland, 2012). Additionally, poplar and many other perennial species have both LFY and AP1 orthologs whose expression coincides with the appearance of flower buds (Rottmann et al., 2000; Benlloch et al., 2007).

1.2 OTHER FLOWERING PATHWAYS

1.2.1 THE VERNALIZATION FLOWERING PATHWAY

Vernalization occurs when several weeks of cold exposure promote subsequent flowering (Michaels and Amasino, 1999; Johanson et al., 2000; Werner et al., 2005; Andrés and Coupland, 2012). Plants with a vernalization requirement, called winter annuals, flower only after exposure to low temperatures (about 4°C) for six to twelve weeks (Andrés and Coupland, 2012). Mechanistically, this pathway entails sustained
cold temperatures triggering the relief of agents that repress flowering, thus facilitating flowering in the spring (Figure 1.4) (Michaels et al., 2003; Werner et al., 2005; Andrés and Coupland, 2012). Without this pathway’s requirement for prolonged cold temperatures, transient temperature fluctuations in the autumn could otherwise induce flowering in the winter (Amasino, 2010). By preventing the formation of flower buds in cold temperatures, vernalization ensures that no flower buds endure the cold conditions of winter (Wigge, 2011). Vernalization cold-length requirements vary among species and sometimes even among accessions within a species (Johanson et al., 2000; Amasino, 2010). Natural variation in cold-length requirement plays a role in regional adaptation (Johanson et al., 2000). Additionally, plant breeders have introduced a vernalization requirement into a number of crops in order to create winter and spring varieties (Johanson et al., 2000).

Unlike Arabidopsis summer annuals, which predominate in warm climates lacking vernalization-promoting cold temperatures, Arabidopsis winter annuals possess active, dominant forms of the genes FLOWERING LOCUS C (FLC) and FRIGIDA (FRI) (Michaels and Amasino, 1999; Michaels and Amasino, 2001; Michaels et al., 2003; Werner et al., 2005; Andrés and Coupland, 2012). At the molecular level, protracted cold temperatures downregulate FLC mRNA (Andrés and Coupland, 2012). When not vernalized, FLC prevents the expression of flowering genes by binding to them in the shoot apical meristem, and higher levels of FLC coincide with later flowering (Michaels and Amasino, 1999; Searle et al., 2006; Andrés and Coupland, 2012). FLC reduces FT expression by binding to the gene’s first intron and reduces FD and SOC1 expression by binding to their promoters (Helliwell et al., 2006; Searle et al., 2006). FLC’s binding to these genes represents an intersection between the vernalization and photoperiodic pathways (Searle et al., 2006; Andrés and Coupland, 2012). The decrease in FLC mRNA brought about by vernalization releases the transcriptional block on the flowering genes (Michaels and Amasino, 1999; Searle et al., 2006; Andrés and Coupland, 2012).

Vernalization downregulates FLC in part by preventing upregulation of FLC by FRI (Michaels and Amasino, 1999). FRI, the other protein associated with winter annuals, enhances FLC transcription, perhaps by modifying its chromatin structure (Johanson et al., 2000; Andrés and Coupland, 2012). Arabidopsis accessions without a vernalization requirement (i.e. summer annuals) carry loss-of-function FRI alleles
containing deletions (Johanson et al., 2000). According to one analysis, these loss-of-function alleles arose independently on at least two occasions (Johanson et al., 2000). These independent events highlight strong selection for the loss-of-function trait in certain environments (Johanson et al., 2000). In terms of geographic distribution, late-flowering ecotypes abound in northern latitudes (Johanson et al., 2000). In eastern and central Europe, however, early-flowering ecotypes—with no vernalization requirement—predominate (Johanson et al., 2000). A vernalization requirement may be advantageous in areas with harsh winter conditions or in locations with very dry summers; earlier flowering allows plants to avoid these conditions (Mitchell-Olds, 1996; Johanson et al., 2000).

Vernalization also downregulates FLC by the action of antisense non-coding RNA and by chromatin modifications at the histone level (Michaels et al., 2004; Sung and Amasino, 2004; Andrés and Coupland, 2012). After vernalization’s initial reduction of FLC expression, a sustained phase of reduced FLC expression ensues (Gendall et al., 2001; Andrés and Coupland, 2012). A variety of agents accomplish this long-term FLC downregulation that facilitates plants’ sensitivity to photoperiod in the spring (Gendall et al., 2001; Andrés and Coupland, 2012). Transcribed from the 3’ region of FLC, the antisense small RNA COOLAIR reaches peak expression 10 days after the cold temperatures of vernalization begin; this peak coincides with reduced FLC expression (Swiezewski et al., 2009; Heo and Sung, 2011; Andrés and Coupland, 2012). COOLAIR transiently silences FLC epigenetically (Swiezewski et al., 2009). Another antisense RNA, COLDAIR, reaches peak expression 20 days after vernalization begins (Heo and Sung, 2011; Andrés and Coupland, 2012). COLDAIR is transcribed from the first intron of FLC and functions in the maintenance/sustained phase of FLC suppression (Heo and Sung, 2011; Andrés and Coupland, 2012).

Forty days after vernalization begins, when COLDAIR and FLC mRNA reach sufficiently low levels, a mechanism involving histone modifications functions in the sustained suppression of FLC expression (Andrés and Coupland, 2012). Histone modifications mediated by the proteins VERNALIZATION INSENSITIVE 3 (VIN3), VERNALIZATION 2 (VRN2) and VERNALIZATION 1 (VRN1) contribute to this process (Gendall et al., 2001; Sung and Amasino, 2004; Searle et al., 2006; Amasino, 2010; Andrés and Coupland, 2012). Transcription of VIN3 occurs 40 days after vernalization
begins (Heo and Sung, 2011; Andrés and Coupland, 2012). VIN3 dimerizes with VERNALIZATION 5 (VRN5) and appears necessary for FLC chromatin modifications (Greb et al., 2007). VRN5 seems to play a role in the maintenance phase by continuing to interact with FLC chromatin after cold ceases (De Lucia et al., 2008; Amasino, 2010). VERNALIZATION 2 (VRN2) also functions in the maintenance phase; in vrn2 loss-of-function mutants, FLC levels diminish during prolonged cold temperatures but fail to remain low after the cold period ends (Gendall et al., 2001; De Lucia et al., 2008; Amasino, 2010).

1.2.2 THE AMBIENT TEMPERATURE FLOWERING PATHWAY

The ambient temperature pathway highlights the importance of temperature for flowering induction. In Arabidopsis, high temperatures—25°C or 27°C rather than the 23°C often used in experiments—stimulate flowering under short photoperiods almost to the same extent that long photoperiods stimulate flowering under normal or low temperatures (Figure 1.5) (Balasubramanian et al., 2006). The mechanism of this pathway involves FT and histone modifications (Blázquez et al., 2003; Samach and Wigge, 2005; Lee et al., 2007; Wigge, 2011). In response to warm temperatures, wild-type Arabidopsis expresses FT at higher-than-normal levels and flowers early (Blázquez et al., 2003; Balasubramanian et al., 2006; Greenup et al., 2009). Unlike in photoperiodic flowering, however, FT induction in response to warm temperature does not depend on CO, as co mutants show no defect in thermal induction (Balasubramanian et al., 2006). The genes FCA and FVE operate in this pathway to increase FT expression (Blázquez et al., 2003; Greenup et al., 2009). While warm temperatures promote flowering, the cool temperature of 16°C delays flowering, likely due to reduced FT expression (Blázquez et al., 2003; Lee et al., 2007). Evidence suggests that FLC and the protein SHORT VEGETATIVE PHASE (SVP) interact under cool temperatures to form a complex that downregulates FT and SOC1 expression (Li et al., 2008; Greenup et al., 2009). svp loss-of-function mutants show heightened FT expression and flower early under both 16°C and 23°C, indicating an inability to detect low temperature (Lee et al., 2007; Li et al., 2008; Greenup et al., 2009).
1.2.3 THE GIBBERELLIN FLOWERING PATHWAY

Gibberellins are versatile hormones that contribute to the growth-related processes of cell division and elongation, developmental stage changes, and floral organ development (Mutasa-Göttgens and Hedden, 2009). In Arabidopsis, gibberellins function in the induction of flowering and bolting, a hallmark of reproductive development characterized by stem elongation (Lang, 1957; Achard et al., 2004; Mutasa-Göttgens and Hedden, 2009). Bolting in Arabidopsis requires the presence of gibberellins, and mutants either lacking gibberellins receptors or expressing deficient gibberellin levels appear dwarfed in all photoperiods (Koornneef and van der Veen, 1980; Griffiths et al., 2006; Mutasa-Görtgens and Hedden, 2009). Gibberellins act at the shoot apex to facilitate the juvenile-to-adult transition by modulating gene expression (Figure 1.6) (Achard et al., 2004; Mutasa-Göttgens and Hedden, 2009). In some species, gibberellins initiate floral induction under non-inductive environmental conditions (Evans, 1964; Wilson et al., 1992; Mutasa-Göttgens and Hedden, 2009). In Arabidopsis, for example, these hormones are essential for the flowering transition under short days because of the plant’s unresponsiveness to photoperiodic inputs and low FT levels under such non-inductive conditions (Wilson et al., 1992; Wigge et al., 2005; Mutasa-Göttgens and Hedden, 2009). Under long-day conditions, however, gibberellins play a less prominent, although still important, role in flowering induction; plants lacking gibberellins receptors flower later even under inductive conditions (Griffiths et al., 2006).

Binding of gibberellins to their receptor, GIBBERELIN INSENSITIVE DWARF 1 (GID1), mediates their functions (Mutasa-Göttgens and Hedden, 2009). Upon binding, a structural change in GID1 takes place, leading to a physical interaction between GID1 and DELLA transcriptional regulators (Mutasa-Göttgens and Hedden, 2009). DELLA transcriptional regulators suppress growth by interacting with growth-promoting transcription factors (Feng et al., 2008; Mutasa-Göttgens and Hedden, 2009). After gibberellins bind to their receptor, the GID1-DELLA interaction promotes the degradation of DELLA proteins via ubiquitin tagging to the proteasome (Fu et al., 2002; Murase et al., 2008; Shimada et al., 2008; Mutasa-Göttgens and Hedden, 2009). Acting through this DELLA-disabling pathway, gibberellins promote SOC1 expression under
short-day conditions (Achard et al., 2004; Mutasa-Göttgens and Hedden, 2009). They also promote expression of *LFY* by acting through GAMYB (Mutasa-Göttgens and Hedden, 2009). However, gibberellins-induced downregulation of DELLA proteins accompanies increased levels of the microRNA miR159, which cleaves GAMYB and therefore reduces *LFY* levels (Achard et al., 2004; Mutasa-Göttgens and Hedden, 2009). This reciprocal interaction comprises a homeostatic mechanism for flowering control under short days (Achard et al., 2004; Mutasa-Göttgens and Hedden, 2009).

1.2.4 THE AUTONOMOUS FLOWERING PATHWAYS

In the autonomous pathways, internal changes independent of external input trigger flowering (Amasino, 2010). The first autonomous pathway, present in Arabidopsis and several other species, indirectly promotes flowering by regulating *FLC* expression (Figure 1.7) (Simpson, 2004; Amasino, 2010). *FLC* inhibits flowering by repressing the activity of genes induced by flowering-promoting pathways (Simpson, 2004). This autonomous pathway regulates *FLC* at the level of RNA processing and epigenetic modification (Simpson, 2004). Thought to downregulate *FLC* via RNA-based mechanisms, genes involved in this pathway include *FCA, FY, FPA, LD, FLD, FLK* and *FVE* (Lee et al., 1994; Simpson and Dean, 2002; Simpson et al., 2003; Simpson, 2004). Loss-of-function mutants for these autonomous pathway genes exhibit upregulated *FLC* mRNA and late flowering under both short and long photoperiods (Koornneef et al., 1998; Simpson and Dean, 2002; Ausín et al., 2004; Amasino, 2010). Other members of this pathway include histone-modifying genes and RNA-silencing genes (He et al., 2003; Wang et al., 2007; Niu et al., 2008; Amasino, 2010).

In a second autonomous pathway, microRNA influences the juvenile-to-adult transition (Figure 1.8) (Amasino, 2010). miR156 promotes the juvenile developmental period, and miR156 overexpression protracts this stage (Wu and Poethig, 2006; Chuck et al., 2007; Amasino, 2010). Expression of this microRNA declines over the plant lifespan, and consequently, advancing age increases the probability of flowering; old plants respond more sensitively to day length (Wu and Poethig, 2006; Chuck et al., 2007; Wigge, 2011). Flowering occurs at the intersection of lowered miR156 and elevated mRNA of *SQUAMOSA PROMOTER BINDING-LIKE* (*SPL*) genes (Wu and Poethig, 2006; Wu et al., 2009). miR156 controls levels of *SPL3* at the post-
transcriptional level, and therefore declining miR156 accompanies increased SPL3 expression (Wu and Poethig, 2006; Wu et al., 2009). SPL upregulation facilitates LFY, AP1, FUL and SOC1 expression (Fornara and Coupland, 2009; Wang et al., 2009; Yamaguchi et al., 2009). Overexpression of miR156, on the other hand, delays FUL and SOC1 expression (Fornara and Coupland, 2009). While miR156 promotes the juvenile stage, miR172 facilitates the transition to reproductive development (Amasino, 2010). Its expression rises over the lifespan, and it inhibits the FT-repressing APETALA2-like (AP2-like) repressors (Aukerman and Sakai, 2003; Chuck et al., 2007; Jung et al., 2007; Mathieu et al., 2009; Wu et al., 2009; Amasino, 2010). miR156 and miR172 engage in reciprocal regulation. miR156 suppresses SPL transcription factors, which themselves elevate miR172 expression; as miRNA156 levels diminish over the plant's life cycle, SPL and miRNA172 levels rise (Wu et al., 2009).

1.3 CONCLUSION

Regulation of flowering is critical to both plant reproductive success and agriculture. Broadening knowledge of photoperiodic flowering control supports efforts to develop crops with better adaptation to diverse environments. Traditional breeding approaches have successfully improved crops such as soybean in the past, and further identification of agronomically useful natural variation in flowering-time genes could help to develop more adaptive, productive germplasm by similar strategies. Newer technologies based on molecular biology, genomics and bioinformatics offer great potential to enhance crops more efficiently, combat the effects of climate change on crops, and more closely meet the agricultural demands of the growing world population. Ultimately, knowledge of flowering regulation supports the long-term agronomic goals of increasing crop yield and breeding superior crop varieties highly adaptive to a wide range of environments.
1.4 LITERATURE CITED


Buzzell RI (1971) Inheritance of a soybean flowering response to fluorescent-daylength conditions. Can J Genet Cytol **13**: 703–707


Chen M, Ni M (2006) RFI2, a RING-domain zinc finger protein, negatively regulates CONSTANS expression and photoperiodic flowering. Plant J **46**: 823–33


Richards RA (2006) Physiological traits used in the breeding of new cultivars for water-scarce environments. Agric Water Manag 80: 197–211


Simpson GG, Dijkwel PP, Quesada V, Henderson I, Dean C (2003) FY is an RNA 3’ end-processing factor that interacts with FCA to control the Arabidopsis floral transition. Cell 113: 777–87


1.5 FIGURES

Figure 1.1. The Arabidopsis photoperiodic flowering pathway
In the photoperiodic pathway, FT and FD form a complex in the shoot apical meristem. FT-FD induces expression of meristem identity genes such as *LFY*, *AP1* and *FUL*. TFL1 also forms a complex with FD, which downregulates expression of meristem identity genes such as *LFY* and *AP1*. 
Figure 1.2. The rice photoperiodic flowering pathway: short-day photoperiod
In the rice photoperiodic pathway, short-day photoperiods induce expression of EHD1 and HD1. EHD1 promotes both RFT1 and HD3A expression, while HD1 promotes HD3A expression.
In the rice photoperiodic pathway, long-day photoperiods promote expression of HD1, GHD7 and the flowering inducer RFT1. Under this photoperiod, HD1 inhibits HD3A expression. Additionally, GHD7 inhibits EHD1 expression, resulting in cessation of EHD1’s induction of HD3A. RFT1 expression is upregulated by the long photoperiod.
Figure 1.4. The vernalization flowering pathway
In the vernalization flowering pathway, 6-12 weeks of cold induce VRN1, VRN2, VIN3, COOLAIR and COLDAIR, which downregulate FLC expression. This downregulation releases FLC’s inhibition of FT, SOC1 and FD expression. The cold exposure also disrupts FRI’s upregulation of FLC expression.
Figure 1.5. The ambient temperature flowering pathway
In the ambient temperature flowering pathway, warm temperatures (25°C or 27°C) induce FCA and FVE, which trigger $FT$ expression. Cool temperatures (16°C) induce formation of the FLC-SVP complex, which inhibits $FT$ expression.
**Figure 1.6. The gibberellin flowering pathway**

In the gibberellin flowering pathway, gibberellins act through GAMYB to induce $LFY$ expression. Additionally, gibberellins’ binding to their receptor, GID1, promotes degradation of DELLA proteins, thereby releasing inhibition of $SOC1$ expression.
In the first autonomous pathway, FCA, FY, FPA, LD, FLD, FLK and FVE downregulate FLC, thereby releasing FLC’s inhibition of FT, SOC1 and FD expression.
Figure 1.8. The autonomous pathway II

In the second autonomous pathway, advancing maturity downregulates miR156 and upregulates miR172. miR156 downregulation releases inhibition of SPL gene expression, thereby promoting expression of LFY, AP1, FUL and SOC1 as well as miR172. SPL and advancing age’s induction of miR172 inhibits AP2-like repressors, resulting in released inhibition of FT expression.
CHAPTER 2
ROLES OF FT AND TFL1 IN ARABIDOPSIS PHOTOPERIODIC FLOWERING CONTROL

2.1 BACKGROUND

2.1.1 FT AND TFL1 REGULATION OF DOWNSTREAM FLOWERING GENES

Several interconnected pathways control flowering in Arabidopsis. These pathways form a gene network that responds to environmental input to induce flowering. The photoperiod-regulated flowering pathway relies on day length to control flowering (Andrés and Coupland, 2012). Two particularly important genes in this pathway are *FLOWERING LOCUS T (FT)* and *TERMINAL FLOWER 1 (TFL1)*. *FT* and *TFL1* encode transcriptional regulators with homology to phosphatidylethanolamine-binding proteins (PEBPs) (Hanzawa et al., 2005). Diverse proteins found in bacteria, animals and plants, PEBPs function in signaling pathways involved in growth and differentiation (Hanzawa et al., 2005; Karlgren et al., 2011). The transcriptional regulators encoded by *FT* and *TFL1* share ~60% amino acid sequence identity but function in an opposite manner (Hanzawa et al., 2005; Hanano and Goto, 2011). *FT* promotes the transition to reproductive development and flowering, while *TFL1* represses this transition (Hanzawa et al., 2005; Hanano and Goto, 2011). Their high sequence homology suggests conserved biochemical action, but much remains unknown about the mechanisms by which *FT* and *TFL1* control downstream flowering gene expression (Hanzawa et al., 2005; Hanano and Goto, 2011).

*FT* expression is induced by the circadian clock-controlled transcription factor *CONSTANS* under flowering-inductive long-day conditions (Wigge, 2011; Andrés and Coupland, 2012). *FT* protein moves from the leaves through the phloem until it reaches the shoot apical meristem (Takada and Goto, 2003; Imaizumi and Kay, 2006; Corbesier et al., 2007; Lin et al., 2007; Wigge, 2011). After reaching the meristem, *FT* binds to the transcription factor FLOWERING LOCUS D (FD) to form a complex that regulates meristem identity genes, resulting in stimulation of flowering (Abe et al., 2005; Hanano and Goto, 2011; Andrés and Coupland, 2012). The FT/FD complex induces expression of
SQUAMOSA BINDING LIKE (SPL) transcription factors, which promote the transcription of meristem identity genes such as LEAFY (LFY), APETALA1 (AP1) and FRUITFULL (FUL) (Wang et al., 2009; Yamaguchi et al., 2009; Andrés and Coupland, 2012; Jung et al., 2012). The actions of meristem identity genes reprogram the primordia to produce reproductive structures instead of vegetative ones (Blázquez et al., 2006; Amasino, 2010).

TFL1 promotes inflorescence meristem identity and suppresses the transition to the reproductive phase by acting at the shoot apical meristem (Ratcliffe et al., 1998; Liljegren et al., 1999; Blázquez et al., 2006; Benlloch et al., 2007). At the molecular level, TFL1 operates by preventing meristem identity genes (such as LFY and AP1) from acting at the center of the shoot apex (Ratcliffe et al., 1999). It both inhibits these floral meristem genes’ expression and prevents the meristem from responding to them (Ratcliffe et al., 1999). However, the molecular mechanism by which TFL1 modulates downstream genes remains largely unknown. TFL1 may suppress LFY, AP1 and other meristem identity genes by partnering with FD, with which it weakly interacts (Abe et al., 2005; Hanano and Goto, 2011). When fused to the transcriptional activator VP16 in wild-type Arabidopsis, TFL1 induces the expression of a number of meristem identity genes, including LFY and AP1 (Hanano and Goto, 2011). Therefore, LFY and AP1 are targets of TFL1 (Hanano and Goto, 2011). However, whether TFL1 directly or indirectly controls expression of these genes remains unclear (Hanano and Goto, 2011).

2.1.2 INDUCTION SYSTEMS

A number of systems to induce gene expression have been developed for use in plants. As opposed to constitutive promoters, which express a target gene at high levels throughout development, chemical-inducible systems allow direct control over a particular gene at a particular point in development (Zuo et al., 2000a). Such systems remain inactive until induced by application of a chemical (Zuo et al., 2000a). Induction systems using steroid hormones (such as dexamethasone and estradiol), ethanol, and heat shock have enjoyed widespread use (Aoyama and Chua, 1997; Matsuhara et al., 2000; Zuo et al., 2000a; Borghi, 2010). These systems generally consist of two separate transcription units. The first unit contains a constitutive promoter (such as 35S) for high expression of a chemically responsive transcription factor (Zuo et al., 2000a; Borghi,
2010). The second unit contains several copies of the transcription factor binding site connected to a minimal promoter (typically 35S) for constitutive expression of the target gene (Zuo et al., 2000a; Borghi, 2010). Application of the chemical inducer may occur by a variety of routes, such as by spraying plants with the chemical or by adding the chemical to soil (Borghi, 2010).

In the glucocorticoid induction system used in this project, the transformation vector pGREEN-0229-35S:GR (pGreen) contains the T-DNA region that is incorporated into the plant genome by Agrobacterium tumefaciens (Agrobacteria)-mediated transformation (Hellens et al., 2000; Yu et al., 2004). In the T-DNA transcription unit containing the target gene, the Cauliflower Mosaic Virus 35S promoter and the hormone-binding domain of the rat glucocorticoid receptor flank opposite ends of the target gene cloning site (Yu et al., 2004). The target gene and glucocorticoid receptor domain are transcribed and translated as a chimeric protein. The helper plasmid pSoup carries the replicase gene required for pGreen to replicate in Agrobacteria (Hellens et al., 2000). This induction system uses post-translational induction. After translation, the chimeric protein consisting of the target protein fused to the glucocorticoid receptor domain remains in the cytoplasm because it associates with regulatory proteins such as Hsp90 (Zuo et al., 2000a; Borghi, 2010). Application of the glucocorticoid hormone dexamethasone disrupts this association and causes the fusion protein to dimerize and translocate to the nucleus, where it affects gene expression (Zuo et al., 2000a; Borghi, 2010). Pitfalls of this system include the large size of the glucocorticoid receptor domain, which could impact native protein function; dexamethasone toxicity on plant tissue; and dexamethasone-induced activation of defense-related genes (Zuo et al., 2000a).

In the estradiol induction system used in this project, the transformation vector pER8 contains the T-DNA region that is incorporated into the plant genome by Agrobacterium-mediated transformation (Zuo et al., 2000b). In the T-DNA transcription unit containing the target gene, the 35S minimal promoter precedes the insertion site to incite high levels of target gene expression (Zuo et al., 2000b). Unlike the glucocorticoid induction system, this system uses transcriptional induction, in which application of the hormone estradiol directly activates expression of a target gene.
Estradiol also differs from dexamethasone in that it appears not to activate defense-related genes or cause toxicity to plant tissue (Zuo et al., 2000b).

2.1.3 EXPERIMENTAL HYPOTHESIS AND AIMS

We hypothesized that FT and TFL1 modulate the same genetic pathways because they regulate flowering in an opposite manner but possess high amino acid sequence homolog and both interact with FD. We sought to investigate which downstream genes TFL1 regulates and whether FT and TFL1 directly regulate the same set of genes. In order to answer these questions, we followed two approaches. The first approach used RNA sequencing to examine gene expression in both TFL1 overexpression and loss-of-function plants. The second approach used two induction systems to identify the immediate targets of FT and TFL1.

2.2 MATERIALS AND METHODS

2.2.1 GENE EXPRESSION ANALYSIS OF 3 ARABIDOPSIS GENOTYPES

Plant material and growth conditions

The 3 Arabidopsis thaliana (Arabidopsis) genotypes studied were wild-type Columbia-0, tfl1-1 loss-of-function mutant, and 35S:TFL1. Seeds were sterilized according to the following procedure. Each set of ~50 seeds from each genotype was placed in its own HiBind® RNA Spin Column (Omega Bio-Tek). 500 µL 70% ethanol was added to each spin column. The columns were then placed on a rotator for 10 minutes to promote sterilization of the entire seed surface. Next, columns were centrifuged using an Eppendorf 5415D® for 1 minute at 13000 rpm, emptied of 70% ethanol flow-through, and then centrifuged once more. 500 µL 100% ethanol was added to each column and the columns were placed on the rotator for 5 minutes. Next, columns were centrifuged at 13000 rpm for 1 minute, emptied of 100% ethanol flow-through, and then centrifuged once more.

Sterilized seeds were sown onto Murashige and Skoog (MS) plates. MS plates were prepared according to the following procedure. A solution containing 4.4 g/L MS basal media (Sigma-Aldrich), 16 g/L sucrose, and 8 g/L agar was adjusted to pH 5.8.
with 1 M KOH and then sterilized by autoclave for 20 minutes at 120°C. After cooling for ~1.5 hours, the MS solution was poured into 25 mm petri dishes (Fisher) under a particle-free clean bench. Within 30 minutes, plates solidified. Sterilized seeds were soaked in autoclaved water and pipetted from the spin columns to MS plates. Three MS plates were devoted to each genotype. Plates were randomized in position and placed in a growth chamber under long-day conditions (LD) (16 hours of light) for 2 weeks. One randomly selected whole seedling was sampled from each plate for RNA preparation, resulting in three replicates of each genotype.

**RNA preparation**

RNA was extracted from plant cells using the E.Z.N.A.® Plant RNA Kit (Omega Bio-Tek). The manufacturer’s protocol was followed with minor modifications. First, ~50 mg of freshly harvested plant tissue from each replicate was placed in an individual 2 mL screw-cap microcentrifuge tube (Denville) and exposed to liquid nitrogen for 15 seconds. A 3/16 inch diameter steel ball bearing (VXB) was then added to each tube. The tubes were placed in a chilled rack that rests inside the TissueLyser II (QIAGEN), which pulverized the tissue for 2 minutes at 30 cycles/second. The rack was returned to the -80 freezer for 10 minutes and then placed in the TissueLyser II for a second round of tissue disruption using the same settings.

Next, cells lysis was performed on each sample by the addition of 500 µL RB buffer containing 20 µL β-mercaptoethanol for RNase denaturation. After thorough mixing using a Vortex-Genie® 1 Touch vortex mixer, each sample was transferred to an individual homogenizer column placed in a 2 mL collection tube and then centrifuged at 13000 rpm for 5 minutes. Each lysate was then transferred to an individual 1.5 mL microcentrifuge tube (Fisher). 1 volume 70% ethanol was added to each tube, and the solutions were mixed with a vortex mixer for 20 seconds. The samples were then transferred to HiBind® RNA Mini Columns (which contain an RNA-binding membrane) and centrifuged at 13000 rpm for 1 minute. Flow-through was discarded and 350 µL RNase Wash Buffer 1 was added to each column. Another 30-second centrifugation at 13000 rpm followed.

Next, DNA was digested using the RNase-free DNase Set (QIAGEN) according to the manufacturer’s instructions with minor modifications. A mixture of 70 µL Buffer
RDD and 10 µL DNase I was added to each column. After 15-minute DNase incubation at room temperature, 500 µL RNA Wash Buffer 1 was added to each column. Columns were then centrifuged at 13000 rpm for 30 seconds. The flow-through was discarded and 700 µL RNA Wash Buffer II was added to each column. Columns were centrifuged again at 13000 rpm for 30 seconds. A second RNA Wash Buffer II and centrifugation step followed. Next, the columns were centrifuged at 13000 rpm for 2 minutes to dry them completely and to remove any trace ethanol. Each column was then transferred to a 1.5 mL microcentrifuge tube, and 50 µL RNase-free water was added to each column membrane. The columns were allowed to sit for 5 minutes and were then centrifuged once more at 13000 rpm for 1 minute. To measure the concentration of purified RNA, a 2 µL sample from each tube was analyzed with the NanoDrop® ND-1000 Spectrophotometer (ThermoScientific®). Samples were also run on a 1% agarose gel to check for RNA quality.

**Gel electrophoresis**

Gel electrophoresis generates an electric field across an agarose gel matrix to separate DNA fragments by size. This technique checks for amplified DNA fragments and indicates their length. It also indicates RNA quality. The 1% agarose gel consisted of a mixture of 0.5 g agarose (Denville) and 50 mL 1x Tris-Acetate-EDTA (TAE) solution. This mixture was heated to its melting point (85°C) in the microwave until becoming homogeneous and clear. Next, 1.5 µL ethidium bromide (EtBr) (Sigma-Aldrich) (10 mg/mL) was immediately added to the mixture. EtBr facilitates DNA visualization by intercalating between DNA bases and fluorescing in response to ultraviolet light. This three-part mixture was then poured into a gel mold containing a well comb. After completely hardening, the gel was placed in a Thermo EC 105 Classic™ electrophoresis tank filled with 1x TAE and 3 µL EtBr.

Next, 1 µL bromophenol DNA marker and 5 µL PCR product were added to individual wells in the gel. 3 µL Quick-Load® 2-Log DNA Ladder (New England Biolabs), which contains a set of standards of known length, was added to an empty lane at the end of the gel. A 130 V electric current then spread across the gel, separating DNA bands by size. Finally, the gel was visualized with a Gel Doc™ XR+ (BIO-RAD) in conjunction with the program Image Lab™.
RNA sequencing analysis

RNA samples (5 µg of total RNA) were submitted to the W.M. Keck Center for Comparative and Functional Genomics (Sequencing and Genotyping Division) for Illumina-based RNA Sequencing. Sequenced reads were aligned to the latest release of Arabidopsis cDNA from the TAIR website (www.arabidopsis.org) using Bowtie alignment software (Langmead et al., 2009). In-house Perl and Python scripts normalized mapped reads according to both RPKM (Mortazavi et al., 2008) and log2 fold change and also combined alternative splice variants. The Bioconductor package edgeR identified differentially expressed genes (p<0.05) (Robinson et al., 2010). Database for Annotation, Visualization and Integrated Discovery (DAVID) was used to perform gene ontology (GO) term analysis (Huang et al., 2009). Microsoft Excel® was then used for data exploration.

2.2.2 GENE CLONING: CONSTRUCTION OF GLUCOCORTICOID INDUCTION SYSTEM

Gene amplification and creation of restriction sites by PCR

Full-length cDNA of FT, TFL1 and TFL1:VP16 was obtained from the following templates. FT (528 bp) was cloned from the plasmid dw16, and TFL1 (534 bp) was cloned from the plasmid dw17 (Table 2.1). TFL1:VP16 (853 bp) was cloned from the plasmid dw15 (provided by Koji Goto, Research Institute for Biological Sciences, Okayama Prefecture, Kaga-gun, Okayama, Japan). The polymerase chain reaction (PCR) was used to amplify the full-length cDNA from the source material.

For each of the three cDNA templates, the PCR mixture consisted of 1 µL (approx. 1 µg) DNA template, 1.5 µL dNTPs (2.5 mM), 10 µL 10x Pfx amplification buffer, 1 µL MgSO4 (50 mM), 1.5 µL both forward and reverse primers (10 µM), 33.1 µL autoclaved water, and 0.4 µL Platinum® Pfx DNA polymerase. The high-fidelity Platinum® Pfx DNA polymerase was used to ensure faithful sequence amplification. The following primer combinations were used: yh315 and yh316 to amplify FT from dw16, yh312 and yh332 to amplify TFL1 from dw17, and yh312 and yh333 to amplify TFL1:VP16 from dw15 (Table 2.2). The transformation vector pGreen used later required the presence of
particular restriction sites at the 5’ and 3’ ends of the three amplified genes. The PCR primers used for template amplification created these restriction sites. Using dw16 as a template, the primer yh315 created an XbaI site immediately preceding the start codon of FT, and the primer yh316 created a SacI site immediately following the coding sequence of TFL1. Using dw17 as a template, the primer yh312 created an XbaI site immediately preceding the start codon of TFL1, and the primer yh332 created an XbaI site immediately following the coding sequence of TFL1. Using dw15 as a template, the primer yh312 created an XbaI site immediately preceding the start codon of TFL1, and the primer yh333 created an XbaI site immediately following the coding sequence of TFL1:VP16. In addition to requiring particular restriction sites, the pGreen vector required that each transgene lack a stop codon. Therefore, the reverse primers yh316, yh332 and yh333 eliminated the stop codons in amplified FT, TFL1 and TFL1:VP16, respectively.

PCR reactions in the Eppendorf Mastercycler® Pro thermal cycler adhered to the following conditions. Prior to cycling, the mixture was heated to 94°C for 2 minutes. Following this initialization, the first step of the three-step cycle proceeded at 94°C for 20 seconds to denature DNA strands. Next, the primer annealing step cooled the mixtures to 50°C, several degrees below the average melting points of all primers. The cycle concluded with a 68°C elongation step for 1 minute, in accordance with the guideline of 1 minute of elongation per 1 kilobase amplicon. The three-step cycle repeated 32 times, after which the reaction concluded with a 5-minute elongation step at 68°C followed by an indefinite hold at 4°C.

After completion of PCR, Poly-A extension was performed on the PCR products. Platinum® Pfx leaves no 3’ A-overhangs, which are required by the destination vector for the amplified genes, pCR®2.1 (Invitrogen). Therefore, a subsequent thermal cycling step added these overhangs: each of the three PCR products was mixed with a solution containing 0.2 µL Native Taq polymerase (New England Biolabs), 3 µL 10x PCR buffer-Mg²⁺, 0.5 µL dNTPs (2.5 mM), and 1 µL MgCl (50 mM). These mixtures were incubated in the thermal cycler for 10 minutes at 72°C. After the 10-minute incubation, 5 µL each Poly-A extended PCR product was run on an agarose gel by electrophoresis to verify amplification of each transgene.
Cloning into pCR®2.1

The PCR products containing FT, TFL1 and TFL1:VP16 were cloned into the pCR®2.1 vector using the Original TA Cloning® Kit with ExpressLink™ T4 DNA ligase (Invitrogen). This vector carries genes for resistance to the selectable markers kanamycin and ampicillin. In accordance with the manufacturer’s protocol, each ligation solution consisted of 1 µL PCR product, 2 µL pCR®2.1 vector, 2 µL 5x T4 ligation buffer, 4 µL water, and 1 µL T4 ligase. These mixtures were incubated at room temperature for 1 hour.

Bacterial transformation by heat shock

After incubation, the pCR®2.1 ligation solutions were used for transformation of One Shot® TOP10 Chemically Competent E. coli cells (Invitrogen) to amplify the plasmids to levels adequate for plasmid purification. A 25 µL aliquot of E. coli cells stored at -80°C was added to each 10 µL ligation solution. After brief agitation, the mixtures were placed on ice for 30 minutes. Next, the solutions were moved to an aluminum heating block for a 30-second heat shock at 42°C. After heat shock, the solutions were returned to ice for 1 minute. 250 µL liquid LB Broth (Lennox) (Sigma-Aldrich) was then added to each solution under a Bunsen burner. Next, these cultures were agitated at 200 rpm for 3 hours in a 37°C incubation chamber. Under a particle-free clean bench, 100 µL each incubated culture was spread on individual LB-agar plates (1.5% agar) bearing the antibiotic kanamycin (50 mg/mL). The plates were left overnight in the 37°C incubation chamber. The following day, colonies were sampled, mixed with 1 µL liquid LB in individual 2 mL microcentrifuge tubes (Fisher), agitated at 200 rpm for 3 hours in the 37°C incubation chamber, and then genotyped by colony PCR and gel electrophoresis.

Colony PCR used the following procedure. For individual colonies representing one of the three pCR®2.1-based plasmids, the PCR mixture consisted of 1 µL (approx. 1 µg) colony culture, 1 µL dNTPs (2.5 mM), 2 µL 10x PCR buffer (New England Biolabs), 0.5 µL both forward and reverse primers (10 µM), 14.9 µL autoclaved water, and 0.1 µL Taq polymerase (New England Biolabs). The mutation-prone but user-friendly Taq polymerase was used because colony PCR does not require high sequence fidelity. In
pCR®2.1, the M13 forward priming site precedes the insert site, while the M13 reverse priming site follows the insert site. Therefore, in conjunction with a gene-specific forward or reverse primer, these M13 primer sites can distinguish forward or backwards orientation of the inserted gene. Primers yh315 and yh41 were used to identify colonies with forward FT orientation, yh312 and yh41 were used to identify colonies with forward TFL1 orientation, and yh312 and yh333 were used to identify colonies with forward TFL1:VP16 orientation. The PCR cycle used a 1-minute elongation time and 50°C annealing temperature. Three colonies, each representing one of the transgenes in pCR®2.1, were individually mixed with 2 mL LB and kanamycin (50 mg/mL) in 2 mL microcentrifuge tubes. These cultures incubated overnight at 200 rpm for 3 hours in the 37°C incubation chamber for amplification to levels adequate for plasmid purification.

**Plasmid purification**

The next day, plasmids were purified from the bacterial cultures using the QIAprep® Spin Miniprep Kit (QIAGEN) in accordance with the manufacturer's instructions. First, 1.5 mL samples of each incubated culture were placed in 2 mL microcentrifuge tubes and spun in an Eppendorf 5415D® centrifuge for 3 minutes to separate bacterial cells from the liquid medium. Following centrifugation, each supernatant was discarded and each pellet was resuspended in 250 µL buffer P1, which contains RNase to degrade any RNA present in the sample. Next, 250 µL P2 lysis buffer was added to each tube. The tubes were then inverted several times to promote cell lysis. Next, 350 µL buffer N3 was added to each tube to neutralize the lysis reaction. The samples were then immediately centrifuged at 13000 rpm for 10 minutes to separate plasmid DNA from cell lipids. After centrifugation, the pellet in each tube was discarded, each supernatant was transferred to a QIAprep® spin column that drains into a collection tube, and each spin column and collection tube were centrifuged at 13000 rpm for 1 minute. This centrifugation step facilitated the binding of plasmid DNA to the spin column membranes. The resulting supernatant in each tube was discarded. 500 µL wash buffer PB was added to each spin column. After another 1-minute centrifugation at 13000 rpm, 750 µL PE wash buffer was added to each membrane. Two more rounds of 1-minute, 13000 rpm centrifugation followed. Finally, each spin column was transferred to separate 1.5 mL microcentrifuge tubes, and 50 µL buffer EB was added to each
column membrane. The columns sat for 5 minutes at room temperature as the buffer soaked through the column membranes. A final 1-minute centrifugation transferred plasmid DNA from the columns to the 1.5 mL microcentrifuge tubes. The plasmids were named as follows: pCR®2.1-XbaI:FT:SacI as dw14, pCR®2.1-XbaI:TFL1:XbaI as dw2, and pCR®2.1-XbaI:TFL1:VP16:XbaI as dw3.

**DNA sequencing**

DNA sequencing verifies that the gene of interest within the purified plasmid carries no mutations. Following purification, 2 µL (~200 ng/µL) each plasmid, along with the primers yh40 and yh41, were submitted to the Core Sequencing Facility at the University of Illinois. The sequenced genes were then compared to the known gene sequences using the alignment program CLUSTALW (http://www.genome.jp/tools/clustalw/) and found to lack mutations.

**Digestion by restriction endonucleases**

Digestion by restriction endonucleases was next carried out on dw14, dw2 and dw3. Restriction enzymes were used to excise a DNA fragment from specific sites in each plasmid to facilitate its insertion into corresponding sites in another plasmid. The restriction digestion solutions each consisted of 3 µg plasmid, 5 µL buffer 4 (New England Biolabs), 0.5 µL bovine serum albumin (BSA) (New England Biolabs), 0.5 µL autoclaved water, and 1.5 µL each restriction enzyme (New England Biolabs). The enzymes XbaI and SacI were used on both dw14 and the transformation vector pGreen. XbaI by itself was used on dw2, dw3 and pGreen. After a 4-hour digestion at 37°C, the five solutions were run on a gel in separate wells. The visualized cut fragments were excised from the gel using a scalpel and purified using the QIAprep® Gel Extraction Kit (QIAGEN) in accordance with the manufacturer's instructions. The buffers in this kit extract a DNA fragment suspended in gel and remove enzyme contamination.

**Subcloning into pGREEN0229-35S:GR**

The vector pGREEN-0229-35S:GR (pGreen) was provided by Elliot Meyerowitz, Division of Biology, California Institute of Technology, Pasadena, California. In pGreen, the Cauliflower Mosaic Virus 35S promoter and the hormone-binding domain of the rat
glucocorticoid receptor flank opposite ends of the cloning site (Hellens et al., 2000; Yu et al., 2004). The genes destined for subcloning into pGreen lacked stop codons to ensure fusion of each transgene with the flanking glucocorticoid receptor domain. pGreen confers resistance to the antibiotic kanamycin in bacteria and to the herbicide BASTA in plants.

The Original TA Cloning® Kit with ExpressLink™ T4 DNA ligase (Invitrogen) was used to subclone genes digested from the pCR® 2.1-based plasmids into pGreen. In accordance with the manufacturer's instructions, mixtures containing 1 µL T4 ligase, 6 µL (50 ng) cut insert (purified from the gel), 2 µL buffer and 2 µL (50 ng) cut pGreen were incubated for 2 hours at room temperature. pGreen cut with XbaI was used in the solutions containing cut dw2 and dw3, while pGreen cut with both XbaI and SacI was used in the solution containing cut dw14.

pGreen, now bearing the genes of interest, was then used to transform Subcloning Efficiency™ DH5α™ Competent Cells (Invitrogen) using the bacterial transformation procedures previously described. The following day, several colonies on LB (with Kanamycin) plates were genotyped by PCR and gel electrophoresis. Primer combinations consisted of the 35S forward primer, which binds to a site on pGreen immediately preceding the insert site, and a gene-specific reverse primer to check for the presence and correct orientation of each insert. Primers yh306 and yh316 were used to identify colonies with forward FT orientation in pGreen, yh306 and yh332 were used to identify colonies with forward TFL1 orientation in pGreen, and yh306 and yh333 were used to identify colonies with forward TFL1:VP16 orientation in pGreen. The PCR cycle used a 1-minute elongation time and 50°C annealing temperature. Three colonies, each representing one of the three transgenes subcloned into pGreen, were cultured overnight with LB and kanamycin (50 mg/mL) for amplification to levels adequate for plasmid purification. After plasmid purification the next day, samples were named as follows: pGreen-FT as dw11, pGreen-TFL1 as dw7, and pGreen-TFL1:VP16 as dw5. A sample of each plasmid was submitted for sequencing with primer yh306.
2.2.3 GENE CLONING: CONSTRUCTION OF ESTRADIOL INDUCTION SYSTEM

**Gene amplification by PCR**

Full-length cDNA of *FT*, *TFL1* and *TFL1:VP16* was obtained from the same templates used by the GR induction system cloning procedure. PCR with Platinum® Pfx DNA polymerase was used to amplify the three genes from the source material as previously described, with the following modifications: the primers yh58 and yh33 were used to amplify *FT* (528bp + 110 bp 3’ UTR) from dw16, yh56 and yh18 were used to amplify *TFL1* (534bp + 88 bp 3’ UTR) from dw17, and yh56 and yh371 were used to amplify *TFL1:VP16* (853 bp) from dw15. These primers left stop codons intact. The PCR cycle used a 55°C annealing temperature and 1-minute extension time. After completion of PCR followed by Poly-A extension, 5 µL each PCR product was run on an agarose gel by electrophoresis to verify amplification of each transgene.

**Cloning into pCR®2.1**

The PCR products containing *FT*, *TFL1* and *TFL1:VP16* were cloned into the pCR®2.1 vector using the Original TA Cloning® Kit with ExpressLink™ T4 DNA ligase (Invitrogen) as previously described. pCR®2.1 carries restriction enzyme cut sites compatible with the transformation vector pER8 used in this induction system.

**Bacterial transformation by heat shock**

After incubation, the pCR®2.1 ligation solutions were used for transformation of One Shot®TOP10 Chemically Competent *E. coli* Cells (Invitrogen) for plasmid amplification to levels adequate for plasmid purification. The bacterial transformation procedures previously described were used once again. Colony PCR used the following procedure. For individual colonies representing one of the three plasmids, the PCR mixture consisted of 1 µL (approx. 1 µg) colony culture, 1 µL dNTPs (2.5 mM), 2 µL 10x PCR buffer (New England Biolabs), 0.5 µL both forward and reverse primers (10 µM), 14.9 µL autoclaved water, and .1 µL Taq polymerase (New England Biolabs). In pCR®2.1, the T7 priming site precedes the insert site. Therefore, in conjunction with a gene-specific forward or reverse primer, this site can distinguish forward or backwards orientation of the inserted gene. Primers yh321 and yh33 were used to identify colonies...
with forward *FT* orientation, yh321 and yh18 were used to identify colonies with forward *TFL1* orientation, and yh321 and yh371 were used to identify colonies with forward *TFL1:VP16* orientation. The PCR cycle used a 1-minute elongation time and 55°C annealing temperature. Three colonies, each representing one of the transgenes in pCR®2.1, were individually mixed with 2 mL LB and kanamycin (50 mg/mL) in 2 mL microcentrifuge tubes. These cultures incubated overnight at 200 rpm for 3 hours in the 37°C incubation chamber for amplification to levels adequate for plasmid purification. The following day, plasmid purification was performed on the three cultures as previously described. Samples of purified plasmid were then submitted for DNA sequencing using the primers yh4 and yh41. The plasmids were named as follows: pCR®2.1-*FT* as dw18, pCR®2.1-*TFL1* as dw19, and pCR®2.1-*TFL1:VP16* as dw20.

**Digestion by restriction endonucleases**

Digestion by restriction endonucleases was next carried out on dw18, dw19 and dw20. As before, the restriction digestion mixtures each contained 3 µg plasmid, 5 µL buffer 4 (New England Biolabs), 0.5 µL bovine serum albumin (BSA) (New England Biolabs), 0.5 µL autoclaved water, and 1.5 µL each restriction enzyme (New England Biolabs). The enzymes Xho1 and Spe1 were used on all three plasmids as well as on the transformation vector pER8. After a 4-hour digestion at 37°C, the four solutions were run on a gel in separate wells. The visualized cut fragments were excised from the gel using a scalpel and purified using the QIAprep® Gel Extraction Kit (QIAGEN) in accordance with the manufacturer’s instructions.

**Subcloning into pER8**

The vector pER8 was provided by Nam-Hai Chua, Laboratory of Plant Molecular Biology, The Rockefeller University, New York, New York. In pER8, the 35S minimal promoter precedes the insertion site (Zuo et al., 2000b). As previously mentioned, the genes destined for subcloning into pER8 contained their native stop codons. pER8 confers resistance to the antibiotic spectinomycin in bacteria and to the antibiotic hygromycin in plants.

The Original TA Cloning® Kit with ExpressLink™ T4 DNA ligase (Invitrogen) was used to subclone genes digested from pCR®2.1 into pER8. The procedure previously
described for this process was employed once again. After ligation, pER8 bearing the genes of interest was used to transform One Shot®TOP10 Chemically Competent E. coli Cells (Invitrogen). The following day, several colonies on LB (with 100 mg/mL spectinomycin) plates were genotyped by PCR to confirm the presence of each transgene within each vector. Primer combinations consisted of a vector-specific forward primer and gene-specific reverse primers to check for the presence and correct orientation of each insert. Primers yh411 and yh33 were used to identify colonies containing $FT$ in pER8, yh411 and yh18 were used to identify colonies containing $TFL1$ in pER8, and yh411 and yh371 were used to identify colonies containing $TFL1:VP16$ in pER8. The PCR cycle used a 1-minute elongation time and 60°C annealing temperature. Three colonies, each representing one of the three transgenes subcloned into pER8, were cultured overnight for amplification to levels adequate for plasmid purification. After plasmid purification the next day, the plasmids were named as follows: pER8-$FT$ as dw23, pER8-$TFL1$ as dw26, and pER8-$TFL1:VP16$ as dw24. A sample of each plasmid was submitted for sequencing with its transgene-specific forward and reverse primers: yh33 and yh58 for dw23, yh56 and yh18 for dw26, and yh56 and yh371 for dw24.

2.2.4 PLANT TRANSFORMATION

**Agrobacteria transformation with pSoup and pGreen**

In the glucocorticoid induction system, the transformation vector pGreen contains the T-DNA region that is incorporated into the plant genome by *Agrobacteria tumefaciens* (Agrobacteria)-mediated transformation. The helper plasmid pSoup carries the replicase gene required for pGreen to replicate in Agrobacteria. pSoup also confers resistance to the antibiotic tetracycline. Agrobacteria were first transformed with pSoup using the following procedure. The Agrobacteria strain GV3101, which bears resistance to the antibiotics gentamycin and rifampicin, was used. 50 µL GV3101 cells stored at -80 °C was mixed with 7 µL purified pSoup plasmid in a 2 mL microcentrifuge tube, which was then placed on ice for 30 minutes. The mixture was next exposed to liquid nitrogen for 1 minute and then moved to the 37°C incubation chamber for 3 minutes. 1 mL liquid LB was then added to the tube, and the culture was incubated for 3 hours in a 28°C chamber with agitation at 200 rpm. Next, this culture was spread on an LB plate containing 100 mg/mL rifampicin, 80 mg/mL gentamycin and 5 mg/mL tetracycline.
The plate was incubated in the dark 28°C incubation chamber for 2 days. One colony appearing on the plate was then sampled, mixed with 3 mL liquid LB (with antibiotics), and incubated overnight in the 28°C chamber with 200 rpm agitation. 1.5 mL of this pSoup Agrobacteria culture was added to a 50 mL screw-cap tube containing 35 mL liquid LB (with antibiotics), which was then incubated with agitation for 6 hours at 28°C. The culture was then centrifuged at 3000 rpm for 10 minutes at 28°C in an Eppendorf 5810R® centrifuge. The supernatant was discarded and the pellet was resuspended in 2 mL 50 mM calcium chloride (CaCl₂), centrifuged again, resuspended in 2 mL 50 mM CaCl₂, and then centrifuged a final time (CaCl₂ facilitated later plasmid uptake by pSoup Agrobacteria). Cells were then resuspended in 1 mL 50 mM CaCl₂. 50 µL aliquots of this Agrobacteria solution were dispensed into pre-chilled 1.5 mL microcentrifuge tubes. These tubes of competent pSoup Agrobacteria were exposed to liquid nitrogen for 15 seconds and then placed in -80°C storage.

dw11, dw7 and dw5 were used to transform pSoup Agrobacteria using the procedure previously described. Colony survival on plates containing the antibiotics gentamycin (80 mg/mL), rifampicin (100 mg/mL), tetracycline (5 mg/mL) and kanamycin (50 mg/mL) indicated successful transformation. Three colonies, each representing one of the three plasmids, were sampled, mixed with 100 µL liquid LB, agitated at 200 rpm for 2 hours in the 28°C chamber, and then genotyped by colony PCR. The PCR conditions and primer combinations previously used for pGreen colony PCR were used once again here. An additional 1 mL LB (with antibiotics) was added to each of the three colony cultures. These cultures were incubated overnight at 200 rpm in the 28°C chamber. The next day, 500 µL each agrobacteria solution was added to 150 mL LB (with antibiotics). These cultures were incubated overnight at 28°C and 200 rpm. The following day they were used for Arabidopsis transformation via floral dip.

**Agrobacteria transformation with pER8**

In the estradiol induction system, the transformation vector pER8 contains the T-DNA region that is incorporated into the plant genome by transformation with Agrobacteria. dw18, dw19 and dw20 were used to transform GV3101 Agrobacteria using the procedure previously described. Colony survival on plates containing the antibiotics gentamycin (80 mg /mL), rifampicin (100 mg/mL) and spectinomycin (100 mg/mL).
indicated successful transformation. Three colonies, each representing one of the three plasmids, were sampled, mixed with 100 μL liquid LB, agitated at 200 rpm for 2 hours in the 28°C chamber, and then genotyped by colony PCR. The PCR conditions and primer combinations previously used for pER8 colony PCR were used once again here. An additional 1 mL LB (with antibiotics) was added to the three colony cultures. These cultures were incubated overnight at 200 rpm in the 28°C chamber. The next day, 500 μL each agrobacteria solution was added to 150 mL LB (with antibiotics). These cultures were incubated overnight at 28°C and 200 rpm. The following day they were used for Arabidopsis transformation via floral dip.

**Arabidopsis transformation via Agrobacterium-mediated floral dip**

Two Arabidopsis genotypes were used for transformation: wild-type Col-0 and *tfl1-1*. Wild-type seeds were sown into 18 5-inch plastic pots and *tfl1-1* seeds were sown into 12 plastic pots. Each pot contained about 20 seeds from either wild-type or *tfl1-1*, and pots were thinned back to 5-10 individuals after seedlings appeared. Wild-type plants were grown under LD in Turner Hall greenhouse room 6E2, where supplementary light established a light period of 16 hours. To retard their early-flowering phenotype, *tfl1-1* plants were grown under short-day conditions (SD) (10 hours of light) in Turner Hall greenhouse room 3E1, where a mechanical blackout curtain established the artificial dark period of 14 hours per day. After bolting, the emerging main stem of each plant was cut at its base to encourage outgrowth of additional flower-producing shoots. Plants were transformed 7 days after cutting of the main stem. For the glucocorticoid induction system, wild-type plants were transformed with Agrobacteria containing dw11, dw7 or dw5; *tfl1-1* plants were transformed with Agrobacteria containing dw7 or dw5. For the estradiol induction system, wild-type plants were transformed with Agrobacteria containing dw18, dw19 or dw20; *tfl1-1* plants were transformed with Agrobacteria containing dw19 and dw20. The floral dip method described below was used for plant transformation (Clough and Bent, 1998).

150 mL each overnight Agrobacteria culture was transferred to 3 50-mL screwcap tubes and then centrifuged at 4000 rpm for 15 minutes at 28°C in the Eppendorf 5810R®. After centrifugation and subsequent disposal of the supernatant in each tube, 50 mL transformation solution—prepared according to Clough and Bent, 1998—was
added to each tube. Next, the tubes were mixed with the vortex mixer to resuspend the pellets. The solution from each tube was poured into individual 25 mm petri dishes (Fisher). Arabidopsis plants were then inverted and dipped for 15 seconds into the petri dishes, with each dish dedicated to three pots. Three pots were dedicated to each of the four transgene constructs. Transformed plants were placed in a dark, humid growth chamber and left overnight. The following day, plants were returned to their original growth conditions. The plants were transformed a second time 7 days later to transform nascent floral buds not present during the first floral dip. After the second transformation, the plants were returned to their original growth conditions, where they remained until producing mature seeds. Brown siliques (seed pods) indicated the appropriate time for seed harvest. Seeds were collected onto sheets of paper, passed through a sieve to remove debris, and placed in coin envelopes kept at room temperature.

2.2.5 ESTABLISHING TRANSGENIC LINES

Plants were screened for positive transformant seed in three stages across three generations: positive transformants were identified in T1, single-insertion lines were identified in T2, and homozygous lines were identified in T3.

Identifying T1 positive transformants

The first step of screening identified positively transformed T1 seed based on seed survival on MS plates (with BASTA or hygromycin). MS plates used for plants transformed with the pGreen constructs contained 10 mg/mL BASTA, while MS plates used for plants transformed with the pER8 constructs contained 10 mg/mL hygromycin. Seeds were sterilized and plated according to the procedures previously described. After approximately 10 days in a growth chamber under LD, transformed seedlings appeared healthy and green, while non-transformed seedlings appeared blanched, small or even dead. For each transgene, 12-20 candidate transformant seedlings were transferred to LC1 soil in individual plastic wells. LC1 soil consists of 1:1:1:1 peat moss:coarse perlite:dolomitic limestone:starter nutrients. After plants began producing main leaves, one leaf from each individual was sampled for genotyping by PCR.
DNA used as a PCR template was extracted from the plant material according to the following procedure. One leaf from each plant was added to its own 2 mL tube containing 50 µL extraction buffer and a 3/16 inch diameter steel ball bearing (VXB). Extraction buffer consisted of 10 mL 1 M TrisHCL, 2.5 mL 0.5 M EDTA, 0.73 g NaCl, 35.5 mL autoclaved water, and 1.25 mL 20% sodium dodecyl sulfate. Tubes containing leaf samples, steel ball bearings and extraction buffer were placed in a TissueLyser II (QIAGEN) for 3-minute agitation at 30 cycles/second and then briefly centrifuged. Next, an additional 150 µL extraction buffer was added to each tube. Tubes were mixed with the vortex mixer and then centrifuged at 13000 rpm for 10 minutes. Each supernatant was transferred to a new tube containing 200 µL isopropanol. After another vortex step, the tubes were placed in a -20°C freezer for 10 minutes and then centrifuged at 13000 rpm for 10 minutes. Each supernatant was discarded and 200 µL 70% ethanol was added to each pellet. The tubes were then mixed with the vortex mixer for several seconds until homogeneity was achieved. Next, tubes were centrifuged at 13000 rpm for 5 minutes and the resulting supernatant was discarded. Tubes were then inverted on a paper towel to dry for 30 minutes. Once dry, the pellets were resuspended in 50 µL autoclaved water, mixed with the vortex mixer for 30 seconds, and centrifuged once more at 13000 rpm for 5 minutes.

Extracted DNA from each plant was then used as a template for PCR. This genotyping step confirmed the presence of a particular transgene in each candidate transformant and checked for contamination by other transgenes. Therefore, primer combinations for each transgene were used on each sampled plant. The following primer combinations were used on plants transformed with the pGreen constructs: yh306 and yh316 to identify the FT transgene, yh306 and yh332 to identify the TFL1 transgene, and yh306 and yh333 to identify the TFL1:VP16 transgene. This PCR cycle used a 1-minute elongation time and 50°C annealing temperature. The following primer combinations were used on plants transformed with the pER8 constructs: yh411 and yh33 to identify the FT transgene, yh411 and yh18 to identify the TFL1 transgene, and yh411 and yh371 to identify the TFL1:VP16 transgene. This PCR cycle used a 1-minute elongation time and 60°C annealing temperature. After PCR verification, plants continued to grow for several weeks in the greenhouse under LD (for wild-type
background) or SD (for tfl1-1 background) until seed production was complete. Seeds from each positive transformant were collected separately.

**Identifying T2 single-insertion lines**

The second screening step analyzed the T2 progeny of the T1 positive transformants. The segregation ratios of the T2 seeds on herbicide- or antibiotic-containing media indicated whether a particular T1 plant harbored a single copy or multiple copies of a transgene. *Agrobacterium*-mediated plant transformations occasionally result in multiple transgene insertions within the plant genome. Single-insertion lines are preferable to multiple-insertion lines because of the additional variability associated with multiple-insertion lines. Lines carrying a single copy of the gene are expected to display a 3:1 segregation ratio of survival:death on the antibiotic medium. T2 seeds from each T1 positive transformant were sterilized and then sown onto MS plates (with BASTA or hygromycin) and placed under LD in a growth chamber. After approximately 10 days, the segregation ratio of seedlings on each plate was analyzed. From each plate with the 3:1 segregation ratio, 10-12 T2 seedlings were transferred to soil and placed in the greenhouse under LD (for wild-type background) or SD (for tfl1-1 background). Several weeks later, their T3 seeds were collected.

**Identifying T3 homozygous lines**

The third and final phase of screening identified homozygous T3 lines, in which both members of a homologous chromosome pair in the diploid Arabidopsis plants contain the transgene. T3 seeds from each T2 line were sown onto separate MS plates containing BASTA or hygromycin. Survival of all seeds from a particular T2 line on the media indicated homozygosity. Six plants from each homozygous plate were transferred to soil and placed in the greenhouse under LD (for wild-type background) or SD (for tfl1-1 background). Several weeks later, their T4 bulked seeds were collected.
2.2.6 TRANSFORMATION OF TRANSGENIC LINES WITH FD

**Gateway® subcloning**

The Gateway® system facilitates rapid, recombination-based transfer of a DNA fragment from an entry vector to an expression vector. It relies on specific att recombination sites instead of restriction enzymes and ligases. The first Gateway® procedure used the Gateway® BP Clonase II™ kit (Invitrogen) to transfer FD from pCR®2.1 to the entry vector pDONR™/Zeo. The second procedure used the Gateway® LR Clonase II™ kit (Invitrogen) to transfer FD from pDONR™/Zeo to the destination expression vector pK2GW7 (Laboratory of Plant Systems Biology, Ghent University, Belgium).

Full-length FD cDNA was first amplified from the plasmid dw39 (pCR®2.1-FD). For amplification from dw39, the PCR mixture consisted of 1 µL (approx. 1 µg) template (dw39), 1.5 µL dNTPs (2.5 mM), 10 µL 10x Pfx amplification buffer, 1 µL MgSO₄ (50 mM), 1.5 µL both forward and reverse primers (10 µL), 33.1 µL autoclaved water, and 0.4 µL Platinum® Pfx DNA polymerase. The primers yh379 and yh380 were used to amplify FD as well as to create attB1 and attB2 sites immediately preceding (yh379) and following (yh380) the coding sequence; the BP reaction required the presence of these sites. The PCR cycle used a 1-minute elongation time and 55°C annealing temperature. After PCR, poly-A extension was performed according to the procedure previously described. Next, FD was cloned into the entry vector pDONR™/Zeo (Invitrogen) using the Gateway® BP reaction. pDONR™/Zeo contains attP1 and attP2 recombination sites compatible with the attB1 and attB2 sites present in the PCR product. This vector carries resistance to the antibiotic zeocin. In accordance with the manufacturer’s protocol, the BP reaction mixture consisted of 4 µL PCR product (FD), 1 µL pDONR™, 2 µL BP Clonase II™ enzyme mix, and 2 µL TE buffer. The reaction was incubated at room temperature for 1 hour. The resulting plasmid, pDONR™/Zeo-FD (dw32), contained attL1 and attL2 recombination sites as a result of the reaction. This plasmid was used to transform Subcloning Efficiency™ DH5α™ Competent Cells for plasmid amplification to levels sufficient for plasmid purification.

After plasmid purification, FD was cloned into the destination expression vector pK2GW7 using the Gateway®LR reaction. pK2Gw7 contains attR1 and attR2 recombination sites compatible with the attL1 and attL2 sites in the entry plasmid.
In accordance with the manufacturer’s protocol, the reaction solution consisted of 4 µL entry vector (dw32), 1 µL destination vector (pK2GW7), 2 µL LR Clonase II™ enzyme mix, and 2 µL TE buffer. The reaction was incubated at room temperature for 1 hour. The resulting plasmid, pK2GW7-FD (dw33), was used to transform Subcloning Efficiency™ DH5α™ Competent Cells (Invitrogen) for plasmid amplification to levels sufficient for plasmid purification.

**Arabidopsis transformation via Agrobacterium-mediated floral dip**

GV3101 Agrobacteria were transformed with dw33 using the methods previously described. Agrobacteria carrying dw33 were then used to introduce FD into three independent pGreen-transformed plant lines in the Col-0 plant background: FT 8-9, TFL1 1-7, and TFL1:VP16 2-8. FD was also introduced into two independent pGreen-transformed plant lines in the tfl1-1 mutant background: TFL1 3-2 and TFL1:VP16 5-4. Agrobacteria carrying dw33 were used to introduce FD into three independent pER8-transformed plant lines in the Col-0 plant background: FT 8-3, TFL1 3-7, and TFL1:VP16 12-10. FD was also introduced into two independent pER8 plant lines in the tfl1-1 mutant background: TFL1 6-5 and TFL1:VP16 5-2.

**2.2.7 INDUCTION**

**GLUCOCORTICOID INDUCTION**

For induction experiments on a particular transgenic line, 30-35 sterilized T3 or T4 seeds were sown onto each of three MS plates containing BASTA (10 mg/mL). Seedlings were grown under LD in a growth chamber for 11 to 18 days. On the morning of induction day, ~20 seedlings from one plate were collected in a 2 mL screw-cap microcentrifuge tube, exposed to liquid nitrogen for 15 seconds, and placed in -80°C storage. Under a particle-free clean bench, each of the remaining two plates was then sprayed with either a mock treatment or a glucocorticoid dexamethasone treatment. The mock treatment consisted of either water or 0.01% (w/v) Tween-20. The dexamethasone treatment consisted of either 5 µM or 30 µM dexamethasone in either water or 0.01% (w/v) Tween-20. After spray treatment, plates were covered with their plastic lids and returned to the growth chamber, where they remained for 7 or 24 hours. Seven or twenty-four hours after induction or mock treatment, ~20 seedlings from each
plate were collected in separate 2 mL screw-cap tubes, exposed to liquid nitrogen for 15 seconds, and placed in -80°C storage.

**ESTRADIOL INDUCTION**

For induction experiments on a particular transgenic line, 30-35 sterilized T3 or T4 seeds were sown onto each of three MS plates containing hygromycin (10 mg/mL). Seedlings were grown under LD in a growth chamber for 12 days. On the morning of day 13, ~20 seedlings from one plate were collected in a 2 mL screw-cap microcentrifuge tube, exposed to liquid nitrogen for 15 seconds, and placed in -80°C storage. Under a particle-free clean bench, each of the remaining two plates was then sprayed with either a mock treatment or an estradiol treatment. The mock treatment consisted of distilled water. The estradiol treatment consisted of 5 µM estradiol. After spray treatment, plates were covered with their plastic lids and returned to the growth chamber, where they remained for 24 hours. 24 hours after induction or mock treatment, ~20 seedlings from each plate were collected in separate 2 mL screw-cap tubes, exposed to liquid nitrogen for 15 seconds, and placed in -80°C storage.

2.2.8 ANALYZING GENE EXPRESSION

**RNA preparation and cDNA synthesis**

After induction, plant tissue samples stored at -80°C were subjected to RNA preparation using the methods previously described. The prepared RNA was then used as a template for cDNA synthesis. cDNA was synthesized using the iScript™ cDNA Synthesis Kit (Bio-Rad). In accordance with the manufacturer’s protocol, each reaction mixture consisted of 1 µg RNA, 4 µL 5x iScript reaction mix, and 1 µL iScript reverse transcriptase, with nuclease-free water added to bring the volume to 20 µL. These mixtures were then incubated in the Mastercycler® Pro under the following conditions: 5 minutes at 25°C, 30 minutes at 42°C, and then 5 minutes at 85°C. To measure the concentration of synthesized cDNA, a 2 µL sample from each reaction product was analyzed with the NanoDrop® ND-1000 Spectrophotometer (ThermoScientific®).
**RT-PCR**

The gene expression of all three samples for a particular induction experiment was measured using the synthesized cDNA as a template for PCR. To check the expression of meristem identity genes potentially controlled by *FT* and *TFL1*, the following primers were used: yh414 and yh415 for *LFY*, yh367 and yh368 for *AP1*, and yh412 and yh 413 for *FUL*. Additionally, for plants transformed with the estradiol induction plasmids, the following primers were used to check the expression of the induced transgenes: yh58 and yh59 for *FT*, yh56 and yh18 for *TFL1*, and yh 56 and yh371 for *TFL1:VP16*. The clathrin primers yh3 and yh4 were used as controls. The PCR cycle used a 1-minute elongation time and 55°C annealing temperature. The resulting PCR products were visualized by gel electrophoresis, and the relative strength of their bands in the gel was analyzed by comparison to PCR products from the primers yh3 and yh4.

**2.3 RESULTS**

**2.3.1 CHANGES IN GENE EXPRESSION PATTERNS UNDER TFL1**

**Number of reads**

The RNA sequencing produced 5.1 to 6.5 million reads per sample, with an average of 5.9 million reads per sample (Table 2.3). Reads were mapped to the Arabidopsis transcriptome (TAIR10 cDNA, updated 12/14/2010), which contained a total of 41,671 transcripts, including 8,123 splice variants. Mapped reads per sample ranged from 5.0 to 6.1 million, with an average of 5.6 million mapped reads (95.4% of total reads) per sample. For further analysis, splicing variants were combined into representative gene models (TAIR10 cDNA representative gene models, updated 1/3/2011).

**Differentially expressed genes**

In total, 1,417 genes were differentially expressed at the p<0.05 level in the 35S:*TFL1* and *tfl1-1* genotypes compared to wild type (Figure 2.1). In 35S:*TFL1*, 281 genes were upregulated and 893 genes were downregulated compared to wild type. Among these genes, 3 known flowering genes were upregulated and 8 were
downregulated. The large number of downregulated genes in 35S:TFL1 suggests that TFL1 represses gene expression. In tfl1-1, 159 genes were upregulated and 475 genes were downregulated compared to wild type. Among these genes, 11 known flowering genes were upregulated and 1 was downregulated. To pinpoint genes potentially induced by TFL1, we identified genes simultaneously upregulated in 35S:TFL1 and downregulated in tfl1-1 (Figure 2.1a). Only 6 genes appeared in this comparison. To pinpoint genes potentially repressed by TFL1, we identified genes simultaneously downregulated in 35S:TFL1 and upregulated in tfl1-1 (Figure 2.1b). Only 4 genes appeared in this comparison. The lack of overlap suggests that this dataset contains secondary gene expression not directly regulated by TFL1; many of the differentially expressed genes may function far downstream of TFL1.

**Differentially expressed flowering genes**

In 35S:TFL1 and tfl1-1, 23 flowering genes were differentially expressed compared to wild type (Figure 2.2, Table 2.4). Most of these genes appeared to be repressed by TFL1, as demonstrated by downregulation of 8 of the 11 differentially expressed flowering genes in 35S:TFL1 but upregulation of only 3 such genes. In 35S:TFL1, TFL1 itself was strongly upregulated (log fold-change 8.70). Among the flowering genes downregulated in 35S:TFL1, LEAFY (LFY) and AGAMOUS-like 14 (AGL14) showed the strongest downregulation, with log fold-change values of -5.58 and -2.88, respectively, followed by AGAMOUS-like 21 (log fold-change -1.97). Previous work has established that LFY functions downstream of FT and that AGL14 and AGL21 function in root development (Yamaguchi et al., 2009; Andrés and Coupland, 2012; Garay-Arroyo et al., 2013). These genes’ downregulation in 35S:TFL1 is consistent with TFL1 acting as a flowering repressor.

Further supporting this trend, 11 of the 12 differentially expressed flowering genes in tfl1-1 were upregulated and only 1 was downregulated. Among the flowering genes upregulated in tfl1-1, APETALA1 (AP1) and APETALA3 (AP3) showed the strongest upregulation, with log fold-change values of 5.54 and 5.46, respectively, followed by SEPALLATA2 (SEP2) (log fold-change 4.96). Previous work has established that AP1 functions downstream of FT and that AP3 and SEP2 function in floral organ formation (Ditta et al., 2004; Blázquez et al., 2006). These genes’ upregulation in tfl1-1
suggests increased flowering gene expression due to the absence of TFL1 in the loss-of-function mutant.

**GO functional analysis**

The 281 genes upregulated in 35S:TFL1, 893 genes downregulated in 35S:TFL1, 159 genes upregulated in tfli-1, and 475 genes downregulated in tfli-1 were next subjected to functional clustering. Functional clustering using DAVID indicated enrichment of gene categories related to transcription and flowering (Table 2.5). Genes upregulated in tfli-1 compared to wild type showed enrichment for the greatest number of transcription- and flowering-related categories, confirming the roles of TFL1 in flowering and inflorescence development. Enrichment of transcription-related categories also appeared in the sets of genes downregulated in 35S:TFL1 and downregulated in tfli-1. Enrichment of defense-related and stress-responsive categories was observed in the sets of genes downregulated in 35S:TFL1 and downregulated in tfli-1, suggesting interaction between the TFL1 pathway and stress-induced flowering pathway. Finally, enrichment of lipid-binding categories appeared in the sets of genes upregulated in 35S:TFL1 and downregulated in 35S:TFL1, consistent with TFL1’s homology to phosphatidylethanolamine-binding proteins.

2.3.2 INDUCTION EXPERIMENT

**Establishing transgenic lines: Glucocorticoid induction system**

*Ti Generation.* Three constructs were used to transform wild-type Col-0: 35S:FT:GR, 35S:TFL1:GR and 35S:TFL1:VP16:GR. Screening of the obtained T1 seeds using the BASTA resistance selectable marker resulted in 36 35S:FT:GR lines, 12 35S:TFL1:GR lines, and 12 35S:TFL1:VP16:GR lines. The existence of each transgene was examined by PCR genotyping, and 26 35S:FT:GR lines, 12 35S:TFL1:GR lines and 12 35S:TFL1:VP16:GR lines were confirmed in this generation (Table 2.6). Two constructs were used to transform tfli-1: 35S:TFL1:GR and 35S:TFL1:VP16:GR. Screening of the T1 seeds for BASTA resistance resulted in 20 35S:TFL1:GR lines and 14 35S:TFL1:VP16:GR lines. PCR genotyping confirmed 15 35S:TFL1:GR lines and 9 35S:TFL1:VP16:GR lines in this generation (Table 2.6). T1 plants were transferred to soil and allowed to self-fertilize.
**T2 Generation.** The T2 seeds were harvested from each T1 plant individually. Of the 26 35S:FT:GR T1 lines in wild type, 17 lines produced sufficient seed for further characterization. BASTA resistance screening showed that 5 lines segregated according to a 3:1 ratio, which suggested that they were single-insertion lines (Table 2.7). Of the 12 35S:TFL1:GR T1 lines in wild type, 9 lines produced sufficient seed for BASTA resistance screening, which showed that 5 lines segregated according to the 3:1 ratio. Of the 12 35S:TFL1:VP16:GR T1 lines in wild type, 5 lines produced sufficient seed for BASTA resistance screening, which showed that 3 lines segregated according to the 3:1 ratio (Table 2.7). Of the 15 35S:TFL1:GR T1 lines in tf1-1, 8 lines produced sufficient seed for BASTA resistance screening, which showed that 5 lines segregated according to the 3:1 ratio. From each of the 5 genotypes, 10-12 BASTA-resistant T2 plants were transferred to soil and allowed to self-fertilize.

**T3 Generation.** The T3 seeds were harvested from each T2 plant individually, and their segregation ratios were assessed for BASTA resistance on MS plates to identify homozygous lines. In wild type, 4 homozygous 35S:FT:GR lines, 3 homozygous 35S:TFL1:GR lines, and 2 homozygous 35S:TFL1:VP16:GR lines were identified (Table 2.8). In tf1-1, 3 homozygous 35S:TFL1:GR lines and 3 homozygous 35S:TFL1:VP16:GR lines were identified (Table 2.8).

**Establishing transgenic lines: Estradiol induction system**

**T1 Generation.** Three constructs were used to transform wild-type Col-0: 35S:FT, 35S:TFL1 and 35S:TFL1:VP16. Screening of the obtained T1 seeds using the hygromycin resistance selectable marker resulted in 23 35S:FT lines, 18 35S:TFL1 lines, and 17 35S:TFL1:VP16 lines. The existence of each transgene was examined by PCR genotyping, and 18 35S:FT lines, 17 35S:TFL1 lines and 12 35S:TFL1:VP16 lines were confirmed in this generation (Table 2.9). Two constructs were used to transform tf1-1: 35S:TFL1 and 35S:TFL1:VP16. Screening of the T1 seeds for hygromycin resistance resulted in 11 35S:TFL1 lines and 15 35S:TFL1:VP16 lines. PCR genotyping confirmed 9 35S:TFL1 lines and 15 35S:TFL1:VP16 lines in this generation (Table 2.9). T1 plants were transferred to soil and allowed to self-fertilize.
T2 Generation. The T2 seeds were harvested from each T1 plant individually. Of the 18 35S:FT T1 lines in wild type, 10 lines produced sufficient seed for further characterization. Hygromycin resistance screening showed that 4 lines segregated according to a 3:1 ratio, which suggested that they were single-insertion lines (Table 2.10). Of the 17 35S:TFL1 T1 lines in wild type, 8 lines produced sufficient seed for hygromycin resistance screening, which showed that 2 lines segregated according to the 3:1 ratio. Of the 12 35S:TFL1:VP16 T1 lines in wild type, 8 lines produced sufficient seed for hygromycin resistance screening, which showed that 5 lines segregated according to the 3:1 ratio. Of the 15 35S:TFL1 T1 lines in tfl1-1, 8 produced sufficient seed for hygromycin resistance screening, which showed that 4 lines segregated according to the 3:1 ratio (Table 2.10). Of the 15 35S:TFL1:VP16 T1 lines in tfl1-1, 9 produced sufficient seed for hygromycin resistance screening, which showed that 5 lines segregated according to the 3:1 ratio. 10-12 Hygromycin-resistant plants from each T2 line were transferred to soil and allowed to self-fertilize.

T3 Generation. The T3 seeds were harvested from each T2 plant individually. Due to a pest infestation in the greenhouse, a large number of T2 plants died. Consequently, fewer than expected T3 seeds were obtained, particularly in wild type. Segregation ratios of the obtained T3 seeds were assessed for hygromycin resistance on MS plates to identify homozygous lines. In wild type, 1 homozygous 35S:FT:GR line, 1 homozygous 35S:TFL1 line, and 2 homozygous 35S:TFL1:VP16 lines were identified (Table 2.11a). In tfl1-1, 3 homozygous 35S:TFL1 lines and 2 homozygous 35S:TFL1:VP16 lines were identified (Table 2.11a). These T3 plants were allowed to self-fertilize, and T4 bulked seed from each line was collected.

Due to the pest infestation that caused the loss of so many transgenic plants in wild type, and the resulting small number of homozygous T3 lines obtained, multiple-insertion T3 individuals from 3 genotypes were allowed to self-fertilize and their T4 bulked seed was collected. The multiple-insertion lines consisted of 2 homozygous 35S:FT lines and 1 homozygous 35S:TFL1 line in wild type as well as 1 homozygous 35S:TFL1 line in tfl1-1 (Table 2.11b).
**Test induction experiments: Glucocorticoid induction system**

As a pilot experiment, the glucocorticoid induction system was tested using 2 genotypes: non-transformed wild-type and a 35S:FT:GR line in the wild-type background. To optimize the induction conditions, different dexamethasone concentrations, solution buffers, plant ages, and growth conditions were examined. The flowering genes LFY, AP1 and FUL were chosen for this pilot experiment because they are known downstream genes of FT.

The first induction was performed on 11-day-old seedlings grown under SD. The induction solution consisted of 5 µM dexamethasone, while the control solution consisted of water. Under these conditions, expression of just one meristem identity gene, FUL, was detected at all time points in both induced (Figure 2.3, lane FT3) and non-induced (lanes FT1, FT2) 35S:FT:GR transgenic plants but not in non-transformed wild type (lanes WT1, WT2, WT3). FUL expression appeared identical across the time points, with no apparent difference between induced and non-induced transgenic plants. The observed FUL expression in both non-induced transgenic plants at time point 0 and time point 7 and in induced transgenic plants at time point 7 suggested that FUL expression occurred independently of dexamethasone treatment, perhaps due to leaky induction of the FT transgene. The positive control for clathrin indicated approximately uniform cDNA concentration across all samples. No visible bands appeared for AP1 and LFY. The absence of AP1 and LFY bands could have resulted from a number of factors. The SD growth condition could have prevented FT from activating downstream genes. Alternatively, the 11-day-old seedlings may have been too immature for FT to activate downstream genes. Finally, the induction solution may have failed to activate FT expression because it lacked a buffer to allow better penetration of dexamethasone into plant tissue.

To clarify whether plant age and induction buffer impacted the results, the second induction experiment was performed 18-day-old seedlings grown under SD. The induction solution consisted of a higher dexamethasone concentration, 30 µM, in 0.01% (w/v) Tween-20 buffer to promote dexamethasone penetration into leaves. The control solution consisted of 0.01% (w/v) Tween-20. Under these conditions, expression of both FUL and AP1 was detected at all time points in both induced (Figure 2.4, lane 4) and non-induced (lanes 2, 3) 35S:FT:GR transgenic plants as well as in young flower buds.
from a wild-type plant approximately 3 weeks old (lane 1) (Figure 2.4). The positive control for clathrin indicated approximately uniform cDNA concentration across all samples, with the weaker clathrin band in lane 4 indicating lower starting cDNA concentration from the dexamethasone-treated plants.

Despite lower cDNA concentration of dexamethasone-treated plants in lane 4, this template’s *FUL* lane 4 band appears as strong as the *FUL* bands in lanes 1-3, which indicates possible induction. No dexamethasone-induced expression was detected in the other meristem identity genes. Unlike in 11-day-old seedlings, *AP1* was expressed in these 18-day-old seedlings. *AP1* expression appeared approximately identical across the time points, with no apparent difference between induced and non-induced transgenic plants (Figure 2.4). This result suggests that the 18-day-old seedlings expressed *LFY* independently of dexamethasone treatment. No *LFY* expression was detected in any sample. The absence of both *AP1* induction and *LFY* expression could have resulted from the SD photoperiod preventing FT from activating downstream genes. In summary, results from the first two experiments suggested that plant age impacted gene expression but not induction and that the Tween-20 buffer solution had no effect on induction.

To clarify whether photoperiodic growth condition impacted the results, the third induction was performed on 11-day-old seedlings grown under flowering-inductive LD. The induction solution consisted of 5 µM dexamethasone, while the control solution consisted of water. Under these conditions, expression of *LFY*, *AP1* and *FUL* was detected at all time points in both induced (Figure 2.5, lane 3) and non-induced (lanes 1, 2) transgenic plants. Expression of the *TIL3* positive control appeared slightly stronger in lane 2, suggesting a higher starting cDNA concentration of non-induced plants sampled at time point 7. Expression of *LFY*, *AP1* and *FUL* appeared similar across the time points, with possible induction appearing in the induced plants (Figure 2.5, lanes 3); however, the clathrin control bands cast some doubt on the PCR accuracy. The expression of *LFY* in transgenic plants grown under LD but not under SD suggests that *LFY* expression requires LD growth condition. This third induction experiment indicated that photoperiod appeared not to influence the failed induction. A possible explanation for the lack of induction under LD is that the dexamethasone solution contained no buffer to promote dexamethasone penetration into plant tissue.
To clarify whether the induction buffer impacted the results under LD, the fourth and final glucocorticoid pilot induction was performed on 13-day-old seedlings grown under LD. The induction solution consisted of 30 µM dexamethasone in 0.01% (w/v) Tween-20 buffer to promote dexamethasone penetration into leaves. The control solution consisted of 0.01% (w/v) Tween-20. Under these conditions, expression of AP1 and FUL was detected at all time points in both induced (Figure 2.6, lane 3) and non-induced (lanes 1, 2) 35S:FT:GR transgenic plants. Expression of these genes appeared stronger in the induced plants sample (lane 3 for both genes); however, the stronger clathrin positive control band in lane 3 suggested a higher starting cDNA concentration of induced plants sampled at time point 7. Therefore, it appears unlikely that induction occurred. No LFY expression was observed in any sample. Because LFY expression was detected under the previous LD induction experiment but not in this one, some error in the PCR procedure may have occurred.

The pilot induction experiments for the glucocorticoid induction system demonstrated leaky induction of the transgene, as evidenced by early expression of the meristem identity genes in the 35S:FT:GR transgenic plants compared to wild type. Additionally, no obvious difference in meristem identity gene expression appeared between induced and non-induced transgenic plants. This result suggested either failed induction of the transgene or indirect control of the transgene over the meristem identity genes. Therefore, we next tested the estradiol induction system, which uses transcriptional rather than post-translational induction. Because this system induces transcription of the transgene (in this case FT), we could simply assay FT expression to determine whether transgene induction was successful.

**Test induction experiments: Estradiol induction system**

The estradiol induction pilot experiment used one genotype, a multiple-insertion 35S:FT line in wild type. The induction solution consisted of 5 µM estradiol, while the control solution consisted of water. Induction was performed on 12-day-old seedlings grown under LD. Twenty-four hours after induction, a portion of the seedlings sprayed with the estradiol treatment and a portion of the seedlings sprayed with the control treatment were collected. The remaining seedlings were sprayed with their respective treatments a second time. Twenty-four hours later, a portion of these seedlings was
collected and the remaining seedlings sprayed again. This cycle repeated two more times. In total, 4 spray treatments were performed. Samples were collected at 5 time points.

*FT* expression was detected in non-induced plants sampled at time point 0 (Figure 2.7, lane 1), in both non-induced and induced plants at time point 24 (lanes 2, 3), and in induced plants at time point 48 (lane 5). *FT* bands were noticeably strongest in lanes 3 and 5, which corresponded to the induced plants. This result indicated successful induction of the transgene. Expression of the clathrin positive control appeared slightly weaker in lane 1, possibly suggesting a lower starting cDNA concentration of non-induced plants sampled at time point 0. *FT* expression was also detected in both induced and non-induced plants sampled at time points 72 and 96 (Figure 2.9). However, this expression again appeared stronger in the induced plants (lanes 2, 4). The non-induced plants sampled at time point 96 (lane 3) seemed to show *FT* expression similar to the induced plants (lanes 2 and 4); however, the stronger clathrin control band for this sample indicated a higher starting cDNA concentration, so its relative *FT* expression was likely lower.

Expression of *FUL* generally appeared similar across the time points, with no clear difference between induced and non-induced 35S:*FT* transgenic plants. However, the *FUL* band corresponding to induced plants sampled 48 hours after induction (lane 5, Figure 2.8) appeared slightly stronger than neighboring bands, which may indicate induction (Figure 2.8, Figure 2.9); additional testing with quantitative reverse-transcription PCR will reveal the accuracy of this observation. The positive control for clathrin indicated approximately uniform cDNA concentration across all samples, with the exception of a slightly stronger band in lane 3 in Figure 2.9 indicating higher starting cDNA concentration from non-induced plants sampled 96 hours after the first induction. This stronger clathrin band corresponds to the stronger *AP1* band in lane 3 in Figure 2.9, suggesting that no induction of this meristem identity gene occurred. Low expression of *AP1* and *LFY* also appeared in both induced and non-induced plants sampled at time points 72 and 96, implying again that the expression of the genes was not caused by induced *FT*.

Both induction systems failed to show clear induction of meristem identity genes. However, the estradiol induction system showed obvious induction of *FT*. This result
suggests that rather than controlling meristem identity genes directly, FT requires some other factor, which appears later in development, in order to modulate downstream gene expression. Because of both FT and TFL1’s known interaction with FD, FD emerged as a strong candidate for this factor. To investigate this possibility, full-length FD cDNA was cloned into the binary vector pK2GW7 (35S:FD), which was used to transform one independent T4 line from each construct in each background from both induction systems. Obtained T1 seeds will be screened for kanamycin and BASTA resistance (glucocorticoid induction system) or kanamycin and hygromycin resistance (estradiol induction system) to identify double transgenics. Induction will then be performed on these plants.

2.4 DISCUSSION

This project sought to determine which downstream genes TFL1 regulates and whether TFL1 and FT directly regulate the same set of genes. We first used RNA sequencing to examine gene expression in TFL1 overexpression and loss-of-function. We then used two induction systems to identify the immediate targets of TFL1 and FT. Previous studies have established that TFL1 and FT regulate the expression of flowering genes. In conjunction with FD, the flowering inducer FT promotes expression of meristem identity genes such as LFY, AP1, FUL, SEPALLATA1 (SEP1) and SEPALLATA3 (SEP3) (Ruiz-García et al., 1997; Abe et al., 2005; Teper-Bammolker and Samach, 2005; Wigge et al., 2005; Hanano and Goto, 2011). The flowering repressor TFL1 inhibits expression of a number of genes induced by FT/FD, including AP1 and LFY (Liljegren et al., 1999; Ratcliffe et al., 1999). TFL1 may suppress LFY, AP1 and other meristem identity genes by partnering with FD, with which it weakly interacts (Abe et al., 2005; Hanano and Goto, 2011). Recent work has shown that TFL1 reduces expression of SEP1 and SEP3 as well (Hanano and Goto, 2011). In agreement with these findings, our RNA sequencing data showed that AP1, SEP1 and SEP3 were upregulated in tfl1-1.

The RNA sequencing data also supports TFL1’s function as a general repressor of downstream genes, including many flowering-related genes. The large number of genes
downregulated in 35S:TFL1 suggests that TFL1 represses gene expression. Also consistent with TFL1 acting as a repressor, the majority of differentially expressed flowering genes were either downregulated in 35S:TFL1 or upregulated in tfl1-1. However, only a few genes were simultaneously upregulated in 35S:TFL1 and downregulated in tfl1-1 or simultaneously downregulated in 35S:TFL1 and upregulated in tfl1-1 (Figure 2.1). The lack of overlap among these gene sets may imply that they contain secondary gene expression not directly controlled by TFL1; many of these differentially expressed genes may function far downstream of TFL1. Both the 35S:TFL1 and tfl1-1 genotypes express aberrant levels of TFL1 throughout development—very high levels in 35S:TFL1 and very low levels in tfl1-1—and such permanently altered gene expression could impact indirect, secondary targets. Therefore, an approach to identify the immediate targets of both TFL1 and FT was devised.

Using the glucocorticoid and estradiol induction systems, we aimed to determine the immediate targets of TFL1 and FT and to compare the effect of TFL1 and FT on downstream gene expression. In both systems, a chemical applied to transgenic plants activates immediate expression of a transgene. The pilot experiments in both systems showed no clear induction of the meristem identity genes LFY, AP1 and FUL; however, induced FT was observed in the estradiol induction system.

The pilot experiments using the glucocorticoid induction system demonstrated leaky induction of the transgene and no obvious difference in downstream LFY, AP1 and FUL expression between induced and non-induced 35S:FT:GR transgenic plants. One possible explanation for this result is that the induction procedure failed to induce the transgene itself. Because this system uses post-translational induction, in which application of dexamethasone causes translocation of the fusion protein to the nucleus, it offers no simple way to detect induction of the transgene. To overcome this problem, we devised a positive control that uses a different transgene with a known direct target in order to assess the effectiveness of the induction procedure. Assaying levels of this direct target would reveal whether the induction procedure actually induced the transgene. We chose CO as the positive control because of its well-established induction of FT (An et al., 2004; Wigge, 2011; Andrés and Coupland, 2012). Efforts to transform Arabidopsis with 35S:CO:GR are ongoing. After induction, expression of FT will be measured in 35S:CO:GR plants to determine whether the induction procedure was
successful. The result of this control experiment could help to optimize the induction conditions used with the 35S:FT:GR, 35S:TFL1 and 35S:TFL1:VP16 transgenic plants.

A second explanation for the lack of LFY, AP1 and FUL induction in the glucocorticoid induction system is that FT controls downstream flowering genes in an indirect manner. The results from the estradiol induction system, which also showed no (or possibly weak) induction of meristem identity genes, lend more support to this hypothesis. Unlike the glucocorticoid induction system, the estradiol induction system’s transcription-based induction allowed us to assay transgene expression to determine whether transgene induction was successful. The observed induction of the transgene FT but lack of strong downstream flowering gene induction in this system suggests that rather than controlling meristem identity genes directly, FT requires some other factor, which appears later in development, in order to modulate downstream gene expression.

Much remains unknown about the mechanisms by which FT and TFL1 control downstream gene expression. However, a number of FT-interacting molecules have been identified. For instance, FT binds to the phospholipid phosphatidylcholine (PC) (Nakamura et al., 2014). Transgenic plants with high levels of PC flower early and demonstrate elevated expression of FT targets such as SOC1 and AP1, while plants with low levels of PC flower late (Nakamura et al., 2014). FT’s binding to PC may therefore comprise a mechanism by which FT controls gene expression (Nakamura et al., 2014). The transcription factor BRANCHED1 (BRC1) also interacts with FT but not with TFL1 (Niwa et al., 2013). In axillary meristems, BRC1 negatively regulates the flowering transition, perhaps by disrupting FT function (Niwa et al., 2013). Finally, FT binds to FT-INTERACTING PROTEIN 1 (FTIP1), which facilitates FT movement from phloem companion cells to sieve elements (Andrés and Coupland, 2012; Liu et al., 2012).

The presence of these FT-interacting molecules may indicate that FT, as well as TFL1, acts in a transcriptional complex with other transcription factors and that the transcriptional activity of the complex is modulated by currently unknown mechanisms. As described earlier, FT and TFL1 also interact with FD. Because FT and TFL1 both interact with FD, FD emerged as a possible factor required by FT and TFL1 in order for flowering gene induction to occur. Our ongoing experiments to transform 35S:FT, 35S:TFL1 and 35S:TFL1:VP16 T4 plants from both induction systems with 35S:FD will reveal whether FT and TFL1 require FD to control meristem identity gene expression. If
induction succeeds in these plants, their plant material will be analyzed with quantitative reverse-transcription PCR and RNA sequencing to identify the immediate targets of TFL1 and FT and to determine whether they directly regulate the same set of genes.

2.5 LITERATURE CITED


Figure 2.1. Differentially expressed genes in 35S:TFL1 and tfl1-1 compared to wild type
Numbers in circles indicate differentially expressed genes in 35S:TFL1 (pink circles) and tfl1-1 (blue circles) compared to wild type. A shows the number of genes upregulated by TFL1 (i.e. simultaneously upregulated in 35S:TFL1 and downregulated in tfl1-1). B shows the number of genes downregulated by TFL1 (i.e. simultaneously downregulated in 35S:TFL1 and upregulated in tfl1-1). Overlapping portions of the circles indicate differentially expressed genes common between comparisons. Numbers in brackets indicate differentially expressed flowering genes. The large number of genes downregulated in 35S:TFL1 suggests that TFL1 represses gene expression. The lack of overlap among the comparisons suggests that this dataset contains secondary gene expression not directly regulated by TFL1.
Figure 2.2. Differentially expressed flowering genes in 35S:TFL1 and tf1-1 compared to wild type

The heat map on the left shows differentially expressed flowering genes in 35S:TFL1 compared to wild type. The heat map on the right shows differentially expressed flowering genes in tf1-1 compared to wild type. Blue color indicates downregulated genes and yellow color indicates upregulated genes. In 35S:TFL1, 8 of the 11 differentially expressed flowering genes are downregulated. In tf1-1, 11 of the 12 differentially expressed flowering genes are upregulated. These observations suggest that TFL1 represses flowering gene expression.
Figure 2.3. Meristem identity gene expression in 11-day-old 35S:FT:GR seedlings grown under SD

The expected size of PCR products is indicated by arrows. T0 is time point 0 (11 days after plating) and T7 is time point 7 (7 hours after induction). *FUL* expression was detected in both non-induced and induced 35S:FT:GR transgenic plants at approximately equal levels at all time points (lanes FT1, FT2 and FT3), suggesting that *FUL* expression occurred independently of dexamethasone treatment. No expression of AP1 or LFY was detected. The SD growth condition, plant age or treatment buffer may have prevented induction of these genes.
Figure 2.4. Meristem identity gene expression in 18-day-old 35S:FT:GR seedlings grown under SD

The expected size of PCR products is indicated by arrows. T0 is time point 0 (18 days after plating) and T7 is time point 7 (7 hours after induction). *FUL* expression was detected in both non-induced and induced 35S:FT:GR transgenic plants at all time points (lanes 2, 3 and 4), with possible induction in dexamethasone-treated plants (lane 4). *AP1* expression was detected in both non-induced and induced transgenic plants at approximately equal levels at all time points (lanes 2, 3 and 4), suggesting that *AP1* expression occurred independently of dexamethasone treatment. No expression of *LFY* was detected. The SD growth condition may have prevented induction of these genes.
Figure 2.5. Meristem identity gene expression in 11-day-old 35S:FT:GR seedlings grown under LD

The expected size of PCR products is indicated by arrows. T0 is time point 0 (11 days after plating) and T7 is time point 7 (7 hours after induction). Expression of LFY, AP1 and FUL was detected in non-induced and induced 35S:FT:GR transgenic plants at approximately equal levels at all time points, suggesting that meristem identity gene expression occurred independently of dexamethasone treatment.
Figure 2.6. Meristem identity gene expression in 13-day-old 35S:FT:GR seedlings grown under LD
The expected size of PCR products is indicated by arrows. T0 is time point 0 (13 days after plating) and T7 is time point 7 (7 hours after induction). Expression of AP1 and FUL was detected in non-induced and induced 3S:FT:GR transgenic plants at approximately equal levels at all time points, suggesting that meristem identity gene expression occurred independently of dexamethasone treatment.

Figure 2.7. FT expression in 35S:FT seedlings grown under LD
The expected size of PCR products is indicated by arrows. T0 is time point 0 (13 days after plating), T24 is time point 24 (24 hours after induction), and T48 is time point 48 (48 hours after induction). FT expression was detected in almost all samples. FT expression appeared strongest in the induced plant samples (lanes 3 and 5), indicating successful induction of the transgene.
Figure 2.8. Meristem identity gene expression in 35S:FT seedlings grown under LD
The expected size of PCR products is indicated by arrows. T0 is time point 0 (13 days after plating), T24 is time point 24 (24 hours after induction), and T48 is time point 48 (48 hours after induction). FUL expression was detected in both non-induced and induced transgenic plants at time points 0, 24 and 48, with no obvious difference in expression among time points. No other meristem identity genes were detected.
Figure 2.9. Meristem identity gene expression in 35S:FT seedlings grown under LD
The expected size of PCR products is indicated by arrows. T72 is time point 72 (72 hours after induction) and T96 is time point 96 (96 hours after induction). AP1 and FUL expression was detected in both non-induced and induced transgenic plants at time points 72 and 96, with no difference in expression among time points. FT expression appeared strongest in the induced plant samples (lanes 2 and 4).

1 = T72 control
2 = T72 estradiol
3 = T96 control
4 = T96 estradiol
### Table 2.1. Plasmids used in this study

<table>
<thead>
<tr>
<th>NAME</th>
<th>DESCRIPTION</th>
<th>ANTIBIOTIC RESISTANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>dw2</td>
<td>Xba: *TFL1:*XbaI in pCR® 2.1, no stop codon</td>
<td>Kanamycin, Ampicillin</td>
</tr>
<tr>
<td>dw3</td>
<td>Xba: *TFL1:*VP16:XbaI in pCR® 2.1, no stop codon</td>
<td>Kanamycin, Ampicillin</td>
</tr>
<tr>
<td>dw5</td>
<td>*TFL1:*VP16 in pGreen</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>dw7</td>
<td><em>TFL1</em> in pGreen</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>dw11</td>
<td><em>FT</em> in pGreen</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>dw14</td>
<td>Xba:*FT:*SacI in pCR® 2.1, no stop codon</td>
<td>Kanamycin, Ampicillin</td>
</tr>
<tr>
<td>dw15</td>
<td>*TFL1:*VP16 in pEntry</td>
<td>Ampicillin</td>
</tr>
<tr>
<td>dw16</td>
<td><em>FT</em> + 3’UTR in pCR® 8</td>
<td>Spectinomycin</td>
</tr>
<tr>
<td>dw17</td>
<td><em>TFL1</em> + 3’UTR in pCR® 8</td>
<td>Spectinomycin</td>
</tr>
<tr>
<td>dw18</td>
<td><em>FT</em> + 3’UTR in pER8</td>
<td>Kanamycin, Ampicillin</td>
</tr>
<tr>
<td>dw19</td>
<td><em>TFL1</em> + 3’UTR in pER8</td>
<td>Spectinomycin</td>
</tr>
<tr>
<td>dw20</td>
<td>*TFL1:*VP16 in pER8</td>
<td>Spectinomycin</td>
</tr>
<tr>
<td>dw32</td>
<td><em>FD</em> in pDONR™/Zeo</td>
<td>Zeocin</td>
</tr>
<tr>
<td>dw33</td>
<td><em>FD</em> in pK2GW7</td>
<td>Spectinomycin</td>
</tr>
<tr>
<td>dw39</td>
<td><em>FD</em> in pCR® 2.1</td>
<td>Kanamycin, Ampicillin</td>
</tr>
</tbody>
</table>

2.7 TABLES
<table>
<thead>
<tr>
<th>NAME</th>
<th>DESCRIPTION</th>
<th>SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>yh3</td>
<td>CLAT forward</td>
<td>AGCATACACTGCGTGCAAAAG</td>
</tr>
<tr>
<td>yh4</td>
<td>CLAT reverse</td>
<td>TCGCCTGTGTCACATATCTC</td>
</tr>
<tr>
<td>yh18</td>
<td>TFL1 reverse</td>
<td>CAGAGTACGGAAATTCAGAACAC</td>
</tr>
<tr>
<td>yh33</td>
<td>FT + 3’UTR reverse</td>
<td>CATCACCCTCTGTTAATCGTATCA</td>
</tr>
<tr>
<td>yh38</td>
<td>TIL3 forward</td>
<td>ATTCAGAAGGTGCTAGAAG</td>
</tr>
<tr>
<td>yh39</td>
<td>TIL3 reverse</td>
<td>TCGAGCAACAGACAG</td>
</tr>
<tr>
<td>yh40</td>
<td>M13 forward</td>
<td>GTAAAAGCAGGCGAC</td>
</tr>
<tr>
<td>yh41</td>
<td>M13 reverse</td>
<td>CAGAGAACAGCTATGAC</td>
</tr>
<tr>
<td>yh56</td>
<td>TFL1 forward</td>
<td>ATGGGAGATATGGGAACAGAGGT</td>
</tr>
<tr>
<td>yh58</td>
<td>FT forward</td>
<td>ATGCTCTAAAAATAGAGAGCCCTCT</td>
</tr>
<tr>
<td>yh59</td>
<td>FT reverse</td>
<td>AAGTCTCTCTTCCGAGC</td>
</tr>
<tr>
<td>yh306</td>
<td>35S forward</td>
<td>GCCATCATTGCGATAAAGAAAG</td>
</tr>
<tr>
<td>yh312</td>
<td>TFL1 forward (+Xbal)</td>
<td>ACTCTAGAATGGGAAATATGGGAACTAGAGT</td>
</tr>
<tr>
<td>yh315</td>
<td>FT forward (+Xbal)</td>
<td>ACTCTAGAATGTCTAAAAATATAAGAGACCTCT</td>
</tr>
<tr>
<td>yh316</td>
<td>FT reverse (+SacI)</td>
<td>TCAGAGCTCAAGTCTTCTTCCGAGCCAC</td>
</tr>
<tr>
<td>yh321</td>
<td>T7 forward</td>
<td>TAATACGACTCAGTAGG</td>
</tr>
<tr>
<td>yh332</td>
<td>TFL1 reverse (+Xbal)</td>
<td>ATTCTAGAAGCTTGCGCACGCGCCCT</td>
</tr>
<tr>
<td>yh333</td>
<td>TFL1:VP16 reverse (+Xbal)</td>
<td>ATTCTAGAACCACGTACTCGTCAATTCAA</td>
</tr>
<tr>
<td>yh367</td>
<td>AP1 forward</td>
<td>AACCAAGGCCACACTATGCC</td>
</tr>
<tr>
<td>yh368</td>
<td>AP1 reverse</td>
<td>ATCATTCCCTCTCATGCCATAG</td>
</tr>
<tr>
<td>yh371</td>
<td>TFL1:VP16 reverse</td>
<td>ATCTACCCCGTACTCGTCAATTCAA</td>
</tr>
<tr>
<td>yh379</td>
<td>FD forward</td>
<td>GGGGACAAATTTTGTACAAAAACAGGCTCAGGAATCTGAC</td>
</tr>
<tr>
<td>yh380</td>
<td>FD reverse</td>
<td>GGGGACAAATTTTGTACAAAAACAGGCTCAGGAATCTGAC</td>
</tr>
<tr>
<td>yh411</td>
<td>pER8 forward</td>
<td>GGATATGTATATGGTGTAATGCA</td>
</tr>
<tr>
<td>yh412</td>
<td>FUL forward</td>
<td>TGCACCCTCAGAGAAGGATTAAGAG</td>
</tr>
<tr>
<td>yh413</td>
<td>FUL reverse</td>
<td>TTCCGTCAGCGAGATGAC</td>
</tr>
</tbody>
</table>

Table 2.2. Primers used in this study
<table>
<thead>
<tr>
<th>Genotype</th>
<th>Replicate</th>
<th>Total Reads</th>
<th>Mapped Reads</th>
<th>Percent Reads Mapped</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild Type</td>
<td>1</td>
<td>6,339,917</td>
<td>6,063,214</td>
<td>0.956</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5,173,405</td>
<td>4,967,734</td>
<td>0.960</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6,447,831</td>
<td>6,079,150</td>
<td>0.943</td>
</tr>
<tr>
<td>tfI1-1</td>
<td>1</td>
<td>5,809,706</td>
<td>5,522,560</td>
<td>0.951</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5,943,350</td>
<td>5,631,280</td>
<td>0.947</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5,911,690</td>
<td>5,656,135</td>
<td>0.957</td>
</tr>
<tr>
<td>35S:TFI1</td>
<td>1</td>
<td>6,154,310</td>
<td>5,880,668</td>
<td>0.956</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5,791,951</td>
<td>5,527,379</td>
<td>0.954</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5,648,223</td>
<td>5,408,056</td>
<td>0.957</td>
</tr>
</tbody>
</table>

Table 2.3. Number of reads obtained by RNA sequencing
<table>
<thead>
<tr>
<th></th>
<th>Wild Type</th>
<th>35S:TFL1</th>
<th>tfl1-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGL14</td>
<td>11.9</td>
<td>1.0</td>
<td>13.8</td>
</tr>
<tr>
<td>SPA4</td>
<td>3.1</td>
<td>5.5</td>
<td>3.8</td>
</tr>
<tr>
<td>ATC</td>
<td>9.3</td>
<td>4.9</td>
<td>12.7</td>
</tr>
<tr>
<td>SPA1</td>
<td>18.5</td>
<td>17.5</td>
<td>16.8</td>
</tr>
<tr>
<td>SPA3</td>
<td>4.7</td>
<td>6.3</td>
<td>5.4</td>
</tr>
<tr>
<td>AGL19, GL19</td>
<td>5.0</td>
<td>1.8</td>
<td>8.1</td>
</tr>
<tr>
<td>AGL21</td>
<td>5.8</td>
<td>0.3</td>
<td>8.3</td>
</tr>
<tr>
<td>TFL1</td>
<td>3.0</td>
<td>3.3</td>
<td>3.9</td>
</tr>
<tr>
<td>AT5G10625</td>
<td>6.0</td>
<td>1.4</td>
<td>7.9</td>
</tr>
<tr>
<td>COL1, ATOCOL1</td>
<td>1.8</td>
<td>2.5</td>
<td>1.4</td>
</tr>
<tr>
<td>LFY, LFY3</td>
<td>0.5</td>
<td>1.1</td>
<td>0.1</td>
</tr>
<tr>
<td>SEP3, AGL9</td>
<td>0.1</td>
<td>0.7</td>
<td>0.4</td>
</tr>
<tr>
<td>STM, BUM1</td>
<td>2.9</td>
<td>3.5</td>
<td>3.3</td>
</tr>
<tr>
<td>AP1, AGL7</td>
<td>0.0</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>SPL3</td>
<td>6.7</td>
<td>12.9</td>
<td>7.6</td>
</tr>
<tr>
<td>ABI5, GAI1</td>
<td>3.3</td>
<td>3.8</td>
<td>1.9</td>
</tr>
<tr>
<td>AGL6</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>SEP2, AGL4</td>
<td>0.1</td>
<td>0.6</td>
<td>0.2</td>
</tr>
<tr>
<td>AP3, ATAP3</td>
<td>0.2</td>
<td>0.6</td>
<td>0.0</td>
</tr>
<tr>
<td>AG</td>
<td>0.0</td>
<td>0.3</td>
<td>0.0</td>
</tr>
<tr>
<td>SEP1, AGL2</td>
<td>0.0</td>
<td>0.6</td>
<td>0.0</td>
</tr>
<tr>
<td>PI</td>
<td>0.3</td>
<td>0.9</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table 2.4. Read counts of differentially expressed flowering genes
Normalized read counts (RPKM) of differentially expressed flowering genes from each replicate are shown.
### Table 2.5. GO-term categories enriched among differentially expressed genes in 35S:TFL1 and tfl1-1

The top 7 GO-term functional categories enriched among upregulated and downregulated genes in 35S:TFL1 and among upregulated and downregulated genes in tfl1-1 are shown. The enrichment score of each category is indicated in parentheses. Functional categories related to transcription and development are indicated in bold. In addition to enrichment of transcription-related categories, enrichment of defense-related, stress-responsive and lipid-binding categories was observed.

<table>
<thead>
<tr>
<th></th>
<th>35S:TFL1 Up</th>
<th>35S:TFL1 Down</th>
<th>tfl1-1 Up</th>
<th>tfl1-1 Down</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lipid binding (4.97)</td>
<td>Response to stimulus (34.74)</td>
<td>Regulation of transcription (9.03)</td>
<td>Defense response (9.35)</td>
<td></td>
</tr>
<tr>
<td>Glycoprotein-associated (4.92)</td>
<td>Cell wall structure (6.40)</td>
<td>Development &amp; flowering (6.15)</td>
<td>Regulation of transcription (8.91)</td>
<td></td>
</tr>
<tr>
<td>Response to stimulus (3.11)</td>
<td>Membrane associated (5.99)</td>
<td>DNA binding &amp; transcription (5.34)</td>
<td>TIFY domain (5.24)</td>
<td></td>
</tr>
<tr>
<td>Membrane binding (2.98)</td>
<td>Response to stress (5.54)</td>
<td>Response to stimulus (2.57)</td>
<td>Response to stress (4.68)</td>
<td></td>
</tr>
<tr>
<td>Membrane structure (1.82)</td>
<td>Regulation of transcription (5.52)</td>
<td>Ion binding &amp; transcription (2.18)</td>
<td>Cell wall associated (4.44)</td>
<td></td>
</tr>
<tr>
<td>Calcium binding (1.77)</td>
<td>Defense response (5.20)</td>
<td>Nutrient storage (1.64)</td>
<td>Calcium binding (4.41)</td>
<td></td>
</tr>
<tr>
<td>Nutrient storage (1.48)</td>
<td>Hormone signaling (5.18)</td>
<td>Seed germination &amp; fruit development (1.19)</td>
<td>Biosynthetic &amp; metabolic control (3.49)</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.6. Numbers of independent T1 transgenic lines in the glucocorticoid induction system**

<table>
<thead>
<tr>
<th></th>
<th>Wild Type</th>
<th>tfl1-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>35S:FT:GR</td>
<td>26</td>
<td>15</td>
</tr>
<tr>
<td>35S:TFL1:GR</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>35S:TFL1:VP16:GR</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.7. Numbers of independent T2 transgenic lines carrying single transgene insertion in the glucocorticoid induction system**

<table>
<thead>
<tr>
<th></th>
<th>Wild Type</th>
<th>tfl1-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>35S:FT:GR</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>35S:TFL1:GR</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>35S:TFL1:VP16:GR</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>35S:TFL1:GR</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>35S:TFL1:VP16:GR</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Wild Type</td>
<td>tfl1-1</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>35S:<em>FT</em>:GR</td>
<td>35S:<em>TFL1</em>:GR</td>
<td>35S:<em>TFL1</em>:VP16:GR</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2.8. Numbers of independent T3 transgenic lines carrying homozygous single transgene insertion in the glucocorticoid induction system

<table>
<thead>
<tr>
<th>Wild Type</th>
<th>tfl1-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>35S:<em>FT</em></td>
<td>35S:<em>TFL1</em></td>
</tr>
<tr>
<td>18</td>
<td>17</td>
</tr>
</tbody>
</table>

Table 2.9. Numbers of independent T1 transgenic lines in the estradiol induction system

<table>
<thead>
<tr>
<th>Wild Type</th>
<th>tfl1-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>35S:<em>FT</em></td>
<td>35S:<em>TFL1</em></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2.10. Numbers of independent T2 transgenic lines carrying single transgene insertion in the estradiol induction system
### Table 2.11

<table>
<thead>
<tr>
<th></th>
<th>Wild Type</th>
<th>tfl1-1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>35S:<em>FT</em></td>
<td>35S:<em>TFL1</em></td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

A) Numbers of independent T3 transgenic lines carrying homozygous single transgene insertion in the estradiol induction system
B) Numbers of independent T3 transgenic lines carrying homozygous multiple transgene insertions in the estradiol induction system
CHAPTER 3
ROLES OF FT AND TFL1 IN SOYBEAN PHOTOPERIODIC FLOWERING CONTROL

3.1 BACKGROUND

3.1.1 FT AND TFL1 FUNCTIONS IN FLOWERING PLANTS

Conserved functions in photoperiodic flowering control and inflorescence architecture

In Arabidopsis, FT acts as a flowering inducer and TFL1 acts as a flowering repressor in the photoperiodic flowering pathway. The functions and amino acid sequences of these transcriptional regulators are highly conserved in flowering plants, including in crop and horticultural species. For instance, rice (Oryza sativa) possesses two FT orthologs, Hd3a and RFT1. Hd3a induces flowering under short days, while RFT1 induces flowering under long days (Kojima et al., 2002; Izawa, 2007; Komiya et al., 2009; Pin and Nilsson, 2012). Additionally, both RFT1 and HD3A independently form complexes with 14-3-3 proteins that interact with OsFD1, rice’s FD counterpart (Taoka et al., 2011; Pin and Nilsson, 2012). Similar to TFL1 in Arabidopsis, the TFL1 orthologs RCN1 and RCN2 in rice repress flowering (Nakagawa et al., 2002). In apple (Malus domestica), the FT orthologs MdFT1 and MdFT2 promote flowering, while the TFL1 ortholog MdTFL1 promotes the juvenile developmental phase and represses flowering (Kotoda and Wada, 2005; Kotoda et al., 2006; Kotoda et al., 2010). Likewise, tomato (Lycopersicon esculentum) carries an FT-like protein, SINGLE FLOWER TRUSS (SFT), that induces flowering and a TFL1-like protein, SELF-PRUNING (SP), that inhibits flowering (Pnueli et al., 1998; Jiang et al., 2013). The pumpkin (Cucurbita maxima) FT orthologs Cm-FTL1 and Cm-FTL2 stimulate flowering, and the peach (Prunus persica) TFL1 ortholog ppTFL1 represses flowering (Lin et al., 2007; Chen and Jiang, 2013). Additionally, rose (Rosa chinensis) contains an FT ortholog, RoFT, highly expressed during the flowering transition (Remay et al., 2009). When overexpressed in dominant form, the rose TFL1 ortholog RoKSN prevents flowering and reduces expression of downstream meristem identity gene such as RoLFY and RoAP1 (Randoux
et al., 2014). Additionally, both RoFT and RoKSN strongly interact with RoFD, the rose FD ortholog (Randoux et al., 2014).

In Arabidopsis, TFL1 also controls the determinacy status of the inflorescence shoot by promoting indeterminate growth (Bradley, 1997). This conserved role in inflorescence architecture appears in a number of species. In tomato, the TFL1 ortholog SP influences stem growth habit. Tomato carrying the dominant SP allele exhibits indeterminate growth characterized by the continuous production of inflorescences and fruit (Pnueli et al., 1998; Jiang et al., 2013). Tomato carrying the recessive allele exhibits determinate growth characterized by the premature cessation of inflorescence production and smaller, bushier plants (Pnueli et al., 1998; Jiang et al., 2013). In both rose and strawberry (Fragaria vesca), non-functional recessive orthologs of TFL1 (RoKSN and FvKSN, respectively) establish the horticulturally desirable trait of continuous flowering habit (Iwata et al., 2012).

**Other functions**

In Arabidopsis, the FT/TFL1 gene family encompasses several functions beyond regulation of flowering time and inflorescence architecture, including contributions to light-induced opening of leaf stomata, sensitivity to ambient temperature, seed germination, and flowering in response to salt stress (Kinoshita et al., 2011; Ryu et al., 2014). Orthologs in other species also play roles in diverse developmental processes. For instance, the two FT orthologs in sugar beet (Beta vulgaris ssp. vulgaris) are involved in vernalization. Prior to vernalization, BvFT1 inhibits flowering by repressing BvFT2 (Pin et al., 2010). Vernalization then triggers flowering by reducing BvFT1 expression and upregulating BvFT2 expression (Pin et al., 2010; Andrés and Coupland, 2012; Pin and Nilsson, 2012). In addition to promoting flowering, a potato (Solanum tuberosum) FT ortholog, StSP6A, functions as a mobile “tuberigen” that induces the photoperiod-sensitive process of tuberization (Navarro et al., 2011).

### 3.1.2 FT AND TFL1 FUNCTIONS IN SOYBEAN

Conforming to the pattern found throughout the plant kingdom, the FT/TFL1 gene family in soybean shows conserved functions in photoperiodic flowering as well as additional functions in other development-related processes. This family includes eleven
FT homologs and several TFL1 homologs (Kong et al., 2010; Tian et al., 2010). The role of two FT homologs, GmFT2a and GmFT5a, in promotion of flowering has been characterized (Kong et al., 2010; Watanabe et al., 2012; Nan et al., 2014). Their expression fluctuates in a circadian fashion and reaches highest levels under flowering-inductive short-day (SD) conditions but remains low under non-inductive long-day (LD) conditions, indicating a response to environment consistent with photoperiod-regulated flowering (Kong et al., 2010; Thakare et al., 2010). In transgenic Arabidopsis, ectopically expressed GmFT2a and GmFT5a accelerate flowering (Kong et al., 2010). In transgenic soybean, ectopically expressed GmFT2a and GmFT5a accelerate flowering under LD and elevate the expression of meristem identity genes such as GmAP1s (a, b, c), GmLFY2, GmSOC1 and GmSOC2 (Nan et al., 2014). Echoing FT’s partnership with FD in Arabidopsis, both GmFT2a and GmFT5a interact with GmFDL19, a soybean FD ortholog (Nan et al., 2014). GmFDL19 binds to the GmAP1 promoter, suggesting a mechanism by which GmFTs and GmFDL19 upregulate expression of a meristem identity gene to induce flowering (Nan et al., 2014).

Expression of both GmFT2a and GmFT5a peaks four hours after dawn (Kong et al., 2010; Watanabe et al., 2012). GmFT2a expression changes rapidly in response to photoperiod changes; moving from SD to LD decreases expression rapidly (Kong et al., 2010). GmFT5a expression, on the other hand, decreases slowly in response to the same photoperiod change and eventually remains at a low, basal level of expression (Kong et al., 2010). The two genes may thus act together to spur flowering under short photoperiods (Kong et al., 2010). Under long photoperiods, however, GmFT5a may act alone to promote flowering independently of photoperiod (Kong et al., 2010). Several dominant alleles of the E loci downregulate expression of FT2a and FT5a under LD, resulting in delayed flowering (Kong et al., 2010; Thakare et al., 2011; Xia et al., 2012; Nan et al., 2014). The recessive E2 locus, which encodes a loss-of-function GmGIa allele, causes early flowering and is associated with elevated FT2a levels (Watanabe et al., 2011). This result suggests that functional GmGIa, encoded by dominant E2, retards flowering by suppressing GmFT2a expression under LD (Watanabe et al., 2011; Nan et al., 2014).
Dt1 is a soybean ortholog of Arabidopsis TFL1 (Liu et al., 2010; Tian et al., 2010; Watanabe et al., 2012). When expressed in the early-flowering Arabidopsis tfl1 loss-of-function mutant, the dominant Dt1 allele reestablishes the wild-type phenotype (Tian et al., 2010). Dominant Dt1 is also associated with indeterminate growth, while the recessive allele is associated with determinate growth (Bernard, 1971; Cober and Morrison, 2010; Tian et al., 2010). Determinate soybean culminates in a shorter height and matures earlier than indeterminate soybean, suggesting Dt1’s role in soybean flowering control (Bernard, 1972; Curtis et al., 2000; Tian et al., 2010). The geographic distribution of dominant and recessive Dt1 alleles may facilitate adaptation to particular photoperiodic environments (Tian et al., 2010). For instance, the dominant allele abounds in northern regions of China, while the recessive allele predominates in southern regions of the country (Tian et al., 2010).

3.1.3 EXPERIMENTAL AIM

The role of the FT homologs GmFT2a and GmFT5a in photoperiod-regulated flowering and the role of the TFL1 homolog Dt1 in inflorescence growth habit have been established; however, the functions of other homologs remain unknown. We sought to characterize the functions of other FT homologs in soybean. In order to investigate these genes’ effect on flowering time, we transformed Arabidopsis with soybean FT homologs and observed the resulting flowering phenotypes.

3.2 MATERIALS AND METHODS

3.2.1 SOYBEAN RNA SEQUENCING

**Plant material and growth conditions**

The 2 soybean genotypes studied were *Glycine max* Williams 82 (PI 518671) and *Glycine soja* (PI 549046). Seeds were grown in rooms 7E2 and 3E1 in the Turner Hall greenhouse. Each room featured a different photoperiod. For long-day (LD) conditions, supplementary light in room 7E2 established a light period of 16 hours. For short-day (SD) conditions, the blackout curtains in room 3E1 established an artificial dark period of 14 hours per day (10 hours of light). Twenty-four days after planting, plants were
sampled at three time points, each separated by eight hours and corresponding to morning, afternoon and evening: T1 (6:30), T3 (14:30) and T5 (22:30). Three to four plants were sampled at each time point under both photoperiod conditions. Each sample consisted of an entire shoot above the cotyledon. Samples were exposed to liquid nitrogen for 15 seconds and then placed in -80°C storage.

**RNA preparation and RNA sequencing data analysis**

RNA was prepared from soybean plant tissue samples stored at -80°C according to the methods previously described. RNA samples were submitted to the W.M. Keck Center for Comparative and Functional Genomics (Sequencing and Genotyping Division) for Illumina-based RNA Sequencing. Reads were aligned to the soybean transcriptome from the phytozome.net database (Schmutz et al., 2010) using Bowtie alignment software (Langmead et al., 2009). In-house Perl and Python scripts normalized mapped reads according to RPKM (Mortazavi et al., 2008). A subset of this data, which already existed in the lab (Wu et al., 2014), was examined for the current project. Microsoft Excel® was used for data exploration and analysis.

3.2.2 GENE CLONING

**RNA preparation and cDNA synthesis**

The prepared RNA from one of the *Glycine soja* T3 short-day samples was used as a template for cDNA synthesis. cDNA was synthesized using the iScript™ cDNA Synthesis Kit (Bio-Rad). In accordance with the manufacturer’s protocol, each reaction mixture consisted of 1 µg RNA, 4 µL 5x iScript reaction mix, and 1 µL iScript reverse transcriptase, with nuclease-free water added to bring the total volume to 20 µL. These mixtures were then incubated in the Mastercycler® Pro under the following conditions: 5 minutes at 25°C, 30 minutes at 42°C, and then 5 minutes at 85°C. To measure the concentration of synthesized cDNA, a 2 µL sample from each reaction product was analyzed with the NanoDrop® ND-1000 Spectrophotometer (ThermoScientific®).

**Gene amplification**

Full-length cDNA of *FT* (528bp + 110 bp 3’ UTR) was obtained from the plasmid dw16 (Table 3.1). Full-length cDNA of *FT2a* (531 bp), *FT2b* (531 bp) and *FT2c* (534 bp)
was obtained from the cDNA library described in the previous section. PCR with Platinum® Pfx DNA polymerase was used to amplify the full-length cDNA from the source material as previously described, with the following modifications: the primers yh58 and yh33 were used to amplify FT from dw16, yh306 and yh385 were used to amplify GsFT2a from the cDNA library, yh306 and yh387 were used to amplify GsFT2b from the cDNA library, and yh306 and yh309 were used to amplify GsFT2c from the cDNA library (Table 3.2). These primers left stop codons intact. The PCR cycle used a 55°C annealing temperature and 1-minute extension time. After completion of PCR followed by Poly-A extension, 5 µL each PCR product was run on an agarose gel by electrophoresis to verify amplification of each transgene.

**Gateway® subcloning**

The PCR products containing full-length, amplified cDNA of FT, GsFT2a, GsFT2b and GsFT2c were cloned into the entry vector pCR®8 (Invitrogen) using the pCR®8/GW/TOPO®TA Cloning® Kit (Invitrogen). This vector carries resistance to the antibiotic spectinomycin. In accordance with the manufacturer’s protocol, 4 µL each PCR product was added to tubes containing 1 µL salt solution and 1 µL pCR®8. These mixtures were incubated at room temperature for 10 minutes. After incubation, the 4 resulting plasmids were used to transform Subcloning Efficiency™ DH5α™ Competent Cells for plasmid amplification to levels sufficient for plasmid purification. The 4 plasmids were named as follows. FT in pCR®8 was designated dw38, GsFT2a in pCR®8 was designated dw35, GsFT2b in pCR®8 was designated dw36, and GsFT2c in pCR®8 was designated dw37.

After plasmid purification, the Gateway® LR reaction was performed using the Gateway® LR Clonase II™ kit (Invitrogen). This recombination reaction transferred each transgene from its respective pCR®8 plasmid into the destination expression vector pEarley100. pEarley100 contains the 35S promoter to induce overexpression of a transgene (Earley et al., 2006). This vector confers resistance to the antibiotic kanamycin in bacteria and to the herbicide BASTA in plants. The 4 reaction solutions consisted of 4 µL entry vector (dw35, dw36, dw37 or dw38), 1 µL destination vector (pEarley100), 2 µL LR Clonase II™ enzyme mix, and 2 µL TE buffer. The reactions were incubated at room temperature for 1 hour and then used to transform Subcloning
Efficiency™ DH5α™ Competent Cells (Invitrogen) for plasmid amplification to levels sufficient for plasmid purification. The resulting plasmids were named as follows. FT in pEarley100 was designated dw30, GsFT2a in pEarley100 was designated dw27, GsFT2b in pEarley100 was designated dw28, and GsFT2c in pEarley100 was designated dw29.

3.2.3 ARABIDOPSIS PLANT TRANSFORMATION

**Plant material and growth conditions**

The LD-grown Arabidopsis FT loss-of-function mutant *ft-10* was transformed with dw27, dw28, dw29 and dw30. Twelve pots containing six seeds each were grown in room 6E2 in the Turner Hall greenhouse, where supplementary light established a light period of sixteen hours.

**Arabidopsis transformation via *Agrobacterium*-mediated floral dip**

GV3101 Agrobacteria were transformed with dw27, dw28, dw29 and dw30 according to the methods previously described. Colony survival on plates containing the antibiotics gentamycin (80 mg/mL), rifampicin (100 mg/mL), and kanamycin (50 mg/mL) indicated successful transformation. Genotyping of plant material by colony PCR then confirmed the presence of the transgenes in particular Agrobacteria colonies. The following primers were used in conjunction with Taq polymerase (New England Biolabs): yh306 and yh33 were used to identify colonies containing dw30, yh306 and yh385 were used to identify colonies containing dw27, yh306 and yh387 were used to identify colonies containing dw28, and yh306 and yh309 were used to identify colonies containing dw29. The PCR cycle used a 1-minute elongation time and 55°C annealing temperature. Agrobacteria carrying either of the 4 plasmids were then used to introduce each transgene separately into *ft-10*. Three pots were dedicated to each of the four transgene constructs. After transformation, plants were returned to room 6E2 in the greenhouse.

3.2.4 ESTABLISHING TRANSGENIC LINES

Approximately 3 weeks after transformation, T1 seeds were collected and then screened on MS plates (with 10 mg/mL BASTA) for positive transformants. The
individuals surviving on the plates (i.e. the candidate positive transformants) were genotyped by PCR to confirm presence of the transgenes and to check for contamination by other transgenes. Extracted DNA from each plant was used as a template for PCR. The following primer combinations were used: yh306 and yh59 to identify the FT transgene, yh306 and yh385 to identify the GsFT2a transgene, yh306 and yh387 to identify the GsFT2b transgene, and yh306 and yh309 to identify the GsFT2c transgene. The PCR cycle used a 1-minute elongation time and 55°C annealing temperature. Six to ten transformed seeds from each line, in addition to untransformed ft-10 seeds and wild-type Col-0 seeds, were then grown under SD to retard the early-flowering phenotype expected in the transgenic plants. Plants were grown in individual plastic inserts filled with LC1 soil in room 3E2 in the Turner Hall greenhouse. The blackout curtains in this room established the artificial dark period of 16 hours per day.

3.2.5 FLOWERING TIME MEASUREMENT

In Arabidopsis, leaf number serves as an indirect measurement of flowering time. Arabidopsis produces two types of leaves: rosette and cauline. Rosette leaves, which grow at the base of the plant, indicate the duration of vegetative growth, with a greater number of rosette leaves signifying a longer vegetative phase of development. The number of rosette leaves generated prior to bolting (i.e. elongation of the main stem) was counted in all plants. Cauline leaves, which grow on the main stem, indicate the duration of reproductive growth, with a greater number of cauline leaves signifying a longer reproductive phase of development. After plants had bolted, the number of cauline leaves was counted every week in all lines until plants stopped producing new leaves. Microsoft Excel® was used to record leaf number data and SAS® was used for statistical analysis.
3.3 RESULTS

3.3.1 mRNA EXPRESSION OF *FT* HOMOLOGS IN WILD AND DOMESTICATED SOYBEAN

To characterize the function of soybean *FT* homologs, we first observed the expression of *FT* homologs in both domesticated (*G. max*) and wild (*G. soja*) soybean at 3 time points under both long-day (LD) and short-day (SD) conditions by analyzing the RNA sequencing data already existing in the lab (Wu et al., 2014). Under non-flowering-inductive LD, expression of all 11 *FT* homologs remained low at each time point in both genotypes, with no gene exceeding 10 RPKM (Figure 3.1a). Under flowering-inductive SD, however, a number of genes demonstrated high expression. In both *G. max* and *G. soja*, the previously characterized flowering inducers *FT2a* and *FT5a* (Kong et al., 2010) were expressed at high levels under SD (Figure 3.1b). Additionally, both *FT2a* and *FT2b* exhibited rhythmic expression with a peak in the afternoon in both genotypes under SD, which indicated photoperiodic response. The most striking difference in *FT* homolog expression between *G. soja* and *G. max* appeared in *FT2c* expression under SD. In *G. max*, *FT2c* expression remained low throughout the day and showed no afternoon peak. In *G. soja*, however, *FT2c* remained high throughout the day and peaked in the afternoon, suggesting a response to photoperiod.

3.3.2 GENE STRUCTURE OF *FT2c*

The difference in *FT2c* expression between *G. max* and *G. soja* drove us to examine the gene structure and polymorphisms of *FT2c*. We found that the structure of *Ft2c* differed between *G. max* and *G. soja*. In the latest soybean gene models (version 9.1 from the Phytozome database), *FT2c* in *G. max* carries a 20 kb insertion between exon 3 and exon 4, potentially rendering the mRNA truncated and the protein nonfunctional (Figure 3.2a). From a *G. soja* cDNA library, we successfully amplified full-length *FT2c* cDNA and found that the gene carries no insertion in *G. soja* (Figure 3.2b) (Wu and Hanzawa, unpublished). This observation, together with the expression data, suggests that *FT2c* is functional in *G. soja* but nonfunctional in *G. max* and that the insertion-deletion (indel) polymorphism in *FT2c* is potentially important for soybean development and evolution.
3.3.3 FUNCTION OF SOYBEAN FT HOMOLOGS IN TRANSGENIC ARABIDOPSIS

To clarify the function of \( GsFT2c \), a transgenic approach in Arabidopsis was carried out. The effects of \( GsFT2c \) and its homeologs \( GsFT2a \) and \( GsFT2b \) on flowering time were compared to the effects of Arabidopsis \( FT \).

Establishing transgenic lines

Four constructs were used to transform the \( FT \) loss-of-function mutant \( ft-10: \) \( 35S:FT \), \( 35S:FT2a \), \( 35S:FT2b \) and \( 35S:FT2c \). Plants were grown under non-inductive SD to retard the expected early-flowering phenotype. Screening of the obtained T1 seeds using the BASTA resistance selectable marker resulted in 9 \( 35S:FT \) lines, 10 \( 35S:FT2a \) lines, 16 \( 35S:FT2b \) lines, and 7 \( 35S:FT2c \) lines. The existence of each transgene was examined by PCR genotyping, and 6 \( 35S:FT \) lines, 9 \( 35S:FT2a \) lines, 8 \( 35S:FT2b \) lines and 6 \( 35S:FT2c \) lines were confirmed in the T1 generation (Table 3.3). T1 plants were transferred to soil and their flowering phenotypes were observed.

Effect of soybean \( FT \) homologs on flowering phenotype

Compared to wild-type and \( ft-10 \) controls, transgenic Arabidopsis carrying soybean \( FT \) homologs and Arabidopsis \( FT \) showed accelerated flowering, as indicated by their reduced numbers of rosette and cauline leaves (Figures 3.3 and 3.4). Wild type (\( n=7 \)) possessed on average 34.75 rosette leaves and 57.43 cauline leaves, and \( ft-10 \) (\( n=8 \)) possessed on average 46.13 rosette leaves and 57.38 cauline leaves. In the \( ft-10 \) background, all transgenic plants showed a significantly different number of rosette and cauline leaves compared to the wild-type and \( ft-10 \) controls at the \( p<0.05 \) level. \( 35S:FT \) (\( n=6 \)) possessed on average 4.33 rosette leaves and 3.17 cauline leaves, \( 35S:GsFT2a \) (\( n=9 \)) possessed on average 10 rosette leaves and 5.22 cauline leaves, \( 35S:GsFT2b \) (\( n=8 \)) possessed on average 9 rosette leaves and 6.75 leaves, and \( 35S:GsFT2c \) (\( n=6 \)) possessed on average 7 rosette leaves and 3.55 cauline leaves (Figure 3.3). Among transgenic plants, the number of rosette leaves did not differ significantly at the \( p<0.05 \) level, with the exception of two comparisons: \( 35S:FT \) and \( 35S:GsFT2a \) (\( p=.0016 \)) and \( 35S:FT \) and \( 35S:GsFT2b \) (\( p=.0097 \)). \( 35S:FT \) and \( 35S:GsFT2c \) were not significantly different (\( p=.1534 \)). The number of cauline leaves was not significantly different among
transgenic plants at the p<0.05 level. The 4 transgenes therefore affected flowering time in a similar fashion. In particular, GsFT2c’s striking induction of flowering in transgenic Arabidopsis suggests that it also promotes flowering in G. soja.

3.4 DISCUSSION

This project sought to characterize the functions of soybean FT homologs, particularly in relation to plant development and evolution. Previous studies have established that two homologs, GmFT2a and GmFT5a, stimulate flowering in response to photoperiodic input (Kong et al., 2010; Nan et al., 2014). The functions of other soybean FT homologs remain unknown. Here, we focused on elucidating the function of GsFT2c in flowering. While the mRNA expression of G. max FT2c remained low at all time points, we observed high G. soja FT2c mRNA expression with a peak in the afternoon under flowering-inductive SD (Figure 3.1). This G. soja FT2c mRNA expression mirrors the photoperiodic response in FT2a expression, which suggests that GsFT2c plays a role in photoperiodic flowering in wild soybean. The gene structure data reinforces this conclusion; while G. max carries a 20kb insertion in FT2c, G. soja carries the intact form (Figure 3.2). Our transgenic approach in Arabidopsis suggests that GsFT2c participates in flowering induction. These observations may indicate that the indel polymorphism in FT2c is related to the evolution of soybean flowering habit. To assess this possibility, we are currently genotyping a large set of G. max and G. soja accessions for the indel. All G. soja varieties examined carry intact FT2c, while most G. max varieties examined carry the indel (Wu and Hanzawa, unpublished). The loss of intact FT2c in G. max may have occurred during the soybean domestication process and may have been a target for selection by humans.

Although FT2c in G. max is expressed at low levels with no photoperiodic response and likely transcribed in truncated form and then translated into a non-functional protein, we cannot exclude the possibility that this gene has evolved a novel function in domesticated soybean. In some species, FT orthologs have developed a repressor function over the course of evolution. For instance, sugar beet contains two FT homologs, BvFT1 and BvFT2, with divergent functions (Pin et al., 2010; Pin and
Nilsson, 2012). An antagonistic pair, they share high sequence similarity but function in an opposite manner (Pin et al., 2010; Pin and Nilsson, 2012). BvFT2 promotes flowering, while BvFT1 represses flowering (Pin et al., 2010; Pin and Nilsson, 2012). BvFT1 inhibits flowering by repressing BvFT2 before the onset of vernalization (Pin et al., 2010). Vernalization then reduces BvFT1 expression and upregulates BvFT2 expression, triggering flowering (Pin et al., 2010; Andrés and Coupland, 2012). Sunflower (Helianthus annuus) contains four FT homologs, one of which underwent a mutation during this species’ evolution that yielded a novel repressor function (Blackman et al., 2010). Wild sunflower carries an in-frame allele of HaFT1, an FT homolog that stimulates flowering in transgenic Arabidopsis (Blackman et al., 2010). Domesticated sunflower, however, carries a frame-shifted HaFT1 allele, which encodes a protein with a sequence 17 amino acids longer than its wild sunflower counterpart (Blackman et al., 2010). In transgenic Arabidopsis, the domesticated HaFT1 allele delays flowering by interfering with another sunflower FT homolog, the inductive long-day expressed HaFT4, in a dominant-negative fashion (Blackman et al., 2010). The widespread prevalence of the HaFT1 frameshift allele among domesticated sunflower varieties suggests that selection played a role in retaining this allele (Blackman et al., 2010). Because FT orthologs possess divergent functions in some species, we are currently investigating the possibility of FT2c function in G. max. An ongoing approach using near-isogenic lines will help to clarify any role that this allele plays in domesticated soybean and confirm the function of intact FT2c in wild soybean.

### 3.5 LITERATURE CITED


**Bernard RL** (1971) Two major genes for time of flowering and maturity in soybeans. Crop Sci 12: 242–244


3.6 FIGURES

Figure 3.1. mRNA abundance of FT homologs under LD and SD in *G. max* and *G. soja*

mRNA abundance of *FT* homologs under LD (A) and SD (B). Samples are from 3 representative

time points: 6:30 (morning), 14:30 (afternoon) and 22:30 (evening). Under LD, no *FT* homologs

were induced. Under SD, *FT2a* (yellow) and *FT5a* (purple) were highly expressed and *FT2a* and

*FT2b* (blue) showed afternoon peaks in *G. max* and *G. soja*. In *G. soja*, *FT2c* (red) was highly

expressed with an afternoon peak under SD only, suggesting a response to photoperiod.
Figure 3.2. FT2c structure in G. max and G. soja
Gene structure of FT2c in G. max (A) and G. soja (B). In G. max, a 20 kb insertion exists between exons 3 and exon 4, potentially rendering the mRNA truncated and the protein nonfunctional. In G. soja, FT2c exists in intact form.

Figure 3.3. Effect of soybean FT homologs on flowering phenotype in Arabidopsis ft-10
Under SD (8 hours light / 16 hours dark), 35S:AtFT, 35S:GsFT2a, 35S:GsFT2b and 35S:GsFT2c accelerated flowering compared to wild-type and ft-10 controls. Error bars depict the standard error of the mean.
Figure 3.4. Effect of soybean FT homologs on flowering phenotype in wild-type Arabidopsis
Under LD (16 hours light / 8 hours dark), 35S:AtFT, 35S:GsFT2a, 35S:GsFT2b and 35S:GsFT2c accelerated flowering compared to wild-type and ft-10 controls. 35S:AtFT and 35S:FT2c appear phenotypically similar to each other, and 35S:GsFT2a and 35S:GsFT2b appear phenotypically similar to each other.
### 3.7 TABLES

<table>
<thead>
<tr>
<th>PLASMID NAME</th>
<th>DESCRIPTION</th>
<th>ANTIBIOTIC RESISTANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>dw27</td>
<td>FT2a in pEarley100</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>dw28</td>
<td>FT2b in pEarley100</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>dw29</td>
<td>FT2c in pEarley100</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>dw30</td>
<td>AtFT in pEarley100</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>dw35</td>
<td>FT2a in pCR®8</td>
<td>Spectinomycin</td>
</tr>
<tr>
<td>dw36</td>
<td>FT2b in pCR®8</td>
<td>Spectinomycin</td>
</tr>
<tr>
<td>dw37</td>
<td>FT2c in pCR®8</td>
<td>Spectinomycin</td>
</tr>
<tr>
<td>dw38</td>
<td>AtFT in pCR®8</td>
<td>Spectinomycin</td>
</tr>
</tbody>
</table>

Table 3.1. Plasmids used in this study

<table>
<thead>
<tr>
<th>NAME</th>
<th>DESCRIPTION</th>
<th>SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>yh33</td>
<td>FT + 3’ UTR reverse</td>
<td>CATCACCCTCGTACTCGTATCA</td>
</tr>
<tr>
<td>yh58</td>
<td>FT forward</td>
<td>AAGTCTTCTTCTCCGCAGC</td>
</tr>
<tr>
<td>yh306</td>
<td>35S forward</td>
<td>GCCATCATTGCAGAAAG</td>
</tr>
<tr>
<td>yh309</td>
<td>GsFT2c reverse</td>
<td>AAAGAGTGTGGGAGAAATTTTCTAG</td>
</tr>
<tr>
<td>yh385</td>
<td>GsFT2a reverse</td>
<td>GAGTGTGGAGATGGCAAT</td>
</tr>
<tr>
<td>yh387</td>
<td>GsFT2b reverse</td>
<td>GTCCCATTATTTTATAAGATCAAAA</td>
</tr>
</tbody>
</table>

Table 3.2. Primers used in this study

Table 3.3. Numbers of independent T1 transgenic lines