ELUCIDATING THE GENETIC MECHANISMS OF FLOWERING AND THE REPRESSION OF FLORAL INITIATION BY FRUIT IN APPLE (*MALUS* × *DOMESTICA* BORKH.)

By

Christopher Charles Gottschalk

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Plant Breeding, Genetics, and Biotechnology - Horticulture - Doctor of Philosophy

2020

ABSTRACT

ELUCIDATING THE GENETIC MECHANISMS OF FLOWERING AND THE REPRESSION OF FLORAL INITIATION BY FRUIT IN APPLE (*MALUS* × *DOMESTICA* BORKH.)

By

Christopher Charles Gottschalk

Many tree fruit crops exhibit yearly cyclical fluctuations in flowering and fruiting, including apple (*Malus* × *domestica* Borkh.). This phenomenon, known as biennial bearing, is an intriguing biological problem, as well as a production limitation of many high-value apple cultivars. The current hypotheses to explain biennial bearing focus on the repressive effects of fruit on the initiation of floral primordia, which would develop into flowers the following year. This floral repressive response could be the result of nutrient competition between the spur apex and the strong sink of the developing fruit, or flowering-inhibitory gibberellins (GAs) that are produced in the seeds of the developing fruit and exported to the spur apex. However, the molecular mechanism by which fruit load and/or GAs represses floral initiation is unknown.

The first aim of my dissertation was to identify the genes involved in the floral initiation pathway. Utilizing transcripts assembled from a transcriptome of the biennial apple cultivar Honeycrisp and the known flowering genes from Arabidopsis, I identified a comprehensive list of flowering-related homologs based on sequence homology, phylogenetic relationship, and syntenic organization. The second aim was to characterize the expression of the flowering-related homologs during the period of floral initiation under crop load conditions that either promote or repress flowering. Homologs of *AGL24/SVP*, *AP1*, *FT*, *LFY*, and *SPLs* were strongly expressed in apices of floral-induced trees, supporting their presumed role as floral promoters. In contrast, a homolog of the floral repressor *TERMINAL FLOWER 1* (*MdTFL1-2*), was strongly

up-regulated in apices of fruit-bearing, floral-repressed trees. Additionally, genes involved in the deactivation of bioactive GAs were strongly up-regulated in the floral-repressed trees.

The timing of floral initiation is also known to vary between cultivars in apple. However, the diversity in flowering gene regulation that underlies this variation is unknown. The third aim of my dissertation was to investigate diversity of gene expression in six different apple species and cultivars. Generally, the expression profiles of the flowering-related genes were similar, suggesting a widely conserved mechanism within the genus. However, a few key genes involved in the regulation of floral initiation and development exhibited differential expression. For example, during the floral initiation period, five of the six genotypes were found to have differentially expressed *MdTFL1s*. This result implicates a potential role for *MdTFL1s* in determining the timing of floral initiation in addition to a role in repressing flowering in response to crop load.

Plant growth regulators (PGRs) offer the potential to control flowering and biennial bearing in established apple plantings. My final aim was to evaluate and improve the efficacy of foliar-applied GA to repress flowering in apple. This study included the evaluation of application timings, GA types and application concentrations, and cultivar-specificity. The most successful PGR trial used GA_3 to inhibit flowering in 'Honeycrisp' when a strong return bloom was expected. As a result of the decreased flowering and fruit load, the harvested fruit in the year following GA application exhibited higher fruit-quality. The results from this collection of studies provide insight into the molecular control of flowering and biennial bearing while demonstrating a practical approach to managing flowering in a high-value apple cultivar.

Dedicated to my loving parents Dawn & Gerard Gottschalk. I also wish to thank my very influential grandparents: Charles & Nancy Curnow, and Charles & Eve Gottschalk.

ACKNOWLEDGEMENTS

I would like to especially thank my major professor, Dr. Steve van Nocker, for his expectational mentorship and for providing me the opportunity to pursue my graduate education. He gave me, a student that did not excel in their undergraduate education, a chance to prove themselves and obtain a Ph.D. He also allowed me endless opportunities to participate in impactful and interesting research, including the pursuit of my own research interests. In particular, I want to acknowledge his willingness to guide and mentor me in the construction of a long-term study to identify cider apples for production in Michigan.

There are many other mentors and scientists I would like to thank for their assistance in my education. My committee members of Drs. M. John Bukovac, Mitch McGrath, and Paolo Sabbatini, helped me develop a more thorough understanding of horticultural principles, plant physiology, and genetics. I would also like to thank the countless times Drs. Joseph Hill, Courtney Hollender, Ben Mansfeld, Shujun Oh, and Joshua VanderWiede came to my assistance. I am also forever grateful for all the collaboration from my counterpart in the lab - Songwen Zhang. We spent many hours working closely together and I am forever grateful for her collaboration. Lastly, I what to thank my fellow graduate students and friends for their support throughout my education: Patrick Abeli, Kristin Andersen, Kathleen Bitter-Rhodes, Charity Goeckeritz, and Emily Pawa.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

Literature Review

Introduction

Many fruit trees such as apple, pear, avocado, olive, and orange exhibit yearly cyclical fluctuations in flowering and fruiting. This phenomenon, termed 'alternate' or 'biennial' bearing, is an intriguing biological problem as well as a production limitation for the tree fruit industry. It is generally accepted that developing fruit inhibits the initiation of flowers, which would otherwise complete development and form fruit the following year. Two current hypotheses to explain biennial bearing focus on (I) nutrient competition between the shoot apex and the subtending fruit, and (II) flower-inhibitory gibberellins (GAs) produced in the developing seeds of the fruit and exported to the spur apex. How this physiology might be driven by or influence the expression of genes that control flowering remains relatively unexplored. In higher plants yet studied, flowering involves the activation of the meristem identity genes *LEAFY* (*LFY*) and *APETELA 1* (*AP1*) in the shoot apex, via the transmissible product of the *FLOWERING LOCUS T (FT)* gene produced in the leaves. A gene related to *FT*, *TERMINAL FLOWERING 1* (*TFL1*), may act in opposition to *FT* to inhibit floral initiation. Recent studies in several tree fruit species suggested that the effects of both crop load and floral-repressive GAs on floral initiation may be ultimately mediated by activity of apple genes homologous to *FT* and/or *TFL1*. However, this has not been rigorously established, due both to the incomplete characterization of these and other genes in the apple (*Malus* × *domestica* Borkh.) genome, and insufficient knowledge of the potential relationships between the underlying physiology and the activity of these genes.

The aim of this dissertation is to identify and describe the molecular and physiological mechanism(s) by which developing fruit inhibit floral initiation in apple. The first step was to rigorously identify flowering-related genes based on homology with known flowering genes from other plants. Phylogenetic techniques were used to identify the number and genomic

organization of the apple homologs of *FT*, *TFL1*, *LEAFY* and *AP1*, as well as other genes with presumed roles in floral initiation, using a recently released, high-quality sequence of the apple genome and exhaustive transcriptional data from the apple shoot apex. The second step was to characterize the variability in timing of floral initiation and development within the *Malus* genus. The third step was to evaluate expression of key flower inducting genes under varying crop load conditions. We used previously generated transcriptional data of the shoot apex and target high-resolution expression assays of leaf tissues to construct an atlas of expression for the identified flowering homologs. The fourth and final step was to evaluate the use of plant growth regulators (PGRs) to control floral initiation and, in turn, biennial bearing.

I hypothesize that developing fruit directly or indirectly repress expression of the *FT* gene in the leaf, leading to a loss of the transmissible floral promotion signal at the apex, and that FT transmission is dependent on the export of photosynthate from the leaves. I further hypothesize that developing fruit promotes expression of *TFL1* in the shoot apex, repressing any positive effect of FT. Finally, I hypothesize that the diversity of apple cultivars to bear annually vs biennially is partially due to the specific timing of floral initiation, in addition to allelic variation within the *FT*, *TFL1*, or other genes that may confer reduced expression and/or activity of these genes. The analyses conducted here provided an updated index of genes that may play central roles in flowering in apple, as well as establish a genetic framework to explain the influence of crop load on floral initiation. The results of this study may lead to new approaches to control biennial bearing in existing plantings and/or provide ideas of how to manipulate flowering, while facilitating the development of improved apple cultivars that are less prone to biennial bearing.

Background and significance

Fruit plays a vital role in the human diet by providing vitamins, sugars, and others nutrients (Seymour et al. 2013). Tree fruit, such as apples, contain some of the highest levels of dietary components (phytonutrients), such as quercetin glycosides, vitamin C, procyanidin, and chlorogenic acid (Boyer and Liu 2004). The cultivated apple is the third highest produced fruit crop worldwide and is the third most valuable fruit crop in the United States (FOA 2014; USDA 2017). However, consistent and sustainable apple production is limited by several aspects, ranging from pre-harvest tree management to post-harvest fruit storage. Many of the obstacles associated with production limitations are increasing in significance as a result of climate change. This limitations include spring frost events, emerging biotic and abiotic pressures, and shifting international trade policies (Cannell and Smith 1986; Lavee 2007).

The control of flowering is a major subject of research for improving commercial apple production due to the importance of precise timing of flowering on crop yield and quality. Three major aspects of flowering – juvenility/precocity, seasonal bloom timing, and biennial bearing - have special significance to apple production. Apple, like other woody perennial plants, transitions through a prolonged, 'juvenile' phase characterized by the inability to flower. This phase can last to over ten years when plants are grown from seed (Visser 1964). Following the juvenile phase, the plant transitions to the adult phase, in which flowering can be triggered by environmental cues such as seasonal temperatures and inductive photoperiod (Corbesier and Coupland 2005). Juvenility is a major limitation for the genetic improvement of apple, because breeders must wait several years after making a cross before the resulting progeny plants have acquired the competency to flower. However, juvenility also impacts the commercial production of apple. When adult-phase shoots are grafted onto rootstocks, as is the common practice for

clonal propagation of apple, the shoot passes through a juvenile-like phase before resuming flowering. The ability of grafted plants to quickly flower and produce a crop is termed precocity, and is a major area of research investigating the interaction between rootstock and grafted scion. In general, rootstocks that limit growth of the scion ('dwarfing rootstocks') provide the greatest precocity, but this effect is variable among rootstock genotypes (Webster et al. 1985), and growers must often wait several years before full production of a new planting is attained.

The second major aspect of flowering impacting apple production is the timing of spring bloom. Bloom timing is important for four primary reasons. First, apples are pollinated by insects, cultivars that bloom very early in the season - when conditions are still relatively cool - may not be adequately pollinated (McGregor 1976; Free 1993). Second, flowers can be easily damaged by early-season frost and freezes (Cannell and Smith 1986; Rodrigo 2000). Third, apples are typically self-incompatible, and so it is imperative to have two or more cultivars blooming synchronously (Lewis and Vincent 1909). Lastly, genotypes that bloom relatively late may be more prone to infection by the fire blight pathogen *Erwinia amylovora,* which prefers hot and humid conditions and easily infects susceptible cultivars through the flowers (Jones 1992). The timing of bloom in the spring is difficult to control as it is strongly influenced by both environment (temperature and/or photoperiod) and genetics (specific genetic alleles within the flowering and dormancy pathway) (Lawson et al. 1995; Liebhard et al. 2003; Celton et al. 2011; Gottschalk et al. 2013; Kurokura et al. 2013).

The third major aspect of flowering that has special significance for apple production is biennial bearing. Biennial bearing is the tendency to flower and produce fruit every other year (Butler 1917; Monselise and Goldschmidt 1982). This trait is mediated at the point of floral initiation, and is characterized by profuse flowering in a given year (the "on-year"), followed by

minimal flowering the subsequent year (the "off-year") (Jonkers 1979; Monselise and Goldschmidt 1982). The resulting variability in production creates instabilities in consumer supplies and income for producers (Williams and Edgerton 1974). Generally, the presence of developing fruit on spurs inhibits floral initiation within the adjacent bourse shoot (Fig. 1.1). Because apple flowers in a two year cycle, with flowers initiated in the first year and completing development (bloom) in the second year, this inhibition results in lack of flowers and fruit the year following a heavy crop load.

Hypotheses of biennial bearing

There are two popular hypotheses to explain the physiology that underlies biennial bearing. The first states that the strong sink strength established by developing fruit during an on-year diverts photosynthates and nutrients away from the shoot apex of the developing spur, where floral initiation would occur (Fig. 1.2). The resulting low concentration of photosynthate in the spur apex then leads to a vegetative fate determination (Wardlaw 1990). In contrast, during an off-year, photosynthates accumulate to relatively high levels, promoting floral initiation. The observation that a high crop load is associated with diminished shoot growth (Quinlan and Preston 1971) demonstrates that there is a competitive interaction between the sink strengths of fruit and bourse shoots, and lends some support to this hypothesis. In addition, it has been shown that floral initiation can be inhibited by removal of bourse shoot leaves, a condition that limits photosynthate availability (Fulford 1966b, Elsysy et al. 2019). In contrast, girdling the branches of mature apple trees, which should result in photosynthate accumulation in the girdled branch and corresponding spur apices, is associated with increased floral initiation (Dennis and Edgerton 1966). These studies manipulated the source and sink balance, suggesting that

photosynthates play a major role in floral initiation. However, they failed to uncouple the potential role of ‘Florigen’, which is also produced in leaves.

The second hypothesis states that excessive floral-repressive GAs are produced in developing seeds during fruit development and are exported up to the spur apex (Fig. 1.2). This hypothesis is supported by studies in the cultivar 'Spencer Seedless’, which naturally produces parthenocarpic (seedless) fruit but can be induced to produce seeds by hand pollination. Under natural conditions, when seedless fruit is produced, this variety tends to bear annually. However, when fruit was manipulated to produce seeds, reduced floral initiation was observed (Neilsen and Dennis 1999). Developing seeds are known to be a rich source of GAs, a class of phytohormone that generally promotes stem elongation and leaf expansion (Olszewski et al. 2002). Apple seeds have been found to contain large quantities of two forms of bioactive GAs, GA_4 and GA_7 (Luckwill et al. 1969). However, because fruit growth is stimulated by seeds via GAs, it has not been clear whether the floral-repressive effects are mediated directly by the GAs or indirectly through the increased sink strength of the larger fruit. Application of exogenous GA can inhibit flower initiation in apple if the GA is applied during the anticipated period of floral initiation; this effect is independent of the presence of fruit, suggesting that it influences flowering more directly (Guttridge 1962; Schmidt et al. 2009). In order for seed-produced GAs to directly inhibit flowering, they would need to be translocated to the site of initiation, the shoot apex (Fig. 1.2). This translocation of GA would be counter to the expected strong flow of solutes and nutrients to the fruit (*i.e.* through the xylem and phloem tissues), transmission of GAs from seed to apex seems unlikely. Direct involvement of transmissible GAs from the fruit also seems unlikely because developing leaves, which are present in both fruiting and non-fruiting bourse shoots, are also known to generate GAs (Grauslund 1972). An argument against transmission of active GAs

is several *GA2 OXIDASE* genes, which catabolize active GAs, are strongly expressed in the base of the fruit pedicel and thus are expected to form a barrier to GA diffusion (Songwen Zhang and Steve van Nocker, unpublished).

The tendency to bear annually or biennially is an inherent characteristic of specific cultivars, indicating a strong genetic basis for this trait. For example, numerous cultivars (*e.g.* Fuji and Honeycrisp) are known to be biennial, whereas others (*e.g.* Gala) typically bear annually (Pellerin et al. 2011). However, environmental and physiological conditions can influence biennial bearing indirectly through effects on flowering and/or crop load. Factors that reduce crop load, such as a hard freeze during bloom or poor pollination, can moderate the amplitude of the bearing cycle, while favorable conditions that promote high crop load, can exacerbate the effect.

Management of biennial bearing

The impact of biennial bearing on production has typically been reduced through the thinning of excess flowers or fruit during on-years (Downing 1900; Dennis 2000; Tromp 2000). Fruit thinning is the primary commercial practice to address subsequent-year flowering (referred to as 'return bloom') but also to maximize fruit size and quality (Link 2000). In regards to biennial bearing, fruit thinning is only effective at promoting a return bloom when carried out before a specific stage of fruit development (Byers and Carbaugh 2002). As a result, commercial apple producers will conduct thinning numerous times across the growing season, addressing the effects of biennial bearing and fruit quality improvement separately.

The application of GA formulations to trees in their off-year has potential to reduce overcropping and the need for thinning. For example, application of GA_3 to five-year old 'Fuji' trees during the anticipated period of floral initiation resulted in a strong (~55%) reduction in

return bloom (Zhang et al. 2016). Bertelsen et al. (2002) reported decreases of ~10 - 20% in the number of flowering buds in 'Pacific Rose' following GA_3 or GA_{4+7} application during an off year. Bertelsen and Tustin (2002), found that GA_3 applied two weeks after full bloom led to a 60% decrease in flowering the following year. Similarly, McArtney and Li (1998) reported that both GA_3 and GA_7 inhibited floral buds on 'Braeburn' by 28 and 38%, respectively. However, repression of return bloom by application of GAs has yet to be implemented as a standard practice, at least in the U.S. This is most likely due to the fact that many other trials have had variable, and often ineffective results, high price of GAs, and the potential loss of a crop by a frost event. For example, Schmidt et al. (2009) found that application of GA_{4+7} to 'Cameo' and 'Fuji' trees in the off-year caused only a small reduction (5% to 30%) in flowering. Year-to-year variation was also reported, where GA_4 applications to the same group of 'Golden Delicious' trees for four consecutive years resulted in decreased flowering only in one year, with no effect or even increases in flowering the remaining years (Greene 1993).

In addition to GAs, synthetic GA biosynthesis inhibitors have been used in attempts to manipulate flowering. The expectation is that, if GAs repress flowering, compounds that interfere with GA biosynthesis or signaling should promote flowering. This hypothesis has been supported through evidence obtained from PGR trials of GA inhibitors. The application of paclobutrazol (PAC) to 'Fuji' in an on year increased flowering by ~25% in the subsequent year (Zhang et al. 2016). PAC, although approved as a PGR for floriculture applications, has yet to be approved for use in apple production by the US EPA. Another inhibitor, Daminozide (Alar), has also been used successfully to promote return bloom (Edgerton and Hoffman 1965; Rogers and Thompson 1968; Looney 1969). However, the use of Alar in apple is now restricted due to its potential to be a human carcinogen (EPA 1989).

Other PGRs have been found to influence flowering in apple, mostly with inconsistent results. 6-benzylaminopurine, a synthetic cytokinin, was used in combination with Alar on blooming branches of 'Baldwin', 'Delicious', and 'McIntosh' resulting in a significant increase to flowering the subsequent year (McLaughlin and Greene 1991). Ethephon, a compound that breaks down in plant tissues to produce ethylene, reduced flowering after an off-year (Bukovac et al. 2006). However, ethephon also increased flowering when a low bloom density was expected after an on-year (Schmidt et al 2009). Naphthaleneacetic acid (NAA), a synthetic auxin, increased return bloom under some conditions but was variable in efficacy over different cultivars and application intervals (McArtney et al. 2007). Combining NAA and Ethephon resulted in a significant increase in floral initiation over thinning alone in biennial bearing 'York' (McArtney et al. 2013). In many of these studies it is not clear whether the increase in bloom was due to the PGR application or due to fruit thinning induced by the PGRs.

Genetics of floral initiation

Ultimately, the mechanism underlying biennial bearing must be governed by the molecular-genetic pathways controlling floral initiation. A working model for the flowering genetic network in apple can be proposed based on the known activities and interactions of flowering genes in the research reference plant *Arabidopsis thaliana* and the known functional conservation of these genes across plant species (Corbesier and Coupland 2005) (Fig. 1.3). In Arabidopsis, the *GIGANTEA* (*GI*) gene mediates output from a circadian clock and cooperates with light quality (wavelength) to regulate the diurnal expression of *CONSTANS* (*CO). CO* promotes expression of the *FT* gene under inductive photoperiods, and, in conjunction with the gibberellin pathway, also activates *SUPPRESSOR OF CONSTANS 1* (*SOC1). SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL)* genes act as additional promoters of *FT, SOC1,*

and the floral meristem identity gene *AP1*. *SPL* gene expression is post-transcriptionally repressed by a class of microRNA, *MIR156/157*, during the very short juvenile-like phase. *FT* expression is primarily repressed by *FLOWERING LOCUS C* (*FLC*), in the absence of vernalizing cold and, thus, mediates the promotive effects of cold temperatures on flowering (vernalization). *FT* and *SOC1* cooperatively activate *AP1* and *LFY*, which together direct floral initiation and flower formation. The *TERMINAL FLOWER 1* (*TFL1*) gene represses *LFY* and *AP1* in the center of the inflorescence meristem, thus preventing formation of a terminal flower and enabling indeterminate flower formation. This effect is the result of *TFL1* competing with *FT* in forming a transcriptional regulator complex with *FLOWERING LOCUS D*.

Functional homologs of these genes have been identified in many plant species. *GI* has been extensively studied in rice for its roles as a major regulator of the circadian rhythm and promoter of additional flowering genes (Matsuzaki et al. 2015; Hayama et al. 2002). *CO* has been studied to understand how sequence and functional divergence has given rise to the varied flowering times as a response to photoperiod across diverse plant species (Griffiths et al. 2003; Campoli et al. 2012). In poplar, two copies of the *FT* gene have been identified (Böhlenius et al., 2006). *FT1* promotes reproductive onset in response to winter temperatures, whereas *FT2* promotes vegetative growth during warm, long days of the growing season (Hsu et al. 2011). Poplar also contains two copies of *TFL1*, *PopCEN1* and *PopCEN2*. Both *TFL1* orthologs control shoot meristem identity by repressing the developmental transition to mature flowering shoots (Mohamed et al., 2010). Developmental regulation of these *TFL1* orthologs was found to be distinct; *PopCEN1* was expressed in shoot tips and vegetative buds, whereas *PopCEN2* was expressed in the stem, leaf blade, petiole and immature inflorescence (Mohamed et al. 2010). Silencing of both *PopCEN* paralogs by RNAi was associated with flowering after only two years

of growth, whereas non-transgenic plants took five years. The catkins of the *CEN* RNAi lines were branch-like in appearance suggesting the apex underwent a reversion towards a vegetative identity during development (Mohamed et al. 2010). A *SOC1* ortholog has been found in both rice and citrus, and is functionally conserved with Arabidopsis (Lee et al. 2004; Tan and Swain 2007). In rice, a suite of *AP1* orthologs have been identified and characterized for their role in meristem identity (Kobayashi et al. 2012). In poplar, a potential functional homolog of *AP1* has been identified as a downstream effector of *FT2* in the flowering pathway (Hsu et al. 2006). Orthologs of *LFY* have been extensively studied in a wide range of plants from tomato to pines, where they are conserved meristem identity genes (Mellerowicz et al. 1998; Molinero-Rosales et al. 1999).

In apple, several of the known Arabidopsis flowering genes have clear structural homologs and exhibit conserved function (Kotoda et al. 2000; Kotoda et al. 2002; Wada et al. 2002; Kotoda et al. 2010; Trankner et al. 2010; Guitton et al. 2012, 2016). An exception is *FLC*; although apple contains several genes showing some sequence homology with *FLC*, a definitive phylogenetic relationship between these genes and *FLC* is still lacking (Porto et al. 2015; Kumar et al. 2016; Peace et al. 2019). In apple, the existence of two *FT*-like homologs has been reported (Kotoda et al. 2010). One *FT* homolog was expressed diurnally in leaves where its expression increased throughout the day and peaked at night (Trankner et al. 2010). Constitutive expression of one of these apple *FT* genes led to an early flowering phenotype in transgenic Arabidopsis, poplar and apple (Trankner et al. 2010). Apple FT protein has also been implicated in transcriptionally regulating a number of different pathways besides flowering, such as cell growth and organ development (Mimida et al. 2011a). Two *TERMINAL FLOWER 1 (TFL1)* homologs, *MdTFL1-1* and *MdTFL1-2*, have been identified in apple (Mimida et al. 2009).

Knock-down experiments targeting *MdTFL1-1* using RNAi resulted in decreased time to flower, reduced vegetative growth and singular terminal flowers (Flachowsky et al. 2012). This suggests that, at least, *MdTFL1-1* is operating as a floral repressor and governor of inflorescence architecture in apple.

Influence of crop load on flowering genes

Research conducted in other perennial fruit trees species on the molecular genetics of biennial bearing suggests mechanism(s) that may also operate in apple. In biennial avocado and citrus, endogenous *FT* expression in leaves during the off-year was markedly higher than that in the on-year, suggesting that crop load mediates its expression (Munoz-Fambuena et al. 2011; Ziv et al. 2014). However in two studies with biennial apples, the level of *FT* expression in the leaves showed no difference between on- and off-years (Kittikorn et al. 2011; Haberman et al. 2016). *MdTFL1* expression has been reported to be higher in apical buds in trees carrying a heavy crop versus trees that were subjected to thinning (Kittikorn et al. 2011). Haberman et al. (2016) found that *MdTFL1-1* expression dramatically decreased during the course of the growing season, and this decrease occurred earlier in the season in on-year trees than in off-year trees, with *MdTFL1-2* significantly increasing in expression later in the season in the on-year trees. This suggests that *MdTFL1-1* might play a role in maintaining the apex in a vegetative state during the early growing season, whereas *MdTFL1-2* might repress flowering in response to high crop load. During off-years, citrus trees showed increased expression of multiple *SPL* genes in the shoot apex (Wu and Poethig 2006; Shalom et al. 2012). Similarly, in apple, two *SPL* genes were expressed to higher levels in apices of trees that were thinned of fruit, relative to non-thinned trees (Guitton et al. 2016). This suggests that the influence of crop load is upstream of these genes.

If the tendency to bear biennially can be attributed to floral-repressive signals originating in the fruit and acting at the shoot apex, then annual bearing could be explained in a number of ways. For example, the repressive signal may be absent or weaker in annual cultivars. Alternatively, the signal may not be routed correctly to the shoot apex. Another possibility is that the signal is transmitted to the shoot apex, but that the apex is not programmed to respond to this signal. There is also the possibility that floral initiation depends on coincidence of a fruit-produced signal and a signal-sensitive phase of apex development, and that this is lost in annual cultivars.

Hypothesis of different floral initiation timings and biennial bearing

I hypothesize that, in biennial cultivars, the period of initiation overlaps with a period of fruit induced influencing of floral promoters/repressors. With annual cultivars, their floral initiation period might fall outside the range of time in which fruit influences the floral promoters/repressors activity. Previous research has identified highly variable timings for floral initiation with variation observed within the same tree, across seasons, and between cultivars (McArtney et al. 2001; Foster et al. 2003; Hoover et al. 2004; Dadpour et al. 2011). As a point of reference for this variation, the cultivar ‘Gala’ has been featured as a prominent cultivar of study for floral initiation. The onset of initiation has been seen to range from as early 72 days after full bloom (DAFB) to as late as 99 DAFB in the following season (McArtney et al. 2001). Another report, from the same orchard, found ranges of initiation between 96 to 109 DAFB in ‘Gala’ trees (Foster et al. 2003). However, all of these measurements are based on documented morphological changes observed in the structure of the meristem.

The morphological changes that is indicate floral initiation, as described by Foster et al. (2003), begins with Stage 1 where the apex is flat (vegetative) but is undergoing an increase in

diameter. This increase in diameter signifies a commitment to floral development. At Stage 2, the apex is no longer flat and appears "domed" shaped, indicating that its identity has transitioned to an reproductive meristem. Identification of associated changes in gene expression at this critical period of time between Stages 1 and 2 could identify the molecular signature of floral initiation and is currently unknown. Foster et al. (2003) also reported observing Stage 1 apices as early as -3 DAFB and as late as 96 DAFB, and 100% percent of the Stage 1 meristems continued through Stage 3, when floral buds begin to form. This suggests that competency to transition can extend for a long period of the growing season. This potential long competency period does not seem to completely align with what is known about how the presence of fruit affects floral initiation, especially in regards to the effective periods of fruit thinning to promote bloom. As reviewed by Jonkers (1979), thinning to promote a return bloom is only effective through the first month following bloom (~30 DAFB). These observations taken together could indicate that the potential for a floral initiation (or repression) signal occurs earlier in the season, before the time in which morphological signs of initiation are evident.

Rationale

Although the phenomenon of biennial bearing has been studied for more than 50 years at the physiological level, there have been few studies at the molecular and genetic levels. Research based in fundamental gene regulation has the potential to quickly advance our understanding of biennial bearing because genetic pathways of flowering have already been drafted in apple. Previous research on the molecular genetics of flowering in apple have been limited by the use of a poor quality reference genome and low-resolution molecular techniques. The experiments following utilized a new high quality reference genome and high-resolution methods to survey the apex and leaves of the bourse shoot for effects of crop load on floral initiation.

A prerequisite step to understand why fruit suppresses flowering in apple is to identify flowering-control genes in the apple genome, and document their expression profiles. Previous studies have already recognized homologs of *AP1, FT, LFY,* and *TFL1*, and have taken preliminary steps to understand their expression and function (Kotoda et al. 2000, 2010; Wada et al. 2002; Mimida et al. 2011b; Flachowsky et al. 2012). However, these studies were obfuscated by the presence of homeologous genes and/or allelic variants, and the use of molecular techniques such as quantitative PCR and microarrays that could not discriminate among them. Although more recent work has begun to employ direct transcriptome sequencing, the analysis to this date was hindered by the poor quality of the previous (pre-2017) apple draft genome (Velasco et al. 2010). A recent release of a new high quality hybrid apple genome (Daccord et al. 2017) coupled with high-resolution RNA-sequencing can provide for an improved census of the flowering genetic network and estimation of expression patterns of these genes.

The *FT* and *TFL1* genes are the focus of the following experiments, given the primary importance of the *FT* gene and its structural and functional conservation. Two *FT* paralogs were reported in apple (Kotoda et al. 2010). Expression of *FT1* was reported to be highest in shoot apices transitioning to flowering, whereas expression of *FT2* was highest in floral organs. Both paralogs exhibited only very low expression levels in mature leaves (Kotoda et al. 2010). If *FT* function is conserved between Arabidopsis and apple, strong expression in apple leaves would be expected. Kotoda et al. (2010) does not specify the developmental stage, age, or type (spur, bourse shoot or elongated shoot) of the mature leaves that were analyzed, and this is an important detail because previous studies showed that various leaf types have different strengths of influence on flowering (Fulford 1965, 1966a, 1966b, Elsysy et al. 2019). In contrast, the study by Haberman et al. (2016) analyzed those leaves that were expected to influence flowering the

most, but did not find a difference in *FT* expression under varying crops loads. However this study used PCR primers that were not expected to discriminate between the potential *FT* paralogs. Expressing apple *FT* to high levels in transgenic apple resulted in early flowering, suggesting that *FT* has potential to promote flowering under natural conditions (Flachowsky et al. 2012; Trankner et al. 2010). However, these experiments were done in juvenile trees, so it is not known whether the influence of *FT* could be blocked by fruit in adult trees. In addition, the dynamics of *FT* trafficking from leaf to apex has yet to be explored in apple.

TFL1, a regulator of shoot architecture in Arabidopsis, has been implicated as a potential floral repressor in apple (Flachowsky et al. 2012). *TFL1* and *FT* are homologous to one another, and one possibility is that these two genes work antagonistically. *TFL1* expression in the apex was reported to increase in spring, peak in June, and then decline through August (Mimida et al. 2011b). This period of expression overlaps with the anticipated period of floral initiation. *TFL1* was also found to be strongly expressed in the apex of juvenile trees (Mimida et al. 2011b), potentially serving as a mechanism to maintain vegetative growth. Apple *TFL1* was claimed to be a floral repressor, because when *TFL1* was silenced in transgenic apple lines the plant flowered early (Flachowsky et al. 2012). However, early flowering was also observed in Arabidopsis when Arabidopsis *TFL1* was silenced, even though the primary role of Arabidopsis *TFL1* is to maintain indeterminate inflorescence architecture. Taken together, this evidence is not sufficient to distinguish between roles in flowering or inflorescence structure. It has been reported that *TFL1* is expressed to higher levels in apices of fruited trees, relative to thinned trees (Kittikorn et al. 2011; Haberman et al. 2016). However, these studies also used low-resolution techniques that were not expected to discriminate between *TFL1* and paralogous genes. These studies also did not consider other potential floral promoters or repressors.

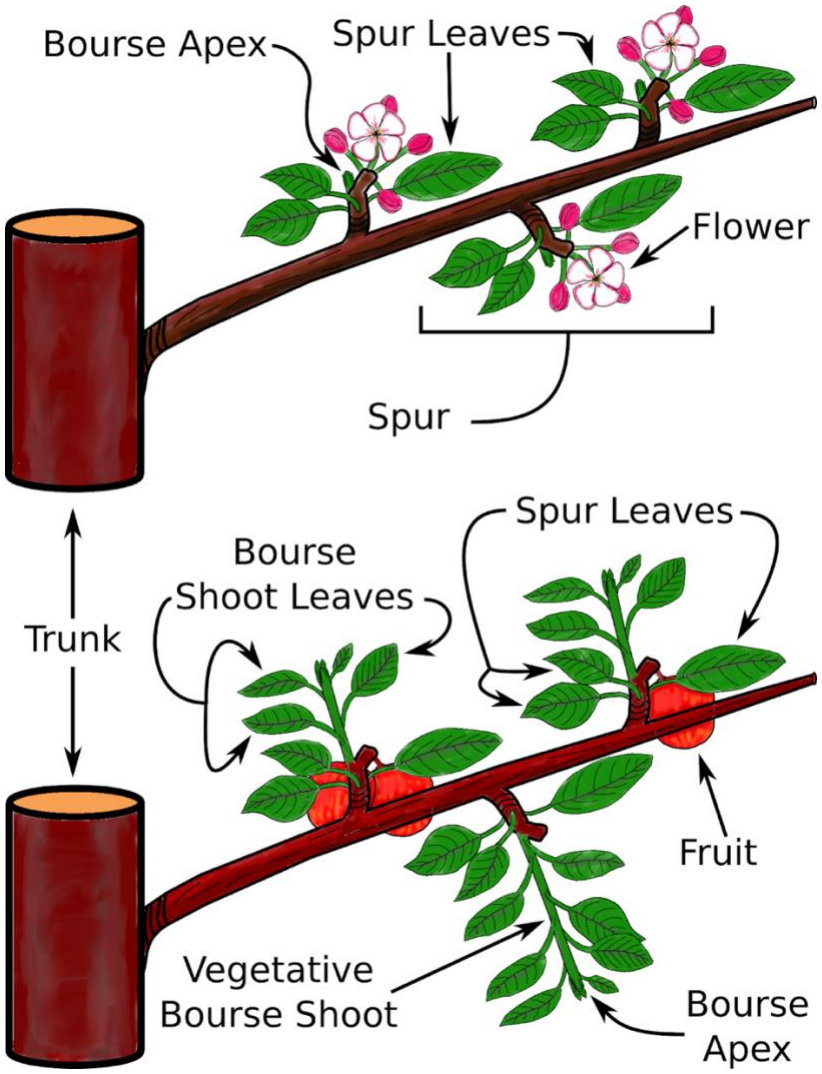


Figure 1.1. Apple shoot architecture. (Top) In the spring, spurs release from dormancy expanding the spur leaves and flowers that were initiated the prior season. (Bottom) After fruit set, the bourse shoot apex begins to grow vegetatively producing bourse shoot leaves while the spur leaves are maintained.

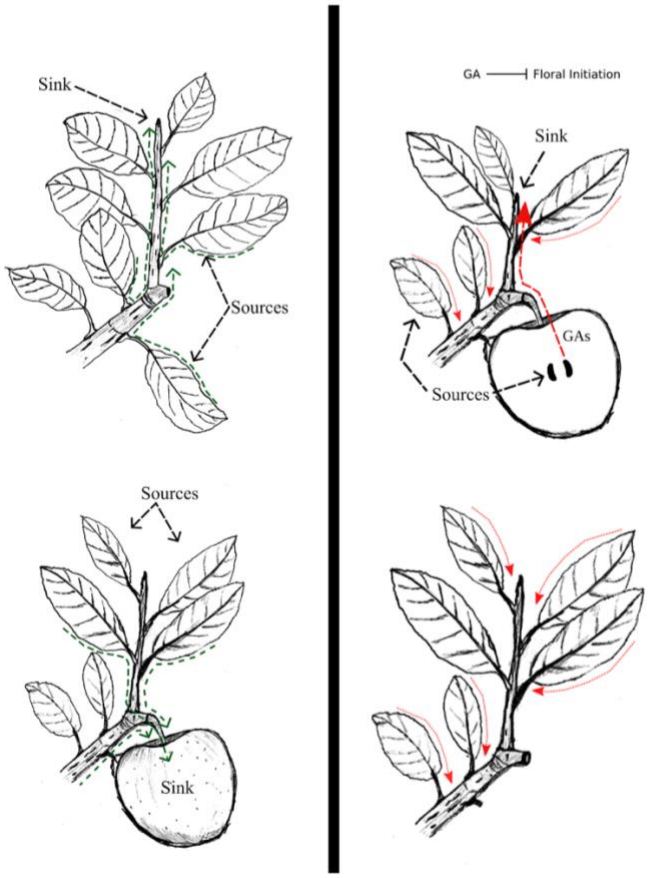


Figure 1.2. Potential roles of fruit in repression of flowering in apple. (Left panel) Nutrient competition. In the presence of developing fruit (bottom), photosynthates are transported from their site of production in the leaf into the fruit, and the deficit of photosynthates in the apex suppresses floral initiation. When no fruit is present (top) the photosynthates are instead transported to the apex, allowing for floral initiation. (Right panel) GA-mediated. In the presence of developing fruit (top) GA produced by the developing seeds is transferred from the fruit to the apex where it suppresses floral initiation. When no fruit is present (bottom) there is not enough endogenously produced GA to repress floral initiation.

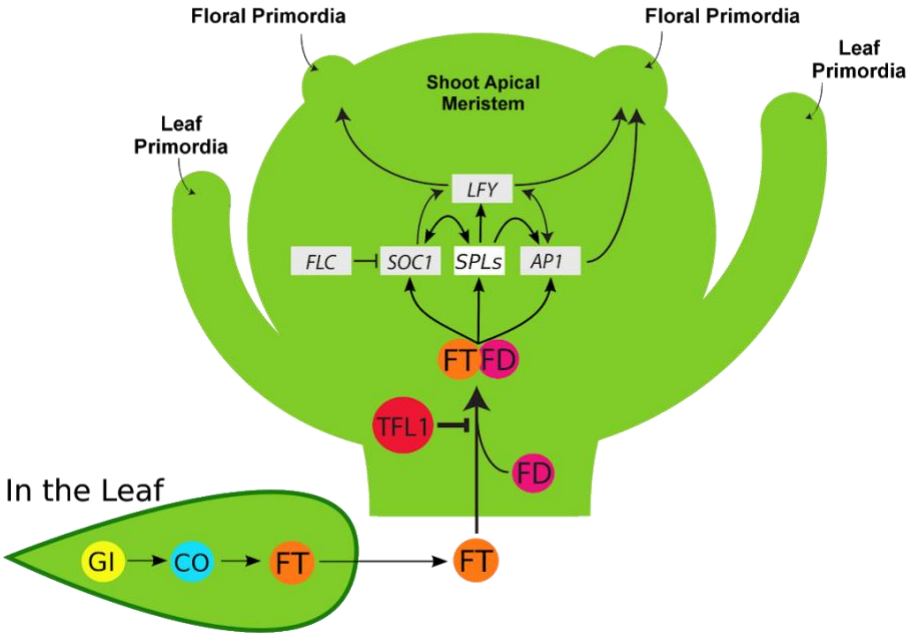


Figure 1.3. Simplified model of the molecular-genetic pathway that regulates floral initiation in Arabidopsis. Inductive photoperiods promotes the production of FT in the leaf, which is transported to the shoot apex. In the shoot apex, FT activates flowering genes such as *LEAFY* and *AP1* cooperatively with the FD protein. TFL1 acts as a competitive antagonist to FT. For a detailed description of the individual role of these genes, see text.

REFERENCES

REFERENCES

Bertelsen MG, Tustin DS and Waagepetersen RP. 2002. Effects of GA_3 and GA_{4+7} on early bud development of apple. J Hort Sci and Biotech 77(1):83-90.

Bertelsen MG and Tustin DS. 2002. Suppression of flower bud formation in light cropping trees of 'Pacific Rose' apple using gibberellin sprays. J Hort Sci and Biotech 77(6): 753-757.

Böhlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH and Nilsson O. 2006. CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. Science 312:1040-1043.

Boyer J and Liu RH. 2004. Apple phytochemicals and their health benefits. Nutrition Journal 3(5).

Bukovac MJ, Sabbatini P and Schwallier PG. 2006. Modifying Alternate Bearing of Spur-Type 'Delicious' Apple with Ethephon. HortScience 41(7):1606-1611.

Butler O. 1917. On the cause of alternate bearing in apple. Bulletin of the Torrey Botanical Club 44(2):85-96.

Byers RE and Carbaugh DH. 2002. Effects of Thinning Time on Yield, Fruit Size, and Return Bloom of 'York' and 'Golden Delicious' Apple Trees. J. of Tree Fruit Production 3(1): 55-62.

Campoli C, Drosse B, Searle I, Coupland G and Korff M. 2012. Functional characterisation of *HvCO1*, the barley (*Hordeum vulgare*) flowering time ortholog of *CONSTANS.* Plant J 69(5):868-880.

Cannell MGR and Smith RI. 1986. Climatic warming, spring budburst and frost damage on trees. J Applied Ecology 23:177-191.

Celton JM, Martinez S, Jammes MJ, Bechti A, Salvi S, Legave JM, and Costes E. 2011. Deciphering the genetic determinism of bud phenology in apple progenies: a new insight into chilling and heat requirement effects on flowering dates and positional candidate genes. New Phytologist 192:378-392.

Corbesier L and Coupland G. 2005. Photoperiodic flowering of Arabidopsis: integrating genetic and physiological approaches to characterization of the floral stimulus. Plant, Cell and Environment 28:54-66.

Daccord N, Celton JM, Linsmith G, Becker C, Choisne N, Schijlen E, van de Geest H, Bianco L, Micheletti D, Velasco R, Di Pierro EA, Gouzy J, Rees JG, Guérif P, Muranty H, Durel CE, Laurens F, Lespinasse Y, Gaillard S, Aubourg S, Quesneville H, Weigel D, van de Weg E, Troggie M, and Bucher E. 2017. High-qualty *de novo* assembly of the apple genome and methylome dynamics of early fruit development. Nature Genetics 49:1099-1106.

Dadpour MR, Movafeghi A, Grigorian W, and Omidi Y. 2011. Determination of floral initiation in *Malus domestica:* A novel morphogenetic approach. Biologia Plantarum 55(2): 243-252.

Dennis FG. 2000. The history of fruit thinning. Plant Growth Regulation 31(1):1-16.

Dennis FG and Edgerton LH. 1966. Effects of gibberellins and ringing upon apple fruit development and flower bud formation. Prco Am Soc Hort Sci 88:14-24.

Downing AJ. 1900. *The fruits and fruit trees of America*. Wiley, New York.

Edgerton LJ and Hoffman MB. 1965. Some physiological response of apple to N-dimethyl amino succinamic acid and other growth regulators. Proc Am Soc Hort Sci 86:28-36.

EPA. 1989. Daminozide (Alar) Pesticide Canceled for Food Use. November 7, 1989. https://www.epa.gov/aboutepa/daminozide-alar-pesticide-canceled-food-uses

Flachowsky H, Szankowski I, Waidmann S, Peil A, Tranker C, and Hanke MV. 2012. The *MdTFL1* gene of apple (*Malus* × *domestica* Borkh.) reduces vegetative growth and generation time. Tree Physiology 32; 1288-1301.

FOA. 2008. Food and Agriculture Organization of the United Nations, FAOSTAT database (FAOSTAT, 2008), http://faostat.fao.org/.

Foster T, Johnston R and Seleznyova A. 2003. A Morphological and Quantitative Characterization of Early Floral Development in Apple (*Malus* × *domestica* Borkh.). Annals of Botany 92:199-206.

Free JB. 1993. Insect pollination of crops, 2nd ed. Academic, London.

Fulford RM. 1965. The morphogenesis of apple buds. I. The activity of the apical meristem. Annal of Botany 29:167-180.

Fulford RM. 1966a. The morphogenesis of apple buds. III. The inception of flower. Annal of Botany 30:207-219.

Fulford RM 1966b. The morphogenesis of apple buds. IV. The effect of fruit. Annal of Botany 30:597-606.

Gottschalk C and van Nocker S. 2013. Diversity in Seasonal Bloom Time and Floral Development among Apple Species and Hybrids. J Amer Soc Hort Sci 138(5):367-374.

Greene DW. 1993. Effects of GA_4 and GA_7 on flower bud formation and russet development on apple. J Hort Sci 68: 171-176.

Griffiths S, Dunford RP, Coupland G and Laurie DA. 2003. The evolution of CONSTANS-like gene families in barley, rice and Arabidopsis. Plant Physiol 131(4):1855-1867.

Grauslund J. 1972. Gibberellins in diffusates from shoots of apple trees. Physiologia Plantarum 27:65-70.

Guitton B, Kelner JJ, Velasco R, Gardiner SE, Chagne D and Costes E. 2012. Genetic Control of biennial bearing in apple. J Exp Botany 63(1):131-149.

Guitton B, Kelner JJ, Celton JM, Sabau X, Renou JP, Chagne D and Costes E. 2016. Analysis of transcripts differentially expressed between fruited and deflowered 'Gala' adult trees: a contribution to biennial bearing understanding in apple. BMC Plant Bio 16:55.

Guttridge CG. 1962. Inhibition of Fruit-Bud Formation in Apple with Gibberellic Acid. Nature 196:1008.

Haberman A, Ackerman M, Crane O, Kelner JJ, Costes E and Samach A. 2016. Different flowering response to various fruit loads in apple cultivars correlates with degree of transcript reaccumulation of a TFL1-encoding gene. The Plant J 87(2): 161-173.

Hayama R, Izawa T and Shimamoto K. 2002. Isolation of Rice Genes Possibly Involved in the Photoperiodic Control of Flowering by a Fluorescent Differential Display Method. Plant & Cell Physiology 43(5):494-504.

Hoover E, De Silva N, McArtney S, and Hirst P. 2004. Bud development and floral morphogenesis in four apple cultivars. J Hort Sci Biotech 79(6): 981-984.

Hsu CY, Liu Y, Luthe DA, and Yuceer C. 2006. Poplar *FT2* Shortens the Juvenile Phase and Promotes Seasonal Flowering. Plant Cell 18(8):1846-1861.

Hsu CY, Adams JP, Kim HJ, No K, Ma CP, Strauss SH, Drnevich J, Vandervelde L, Ellis JD, Rice BM, Wickett N, Gunter LE, Tuskan GA, Brunner AM, Page GP, Barakat A, Carlson JE, dePamphilis CW, Luthe DS and Yuceer C. 2011. *Flowering Locus T* duplication coordinates reproductive and vegetative growth in perennial poplar. PNAS 108(26):10756-10761.

Jones AL. 1992. Evaluation of the Computer Model MARYBLYT for Predicting Fire Blight Blossom Infection on Apple in Michigan. Plant Disease 76(4): 344-347.

Jonkers H. 1979. Biennial bearing in apple and pear: a literature survey. Scientia Horticulturae 11:303-317.

Kittikorn M, Okawa K, Ohara H, Kotoda N, Wada M, Yokoyama M, Ifuku O, Yoshida S, and Kondo S. 2011. Effects of fruit load, shading, and 9,10-ketol-octadecadienoic acid (KODA) application on MdTFL1 and MDFT1 genes in apple buds. Plant Growth Regulator 64:75-81.

Kobayashi K, Yasuno N, Sato Y, Yoda M, Yamazaki R, Kimizu M, Yoshida H, Nagamura Y, and Kyozuka J. 2012. Inflorescence meristem identity in rice is specified by overlapping functions of three AP1/FUL-like MADS box genes and PAP2, a SEPALLLATA MADS box gene. Plant Cell 24(5):1848-59.

Kotoda N, Wada M, Komori S, Kidou S, Abe K, Masuda T and Soejima J. 2000. Expression Pattern of Homologues of Floral Meristem Identity Genes *LFY* and *AP1* during Flower Development in Apple. J Amer Soc Hort Sci 125(4):398-403.

Kotoda N, Wada M, Kusaba S, Kano-Murakami Y, Masuda T and Soejima J. 2002. Overexpression of *MdMADS5*, and *APETALA1-Like* gene of apple, causes early flowering in transgenic Arabidopsis. Plant Sci 162:679-687.

Kotoda N, Hayashi H, Suzuki M, Igarashi M, Hatsuyama Y, Kidou S, Igasaki T, Nishiguchi M, Yano K, Shimizu T, Takahashi S, Iwanami H, Moriya S and Abe K. 2010. Molecular Characterization of *FLOWERING LOCUS T-Like* Genes of Apple (*Malus* × *domestica* Borkh.). Plant Cell Physiol 51(4):561-575.

Kumar G, Arya P, Gupta K, Randhawa V, Acharya V and Singh AK. 2016. Comparative phylogenetic analysis and transcriptional profiling of MADS-box gene family identified *DAM* and *FLC*-like genes in apple (*Malus* × *domestica).* Sci Reports 6:20695.

Kurokura T, Mimida N, Battey NH and Hytönen T. 2013. The regulation of seasonal flowering in the Rosaceae. J Exp Bot 64 (14): 4131-4141.

Lawson DM, Hemmat M, and Weeden NF. 1995. The use of molecular markers to analyze the inheritance of Morphological and developmental traits in apple. J Amer Soc Hort Sci 120(3): 532-537.

Lavee S. 2007. Biennial Bearing in Olive (*Olea europaea*). Annales, Series Historia Naturalis 17(1):101-112.

Lee SY, Kim J, Han JJ, Han MJ and An GH. 2004. Functional analyses of the flowering time gene *OsMADS50*, the putative *SUPPRESSOR of OVEREXPRESSION of CO1*/*AGAMOUS*-like 20 (*SOC1/AGL20*) ortholog in rice. Plant J38(5):754-764.

Lewis C.I. and Vincent C.C. Pollination of the Apple. 1909. Oregon Agricultural College Experiment Station Bulletin no. 104.

Liebhard R, Kellerhals M, Pfammatter W, Jertmini M, and Gessler C. 2003. Mapping quantitative physiological traits in apple (*Malus* × *domestica* Borkh.). Plant Molecular Bio52:511-526.

Link H. 2000. Significance of flower and fruit thinning on fruit quality. Plant Growth Regulation 31:18-26.

Looney NE. 1969. Importance of time and concentration of annual applications of succinic acid 2,2-dimethyl hydrazide (Alar) to pome fruits. I. McIntosh and Spartan apples. Can J Plant Sci 49:489-494.

Luckwill LC, Weaver P and MacMillan J. 1969. Gibberellins and other growth hormones in apple seeds. J Hort Sci 44:413-424.

Matsuzaki J, Kawahara Y and Izawa T. 2015. Punctual Transcriptional Regulation by the Rice Circadian Clock under Fluctuating Field Conditions. Plant Cell 27(3):633-648.

McArtney S, and Li S. 1998. Selective Inhibition of Flowering on 'Braeburn' Apple Treees with Gibberellins. HortScience 33(4): 699-700.

McArtney S, Hoover E, Hirst P and Booking IR. 2001. Seasonal variation in the onset and duration of flower development in 'Royal Gala' apple buds. J Hort Sci Biotech 76(5): 536-540.

McArtney S, Unrath D, Obermiller JD and Green A. 2007. Naphthaleneacetic Acid, Ethephon, and Gibberellin A4 + A7 Have Variable Effects on Flesh Firmness and Return Bloom of Apple. HortTechnology 17(1):32-38.

McArtney S, Greene D, Schmidt T and Yuan R. 2013. Naphthaleneacetic acid and Ethephon Are Florigenic in the Biennial Apple Cultivars Golden Delicious and York Imperial. HortScience 48(6):742-746.

McGregor SE. 1976. Insect Pollination of Cultivated Crop Plants. United States Department of Agriculture.

McLaughlin JM and Greene DW. 1991. Fruit and Hormones Influence Flowering of Apple. I. Effects of Hormones. J Amer Soc Hort Sci 166(3):446-449.

Mellerowicz EJ, Horgan K, Walden A, Coker A and Walter C. 1998. *PRFLL* – a *Pinus radiata* homologue of *FLORICAULA* and *LEAFY* is expressed in buds containing vegetative shoot and undifferentiated male cone primordia. Planta 206(4): 619-629.

Mimida N, Kotoda N, Ueda T, Igarashi M, Hatsuyama Y, Iwanami H, Moriya S and Kazuyuki A. 2009. Four TFL1/CEN-Like Genes on Distinct Linkage Groups Show Different Expression Patterns to Regulate Vegetative and Reproductive Development in Apple (Malus × domestica BorkH.). Plant Cell Physiol 50(2):394-412.

Mimida N, Kidou SI, Iwanami H, Moriya S, Abe K, Voogd C, Varkonyi-Gasic E, Kotoda N, Nasholm T. 2011a. Apple FLOWERING LOCUS T proteins interact with transcription factors implicated in cell growth and organ development. Tree Physiology 31(5): 555-566.

Mimida N, Ureshino A, Tanka N, Shigeta N, Sato N, Moriya-Tanaka Y, Iwanami H, Honda C, Suzuki A, Komori S and Wada M. 2011b. Expression patterns of several floral gene during flower initiation in apical buds of apple (*Malus x domestic* Borkh.) revealed by in situ hybridization. Plant Cell Rep 30:1485-1492.

Mohamed R, Wang CT, Ma C, Shevchenko O, Dye SJ, Puzey JR, Etherington E, Sheng XY, Meilan R, Strauss SH and Brunner AM. 2010. *Populus CEN/TFL1* regulates first onset of flowering, axillary meristem identity and dormancy release in *Populus*. Plant J 62(4):647-688.

Molinero-Rosales N, Jamilena M, Zurita S, Gomez P, Capel J and Lozano R. 1999. *FALSIFLORA*, the tomato orthologue of *FLORICAULA* and *LEAFY*, controls flowering time and floral meristem identity. Plant J 20(6):685-693.

Monselise SP and Goldschmidt EE. 1982. Alternate bearing in fruit trees. Hort Rev 4:128-173.

Munoz-Fambuena N, Mesejo C, Gonzalez-Max MC, Primo-Millo E, Agusti M and Iglesias DJ. 2011. Fruit regulates seasonal expression of flowering genes in alternate-bearing 'Moncada' mandarin. Annals of Botany 108:511-519.

Neilsen JC and Dennis FG. 1999. Effects of seed number, fruit removal, bourse shoot length and crop density on flowering in 'Spencer Seedless' apple. Acta Hort 527:137-146.

Olszewski N, Sun T, and Gubler F. 2002. Gibberellin Signaling: Biosynthesis, Catabolism, and Response Pathways. The Plant Cell S61-S80.

Pellerin BP, Buszard D, Iron D, Embree CG, Marini RP, Nichols DS, Neilsen GH and Neilsen D. 2011. A Theory of Blossom Thinning to Consider Maximum Annual Flower Bud Numbers on Biennial Apple Trees. HortScience 46(1):40-42.

Porto DP, Brueneau M, Perini P, Anzanello R, Renou JP, dos Santos HP, Failho FB and Revers LF. 2015. Transcription profiling of the chilling requirement for bud break in apples: a putative role for FLC-like genes. J. of Experimental Botany 66(9):2659-2672.

Quinlan JD and Preston AP. 1971. Influence of shoot competition on fruit retention and cropping of apple trees. J Hort Sci Biot 46:525-534.

Rodrigo J. 2000. Spring frosts in deciduous fruit trees – morphological damage and flower hardiness. Scientia Horticulturae 85:155-173.

Rogers BL and Thompson AH. 1968. Growth and fruiting response of young apple and pear trees to annual applications of succinic acid 2,2-dimethyl hydrazide. Proc Am Soc Hort Sci 93: 16-24.

Schmidt T, Elfving DC, Mcferson JR and Whiting MD. 2009. Crop Load Overwhelms Effects of Gibberellic Acid and Ethephon on Floral Initiation in Apple. HortScience 44(7):1900-1906.

Seymour GB, Ostergaard L, Chapman NH, Knapp S and Martin C. 2013. Fruit Development and Ripening. Annual Rev of Plant Bio 64:219-241.

Shalom L, Samuels S, Zur N, Shlizerman L, Zemach H, Weissberg M, Ophir R, Blumwald E and Sadka A. 2012. Alternate Bearing in Citrus: Changes in the Expression of Flowering Control Genes and in Global Gene Expression in On- versus Off-Crop Trees. PloS One 7(10).

Tan FC and Swain SM. 2007. Functional characterization of *AP1, SOC1* and *WUS* homologues from citrus (*Citrus sinensis*). Physiologia Plantarum 131(3):481-495.

Trankner C, Lehmann S, Hoenicka H, Hanke M, Fladung M, Lenhardt D, Dunemann F, Gau A, Schlangen K, Malnoy M and Flachowsky H. 2010. Over-expression of an *FT*-homologous gene of apple induces early flowering in annual and perennial plants. Planta 232:1309-1324.

Tromp J. 2000. Flower-bud formation in pome fruits as affected by fruit thinning. Plant Growth Regulation 31:27-34.

USDA. 2017. National Agricultural Statistics Service: Noncitrus Fruits and Nuts 2016 Summary. Washington, DC: U.S. Department of Agriculture.

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagne D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouze P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel CE, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F and Viola R. 2010. The genome of the domesticated apple (*Malus* × *domestica* Borkh.). Nature Genetics 42:833–839.

Visser T. 1964. Juvenile Phase and Growth of Apple and Pear Seedlings. Euphytica 1964:119-129.

Wada M, Cao Q, Kotoda N, Soejima J and Masuda T. 2002. Apple has two orthologues of *FLORICAULA/LEAFY* involved in flowering. Plant Molecular Biology 49:567-577.

Wardlaw IF. 1990. The control of carbon partitioning in plants. New Phytologist 116:341-381.

Webster AD, Oehl VH, Jackson JE, and Jones OP. 1985. The orchard establishment, growth and precocity of four micropropagated apple scion cultivars. J Hort Sci 60:169-180.

Williams MW and Edgerton LJ. 1974. Biennial bearing of apple trees. Proc XIX Intern Hort Congr Warsaw 3:343-352.

Wu G and Poethig RS. 2006. Temporal Regulation of Shoot Development in Arabidopsis Thaliana by Mir156 and Its Target SPL3. Development 133(18):3539-3547.

Zhang S, Zhang D, Fan S, Du L, Shen Y, Xing L, Li Y, Ma J, and Han M. 2016. Effect of exogenous hormones, and flowering-associated genes in 'Fuji' apple (*Malus domestica* Borkh.) Plant Physio Biochem 107: 178-186.

Ziv D, Zviran T, Zezak O, Samach A and Irihimovitch V. 2014. Expression Profiling of FLOWERING LOCUS T-like Gene in Alternate Bearing 'Hass' Avocado Trees Suggests a Role of PaFT in Avocado Flower Induction. PloS One 9(10).

CHAPTER TWO

Genetic mechanisms associated with floral initiation and the repressive effect of fruit on flowering in apple (*Malus* × *domestica* Borkh.)

Submitted.

ABSTRACT

Many apple cultivars are subject to biennial fluctuations in flowering and fruiting. It is believed that this phenomenon is caused by a repressive effect of developing fruit on the initiation of flowers in the apex of proximal bourse shoots. However, the genetic pathways of floral initiation are incompletely described in apple, and the biological nature of floral repression by fruit is currently unknown. In this study, we characterized the transcriptional landscape of bourse shoot apices in the biennial cultivar, 'Honeycrisp', during the period of floral initiation, in trees bearing a high fruit load and in trees without fruit. Trees with high fruit load produced almost exclusively vegetative growth in the subsequent year, whereas the trees without fruit produced flowers on the majority of the potential flowering nodes. Using RNA-based sequence data, we documented gene expression at high resolution, identifying >11,000 transcripts that had not been previously annotated, and characterized expression profiles associated with vegetative growth and flowering. We also conducted a census of genes related to known flowering genes, organized the phylogenetic and syntenic relationships of these genes, and compared expression among homeologs. Several genes closely related to *AP1*, *FT*, *FUL*, *LFY*, and *SPLs* were more strongly expressed in apices from non-bearing, floral-determined trees, consistent with their presumed floral-promotive roles. In contrast, a homolog of *TFL1* exhibited strong and persistent up-regulation only in apices from bearing, vegetative-determined trees, suggesting a role in floral repression. Additionally, we identified four *GIBBERELLIC ACID (GA) 2 OXIDASE* genes that were expressed to relatively high levels in apices from bearing trees. These results define the flowering-related transcriptional landscape in apple, and strongly support previous studies implicating both gibberellins and *TFL1* as key components in repression of flowering by fruit.

INTRODUCTION

In many tree fruits and nuts, flowering follows a biennial cycle, with maximal and minimal flowering alternating yearly [1-3]. This phenomenon, termed biennial (alternate) bearing, is both an intriguing biological phenomenon and a significant limitation for the production of many horticultural crops. In commercial (domesticated) apple, similar to many other tree fruit species, flowering spans two growing seasons. In the first growing season, floral meristems initiate at the tips of condensed shoots called bourse shoots [4-7]. The floral meristems develop during the remainder of the growing season and arrest in a partially developed state before the winter dormant period. In early spring of the subsequent growing season, flowers complete development, culminating in bloom shortly after release from dormancy. This two-year cycle leads to an overlap between the period of fruit development (from the previous season's flowers) and the period of floral initiation (current season). At least for domesticated apple, it is generally acknowledged that the presence of developing fruit inhibits floral initiation within the adjacent bourse shoot. Several ideas have been offered to explain how developing fruit might repress floral initiation. For example, gibberellins (GAs) have been shown to repress floral initiation in apple [8], and as developing fruit contain relatively high concentrations of gibberellins [9], it is thought that diffusion of GAs from the fruit to shoot apex could underlie floral repression [8]. It has also been hypothesized the biennial bearing results from diversion of photosynthate from the apex to the developing fruit due to the potentially higher sink strength of the fruit [3,10].

The repressive effect of fruit should ultimately be reflected in expression of floral-promotive genes at the shoot apex. Previous studies have identified putative molecular components of the flowering pathway in apple based on apparent homology with well-studied

flowering genes such as *APETALA 1* (*AP1*), *LEAFY* (*LFY*), *SUPPRESSOR OF CONSTANS* (*SOC1*), *FRUITFULL* (*FUL*), *FLOWERING LOCUS T* (*FT*), and *TERMINAL FLOWER 1* (*TFL1*). Homologs of the floral promoters *LFY* (designated *MdAFL2), FUL (MdFUL/MdMADS2),* and *FT (MdFT1)* were reported to exhibit increased expression within the apex during the anticipated period of floral induction [11-18]. In contrast, two homologs of the floral repressor *TFL1* (*MdTFL1-1/2*) were reported to exhibit rapidly decreasing expression either prior to or during the floral induction and initiation period [13-14,16-19]. During floral initiation, *MdFUL,* as well as homologs of the floral promoters *AP1* (*MdAP1a* and *MdAP1b*), and *SOC1* (*MdSOC1*), exhibited increasing expression [14,16-18,20]. Following floral initiation, the expression of homologs of *FT* (*MdFT2*), *LFY* (*MdAFL1*), and *AP1* (*MdAP1*) were either maintained at a relatively high level or were further increased coinciding with floral development [11,13-14,16-18,20]. During floral development and thereafter, *MdTFL1-2* expression increased [14,18-19]. Various homologs of *AP1, LFY, FT,* and *TFL1* have been additionally examined for flowering function by manipulating their expression in transgenic Arabidopsis or apple [11-12,16, 19; 21-24].

Relatively few studies have investigated the effect of fruit load on the expression of specific, presumed flowering genes in apple or other tree fruit species. In biennial-bearing avocado and citrus cultivars, *FT*-like genes were expressed in fully developed adult leaves during the period of floral initiation, and this expression was found to be significantly higher in the bearing year [25-26]. However, in the biennial-bearing apple cultivar ‘Red Delicious’, *MdFT1* expression in the leaves was found to be similar between non-bearing and bearing years [18]. On the other hand, a PCR-based study suggested that the expression of a *MdFT2* was higher in apical buds of apple trees carrying a high fruit load compared with trees with no fruit [17,27].

Haberman et al. (2016) reported that *MdTFL1-1* expression in bourse shoot apices decreased during the course of the growing season, and that the decrease in *MdTFL1-1* expression was more rapid in trees carrying a high fruit load compared to low fruit load. In addition, they found that *MdTFL1-2* expression significantly increased relatively late in the season, but only in the high fruit load trees. This pattern was interpreted as suggesting that *MdTFL1-1* might play a role in maintaining the apex in a vegetative state early in the growing season, whereas *MdTFL1-2* might repress flowering in response to high fruit load [18].

Although these previous studies documenting gene expression in the apex have provided a solid blueprint for the advanced molecular study of flowering and alternate bearing in apple, they have focused on a limited number of anticipated landmark genes. In this study, as a subsequent step to understanding the genetic basis of floral repression by fruit in apple, we carried out an extensive census of flowering-related genes, a comprehensive analysis of gene expression in the bourse shoot apex during the transition to floral initiation, and evaluation of the effect of fruit load on the expression of flowering-related genes. Ultimately, this work should provide a deeper understanding of the endogenous mechanism(s) responsible for floral initiation and alternate bearing. This, in turn, may facilitate the development of approaches to control flowering in commercial operations, and the development of new cultivars less prone to alternate bearing.

MATERIALS AND METHODS

Plant materials, growth conditions and field experimental design

Field experiments were conducted at the Michigan State University (MSU) Clarksville Research Center (Field: 42°52'28.91"N, 85°16'8.15"W – Station: 42°52'24.20"N, 85°15'30.81"W) located in Clarksville, MI. Trees were managed in accordance with standard

commercial practices for disease, insect, and weed control. ‘Honeycrisp' trees had been established for five years as grafts on Nic 29® rootstocks. The date of full bloom was defined as the date in which the maximum numbers of flowers were open but had not reached anthesis. Six trees were chosen that showed at least 80% bloom density, defined as the percentage of nodes on one-year-old shoots that showed flower clusters. For each tree, six branches, each between 4 and 6 cm diameter at the base, were selected and randomly assigned for apex collection dates (five branches) or for observation of flowering the following spring (one branch). Plants were randomly assigned as three replicate pairs, with each pair comprising one plant that was subjected to removal of all flowers, and one plant that was left untouched. Collections were made at 2 days after full bloom (DAFB), 15-17 DAFB, 35-38 DAFB, 49-52 DAFB, and 72-75 DAFB. On each collection date, dominant buds immediately subtending the position of flower clusters or former cluster position, or the apex of actively growing shoots originating from this position, were removed using a razor blade, immediately frozen in liquid N2, and transferred to storage at -80°C.

Nucleic acid preparation, sequencing, and data analyses

RNA was isolated from frozen apex samples using the method of Gasic et al. (2004) with the exception that spermine was substituted for spermidine in the extraction buffer, followed by a final 'clean-up' step using a commercial kit (RNeasy Mini; QIAGEN, Germantown, MD). RNA quality and quantification was analyzed by the use of a Nanodrop 2000c (Thermo Scientific, Waltham, MA) and electrophoresis (2100 Bioanalyzer; Agilent, Santa Clara, CA). Library preparation and sequencing used the Illumina (San Diego, CA) platform and TruSeq platform with 101-b paired-end protocols, starting with 1 ug of total RNA from each sample. The raw sequence files were processed with fastq-mcf [29] using the parameters -t 0.10 -p 15 -l 20 -q 25

to remove adapter sequences, very short reads, and terminal bases with a Phred score below 25. The number of read pairs generated is shown in S1 Table.

Reference-based transcriptome assembly

Sequence reads were aligned to v1.1 of the GDDH13 reference sequence [30] and splice junctions were identified using the program HISAT2 (v.2.1.0) invoking the --dta-cufflinks and --un-conc-gz options [31]. The --un-conc-gz option was invoked to capture reads that failed to map to the reference genome. HISAT2 was operated using the default maximum and minimum mismatch penalties of six and two, respectively. Alignment metrics are shown in S1 Table. Transcript models were assembled using StringTie (v.1.3.3) using default parameters, including the -G option for use of a reference annotation as described [32-33]. Transcript models generated for each sample library were reduced to a consensus set of transcript models using the StringTie –merge function. The program Cuffquant (v.2.2.1; included in the Cufflinks suite [34]) was then used to calculate sequence read counts for each transcript model, and significant differentially expressed genes and isoforms were identified by the use of Cuffdiff [35]. Metrics for assessing read mapping and transcriptome assembly were obtained using RNA-SeQC (v.1.1.8 [36]) and GFF utilities suite [37], respectively.

Identification of novel 'Honeycrisp' reference-based transcripts

Novel transcripts contained within the reference-based transcriptome were identified by comparing the reference genome gene models (retrieved from https://iris.angers.inra.fr/gddh13/the-apple-genome-downloads.html as gene_models_20170606.gff3) and the 'Honeycrisp' reference-based transcript models (S1 File) using the gffcompare (v.0.9.12) software package within the GFF utilities [37]. The resulting annotated gtf file was filtered for classification codes associated with non-isoform-like transcript

features and/or assembly errors (classification codes: e, i, o, u, x, y) and removal of transcripts with lengths <200b. This subset of transcripts was then analyzed for protein-coding capacity using the software programs CPC2 (beta version [38]), PLEK (v1.2 [39]), and CPAT (v.1.2.4 [40]). For CPC2 and CPAT, the coding potential probability was set to ≥0.5 to assign a transcript as coding and ≤0.5 as noncoding/ambiguous. For PLEK, the coding or noncoding/ambiguous determination was assigned by the program's default parameters. The final coding definition of a transcript was based on an agreement between at least two of the programs. Detailed transcript information can be found in S2 - S4 Tables.

De novo assembly of unmapped reads

Reads that were unmapped by HISAT2 were assembled into contiguous sequences using the Trinity de novo assembler with default settings [41]. The resulting FASTA file (S2 File) containing the de novo assembled transcripts was then used as an input to construct consensus gene models using the python program Trinity_gene_splice_modeler.py provided by the Trinity suite (S3 File). The python script produced a consensus FASTA file containing gene models and a corresponding GTF file. The unmapped read files were then realigned to the consensus gene FASTA file using HISAT2 invoking the --dta-cufflinks options and using the previously generated GTF file. Alignments were then processed through the same Cufflinks pipeline used in the referenced-based transcriptome assembly. The initial output comprised ~250,000 sequences corresponding to ~92,000 distinct loci. Because most output sequences appeared to be sequencing or assembly artifacts, we limited further consideration to contigs representing putative transcripts that were likely to be strongly expressed (upper 10th percentile based on FPKM, and expressed in at least three samples) and that had coding potential (determined as described above for novel reference-based transcript models) (S5 - S7 Tables).

General transcriptome annotation

Transcript sequences were annotated based on sequence homology to Arabidopsis open reading frame translations (TAIR10; TAIR10_pep_20101214_updated 2012-04-16, [42]) using the BLASTx module from NCBI [43] with an Expect (E)-value cutoff of $1e^{-11}$. Homologous sequences were then used as queries to identify similar transcripts within the 'Honeycrisp' transcriptome, using the tBLASTx module. Gene model sequences generated from the de novo assembly were annotated by aligning sequences to the nr NCBI database (downloaded on 2018-09-18, [44]) and the 'Honeycrisp' reference-based transcriptome using the BLASTx and BLASTn modules, respectively. A minimum E-value of $1e^{-10}$ and a max_target_seqs of 1 were used.

Data accessibility

Raw sequence libraries can be downloaded from the National Center of Biotechnology Information Short Read Archive under biosample SAMN04239699. Our constructed reference-based transcriptome annotation (S1 File), de novo transcript FASTA and annotation (S2 and S3 Files), phylogenies of 125 flowering genes, and differential expression data files (S4 and S5 Files) can be retrieved from the Dryad repository https://doi.org/10.5061/dryad.fn2z34tr5. Computation protocols used in this study can be retrieved from the Protocols.io repository https://dx.doi.org/10.17504/protocols.io.bp54mq8w.

Identification of apple flowering genes

To identify potential homologs of flowering genes, we indexed genes from Arabidopsis (TAIR10) annotated with potential roles in flowering: Gene Ontology terms 0048438 ('floral whorl development'), 0009908 ('flower development'), 0009910 ('negative regulation of flower development'), 0009911 ('positive regulation of flower development'), 0048578 ('positive

regulation of long-day photoperiodism, flowering'), 0010220 ('positive regulation of vernalization response'), 0009909 ('regulation of flower development'), 0048510 ('regulation of timing of transition from vegetative to reproductive phase'), 0010321 ('regulation of vegetative phase change'), 0010228 ('vegetative to reproductive phase transition of meristem'), 0010048 ('vernalization response'), and 0010093 (‘specification of floral organ identity’). This set of 437 genes was manually curated to omit those without strong functional evidence for a direct role in flowering. The curated subset contained 180 genes. Conceptual translations of the corresponding representative gene models were obtained from TAIR (TAIR10_pep_20110103_representative_gene_model_updated) and used as queries to search open reading frame translations of our mapped-assembled and de novo-assembled transcript models (BLASTp) using an E-value cutoff of $1e^{-12}$. The open reading frame translations of the Honeycrisp assembled transcript models were identified using TransDecoder (v5.5.0;https://github.com/TransDecoder). All identified transcript translations were then used as queries to search the Arabidopsis representative gene model translations. Those transcripts that reciprocally identified their original Arabidopsis query were defined as reciprocal homologs. For phylogenetic analyses of the 16 intensively studied flowering gene families, we considered only the 25 highest-scoring apple transcript translations and only the 25 highest-scoring Arabidopsis gene translations identified with each apple sequence query. Phylogenetic trees were then constructed using the ETE3 toolkit (v.3.1.1) build function invoking the standard_fasttree workflow under default settings [45-47]. Collinearity among identified flowering genes was performed using the MCScanX toolkit following the manual’s instructions and the use of default parameters [48]. Graphics to illustrate the collinear relationships between homeologous

chromosomes and flowering genes identified by MCScan X were generated using Circos (v.0.69-6) program package [49].

Gene expression analysis

Estimated expression levels for homologs of flowering-related genes/transcripts were obtained from Cuffnorm and Cuffdiff output. Heat maps were created and expression profiles were clustered using R statistical software (v.3.5.2 [50]) and the CummeRbund (v.2.24.0 [51]) package. Expression profiles of homologous flowering-related genes that exhibited significant changes in expression were clustered using a K-means approach by the csCluster command of CummeRbund. Cluster expression pattern was then defined by the general trend of the modal expression pattern. Venn diagram of differentially expressed genes was created using an online tool (http://bioinformatics.psb.ugent.be/webtools/Venn/). Co-expressed gene modules were identified using the weighted gene co-expression network analysis (WGCNA) R package [52], following the analysis methodology outlined by Zhang and Horvath (2005) and using normalized gene expression (FPKM) as calculated by cufflinks.

TaqMan® qRT-PCR

Confirmation of MdTFL1 gene expression was determined using a two-step quantitative polymerase chain reaction (qRT-PCR). Primers and probes were designed from sequences assembled in our transcriptome and aligned to apple nucleotide sequences maintained by the NCBI (taxid: 3750). The primers and probes were designed in a previous study [54] and were based on the specificity to selected target sequence and overlapped of an exon junction (S8 Table). An apple homolog of ACTIN served as an internal control. The reactions were performed using an Agilent Technologies Stratagene Mx3005P (Santa Clara, CA) qPCR machine with cDNA derived from the RNA samples prepared for RNA-seq. Each reaction consisted

TaqManTM Gene Expression Master Mix (10 µl), 5x diluted cDNA template (2 µl), forward (1 µl) and reverse primers (1 µl), and probe (1 µl) for ACTIN, the primer-probe assay for the gene of interest (1 µl), and ddH2O (4 µl). The thermal profile for TaqManTM assay followed the instructions provided with the Agilent machine.

RESULTS AND DISCUSSION

Effects of reducing fruit load on floral initiation

As a physiological and molecular model for biennial bearing in apple, we focused on the popular commercial cultivar 'Honeycrisp', which can exhibit extreme biennial tendency under production conditions [55]. At full bloom in early spring, we selected three paired sets of trees with high bloom density and removed all flowers from one tree from each pair. This floral thinning treatment had a strong effect on initiation of new flowers, as evidenced by observed bloom density in the spring of the second year of the study (Fig 2.1A). Those trees that were thinned of flowers produced floral shoots at an average of ~52% (range 39-82%) of potential flowering nodes, whereas the non-thinned control trees produced almost exclusively vegetative shoots (Fig 2.1B).

Transcriptome assembly and characterization

Based on anatomical characterization of the bourse shoot meristem in the apple cultivar used in this study, floral meristems began to be initiated approximately two weeks after full bloom (data not shown). This is consistent with observations of the bourse shoot meristem in other apple cultivars [4-7]. Based on this, we sampled the bourse shoot apex from the thinned and non-thinned trees at approximately 2, 15, 35, 50, and 70 DAFB. Dissected apices were subjected to high-throughput RNA-based sequencing, yielding a total of ~390 million paired reads. These were aligned to a recently published reference genome sequence (GDDH13 v.1.1)

assembled from a doubled-haploid individual generated from ‘Golden Delicious’ [30]. We obtained a mean alignment rate of 90.1%, with 96.3% of the aligned reads mapping within annotated intragenic regions (S1 Table). Transcript models were then assembled from aligned reads using the StringTie transcript assembler [33].

The recent availability of a high-quality apple genome sequence and exhaustive depth of our transcriptional data provided the opportunity to document genes expressed in the apple bourse shoot apex with high-resolution and accuracy. Our reference-based transcriptome assembly cataloged a total of 104,690 transcripts arising from 58,452 loci (Table 2.1; S1 File). This extends considerably the previously annotated gene content of the GDDH13 genome, which was based on nine RNA-seq libraries representing diverse structures, including the shoot apex, along with cDNAs and expressed sequence tags (ESTs) cataloged in NCBI databases. Our sequence and assembly results complemented the reference annotation with the identification of an additional 11,264 novel transcriptional models and 39,227 'Honeycrisp'-specific isoforms of annotated transcripts (~10.8% and ~37.5% of the assembled transcripts, respectively; Fig 2.2A). These novel transcripts comprised 23,034 novel exons and originated from 8,753 previously unidentified loci. We further characterized these novel transcripts in terms of length, expression level, coding potential, genomic organization, and homology with known, expressed genes (Fig 2.2B; S2 - S4 Tables).

The majority of these transcripts (81.2%; 9,096) were expressed (FPKM > 1; TPM 1.19 - 1.72; Fig 2.2B; S3 Table). Of those expressed transcripts, 63.7% were predicted to encode proteins. About 46% of the expressed-coding transcripts were located in previously annotated intergenic regions. The remaining 56% showed some positional overlap with previously annotated genes (Fig 2.2B). In total, 73.8% (8,310) of the novel transcripts showed significant

(E-value < 1e-10) nucleotide sequence homology to previously cataloged, expressed genes from *Malus spp.* (Fig 2.2C). These genes included 163 distinct loci encoding the *M. floribunda HcrVf*-like and *M.* x *domestica Rvi15* apple scab (*Venturia inaequalis*) resistance genes, and 159 loci encoding the *M.* x *robusta* fire blight (*Erwinia amylovora*) resistance genes. A total of 1,304 reference-mapped transcripts exhibited no significant homology to any sequence cataloged in the NCBI nt database (S2 Table).

Reads that did not align with the reference genome may represent sequence from uncharted segments of the apple genome including extrachromosomal DNAs or loci that are extremely diverged between GDDH13 and 'Honeycrisp', or may be derived from exogenous biota. We assembled unmapped reads *de novo* into contiguous sequences (S2 and S3 Files) (see *Methods*), and evaluated the potential of the contigs to represent authentic apple transcripts. A total of 5,542 potential transcripts, representing 4,737 gene models, showed apparent expression values >100 FPKM in at least three of the sequencing libraries. About 39% of this subset of *de novo* transcripts were predicted to encode proteins (S5 Table; S1 Fig A). About 75% of the 5,542 strongly expressed potential transcripts displayed significant homology to cataloged *Malus* sequences, and another 15% to sequences from related Rosaceae genera (S5 Table; S1 Fig B).

Identification of flowering gene homologs

Although genes with anticipated roles in flowering have previously been identified in apple, there has often been confusion and conflicting reports regarding gene identity, copy number, and expression pattern. This is most likely due to the existence of closely related orthologs for some of these genes, the heterogeneous and paleo-allopolyploid nature of the apple genome, and the inability of some previous approaches to discriminate among closely related sequences. The ~40 billion bases of transcriptional sequence data from the shoot apex analyzed

in this project, as well as our identification of novel genes, provided the opportunity to resolve gene identities and estimate orthologous relationships. We identified a set of 180 Arabidopsis genes with flowering-related annotations (see *Methods*), and searched the combined GDDH13 / 'Honeycrisp' transcriptome for expressed sequences with significant homology (E-value < 1e-12). In each case, the open reading frame translation from the primary designated transcript of the Arabidopsis gene was used to query the primary translations from both the annotated and novel reference-based transcriptional models, as well as the ORF-containing *de novo* transcriptional models. The highest-scoring, matching sequences were then used reciprocally to query a comprehensive database of open reading frame translations from Arabidopsis. Using this approach, we identified a total of 321 apple counterparts to 125 Arabidopsis genes. For further discussion, we refer to this collection as 'flowering gene homologs'. Three of the identified apple genes had not previously been annotated in the GDDH13 reference genome (S9 Table).

At least 106 of the 125 Arabidopsis flowering genes had multiple homologs. Previous research indicates that genes in apple generally exist as duplicates as a result of an ancient whole-genome duplication [56]. We analyzed genomic synteny for all of the 125 flowering gene families (S9 Table; Fig 2.3). Based on chromosomal positions, a simple genome duplication appears to have contributed to family expansion for at least 86 of these 106 Arabidopsis genes, and tandem duplication contributed to expansion for at least 14 (S10 Table).

We identified 55 Arabidopsis flowering-related genes lacking a reciprocal homolog in the combined reference shoot apex transcriptome. These unrepresented genes included several functioning in the Arabidopsis vernalization-response pathway, including *FLOWERING LOCUS C* (*FLC*) and its sibling *MADS AFFECTING FLOWERING* (*MAF*), *FRIGIDA* (*FRI*), and *VERNALIZATION INSENSITIVE 3* (*VIN3*). This result is consistent with the apparently cold-

independent initiation of flowers during the summer period in apple [57]. The gene previously described as an *FLC* homolog (MD09G1009100) by Takeuchi et al. (2018) and Nishiyama et al. (2019) was found to be not closely related to *FLC* in our study (S2 Fig [d] and not shown). Other Arabidopsis flowering genes without clear apple representatives included 16 additional members of the *AGAMOUS*-like (*AGL*) MADS-box gene superfamily. These results are consistent with the observed rapid evolution and diversification of the large MADS-box genes observed in apple and other plants [60-61].

We focused further study on a subset of flowering genes that have been intensively studied both in Arabidopsis and other plants [62]. Apart from *FLC*, this subset included *AGL24, AP1, FD, FUL, FT, LFY, SOC1, SPL3, SPL4, SPL5, SPL9, SPL15, SVP, TSF,* and *TFL1*. (Table 2.2; S2 Fig [a-n]; Fig 2.3). In Arabidopsis, *FT* is transcribed in the leaves along with its paralog *TSF* and translocated to the apex[Andrés and Coupland 2012]. In the apex, FT forms a complex with FD which activates transcription of *AP1* and *SPL3/4/5* directly [62]. The FD/FT complex also indirectly activates the expression of *FUL* and *SOC1* [62]. In addition, *FUL* and *SOC1* expression is reinforced by *SPL9/15* [63]. This collective network promotes a phase change within the apex leading to floral initiation. *SOC1* and *AGL24* form a positive-feedback loop, promoting one another's expression along with promoting *LFY* expression [64]. *LFY* expression is also directly promoted by *SPL3/4/5* and indirectly by *AP1* establishing floral meristem identity [63]. Negative regulators of this process are *SVP* and *TFL1*. *SVP* represses *FT* expression, whereas TFL1 competes with FT for complex formation with FD [62].

We reconstructed phylogenies for these genes, including the most homologous genes from both Arabidopsis and apple, and generated un-rooted trees (Fig 2.3). Apple genes related to *AGL24/SVP, AP1/FUL, FD, FT/TFL1, LFY, SOC1, SPL4/5,* and *SPL9/15* were included in well-

defined (>90% bootstrap replicates) clades. The majority of these apple genes existed as pairs on homeologous chromosomes, as anticipated. Additional homologs likely resulted from tandem duplication, as evidenced by their close proximity (e.g., the *AGL24/SVP* clade pair *MD15G1384500*/*MD15G1384600*). Apple was previously found to contain two homologs of *FT*, one positioned on Chr. 4 (*MdFT2*) and the other on Chr. 12 (*MdFT1)* [16]. The GDDH13 genome contains only *MdFT1*. We assembled sequence reads that did not map to the GDDH13 genome (*see Methods*) and were able to identify a *MdFT2*-like transcript (*FT-like de novo*) (Table 2.2; S2 Fig [a-n]; Fig 2.3). This result suggests that the GDDH13 genome sequence is incomplete for Chr. 4 or that the GDDH13 doubled-haploid genotype lacks *MdFT2*. As anticipated from the reciprocal homology results (above), individual members of the *FLC/MAF* family showed no specific phylogenetic relationships with apple genes, although a group of three apple genes were often (88% of bootstrap replicates) placed into a clade with the *FLC/MAF* family (S2 Fig).

Transcriptional analysis of the apple shoot apex during the floral transition

To gain insight into genetic pathway(s) associated with flowering in apple, we examined changes in gene expression occurring in the bourse shoot apex in the set of flowering-induced (thinned) trees spanning 2 DAFB to 70 DAFB. Because an appreciable fraction (~48%) of the apices did not initiate flowers during the year of this study (Fig 2.1), this set of genes represent those associated with vegetative apex activity (i.e. continued production of leaf primordia), as well as the transition to flowering. We identified a total of 12,661 reference-mapped genes, including ~100 flowering gene homologs, that exhibited significant changes in expression in at least one pairwise comparison among the five developmental stages evaluated.

To define transcriptional programs potentially involving the ~100 flowering gene homologs, we clustered their expression profiles using a K-means approach (k = 5) (Fig 2.4). Clusters 1-3 represented genes that showed decreases in expression at some point during the period, consistent with a floral repressive role, or expression largely limited to vegetative phase. Cluster 1 genes (n = 18) generally showed a strong decrease in expression at the earliest studied interval, between 2 and 15 DAFB, with little or no expression change at later time points. This cluster included a homolog of *FD* (Fig 2.4). In contrast, Cluster 2 (n = 19) genes showed progressively decreasing expression over the course of the season. This cluster included *MdSOC1a*, as well as homologs of *AGL24/SVP, SPL4/5,* and *SPL9/15*. Most of the genes in Cluster 3 (n = 10) showed a strong decrease in expression at the earliest studied interval, between 2 and 15 DAFB, and continued decreasing expression at the later time points. This cluster included both *MdTFL1-1* and *MdTFL1-2*, although we noted that *MdTFL1-1* was upregulated between 2 and 15 DAFB (Fig 2.4). The strong decrease in expression of these two *TFL1* homologs during the anticipated period for floral initiation has previously been reported [13-14,16-19,65-66].

Genes in Clusters 4 and 5 showed generally increasing expression across the entire study period, suggesting promotive roles in flowering or expression domains linked with the floral phase. Cluster 4 genes (n = 48) showed steadily increasing expression across the period. These included *MdFT1*, as well as homologs of *AGL24/SVP*, *SPL3/4/5,* and *AP1/FUL*. Cluster 5 contained only five genes, and these were characterized by a generally more substantial increase in expression over the season. This cluster included *MdAFL1*, as well as a homolog of *AP1/FUL* and two homologs of *AGL24/SVP*. The increasing expression of *MdAFL1* and the *AP1/FUL* homolog reflects the increased expression of their counterparts during flowering in Arabidopsis.

Expression of the two *AGL24/SVP* homologs was analogous with that of *AGL24* in Arabidopsis during the transition to a reproductive meristem [62].

Transcriptional response to the presence of a fruit load

To identify genetic mechanisms that may be specifically involved in the repression of flowering by developing fruit, we compared gene expression between apices from the thinned (flowering-induced) and non-thinned (non-induced) trees at each time point over the study. At the 15 and 35 DAFB sampling times, fruit had reached ~10 mm and ~20 mm in diameter, respectively. At 50 DAFB, fruit had reached ~ 30 mm in diameter, and at 70 DAFB, fruit was ~40 mm in diameter. At 70 DAFB, seeds and embryos were still immature, but had reached their final size. Fruit and seed reached maturity at ~120-130 DAFB (not shown). We identified a total of 6,595 genes that were differentially expressed between the two conditions at one or more time points. Of these, 55 were included in the defined set of flowering gene homologs (S11 Table; Fig 2.5).

K-means clustering identified five modal expression patterns. Genes in Cluster 1 were generally expressed to higher levels in non-thinned apices at later time points (50 and 70 DAFB) and thus could represent downstream floral repressors or genes expressed in the vegetative tissues of the apex. This expression pattern was exemplified by the *AGL24/SVP* homolog MD15G1384600 (Fig 2.4). This result suggests that the function of MD15G1384600 could be similar to *SVP* in maintaining vegetative identity [62].

Cluster 2 genes were generally expressed to higher levels in thinned apices at the earliest time points (2 and 15 DAFB) and could represent early flowering promoters. An example included in this cluster is the *AP1/FUL*-related gene, MD06G1204400. Genes in Clusters 3 and 4 showed generally increasing expression in thinned apices, relative to non-thinned apices,

throughout the study period. Cluster 3 (higher expression in non-thinned apices only at the earliest time points) could represent early flowering repressors or genes expressed early in the vegetative tissues. This cluster contained the *SPL4/5* homolog, MD03G1230600. Cluster 4 (higher expression in thinned apices at the latest time point) might represent genes acting as promoters late in flowering, including floral development, or genes expressed in floral tissues. This cluster contained homologs of *SPL3/4/5* and *SPL9/15*, as well as the *AP1/FUL* homolog *MdMADS2.1* (Fig 2.5), which we had also found to increase in absolute expression over the season (Fig 2.4). Cluster 4 additionally included *MdTFL1-1*, which we had also found to show a strong decrease in absolute expression after 15 DAFB (Fig 2.4). This suggests that the presence of fruit promotes the seasonal decrease in expression of *MdTFL1-1*, as previously observed by Haberman et al. (2016).

The final cluster, Cluster 5, contained a small group of genes (76) that showed greatly reduced expression in thinned apices, relative to non-thinned apices, at 15 DAFB and 70 DAFB (Fig 2.5). *MdTFL1-2* was the sole flowering gene homolog included in this group. Like *MdTFL1-1*, *MdTFL1-2* showed a decrease in expression throughout the season in flowering-induced apices, and the observed differential expression pattern suggests that the presence of fruit counteracts this seasonal decrease. This was also previously observed by Haberman et al. (2016).

MdTFL1-2 expression profiling and identification of co-expressed genes

This expression pattern of *MdTFL1-2* as reported previously by other groups, and here determined by RNA-seq, suggests this could be a key gene in regulating floral repression in the presence of fruit on the bourse shoot. We carried out qRT-PCR to quantify relative expression of both *MdTFL1-1* and *MdTFL1-2* to confirm the expression trend observed in the RNA-seq results.

The results were generally consistent between the two approaches (Fig 2.6). Haberman et al. (2016) previously reported that *MdTFL1-2* increased in expression between ~30 and ~60 DAFB in fruit bearing spurs. Our observations are distinct from those of Haberman et al. (2016), as our results indicate that fruit promotes significant increased expression of *MdTFL1-2* as early as 15 DAFB. This early seasonal expression of *MdTFL1-2* overlaps with the period of floral induction/initiation in apple [4-7], and argues for a direct role for *MdTFL1-2* in repressing floral initiation, rather than a conceivable function in governing inflorescence architecture once initiation has occurred [54].

MdTFL1-2 is expected to act in transcriptional regulation of flowering. Genes expressed similarly with *MdTFL1-2* (i.e., more strongly in non-thinned apices) could represent upstream promoters of *MdTFL1-2* expression or downstream positive targets. Considering only flowering gene homologs, in addition to the *AGL24/SVP* homolog MD15G1384600, these included homologs of *AGL16*, *BLH8*, *EFM*, *GPRI1*, *LSN*, and one homolog of *RAP2.7* (Fig 2.5). Conversely, genes expressed in a reciprocal manner to *MdTFL1-2* (i.e., less strongly in non-thinned apices), could represent upstream repressors of *MdTFL1-2* expression or downstream negative targets. This included homologs of *AGL6*, *MdMADS2.1*, *RAP2.7*, and various *SPLs* (Fig 2.5). In Arabidopsis, *AP1* represses *TFL1* expression, and this finding is consistent with a conserved function of *MdMADS2.1* in apple [67].

We also searched the subset of genes assigned to Cluster 5 for other potential upstream promoters or downstream regulatory targets of *MdTFL1-2* (Fig 2.5; S12 Table). Of the 75 other genes assigned to this cluster, four would encode transcriptional regulator-like proteins. These included MD05G1203300, a homolog of *FERTILIZATION INDEPENDENT ENDOSPERM* (*FIE*). *FIE* encodes a component of a POLYCOMB REPRESSOR COMPLEX 2 (PRC2) protein

that represses flowering and floral development in Arabidopsis [68]. In our study, this *FIE* homolog was expressed to higher levels in the non-thinned apices, relative to thinned, from 15 DAFB thru 70 DAFB (S3 Fig).

We employed a second, independent method to identify genes that could represent upstream promoters of *MdTFL1-2* or downstream targets through the identification of co-expression networks using the weighted gene correlation network analysis (WGCNA) approach [52]. This resolved 28 modules of co-expressed genes, with *MdTFL1-2* assigned to a module containing 200 genes (Fig 2.5C; S13 Table). This module contained homologs of *APL, EFM, SPL3,* and *ATH1* (S14 Table). We also identified a module with a strong negative correlation to *MdTFL1-2's* module (Fig 2.5C; S15 Table). Here, we identified a total of 93 genes, including a distinct homolog of *ATH1* (S15 Table).

Expression profiles of GA2ox and GA20ox genes

In our previous study of the mechanisms of the repression of flowering by GAs in apple, we found that *MdTFL1-2* was rapidly (within 2 days) upregulated in the shoot apex in response to exogenous GA_{4+7} [54]. In that study, we also found that exogenous GA resulted in the rapid upregulation of four genes classified as *GA2 OXIDASE* (*GA2ox*). Interestingly, all of the four *GA2ox* genes were included in the set of 6,595 genes differentially expressed in response to fruit load. A heat map of expression of these and additional *GA2ox* genes identified by Zhang et al. (2019) is shown in Fig 2.5E. The four *GA2ox* genes identified as differentially expressed shared a general pattern of higher expression in the thinned, relative to non-thinned, apices very early in the season (2 DAFB). Interestingly, as the season progressed, these genes showed higher expression in the non-thinned apices. This is consistent with previous studies by Guitton et al. (2016) and Habermann et al. (2016) showing that two of these four genes, MD05G1207000 and

MD10G1194100, were expressed to higher levels in non-thinned samples at a similar sampling date (48 DAFB) as in our study. If cellular GA levels promote expression of these *GA2ox* genes, then the strong shift to higher expression in non-thinned apices could reflect increased GA levels in the apex, potentially driven by the presence of fruit.

In the previous study [54] we also documented that exogenous GA resulted in rapid downregulation of several genes encoding *GA20 OXIDASES* (*GA20ox*), which participate in GA biosynthesis and are recognized to be subject to feedback repression in many contexts in various plants. Here, we observed that the *GA20ox* homolog MD01G1192100 was expressed to relatively higher levels in thinned apices at all time points (Fig 2.5B). Habermann et al. (2016) also reported higher expression of specific *GA20ox* homologs in thinned apices, including MD01G1192100. Thus, this observation might reflect lowered levels of bioactive GAs in thinned apices.

Divergent expression patterns of homeologous gene pairs

The differential regulation of the apple *TFL1-1* and *TFL1-2* genes is an interesting example of functional divergence of ancestrally related genes. Although expression of many of the key flowering gene homologs could not be reliably estimated, we identified several additional cases in which apparent gene duplication and/or gene family expansion was associated with distinctions in expression (Fig 2.7). For example, the *AP1/FUL* homeologs MD06G1204400 and MD14G1215700 showed distinct absolute expression patterns across the season, with MD06G1204400 increasing strikingly and MD14G1215700 remaining relatively constant. In contrast, the *SPL4/5* gene MD03G1230600 exhibited a strong decrease in expression as the season progressed, while expression of the homeologous MD11G1251800 stayed

relatively constant. A third example was the *FD* homolog MD15G1230800, which was strongly increased at later time points, while its homeolog MD02G1125100 was not (Fig 2.7).

Expression of related genes was also differentially influenced by fruit load in several cases. Besides *TFL1-1/TFL1-2*, *MdAFL1* was more strongly expressed in the non-thinned apices at several time points, whereas *MdAFL2* was more weakly expressed. The distinctions in expression of these homologs of the intensively studied flowering genes underscores the importance of thoroughly indexing the genomic content and rigorously establishing phylogenetic relationships. Future characterization of function of these gene pairs can provide novel insight into the conserved and/or divergent genetic mechanism(s) that underlie their role in flowering in apple.

Table 2.1. Sequencing and transcriptome assembly statistics.

Measured Statistic	**Value**
Bases Sequenced	39,201,481,600
Total Sequence Reads	392,014,816
Mean Overall Read Mapping Rate	90.10%
Total Number of Transcript Models	104,690
Total Number of Genes	58,452
Novel Exons	23,034
Novel Introns	11,152
Novel Loci	8,753
Average Transcripts per Loci	3.4
Average Transcript Length (bases)	1,617

Table 2.2 Identified homologs of Arabidopsis flowering genes in apple.

Gene Family			Reference Locus	Clade/Alias	Citation
AGL24 ***SVP***			MD01G1038600	*AGL24/SVP*	
			MD08G1197300	*AGL24/SVP*	
			MD15G1313200	*AGL24/SVP*	
			MD15G1384500	*AGL24/SVP*	
			MD15G1384600	*AGL24/SVP*	
AP1 ***CAL*** ***FUL*** ***AGL79***	***AP1*** ***CAL***		MD13G1059200	*MdMADS5*	Kotoda et al. 2010
			MD16G1058500	*AP1*	
			MD06G1204400	*AP1/FUL fam*	
			MD14G1215700	*MdMADS2.1*	Cevik et al. 2010
			MD17G1065500	*AP1/FUL fam*	
FD ***BZIP27***			MD02G1125100	*FD*	
			MD15G1008300	*FD*	
			MD15G1240800	*FD*	
FT ***TSF*** ***AT5G62040*** ***ATC*** ***TFL1***	***FT*** ***TSF***		MD12G1262000	*MdFT1*	Kotoda et al. 2010
			FT-like *de novo*	*MdFT2*	
	TFL1 ***ATC***	***ATC***	MD03G1143000	*ATC*	
			MD11G1163500	*ATC*	
			MD12G1023900	*MdTFL1-1*	Kotoda and Wada 2005; Hättasch et al 2008; Mimida et al. 2009
			MD14G1021100	*MdTFL1-2*	
			MD01G1198400	*FT/TFL1 fam*	
			MD07G1265900	*FT/TFL1 fam*	
LFY			MD06G1129500	*MdAFL1*	Wada et al. 2002
			MD14G1146700	*MdAFL2*	
SOC1			MD02G1197400	*MdSOC1a*	Kotoda et al. 2010
			MD07G1123600	*MdSOC1b*	
SPL3 ***SPL4*** ***SPL5***	***SPL4*** ***SPL5***		MD03G1230600	*SPL4/5*	
			MD11G1251800	*SPL4/5*	
			MD05G1312300	*SPL3/4/5*	
			MD09G1244200	*SPL3/4/5*	
			MD10G1291800	*SPL3/4/5*	
			MD17G1236000	*SPL3/4/5*	
SPL9 ***SPL15***			MD12G1060000	*SPL9/15*	
			MD12G1060200	*SPL9/15*	
			MD14G1060200	*SPL9/15*	

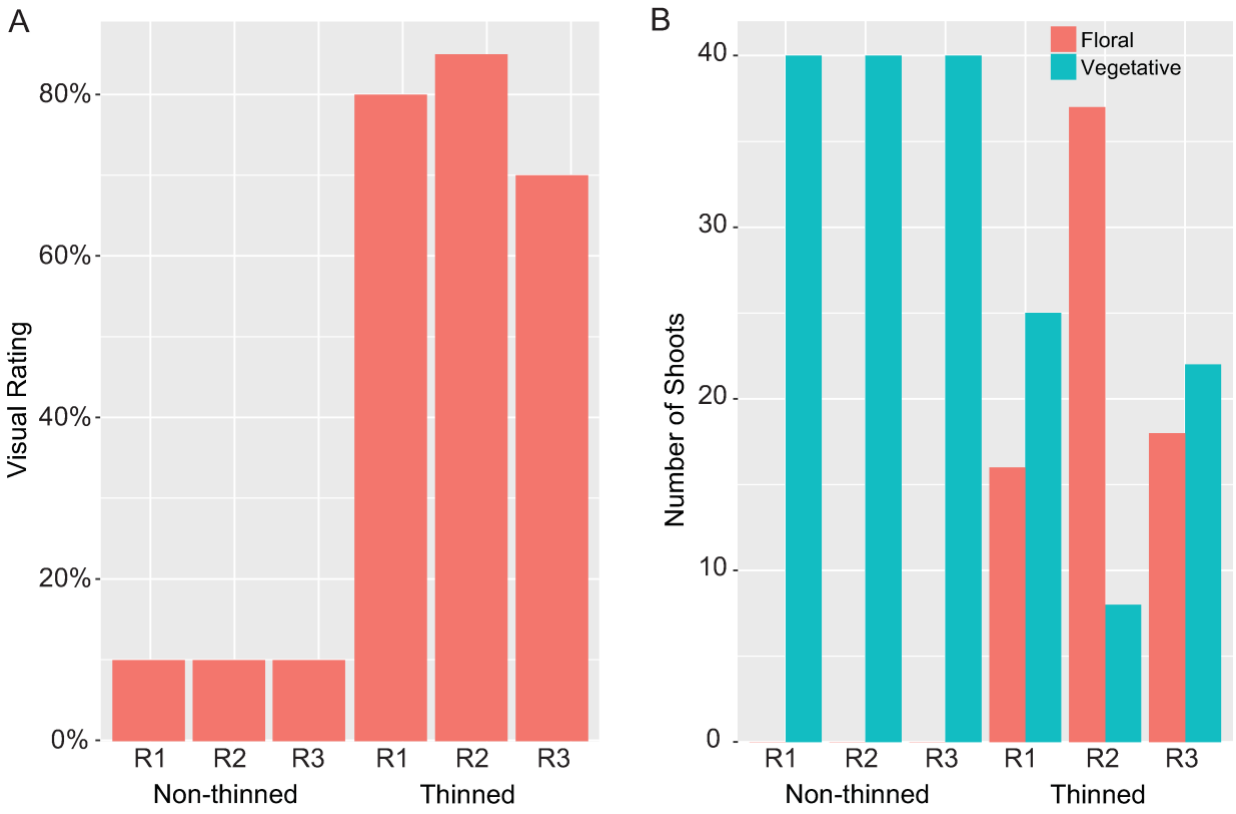


Figure 2.1. Effect of flower removal on subsequent-year flowering. At full bloom, trees were thinned of flowers or were left non-thinned, and the fraction of flower-bearing shoots, relative to total shoots arising from spur structures, was evaluated the following spring. (A) Visual estimation of floral density. The rating scale extends from 0% (no obvious flowers) to 100% (abundant flowers). Visual density was estimated by two, independent, trained observers (correlation p value < 0.05). (B) Quantification based on sampling a minimum of 40 spur shoots designated for evaluation prior to flower removal. Graphs show the results from three biological replicates (R1-R3).

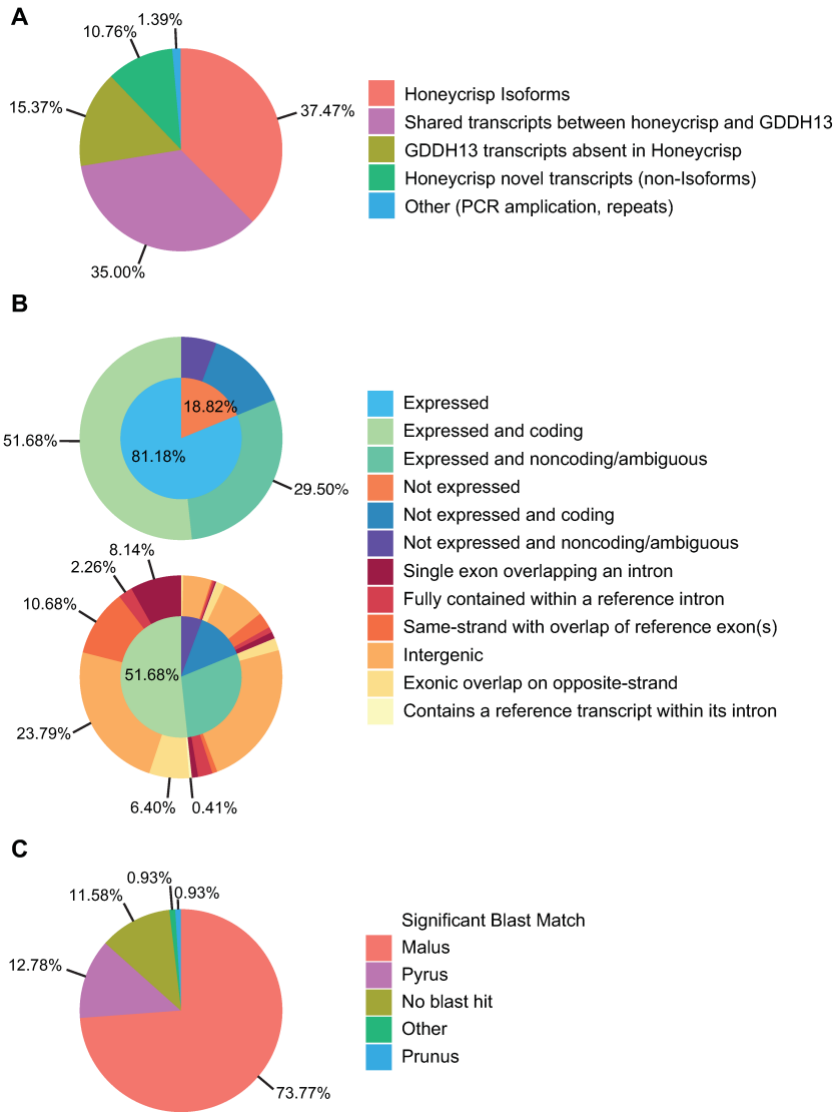


Figure 2.2. Characterization of the reference-mapped assembled 'Honeycrisp' transcriptome. (A) Proportions of the assigned organization of transcripts assembled from reads that mapped to the reference genome. (B) Upper pie chart illustrates the proportion of novel (non-isoform) transcripts that were expressed (FPKM > 1) and were either predicted to be protein coding or noncoding/ambiguous. Lower pie chart further characterizes the expressed and coding novel transcripts by their genomic organization. (C) The distribution of assigned BLASTn match results.

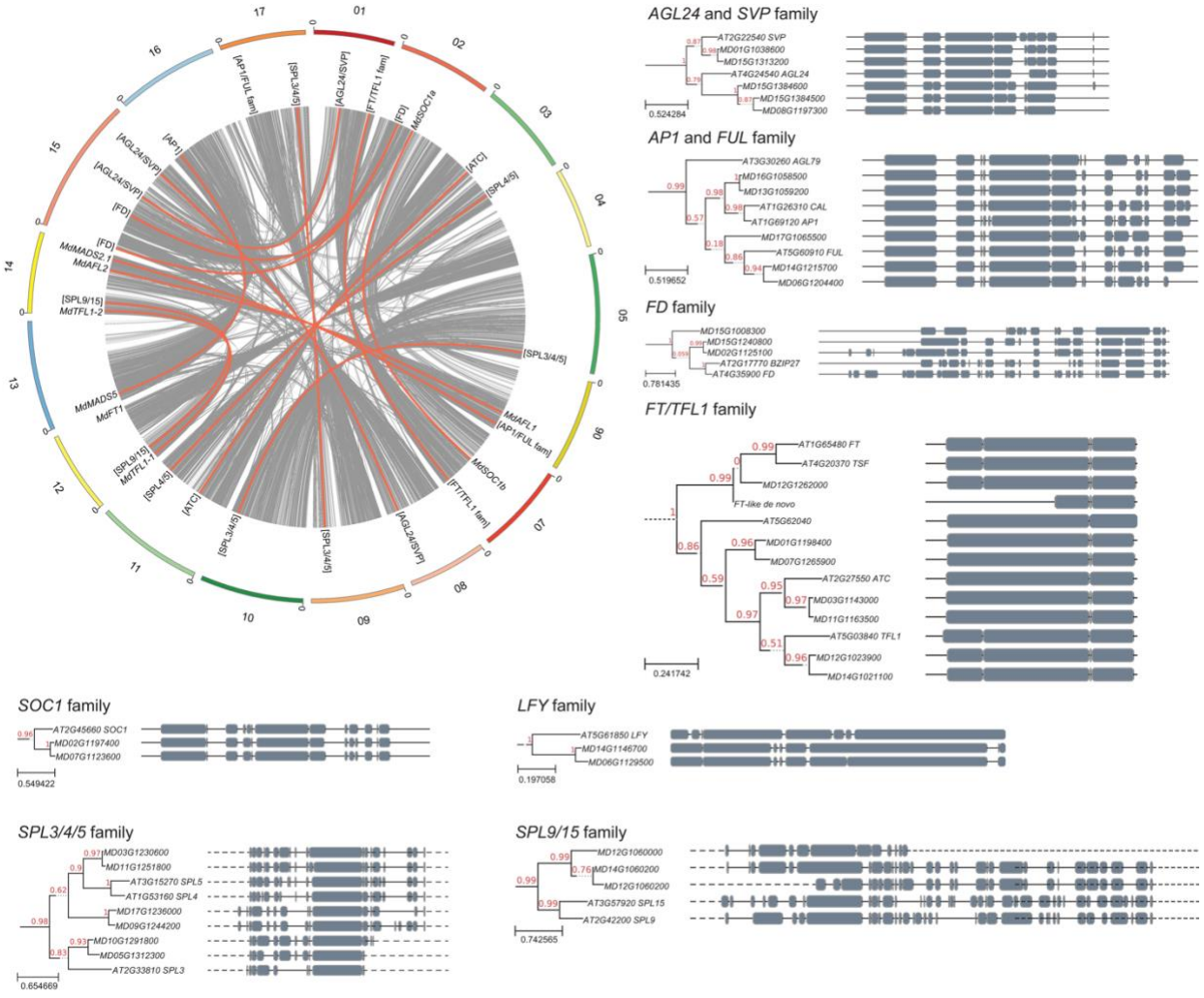


Figure 2.3. Analysis of genomic collinearity and phylogeny for 15 flowering gene families.

The circle plot depicts the position of flowering genes in the apple genome and their collinear relationships (red links). The background grey links represent the complete genomic collinearity within the apple genome.

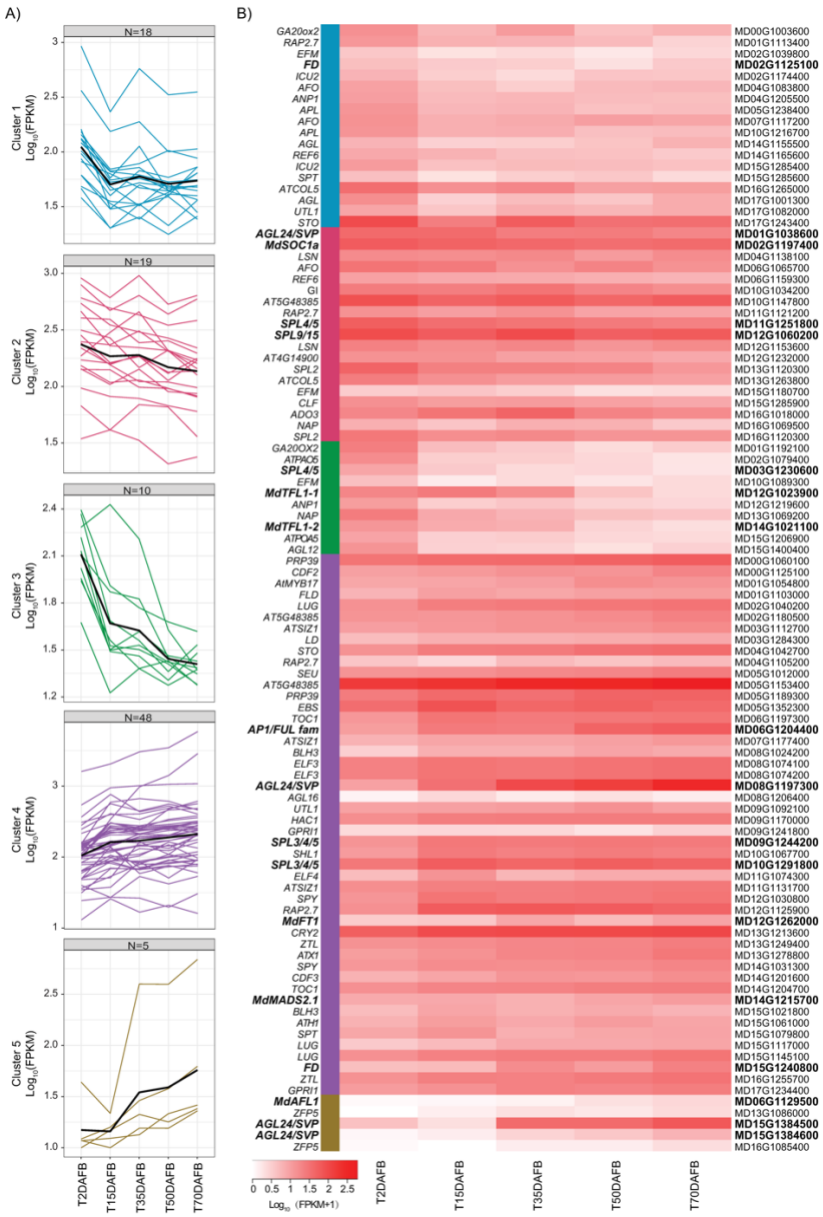


Figure 2.4. Expression patterns of flowering gene homologs in flowering-induced apices. A) Expression plots for each of the five clusters. The average expression pattern is represented by a black line. B) Heatmap and K-means clustering of expression values sorted by cluster assignment for the flowering gene homologs that exhibited significant changes in expression across the study period.

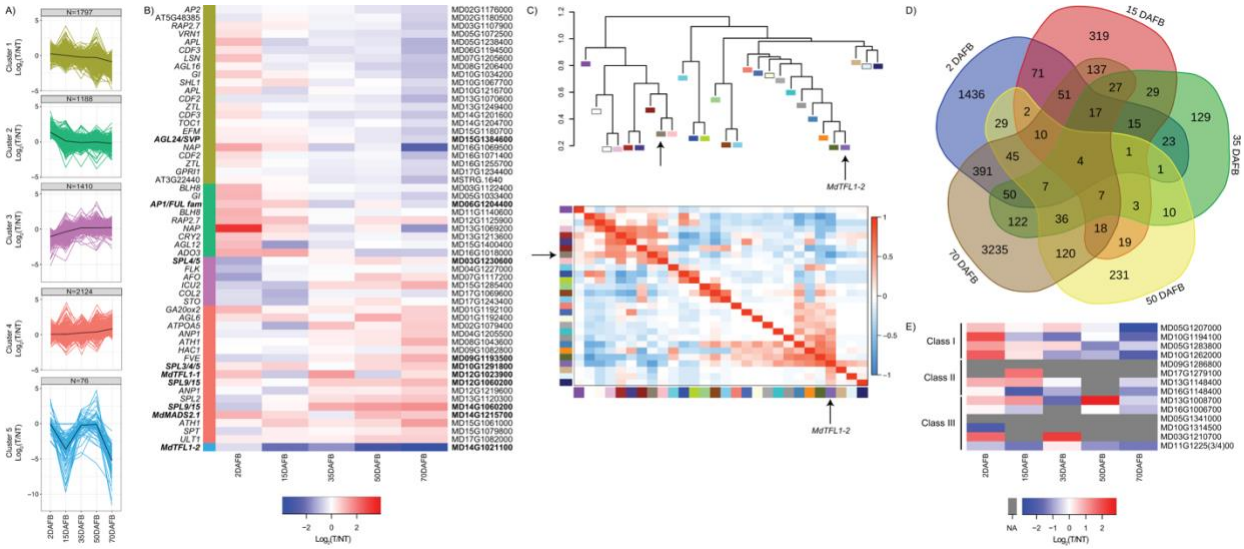


Figure 2.5. Heatmap of differential expression of genes between thinned and non-thinned apices across the study period. A) K-means clusters of all differentially expressed genes between the two conditions at each time point. B) Heatmap of the fold-difference in expression of flowering gene homologs at each sample date. Homologs of the intensively studied flowering genes are indicated in bold text. C) Dendrogram of co-expressed gene modules and a heatmap of the correlation between modules. D) Overlap among differentially expressed genes at each sample date. E) Heatmap of fold change in expression of *GA2ox* homologs.

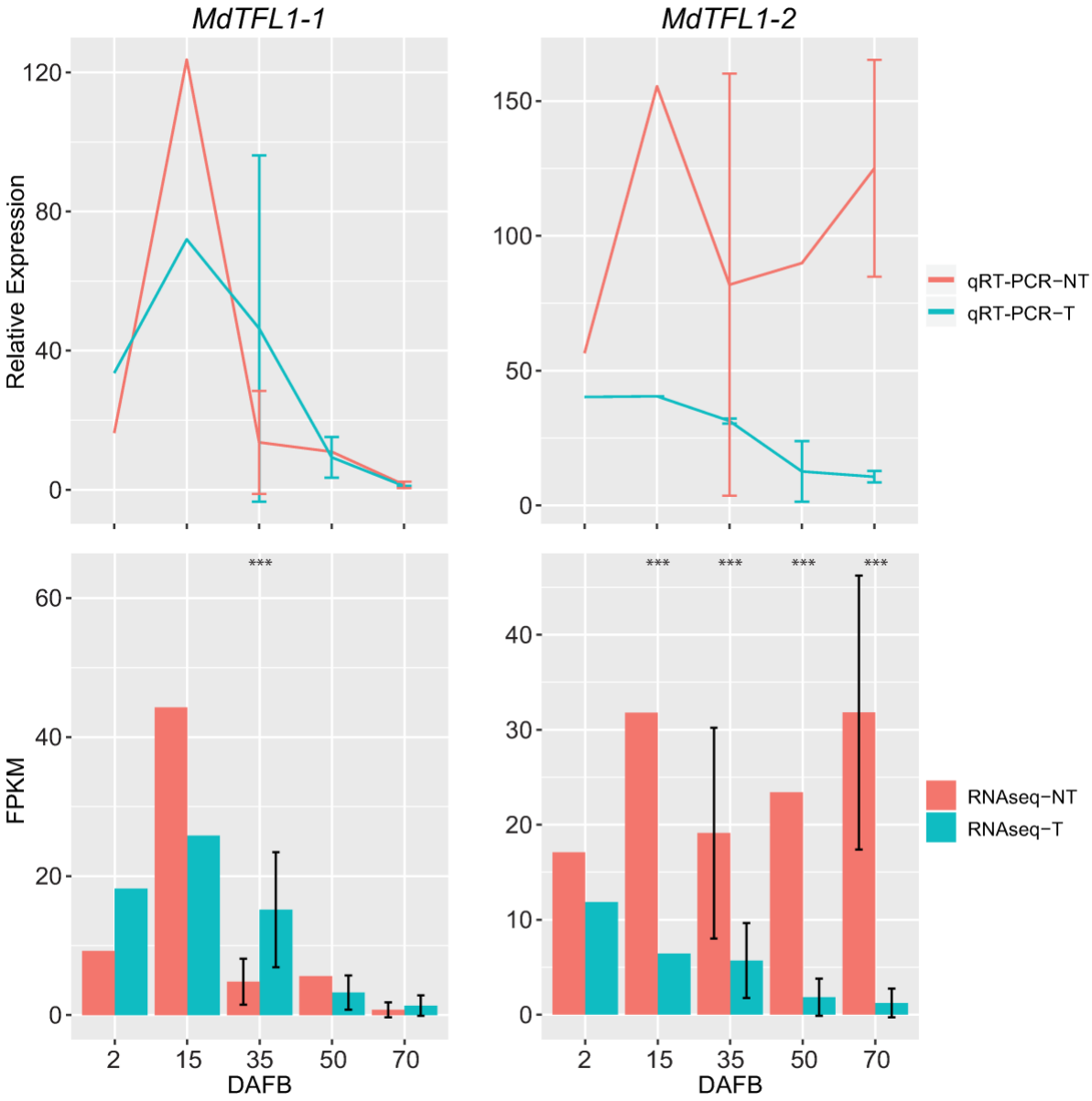


Figure 2.6. **Expression profiles of *MdTFL1-1* and *MdTFL1-2* calculated from qPCR (upper panels) and RNAseq (lower panels).** The correlation between the relative expression and FPKM values for *MdTFL1-1* and *MdTFL1-2* had R^2 values of 0.97 and 0.99, respectively. Triple asterisks (***) indicate a q value <0.001. NT = Non-thinned and T = Thinned

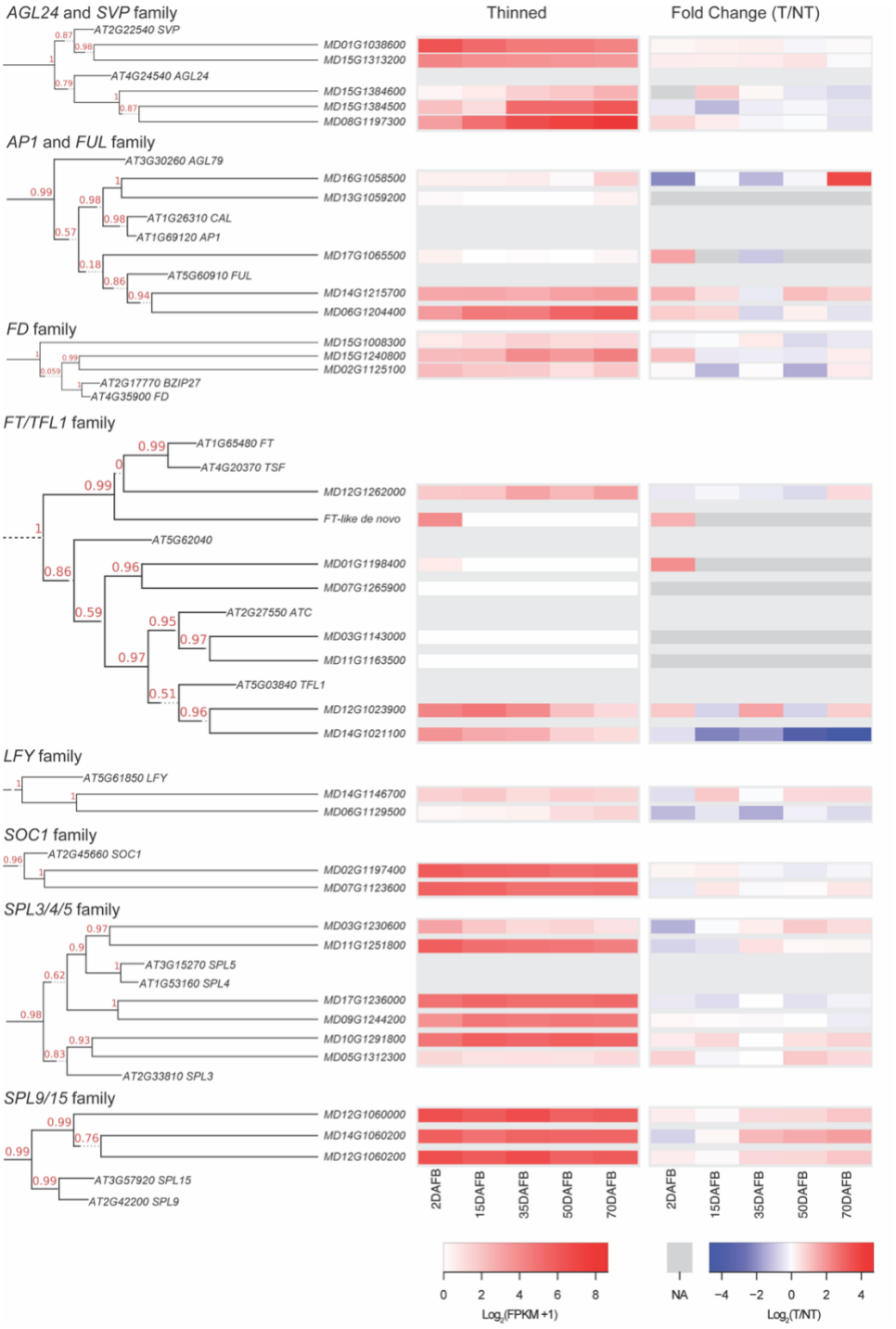


Figure 2.7. Heatmap of expression of homeologous gene pairs under various crop load conditions and the resulting fold-change between conditions across the study period.

REFERENCES

REFERENCES

1. Butler O. On the cause of alternate bearing in apple. Bulletin of the Torrey Botanical Club. 1917;44(2): 85-96.

2. Jonkers H. Biennial bearing in apple and pear: a literature survey. Sci Hortic. 1979;11: 303-317.

3. Monselise SP, Goldschmidt EE. Alternate bearing in fruit trees. In: Janick J, editor. Horticultural Reviews. Westport, CT: AVI Publishing Company Inc; 1982. pp.128-173.

4. McArtney S, Hoover E, Hirst P, Booking IR. Seasonal variation in the onset and duration of flower development in 'Royal Gala' apple buds. J Hort Sci Biotech. 2001;76(5): 536-540.

5. Foster T, Johnston R, Seleznyova A. A morphological and quantitative characterization of early floral development in Apple (Malus x domesitca Borhk.). Ann Bot. 2003;92: 199-206.

6. Hoover E, De Silva N, McArtney S, Hirst P. Bud development and floral morphogenesis in four apple cultivars. J Hort Sci Biotech. 2004;79(6): 981-984.

7. Dadpour MR, Movafeghi A, Grigorian W, Omidi Y. Determination of floral initiation in Malus domestica: A novel morphogenetic approach. Biol Plant. 2011;55(2): 243-252.

8. Dennis FG, Neilsen JC. Physiological factors affecting biennial bearing in tree fruit: the role of seeds in apple. HortTechnology. 1999;9(3): 317-322.

9. Luckwill LC, Weaver P, MacMillan J. Gibberellins and other growth hormones in apple seeds. J Hort Sci. 1969;44: 413–424.

10. Wardlaw IF. The control of carbon partitioning in plants. New Phytol. 1990;116: 341-381.

11. Wada N, Cao Q, Kotoda N, Soejima J, Masuda T. Apple has two orthologues of FLORICAULA/LEAFY involved in flowering. Plant Mol Bio. 2002;49: 567-577.

12. Kotoda N, Wada M, Kusaba S, Kano-Murakami Y, Masuda T, Soejima J. Overexpression of MdMADS5, an APETALA1-like gene of apple, causes early flowering in transgenic Arabidopsis. Plant Sci. 2002;162(5): 679-687.

13. Kotoda N, Wada M. MdTFL1, a TFL1-line gene of apple, retards the transition from the vegetative to reproductive phase in transgenic Arabidopsis. Plant Sci. 2005;168(1): 95-104.

14. Hättasch C, Flachowsky H, Kapturska D, Hanke M. Isolation of flowering genes and seasonal changes in their transcript levels related to flower induction and initiation in apple (Malus domestica). Tree Physiol. 2008;28: 1459-1466.

15. Cevik V, Ryder CD, Popovich A, Manning K, King Gj, Seymour GB. A FRUITFUL-like gene is associated with genetic variation for fruit flesh firmness in apple (Malus domestica Borkh.). Tree Genet Genomes. 2010;6: 271-279.

16. Kotoda N, Hayashi H, Suzuki M, Igarashi M, Hatsuyama Y, Kidou S, et al. Molecular characterization of FLOWERING LOCUS T-Like gene in apple (Malus × domestica Borkh.). Plant Cell Physiol. 2010;51(4): 561-575.

17. Guitton B, Kelner JJ, Celton JM, Sabau X, Renou JP, Chagné D, et al. Analysis of transcripts differentially expressed between fruited and deflowered 'Gala' adult trees: a contribution to biennial bearing understanding in apple. BMC Plant Biol. 2016;16: 55.

18. Haberman A, Ackerman M, Crane O, Kelner JJ, Costes E, Samach A. Different flowering response to various fruit loads in apple cultivars correlates with degree of transcript reaccumulation of a TFL1-encoding gene. Plant J. 2016;87: 161-173.

19. Mimida N, Kotoda N, Ueda T, Igarashi M, Hatsuyama Y, Iwanami H, et al. Four TFL1/CEN-like genes on distinct linkage groups show different expression patterns to regulate vegetative and reproductive development in apple (Malus × domestica Borkh.). Plant Cell Physiol. 2009;50(2): 394-412.

20. Kotoda N, Wada M, Komori S, Kidou S, Abe K, Masuda T, et al. Expression pattern of homologues of floral meristem identity Genes LFY and AP1 during flower development in apple. J Amer Soc Hort Sci. 2000;125(4): 398-403.

21. Li WM, Tao Y, Yao YX, Hao YJ, You CX. Ectopic over-expression of two apple Flowering Locus T homologues, MdFT1 and MdFT2, reduces juvenile phase in Arabidopsis. Biol Plant. 2010;54(4): 639-646.

22. Tränkner C, Lehmann S, Hoenicka H, Hanke M, Fladung M, Lenhardt D, et al. Over-expression of an FT-homologous gene of apple induces early flowering in annual and perennial plants. Planta. 2010;232(6): 1309-1324.

23. Mimida N, Kidou S, Iwanami H, Moriya S, Abe K, Voogd C, et al. Apple FLOWERING LOCUS T proteins interact with transcription factors implicated in cell growth and organ development. Tree Physiol. 2011;31(5): 555-566.

24. Flachowsky H, Szankowski I, Waidmann S, Peil A, Tranker C, Hanke MV. The MdTFL1 gene of apple (Malus × domestica Borkh.) reduces vegetative growth and generation time. Tree Physiol. 2012;32: 1288-1301.

25. Munoz-Fambuena N, Mesejo C, Gonzalez-Max MC, Primo-Millo E, Agusti M, Iglesias DJ. Fruit regulates seasonal expression of flowering genes in alternate-bearing 'Moncada' mandarin. Ann Bot. 2011;108: 511-519.

26. Ziv D, Zviran T, Zezak O, Samach A, Irihimovitch V. Expression profiling of FLOWERING LOCUS T-like gene in alternate bearing 'Hass' avocado trees suggests a role of PaFT in avocado flower induction. PloS One. 2014;9(10): e110613.

27. Kittikorn M, Okawa K, Ohara H, Kotoda N, Wada M, Yokoyama M, et al. Effects of fruit load, shading, and 9,10-ketol-octadecadienoic acid (KODA) application on MdTFL1 and MdFT1 genes in apple buds. Plant Growth Regul. 2011;64: 75-81.

28. Gasic K, Hernandez A, Korban SS. RNA extraction from different apple tissues rich in polyphenols and polysaccharides for cDNA library construction. Plant Mol Bio Rep. 2004;22: 437a-437g.

29. Aronesty E. ea-utils : "Command-line tools for processing biological sequencing data"; https://github.com/ExpressionAnalysis/ea-utils 2011.

30. Daccord N, Celton JM, Linsmith G, Becker C, Choisne N, Schijlen E, et al. High-quality de-novo assembly of the apple genome and methylome dynamics of early fruit development. Nat Genet. 2017;49: 1099-1106.

31. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12: 357-360.

32. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotech. 2015;33: 290-295.

33. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc. 2016;11: 1650-1667.

34. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7: 562-578.

35. Trapnell C, Williams B, Pertea G, Mortazavi A, Kwan G, van Baren J, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotech. 2010;28(5): 511-515.

36. Deluca DS, Levin JZ, Sivachenko A, Fennell T, Nazaire MD, Williams C, et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics. 2012;28(11): 1530-1532.

37. Pertea G, Pertea M. GFF Utilities: GffRead and GffCompare. F1000 Research. 2020;9: 304.

38. Yu-Jian K, De-Chang Y, Lei K, Mei H, Yu-Qi M, Liping W, et al. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. Nucleic Acids Res. 2017;45(W1): W12–W16.

39. Li A, Zhang J, Zhou Z. PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. BMC bioinformatics. 2014;15: 311.

40. Wang L, Park HJ, Dasari S, Wang S, Kocher JP, Li W. CPAT: Coding-potential assessment tool using an alignment-free logistic regression model. Nucleic Acids Res. 2013;41(6): e74.

41. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013;8(8): 1494-1512.

42. Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, et al. The Arabidopsis Information Resource: Making and mining the "gold standard" annotated reference plant genome. Genesis. 2015;53: 474-485.

43. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics 2008; 10:421.

44. Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, et al. The NCBI BioSystems database. Nucleic Acids Res. 2010;38: D492-D496.

45. Huerta-Cepas J, Serra F, Bork P. ETE 3: Reconstruction, analysis and visualization of phylogenomic data. Mol Biol Evol. 2016;33(6): 1635-1638.

46. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7: 539.

47. Price MN, Dehal PS, Arkin AP. FastTree 2 - approximately maximum-likelihood trees for large alignments. PLoS One 2010;5(3): e9490.

48. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7): e49.

49. Krzywinski MI, Schein JE, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: An information aesthetic for comparative genomics. Genome Res. 2009;19: 1639-1645.

50. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www. R-project.org. 2013.

51. Goff L, Trapnell C, Kelley D. cummeRbund: Analysis, exploration, manipulation, and visualization of Cufflinks high-throughput sequencing data. 2013;R package version 2.22.0.

52. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9: 559.

53. Zhang B, Horvath S. A General Framework for Weighted Gene Co-Expression Network Analysis. Stat Appl Genet Mol. 2005;4(1): 17.

54. Zhang S, Gottschalk C, van Nocker S. Genetic mechanisms in the repression of flowering by gibberellins in apple (Malus × domestica Borkh.). BMC Genomics. 2019; 20:747.

55. Embree CG, Myra MTD, Nichols DS, Wright AH. Effect of blossom density and crop load on growth, fruit quality, and return bloom in 'Honeycrisp' apple. HortScience. 2007;42(7): 1622-1625.

56. Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, et al. The genome of the domesticated apple (Malus × domestica Borkh.). Nat Genet. 2010;42(10): 833–839.

57. Peace CP, Bianco L, Troggio M, van de Weg E, Howard NP, Cornille A, et al. Apple whole genome sequences: recent advances and new prospects. Hortic Res. 2019;6: 59.

58. Takeuchi T, Matsushita MC, Nishiyama S, Yamane H, Banno K, Tao R. RNA-sequencing analysis identifies genes associated with chilling-mediated endodormancy release in apple. J Am Soc Hort Sci. 2018;143(3): 194-206.

59. Nishiyama S, Matsushita MC, Yamane H, Honda C, Okada K, Tamada Y, et al. Functional and expressional analysis of apple FLC-like in relation to dormancy progress and flower bud development. Tree Physiol. 2019;tpz111.

60. Ng M, Yanofsky MF. Function and evolution of the plant MADS-box gene family. Nat Rev Genet. 2001;2: 186-195.

61. Tian Y, Dong Q, Ji Z, Chi F, Cong P, Zhou Z. Genome-wide identification and analysis of the MADS-box gene family in apple. Gene. 2015;555(2): 277-290.

62. Andrés F, Coupland G. The genetic basis of flowering responses to seasonal cues. Nat Rev Genet. 2012;13: 627-639.

63. Xu M, Hu T, Zhao J, Park M, Earley KW, et al. Developmental functions of miR156-regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) genes in Arabidopsis thaliana. PLoS Genet. 2016;12(8): e1006263.

64. Liu C, Chen H, Er HL, Soo HM, Kumar PP, et al. Direct interaction of AGL24 and SOC1 integrates flowering signals in Arabidopsis. Development. 2008;135: 1481-1491.

65. Bradley D, Ratcliffe O, Vincent C, Carpenter R, Coen E. Inflorescence Commitment and Architecture in Arabidopsis. Science. 1997;275(5296): 80-83.

66. Ratcliffe OJ, Bradley DJ, Coen ES. Separation of the shoot and floral identity in Arabidopsis. Development. 1999;126:1109-1120.

67. Goslin K, Zheng B, Serrano-Mislata A, Rae L, Ryan PT, Kwaśniewska K, et al. Transcription factor interplay between LEAFY and APETALA/CAULIFLOWER during floral initiation. Plant Physiol. 2017;174(2):1097-1109.

68. Kinoshita T, Harada JJ, Goldberg RB, Fischer RL. Polycomb repression of flowering during early plant development. PNAS. 2001;98(24):14156-14161.

CHAPTER THREE
Molecular basis of diversity in floral initiation and development in *Malus*

ABSTRACT

An important limitation to commercial apple production is freezing temperatures during the spring blooming period which can damage floral organs. Members of the *Malus* genus, including the domesticated apple, wild species, and hybrids all bloom in spring, and exhibit wide variation in spring bloom date. Although variation in seasonal bloom time is thought to be strongly influenced by genotypic differences in the chilling requirement, contributions of other factors have not been well studied. Here, we documented the seasonal timing of floral initiation and early floral development for three extreme early-blooming and three extreme late-blooming apple accessions through transcriptional profiling. The generated transcriptomic data also provided an opportunity to thoroughly identify homologs of floral initiation and developmental genes within *Malus*. The floral initiation genes were found to be well-conserved between accessions, whereas the floral development genes were less conserved. Transcript profiling identified shared differential regulation of the *MdTFL1* genes between five of the six accessions in July, which is the anticipated peak timing of floral initiation. This result suggests *MdTFL1* has a role in regulating the onset of floral initiation. In October, the expression of floral development homologs was relatively consistent across cultivars, except for *SEP1/2* and *SEP3.* These two genes were highly expressed in one accession of *M. angustifolia* relative to the other accessions. This result correlates with *M. angustifolia's* delayed floral development, where the anatomical dissections indicated the accession had just entered the sepal initiation stage. Taken together, these results suggest that genotypic variation in floral initiation and early development differs among the members of the *Malus* genus. Our results also indicate that floral initiation timing, rate of development, and stage at which the apices enter dormancy has no correlation with the respective spring bloom time.

INTRODUCTION

The domesticated apple (*Malus* × *domestica* Borkh.) is an important source of nutrients and energy throughout the temperate regions of the world. Similar to other perennial tree fruits, domesticated apple exhibits a two-season flowering cycle. Floral meristems are initiated and begin to develop during the first growing season, remain dormant during the winter, and complete development and bloom the following spring. Commercial apple cultivars and wild *Malus* species can exhibit a wide range of spring bloom times with some cases spanning greater than 20 days (Gottschalk and van Nocker, 2013). This natural variation in bloom time can have negative impacts on commercial apple production. For example, apple typically outcrosses, thus growers need to select cultivars that bloom synchronously to ensure adequate cross-pollination (Dennis, 2003; Ramírez and Davenport, 2013). In addition, cultivars that bloom late in the spring are prone to the devastating disease fire blight, as the bacterial pathogen agent, *Erwinia amylovora* enters the plant through open flowers in warm and humid conditions (Spotte et al., 1976; Jones, 1992). Lastly, early-blooming apple cultivars are susceptible to freezing injury from ‘spring frosts' that occur in temperate growing regions (Cannell and Smith, 1986; Rodrigo, 2000; Campoy et al., 2011). Spring freezing injury also impacts other tree fruit crops, and is a major determinant of production for *Prunus* species. Tree fruit crop loss due to spring freeze events is becoming of greater concern as climate change further exacerbates extreme and unpredictable weather patterns (Cannell and Smith 1986; Legave et al., 2007; Unterberger et al., 2018). These production considerations underscore the importance of developing improved cultivars with bloom times tailored to specific environments.

In apple, bloom time is conditioned by both genotype and environment (Hauagge and Cummins, 1991; Lawson et al., 1995; Liebhard et al., 2003; Celton et al., 2011; Gottschalk and

van Nocker, 2013). Numerous quantitative trait loci (QTL) influencing bloom time have been identified in apple (Lawson et al., 1995; Liebhard et al., 2003; Celton et al., 2011). The identified loci were generally distinct between different studies, and none were found to have a major effect. In other Rosaceous fruit trees, including peach (*Prunus persica* (L.) Batsch), cherry (*Prunus cerasus* L.), and almond (*Prunus dulcis* (Mill.) D. A. Wade), large-effect QTLs for bloom time have been identified (Ballester et al., 2001; Fan et al., 2010; Sánchez-Pérez et al., 2012; Castede et al., 2014; Cai et al., 2018). The high degree of synteny among *Prunus* species and the colocalization of major QTLs on linkage groups 1 and 4 suggest there may be common mechanisms controlling bloom time in these species (Dirlewanger et al., 2004; Fan et al., 2010).

In domesticated apple, the inflorescence is initiated terminally on shoots early in the growing season. The meristem differentiates into a terminal and lateral floral meristems with subtending bracts and bractlets, respectively. Floral meristems continue to develop throughout the summer, and arrest in an incompletely developed state prior to winter dormancy. Following an extended period of chilling, warm temperatures in early spring promote final development (Foster et al., 2003). At the developmental level, the seasonal time of spring bloom could be conditioned by the timing of floral meristem initiation, rate of development prior to and during winter, amount of chilling required for exit from dormancy, and the rate of development in spring. The timing of floral initiation has been reported to be variable among *M. x domestica* cultivars (McArtney et al., 2001; Foster et al., 2003; Hoover et al., 2004; Dadoup et al., 2011). Wide natural variation also exists for chilling requirements (Hauagge and Cummins, 1991). Although chilling requirement obviously plays an important role in bloom time in apple, the extent to which other factors contribute to bloom time has not been addressed. Furthermore,

there is only limited information on natural variation in floral initiation and development among apple cultivars and *Malus* species.

We previously documented natural diversity in spring bloom time among nearly 1,800 distinct Malus accessions, comprising 31 wild species, and >1,000 domestic cultivars and hybrids, maintained at the USDA Agricultural Research Service Plant Genetic Resources Unit (Geneva, NY) (Gottschalk and van Nocker, 2013). This resulted in the identification of subsets of accessions with extreme early or extreme late bloom times. In the following study, we combined anatomical observations and transcriptional profiling studies for six representative, from the extreme-early or extreme-late bloom groups to investigate the potential relationship between bloom time, timing of floral initiation, and rate of development. Representative apices were collected at seasonal timings that were anticipated to represent periods of floral initiation and development. Two collection dates that correspond to the periods of floral initiation and development were also transcriptional profiled to identify transcriptional differences between the genotypes.

METHODS

Plant material and sample collection

Shoot apices were randomly collected from six apple accessions the belong to extreme late and early bloom groups (Gottschalk and van Nocker 2013). The specimen trees were located at the USDA Agricultural Research Service Plant Genetic Resources Unit (USDA ARS PGRU) in Geneva, NY (Table 3.1). Apex collections occurred on July 25th, August 28th, October 5th, and November 8th, 2018, as well as on February 26th, 2019. Bud scales, when present, were removed. Pools of 25-35 apices/buds were collected for RNA extraction and were immediately frozen in liquid nitrogen and stored at -80℃.

RNA extraction and sequencing

Frozen apices were ground into a fine powder using a mortar and pestle under liquid nitrogen, and RNA was extracted from the frozen ground tissue using a CTAB-based extraction method (Gasic et al. 2004). Spermidine was substituted for spermine in the extraction buffer. Extracted RNA was then further purified using a commercial kit (RNeasy Plant; Qiagen). Purified RNA was assessed for quality using 1.2% formaldehyde gel electrophoresis, and quantified by spectroscopy (NanoDrop 2000c, ThermoScientific, Waltham, MA USA). RNA libraries were prepared and sequenced by Novogene (Sacramento, CA USA) on the Illumina HiSeq platform generating paired-end, 150-base reads. The sequencing generated a total of ~1.63 billion reads, with an average of ~25.9 million reads per library.

Transcriptome assembly and analysis

Computational quality filtering of raw sequences was performed by Novogene and included removal of adapter sequence, reads containing Ns >10% of read length, and reads with Phred quality (Q)-scores ≤5. Filtered reads were then aligned to a draft *Malus* × *domestica* genome sequence generated from a doubled haploid individual (GDDH13v.1.1; Daccord et al., 2017) using the HISAT2 (v.2.1.0) aligner with the -dta-cufflinks parameter (Kim et al., 2015). Transcript models were constructed using the alignment files and the transcriptome assembler StringTie (v.1.3.3) with default parameters (Pertea et al., 2015). The StringTie -G option was invoked to guide the construction based on externally supplied transcriptomes, which consisted of annotated transcripts from the GDDH13v.1.1 reference genome supplemented with additional novel loci identified in a previous study of flowering in apple (Gottschalk et al. Chapter 2). Transcriptome comparisons were conducted using the GFFcompare feature within the GFF Utilities package (v.0.11.2; Pertea and Pertea, 2020). Determination of a gene to be in an

expressed state or basally expressed was calculated using a hierarchical Bayesian mixture model using FPKM values retrieved from the output of Cuffnorm (v.2.2.1 Goff et al., 2019) and executed through the R (v.3.6.2; R Core Team 2013) program Zigzag (v.0.1.0; Thompson et al. 2020). Genes with a calculated probability ≥0.8 were determined to be in an expressed state. Plots generated to display the intersections of expressed genes between accessions were generated using the R program UpSet (v.1.4.0; Lex et al., 2014). Differential expression analysis used the multifactor comparisons scheme in the R program edgeR (v.3.11; McCarthy et al., 2012) with raw gene counts that were retrieved from denormalized Cuffnorm outputs. Differentially expressed genes (DEGs) were identified with FDR adjusted P-value <0.05 and a log fold change of >2 or <-2 for each comparison.

Analysis of expression for landmark flowering genes

Previously identified apple homologs of floral initiation genes were identified as described by Gottschalk et al. (Chapter 2). In brief, we indexed genes GO term annotations related to potential roles in flowering from Arabidopsis (TAIR10):GO terms 0048438 ('floral whorl development'), 0009908 ('flower development'), 0009910 ('negative regulation of flower development'), 0009911 ('positive regulation of flower development'), 0048578 ('positive regulation of long-day photoperiodism, flowering'), 0010220 ('positive regulation of vernalization response'), 0009909 ('regulation of flower development'), 0048510 ('regulation of timing of transition from vegetative to reproductive phase'), 0010321 ('regulation of vegetative phase change'), 0010228 ('vegetative to reproductive phase transition of meristem'), 0010048 ('vernalization response'), and 0010093 (‘specification of floral organ identity’). This data set contained 437 genes which were then manually curated to only those genes that have been intensively studied for roles in the flowering pathway. This subset included *AG*, *AGL24*, *AP1*-

AP3, *FD*, *FLC*, *FT*, *FUL*, *LFY*, *LSN*, *PI*, *SEP1-SEP4*, *SOC1*, *SPL3-SPL5*, *SPL9*, *SPL15*, *SUP*, *SVP*, *TFL1*, *TSF*, and *UFO*. Conceptual translations of the corresponding representative gene models were obtained from TAIR (TAIR10_pep_20110103_representative_gene_model_updated) and used as queries to search genotype-specific protein databases for each apple genotype in this study, using BLASTp and an expect (E) value cutoff of $1e^{-12}$ (Camacho et al., 2008).

For each of the six genotypes, protein databases were generated using the following procedure. A consensus set of transcript models were collected from a combined transcriptome representing the July and October datasets, as well as a single library of pooled RNA from June, August, and November for each genotype separately. The resulting transcript model annotations were then merged with the GDDH13 reference gene model annotation (gene_models_20170606.gtf) using StringTie -merge option. Nucleotide FASTA sequences were then extracted from the merged transcriptome using the GFF utilities package (Pertea and Pertea, 2020). For transcript models corresponding to annotated transcripts in GDDH13, a single corresponding GDDH13 protein model was adopted. For transcript models assembled *de novo* by StringTie that were not organized within GDDH13 reference loci had all their corresponding transcript sequences extracted and retained for analysis. The extracted nucleotide sequences were then processed by TranDecoder (v.5.5.0; https://github.com/TransDecoder) to construct open reading frame translations. TransDecoder often predicts multiple open reading frame products from a single transcript model, which resulted in redundancy for non-reference *de novo* transcript models. In this case, all translations were retained for each genotype-specific protein sequence dataset.

For each genotype-specific protein dataset, individual peptide sequences were used as a query to search against all Arabidopsis representative gene model translations using Blastp. The top 50 returned blast hits were then used as queries to blast back into the apple peptide sequences (*i.e.* reciprocal BLAST). The 50 highest-scoring apple transcript translations and the 25 highest-scoring Arabidopsis gene translations identified with each apple sequence query were then used for phylogenetic tree construction. Phylogenetic trees were constructed using the ETE3 toolkit (v.3.1.1) build function invoking the standard_fasttree workflow under default settings (Huerta-Cepas et al., 2016; Sievers et al., 2011; Price et al., 2010). Representative flowering homologs were then identified through manual curation of each phylogenetic tree.

RESULTS

Sequencing and transcriptome assembly results

To determine if potential distinctions in seasonal floral initiation and development exist and may influence spring bloom, we carried out transcriptional profiling of the developing shoot apices focusing on the expression of flowering-related genes. Apices collected on July 25th, and Oct 5th, representing anticipated stages of early and late floral developmental, were subjected to RNA extraction and sequencing. For each accession per time point, between 5.7 and 10 high-quality gigabases of sequence were generated. Sequence reads were then aligned to a draft *Malus* genome sequence generated from a doubled-haploid clone of 'Golden Delicious' (GDDH13; Daccord et al., 2017), and transcripts were modeled based on established computational approaches. As anticipated, given that the reference sequence GDDH13 was derived from *M.* × *domestica*, sequences generated from the *M.* × *domestica* accessions 'Anna' and 'Koningszuur', and *M. sylvestris* exhibited the highest alignment rates (>90%). The remaining wild species exhibited lower alignment rates, with *M. angustifolia* having the lowest (~74%). The similarly

high alignment rates of the *M.* × *domestica* cultivars and *M. sylvestris* are indicative of the domestication history of *M.* × *domestica*, of which *M. sylvestris* is a progenitor (Velasco et al. 2010). The lower alignment rates of the wild species reflects their evolutionary distance from *M.* × *domestica* (Nikiforova et al., 2013).

In a previous study, we defined a clear subset of genes associated with floral initiation (Gottschalk et al., Chapter 2). The previous census of flowering genes was incomplete as it did not catalog genes involved in floral development. To identify genes with possible conserved function in the development of floral meristems and floral organ primordia, we selected a subset of well-studied floral development genes in Arabidopsis and identified their homologs in these *Malus* transcriptomes. We identified apple genes related to *AG, SHP1/2, AP2, AP3, LSN, PI/TOE3, SEP1/2, SEP3, SUP,* and *UFO* that had significant blast homology and belonged to well-defined (>90% bootstrapped) clades for each accession. Seven of the nine families of floral development genes were found to contain pairs of apple homologs (Table 3.2). This result was anticipated due to the paleopolyploid origins of apple (Velasco et al., 2010). We identified the families of *AP3/PI* and *LSN* to have undergone an expansion event, with each family containing three homologous copies (Table 3.2). Of note, we identified a novel copy of an *FD* homolog unique to *M. orthocarpa* (Table 3.2). Additionally, *SUP* homologs were absent from most accessions except *M. angustifolia* PI589789 (two copies), and 'Koningszuur' (one copy) (Table 3.2). Both *M. angustifolia* accessions lacked homologs of *AG* (Table 3.2). Generally, the floral initiation genes were more conserved across the transcriptomes of the accessions than floral development genes.

Determining floral genes expressed state

To identify molecular pathways associated with floral initiation and development we tracked the expression for a set of genes that we previously identified as indices for floral initiation in *M.× domestica* cultivar 'Honeycrisp' (Gottschalk et al., Chapter 2) and the genes we identified here with potential roles in flower development. In total, we identified 93 flowering genes of interest. To infer the probability of their active expression, we utilized a Bayesian statistical model to compare gene expression across each of the six assembled transcriptomes independently. With this information, we then assign an expression state classification to all 93 flowering-associated homologs per accession. Any gene with a probability for expression of ≥0.8 was deemed as being in an "expressed state" as opposed to a "basal expression". Of the 93 genes surveyed, ten were found to be considered as basally expressed in all six accessions. These genes included homologs of *AG*, *AGL12*, *AP3/PI*, *SEP1/2*, *SEP3*, *SHP1/2*, *SUP*, *UFO*, and the novel *FD* transcript identified in *M. orthocarpa.* We also identified 19 genes that were deemed as expressed in one or more accessions but not in all six accessions (Table 3.3). In this subset, homologs of known roles in the promotion or repression of flowering in apple were identified. These homologs included the floral promoters *MdAFL1*, *MdMADS2.1*, and *MdMADS5* and the floral repressors *MdTFL1-1* and *MdTFL1-2.* Of the expressed genes, a total of 67 were shared across all the accessions (Fig 3.1). Also none of the expressed genes were uniquely associated with either bloom time group or were associated with accessions of the same species (*M. angustifolias* and *M. × domestica* accessions). *M. orthocarpa* and the two *M. angustifolias* were found to have a six uniquely expressed genes. For *M. orthocarpa*, the uniquely expressed genes were homologs of *AG*, *AP1*, *SEP1/2*, and *UFO.* For the two *M. angustifolias*, the uniquely expressed genes were homologs of *SUP* (PI 589763) and *MdTFL1-1* (PI 613880).

Expression patterns of flowering homologs in July

With the subset of expressed flowering genes identified, we then evaluated their expression levels across accessions during the July sample date (Fig 3.2). Hierarchical clustering was performed on the six accessions, resulting in two clades that contained both *M. angustifolia* accessions and a clade containing the other four accessions. The two *M.* x *domestica* accessions also clustered together within the larger of the two clades. Differential gene expression was then calculated for comparisons between each accession at the July sample date. The two *M. angustifolia* accessions generally exhibited a similar expression profiles, but noticeable expression differences were observed with two homologs of *AGL24/SVP,* a homolog of *SHP1/2,* and *MdTFL1-1.* These four genes were also found to meet our threshold to be classified as differentially expressed (DE) between the two genotypes. The two *AGL24/SVP* homologs were more highly expressed in *M. angustifolia* PI 589763, whereas the *SHP1/2* homolog and *MdTFL1-1* were more highly expressed in *M. angustifolia* PI 613880. The cluster analysis identified a separation of *M. orthocarpa* from a clade containing *M. sylvestris* and the two *M.* x *domestica* cultivars. The homologs that differentiated *M. orthocarpa* were *AGL24/SVP, FUL, LSNs, MdAFL1, MdFT1, MdMADS2.1, MdTFL1s, RAP2.7, and SPL4/5.* However, the only differentially expressed homolog shared between *M. orthocarpa, M. sylvestris,* and two *M.* x *domestica* accessions was *MdMADS5*. *MdMADS5* was expressed more highly in the *M. orthocarpa* relative to the other three accessions. However, *MdMADS5* was predicted to not be in an "expressed state" in *M. orthocarpa, M. sylvestris,* and two *M.* x *domestica* accessions (Table 3.3). In the subclade containing *M. sylvestris* and *M.* x *domestica* accessions, the expression profiles were strikingly similar. Although, *COL2* and *MdTFL1-2* exhibited distinct expression profiles among the accessions and were found to be DE. In this comparison, *COL2* was more

highly expressed in the two *M.* x *domestica* accessions and *MdTFL1-2* was more highly expressed in *M. sylvestris.* The only DE homologs between the two *M.* x *domestica* accessions were *ANP1* and *MdTFL1-2. ANP1* was more highly expressed in 'Anna' and *MdTFL1-2* was more highly expressed by 'Koningszuur'.

Expression patterns of flowering homologs in October

It was anticipated that by October most reproductive apices had initiated flowers and were then undergoing differentiation. Transcriptional profiling was conducted from this sample date using the same core set of flowering-related genes. Hierarchical clustering applied to the expression profiles for the accession resulting in the formation of two distinct clades (Fig 3.3). The first clade contained *M. orthocarpa, M. sylvestris,* and 'Koningszuur', and the second clade contained the two *M. angustifolia* accessions and 'Anna'. Differential gene expression was then calculated between each accession. Within the first clade, there were no DE homologs that could explained the separation of *M. orthocarpa, M. sylvestris,* and 'Koningszuur'. In the second clade, both homologs *AGs*, a homologs of *AP3/PI,* and *SUP* were found to be DE between 'Anna' and the two *M. angustifolia* accessions. However, both of the *AGs* homologs were previously identified as being "basally expressed" in all three accessions. The classification of the *AG* loci as being DE is most likely an artifact of extreme low read coverage at those loci. In this comparison, *AP3/PI* homolog was more highly expressed in 'Anna' relative to the *M. angustifolia* accessions. The *SUP* homolog was more highly expressed in the two *M. angustifolia* accessions relative to Anna. Between the two *M. angustifolia* accessions, homologs of *AGL16, SHP1/2,* both *UFOs,* and *MdTFL1-1* were found to be DE. Here, expression of the *AGL16* homolog and *MdTFL1-1* was higher in *M. angustifolia* PI 589763. The homologs of *SHP1/2* and two *UFOs* were more highly expressed in *M. angustifolia* PI 613880. Additionally, when

comparing the two three-member clades against one another no DE homologs were identified. Although, when interpreting the expression values for the *SEP1/2* and *SEP3* homologs a distinction between the two clades is observed (Fig 3.3). When inspecting the DE results, we found DE in these two genes were associated with only *M. angustifolia* PI 613880 compared to the three members of the other clade (*M. orthocarpa*, *M. sylvestris*, and 'Koningszuur').

DISCUSSION

The transcriptomes that we assembled suggested a well-conserved exome across the genus, even considering the evolutionary distance of some of the species from *Malus* × *domestica*. In particular, *M. angustifolia* is the most distant relative to *M.* x *domestica* included in our analysis and it yielded read alignments of ~74% (Nikiforova et al., 2013). In comparison, the more related *M. sylvestris* exhibited an alignment rate (>90%) similar to those observed with the *M.* x *domestica* accessions (Velasco et al. 2010; Nikiforova et al., 2013). Without individual genome sequences available for each species, we cannot definitively measure expression as a considerable percentage of reads fail to map to the reference genome. This limitation can be attributed to potential sequence divergence of these wild species and potential variations in gene copy number.

The phylogenetic analysis of flowering genes among these accessions identified many conserved homologs of floral initiation associated genes. In contrast, the floral development genes were less conserved. For example, homologs of *AG*, *SHP1/2*, *SEP3*, and *SUP* were inconsistently identified between the accessions. Generally, the two *M.* x *domestica* accessions studied contained homologs of these genes whereas, they were absent in the wild species. This suggests that there is a divergence in gene content or appreciable sequence variation of these floral development genes within the *Malus* genus. We also identified an expansion of the *AP3/PI*

and *LSN* gene families from the anticipated two copy organization to three paralogs. The expansion of these families could result in neo- or subfunctionalization.

The expression profiling of the floral initiation and developmental genes illustrates a relatively well-conserved flowering pathway between the accessions. The major difference at the July sample date was the DE of the *MdTFL1s*, which were identified as being DE between the *M. angustifolia* accessions, *M. sylvestris,* and the two *M.* x *domestica* cultivars. *MdTFL1s* have been previously reported to be associated with and function as floral repressors in apple (Kotoda and Wada, 2005; Hättasch et al., 200; Mimida et al., 2009; Guitton et al., 2016; Haberman et al., 2016; Gottschalk et al., Chapter 2). The finding that *MdTFL1s* are DE between these accessions suggests that *MdTFL1* could have a role in the determining the onset of floral initiation. The *MdTFL1s* generally exhibited lower expression at the October sample date, which is consistent with previous finding that *MdTFL1* expression decreases throughout the period of floral initiation (Gottschalk et al. Chapter 2). We also found no consistent DE homologs between the accessions at the October sampling date. This suggests a diverse transcriptome during the floral developmental stage. However, the most striking difference was the DE of the *SEP1/2* and *SEP3* homologs associated with *M. angustifolia* PI 613880 compared to *M. orthocarpa, M. sylvestris,* and 'Koningszuur'. These genes are required for the development of the B and C floral organ whorls (Pelaz et al., 2000). The finding that *SEP1/2* and *SEP3* are DE at this time point suggests that *M. angustifolia* PI 613880 is developing its B and C whorls, whereas the other accessions have either already developed their whorls or have not yet reached that stage of development (data not shown).

The results obtained from this study also indicate that the timing of floral initiation and pace of floral development do not influence spring bloom times. The transcriptomes analyzed

exhibited gene expression profiles that were strikingly similar both in and between spring bloom time groups. Furthermore, none of the flowering genes were differentially expressed and associated with a respective bloom group. The differences in expression that were identified only support the previously identified variation of the flowering processes in *M.* x *domestica* (McArtney et al., 2001; Foster et al., 2003; Hoover et al., 2004; Dadoup et al., 2011). However, this results indicate the variation extends beyond *M.* x *domestica* and exists within the genus at-large. Ultimately, the findings we present here support a conclusion that the flowering processes in apple are not a determining factor nor influential in spring bloom time difference.

Table 3.1. Tissue collections for anatomical (D) and transcriptome (T) analyses from selected early and late-blooming *Malus* species from the USDA ARS PGRU *Malus* germplasm collection.

			Collection Dates and Purpose of Tissue				
Bloom Group	**Cultivar name**	**PI number**	7/25/18	8/28/18	10/5/18	11/8/18	2/26/19
Extreme early bloom							
M. orthocarpa		589392	A, T	A	A, T	A	A
M. sylvestris		633824	A, T	A	A, T	A	A
M. x *domestica*	Anna	280400	A, T	A	A, T	A	A
Extreme late bloom							
M. angustifolia		589763	A, T	A	A, T	A	A
M. angustifolia		613880	A, T	A	A, T	A	A
M. x *domestica*	Koningszuur	188517	A, T	A	A, T	A	A

Table 3.2. Identified homologs of Arabidopsis floral initiation and development genes in apple. "+" indicates the presence of the homolog in the respective transcriptome. NA indicates the absence of the homolog in the respective transcriptome.

				Presence in Transcriptome					
Family		**Locus**	**Clade Assignment**	***M. angustifolia* 589763**	***M. angustifolia* 613880**	**Anna**	**Koningszuur**	***M. orthocarpa***	***M. slyvestris***
AGL24 SVP		MD01G1038600	*AGL24/SVP*	+	+	+	+	+	+
		MD08G1197300	*AGL24/SVP*	+	+	+	+	+	+
		MD15G1313200	*AGL24/SVP*	+	+	+	+	+	+
		MD15G1384500	*AGL24/SVP*	+	+	+	+	+	+
		MD15G1384600	*AGL24/SVP*	NA	+	+	+	+	+
AP1 CAL FUL AGL79	***AP1 CAL***	MD13G1059200	*MdMADS5*	+	+	+	+	+	+
		MD16G1058500	*AP1*	+	+	+	+	+	+
		MD06G1204400	*FUL*	+	+	+	+	+	+
		MD14G1215700	*MdMADS2.1*	+	+	+	+	+	+
FD BZIP27		MD02G1125100	*FD*	+	+	+	+	+	+
		MD15G1008300	*FD*	NA	+	+	+	+	+
		MD15G1240800	*FD*	+	+	+	+	+	+
		MSTRG.33471.1	*FD*	NA	NA	NA	NA	+	NA
FT TSF		MD12G1262000	*MdFT1*	+	+	+	+	+	+
TFL1 ATC		MD12G1023900	*MdTFL1-1*	+	+	+	+	+	+
		MD14G1021100	*MdTFL1-2*	+	+	+	+	+	+
LFY		MD06G1129500	*MdAFL1*	+	+	+	+	+	+
		MD14G1146700	*MdAFL2*	+	+	+	+	+	+
SOC1		MD02G1197400	*MdSOC1a*	+	+	+	+	+	+
		MD07G1123600	*MdSOC1b*	+	+	+	+	+	+
SPL3 SPL4 SPL5	***SPL3***	MD05G1312300	*SPL3*	NA	+	+	+	NA	+
		MD10G1291800	*SPL3*	NA	+	+	+	NA	+
	SPL4 SPL5	MD03G1230600	*SPL4/5*	+	+	+	+	+	+
		MD11G1251800	*SPL4/5*	+	+	+	+	+	+
		MD09G1244200	*SPL4/5*	+	+	+	+	+	NA
		MD17G1236000	*SPL4/5*	+	+	+	+	+	NA
SPL9 SPL15		MD12G1060000	*SPL9/15*	+	+	+	+	+	+
		MD12G1060200	*SPL9/15*	+	+	+	+	+	+

Table 3.2. (cont'd)

		MD14G1060200	*SPL9/15*	+	+	+	+	+	+
AG ***SHP1*** ***SHP2***	***AG***	MD05G1293700	*AG*	NA	NA	+	+	NA	+
		MD10G1271000	*AG*	NA	NA	NA	+	+	+
	SHP1 ***SHP2***	MD09G1155200	*SHP1/2*	NA	NA	+	+	+	+
		MD17G1141300	*SHP1/3*	+	NA	NA	NA	NA	+
AP2 ***TOE3***		MD02G1176000	*AP2/TOE3*	+	+	+	+	+	+
		MD15G1286400	*AP2/TOE3*	+	+	+	+	+	+
AP3 ***PI***		MD02G1136500	*AP3/PI*	+	+	+	+	+	+
		MD08G1021300	*AP3/PI*	+	+	+	+	+	+
		MD15G1250200	*AP3/PI*	NA	NA	+	+	+	+
LSN		MD04G1138100	*LSN*	+	+	+	+	+	+
		MD07G1205600	*LSN*	+	+	+	+	+	+
		MD12G1153600	*LSN*	+	+	+	+	+	+
SEP1 ***SEP2*** ***SEP3*** ***SEP4***	***SEP1*** ***SEP2***	MD09G1073900	*SEP1/2*	+	+	+	+	+	+
		MD17G1065400	*SEP1/2*	+	NA	+	+	+	+
	SEP3	MD13G1121500	*SEP3*	+	+	+	+	+	+
		MD16G1121800	*SEP3*	NA	NA	NA	+	+	+
SUP ***ZPF11***		MD01G1139000	*SUP*	+	NA	NA	NA	NA	NA
		MD07G1203300	*SUP*	+	NA	NA	+	NA	NA
UFO		MD05G1246800	*UFO*	+	+	+	+	NA	+
		MD10G1227400	*UFO*	+	NA	+	+	+	+

Table 3.3. Uniquely expressed states of flowering homologs. NE – predicted to be expressed at a basal level or not expressed.

		Expression Probability					
Gene ID	**Flowering Homolog**	***M. sylvestris***	***M. orthocarpa***	**Koningszuur**	**Anna**	***M. angustifolia* 589763**	***M. angustifolia* 613880**
MD10G1271000	*AG*	NE	Expressed	NE	NE	NE	NE
MD08G1206400	*AGL16*	Expressed	Expressed	Expressed	Expressed	NE	NE
MD16G1058500	*AP1*	NE	Expressed	NE	NE	NE	NE
MD15G1008300	*FD*	Expressed	Expressed	Expressed	Expressed	NE	Expressed
MD01G1192100	*GA20OX2*	Expressed	Expressed	Expressed	NE	Expressed	Expressed
MD06G1129500	*MdAFL1*	NE	Expressed	Expressed	Expressed	Expressed	Expressed
MD14G1215700	*MdMADS2.1*	Expressed	Expressed	Expressed	Expressed	Expressed	NE
MD13G1059200	*MdMADS5*	NE	Expressed	NE	NE	Expressed	NE
MD12G1023900	*MdTFL1-1*	NE	NE	NE	NE	NE	Expressed
MD14G1021100	*MdTFL1-2*	Expressed	NE	Expressed	NE	Expressed	NE
MD09G1073900	*SEP1/2*	NE	Expressed	NE	NE	NE	NE
MD13G1121500	*SEP3*	Expressed	Expressed	Expressed	NE	NE	NE
MD10G1067700	*SHL1*	Expressed	Expressed	Expressed	Expressed	Expressed	Expressed
MD09G1155200	*SHP1/2*	Expressed	Expressed	Expressed	Expressed	NE	Expressed
MD13G1120300	*SPL2*	Expressed	Expressed	Expressed	Expressed	Expressed	Expressed
MD05G1312300	*SPL3*	Expressed	NE	Expressed	Expressed	NE	NE
MD01G1139000	*SUP*	NE	NE	NE	NE	Expressed	NE
MD14G1204700	*TOC1*	Expressed	Expressed	Expressed	Expressed	Expressed	Expressed
MD10G1227400	*UFO*	NE	Expressed	NE	NE	NE	NE

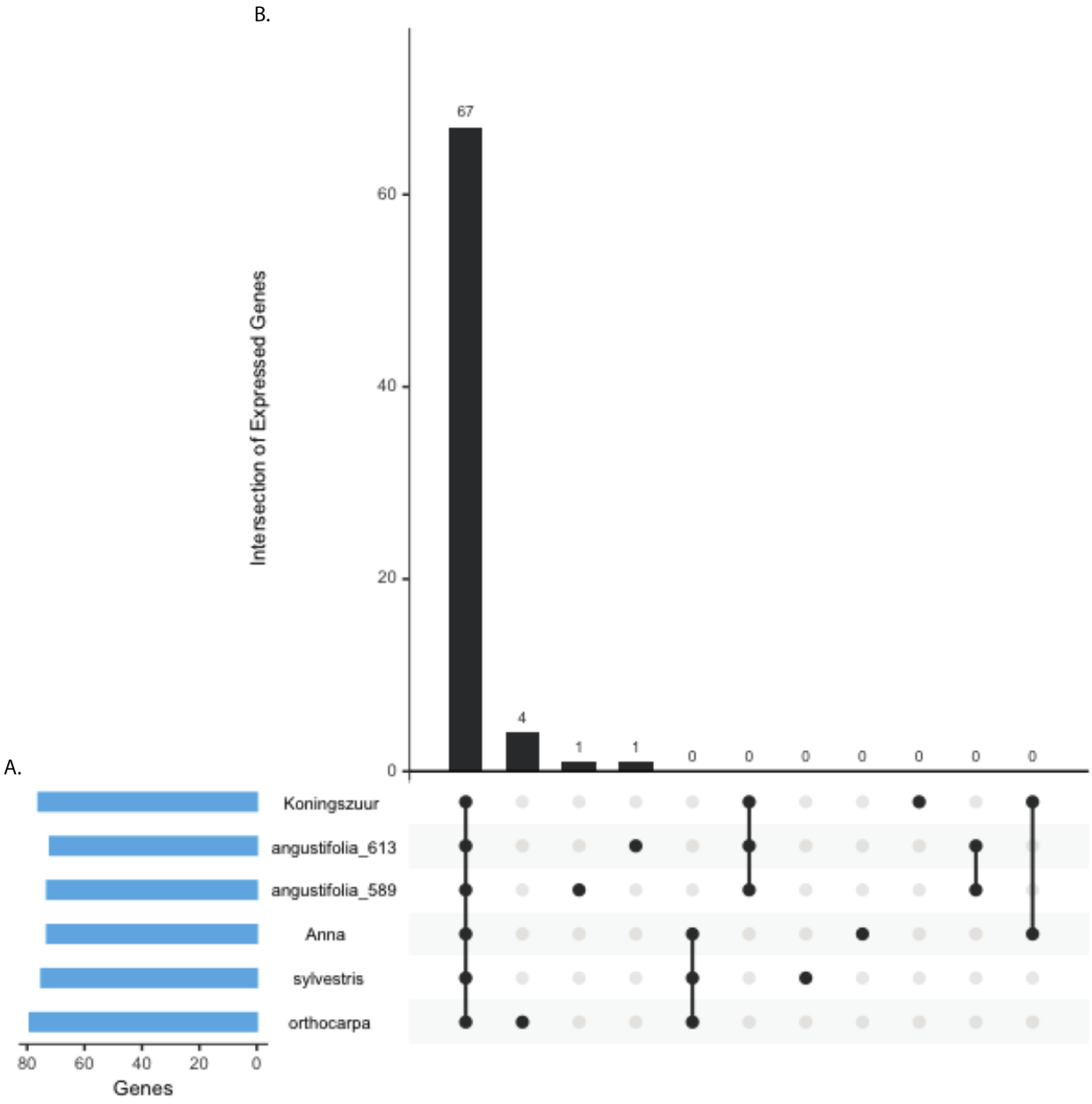


Figure 3.1. Intersections of expressed flowering genes in the transcriptomes of each accession. A) Total number of flowering genes that are expressed in each accession's transcriptome. B) The intersection of the number of expressed flowering genes between accessions. Each column of the histogram represents the number of genes that correspond with the intersection displayed as dots connected by lines below the X-axis.

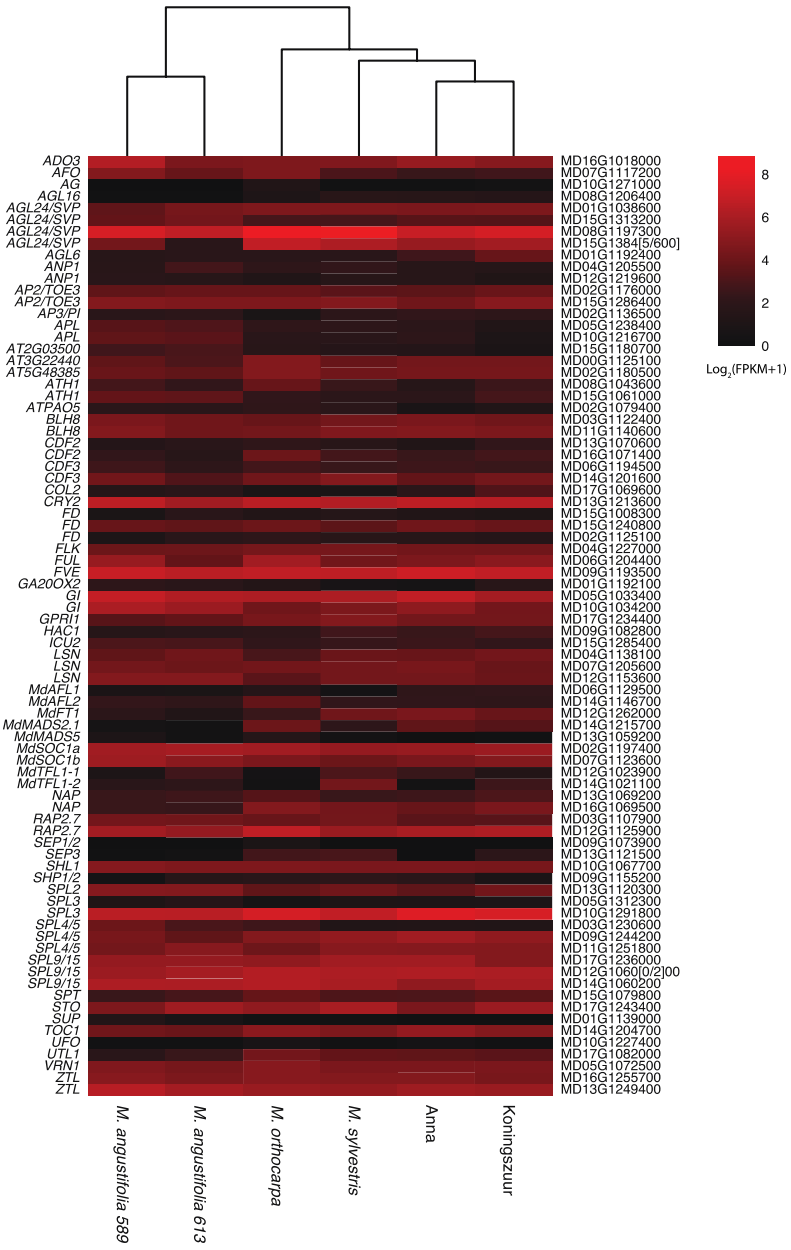


Figure 3.2. Heatmap of the expression of flowering genes in July. Homolog symbol as displayed on the left side of the heatmap and gene names on the right. Gene names that represent a locus containing more than one gene are identified by the two potential integers listed within brackets. General expression trends across all genes were utilized to perform hierarchical clusters of similarity between accessions.

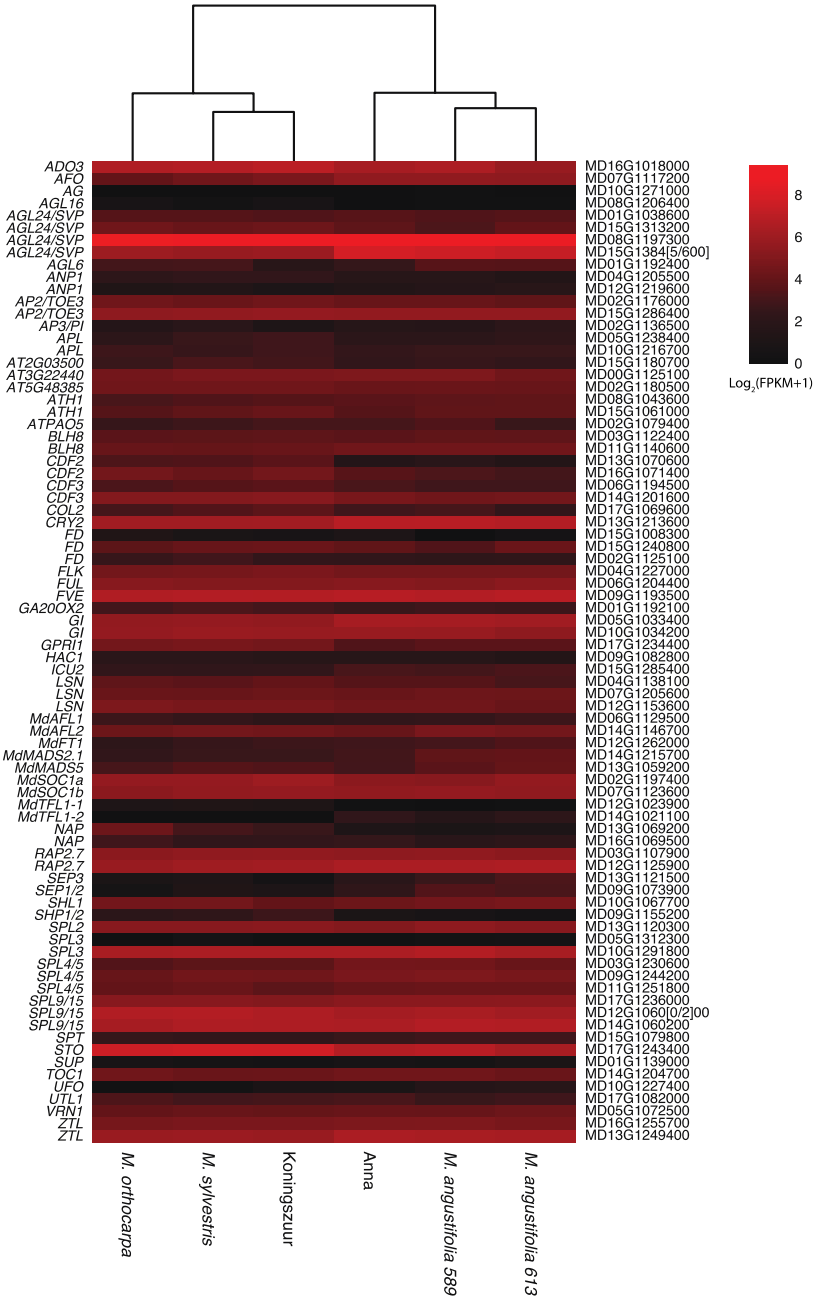


Figure 3.3. Heatmap of the expression of flowering genes in October. Homolog symbol as displayed on the left side of the heatmap and gene names on the right. Gene names that represent a locus containing more than one gene are identified by the two potential integers listed within brackets. General expression trends across all genes were utilized to perform hierarchical clusters of similarity between accessions.

REFERENCES

REFERENCES

Andres F and Coupland G. 2012. The genetic basis of flowering in response to seasonal cues. Nat Rev Genet 13(9): 627-639.

Ballester J, Socias I Company R, Arus P, and De Vicente MC. (2001). Genetic mapping of a major gene delaying blooming time in almond. Plant Breeding, 120(3): 268–270.

Blažek J, Paprštein F, and Křelinová J. (2015). Heritability of flowering time within apple progeny. Czech Journal of Genetics and Plant Breeding, 51(1), 23–28.

Buban T and Faust M. (1982). Flower bud induction in apple trees: internal control and differentiation. Horticultural Reviews, 4(6), 174–203.

Cai L, Stegmeir T, Sebolt A, Zheng C, Bink MCAM, and Iezzoni A. (2018). Identification of bloom date QTLs and haplotype analysis in tetraploid sour cherry (*Prunus cerasus*). Tree Genetics and Genomes, 14(2): 1-11.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics 2008; 10:421.

Campoy J, Ruiz D, and Ega J. 2011. Dormancy in temperate fruit trees in a global warming context: A review. Scientia Horticulturae 130(2): 357-372.

Cannell MGR and Smith RI. 1986. Climatic warming, spring budburst and frost damage on trees. J Applied Ecology 23:177-191.

Castede S, Campoy A, Garcia J, Dantec L, Lafargue M, et al. (2014). Genetic determinism of phenological traits highly affected by climate change in *Prunus avium*: flowering date dissected into chilling and heat requirements. New Phytologist, 202: 703–715.

Celton JM, Martinez S, Jammes MJ, Bechti A, Salvi S, et al. 2011. Deciphering the genetic determinism of bud phenology in apple progenies: a new insight into chilling and heat requirement effects on flowering dates and positional candidate genes. New Phytologist 192:378-392.

Daccord N, Celton J, Linsmith G, Becker G, Choisne N, et al. 2017. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. Nature Genetics 49:1099-1106.

Dennis F. 2003 Flowering, pollination and fruit set and development. Apples: botany, production and uses. Ed. D.C. Ferree and I.J. Warrington. Cambridge: CABI Publishing: 153-166.

Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldere F, Cosson P, et al. 2004. Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proceedings of the National Academy of Sciences, 101(26): 9891–9896.

Doukhanina EV, Chen S, van der Zalm E, Godzik A, Reed J, and Dickman MB. 2006. Identification and Functional Characterization of the BAG Protein Family in *Arabidopsis thaliana.* J Biological Chem 281(27): 18793-18801.

Falavigna VS, Porto DD, Buffon V, Margis-Pinheiro M, Pasquali G, and Revers LF. 2014. Differential Transcriptional Profiles of Dormancy-Related Genes in Apple Buds. Plant Mol Biol Rep 32:796-813.

Fan S, Bielenberg DG, Zhebentyayeva TN, Reighard GL, Okie WR, et al. 2010. Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). New Phytologist, 185: 917–930.

Foster T, Johnston R, and Seleznyova A. 2003. A Morphological and Quantitative Characterization of Early Floral Development in Apple (Malus × domestica Borkh.). Annals of Botany 92:199-206.

Gasic K, Hernandez A, and Korban SS. 2004. RNA Extraction From Different Apple Tissues Rich in Polyphenols and Polysaccharides for cDNA Library Construction. Plant Mol Biol Rep 22: 437a-437g.

Goff L, Trapnell C, and Kelley D. 2019. cummeRbund: Analysis, exploration, manipulation, and visualization of Cufflinks high-throughput sequencing data. R package version 2.24.0.

Gómez-Mena C, Piñeiro M, Franco-Zorrilla JM, Salinas J, Couland G, and Martínez-Zapater JM. 2001. Early bolting in short days: An Arabidopsis Mutation That Causes Early Flowering and Partially Suppresses the Floral Phenotype of LEAFY. Plant Cell 13: 1011-1024.

Gottschalk C and van Nocker S. 2013. Diversity in Seasonal Bloom Time and Floral Development among Apple Species and Hybrids. J Amer Soc Hort Sci 138(5):367-374.

Gottschalk C, Zhang S, Schwallier P, Rogers S, Bukovac MJ, and van Nocker S. Genetic mechanisms associated with the repressive effect of fruit on floral initiation in apple (*Malus* × *domestica* Borkh). Submitted

Gou D, Zhang Z, Wang X, Han X, Wei B, et al. 2015. The WRKY Transcription Factor WRKY71/EXB1 Controls Shoot Branching by Transcriptionally Regulating *RAX* Genes in Arabidopsis. Plant Cell 27: 3112-3127.

Grainger J. 1939. Studies upon the time of flowering of plants: anatomical, floristic and phenological aspects of the problem. Annals of Applied Biology 26, 684–704.

Guitton B, Kelner JJ, Celton JM, Sabau X, Renou JP, Chagné D, et al. Analysis of transcripts differentially expressed between fruited and deflowered 'Gala' adult trees: a contribution to biennial bearing understanding in apple. BMC Plant Biol. 2016;16: 55.

Hauagge R and Cummins JN. 1991. Seasonal Variation in Intensity of Bud Dormancy in Apple Cultivars and Related *Malus* Species. J Amer Soc Hort Sci 116(1): 107-115.

Haas B. TransDecoder (find coding regions within transcripts). http://github.com/TransDecoder. Accessed September 22, 2020.

Haberman A, Ackerman M, Crane O, Kelner JJ, Costes E, Samach A. Different flowering response to various fruit loads in apple cultivars correlates with degree of transcript reaccumulation of a TFL1-encoding gene. Plant J. 2016;87: 161-173.

Hättasch C, Flachowsky H, Kapturska D, Hanke M. Isolation of flowering genes and seasonal changes in their transcript levels related to flower induction and initiation in apple (Malus domestica). Tree Physiol. 2008;28: 1459-1466.

Huerta-Cepas J, Serra F, Bork P. ETE 3: Reconstruction, analysis and visualization of phylogenomic data. Mol Biol Evol. 2016;33(6): 1635-1638.

Hoover E, De Silva N, McArtney S, and Hirst P. 2004. Bud development and floral morphogenesis in four apple cultivars. J Hort Sci Biotech 79(6): 981-984.

Jones AL. 1992. Evaluation of the Computer Model MARYBLYT for Predicting Fire Blight Blossom Infaction on Apple in Michigan. Plant Disease 76(4): 344-347.

Kim D, Langmead B, and Salzberg SL. 2015. HISAT: a gast spliced aligner with low memory requirements. Nature Methods 12: 357-360.

Kotoda N and Wada M. MdTFL1, a TFL1-line gene of apple, retards the transition from the vegetative to reproductive phase in transgenic Arabidopsis. Plant Sci. 2005;168(1): 95- 628 104.

Lawson DM, Hemmat M, and Weeden NF. 1995. The use of molecular markers to analyze the inheritance of morphological and developmental traits in apple. J Amer Soc Hort Sci 120(3): 532-537.

Legave JM, Farrera I, Almeras T, and Calleja M. 2007. Selecting models of apple flowering times and understanding how global warming has had an impact on this trait. J Hort Sci and Biotech 83(1): 76-84.

Lex A, Gehlenborg N, Strobelt H, Vuillemot R, and Pfister H. 2014. UpSet: Visualization of Intersecting Sets. IEEE Trans Vis Comput Graph 20(12):1983-1992.

Liebhard R, Kellerhals M, Pfammatter W, Jertmini M, and Gessler C. 2003. Mapping quantitative physiological traits in apple (*Malus* × *domestica* Borkh.). Plant Molecular Bio 52:511-526.

Kiyosue T and Wada M. 2001. LKP1 (LOV kelch protein 1): a factor involved in the regulation of flowering time in Arabidopsis. Plant J 23(6): 807-815.

Kotoda N, Wada M, Komori S, Kiduo S, Abe K, et al. 2000. Expression Pattern of Homologues of Floral Meristem Identity Genes *LFY* and *AP1* during Flower Development in Apple. J Amer Soc Hort Sci 125(4): 398-403.

McArtney S, Hoover E, Hirst P, and Booking IR. 2001. Seasonal variation in the onset and duration of flower development in 'Royal Gala' apple buds. J Hort Sci Biotech 76(5): 536-540.

McCarthy DJ, Chen Y, and Smyth GK. 2012. Differential expression analysis of multifactor RNA-seq experiments with respect to biological variation. Nucleic Acids Research 40(10): 4288-4297.

Mimida N, Kotoda N, Ueda T, Igarashi M, Hatsuyama Y, Iwanami H, et al. Four TFL1/CEN-like genes on distinct linkage groups show different expression patterns to regulate vegetative and reproductive development in apple (Malus × domestica Borkh.). Plant Cell Physiol. 2009;50(2): 394-412.

Nikiforova SV, Cavalieri D, Velasco R, and Goremykin V. 2013. Phylogenetic analysis of 47 chloroplast genomes clarifies the contribution of wild species to the domesticated apple maternal line. Molecular Biology and Evolution 30(8): 1751-1760.

Pelaz S, Ditta GS, Baumann E, Wiseman E, and Yanofsky MF. 2000. B and C floral organ identity functions require SEPALLATA MADS-box genes. Nature 405(6783): 200-203.

Pertea M, Pertea GM, Antonescu CM, Chang TC, Medell JT, and Salzberg SL. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature Biotech 33: 290-295.

Pertea G and Pertea M. 2020. GFF Utilities: GFFRead and GFFCompare. F1000 Research 9: 304.

Piñeiro M, Gómez-Mena C, Schaffer R, Martínez-Zapater JM, and Coupland G. 2003. EARLY BOLTING IN SHORT DAYS Is Related to Chromatin Remodeling Factors and REgulates Flowering in Arabidopsis by Repressing *FT.* Plant Cell 15: 1552-1562.

Price MN, Dehal PS, Arkin AP. FastTree 2 - approximately maximum-likelihood trees for large alignments. PLoS One 2010;5(3): e9490.

R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www. R-project.org. 2013.

Ramírez F and Davenport TL. 2013. Apple pollination: A review. Scientia Horticulturae 162: 188-203.

Rodrigo J. 2000. Spring frosts in deciduous fruit trees - morphological damage and flower hardiness. Scientia Horticulturae 85(3): 155-173.

Sánchez-Pérez R, Dicenta F, and Martínez-Gómez P. (2012). Inheritance of chilling and heat requirements for flowering in almond and QTL analysis. Tree Genetics & Genomes, 8: 379–389.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7: 539.

Spotte RA, Stang EJ, and Ferree DC. 1976. Effect of overtree misting for bloom delay on incidence of fire blight in apple. Plant Dis. Rep. 60: 329–330.

Spurr AR. 1969. A low-viscosity epoxy resin embedding medium for electron microscopy. Journal of Ultrastructure Research, (26), 31–43.

Thompson A, May MR, Moore BR, and Kopp A. 2020. A hierarchical bayesian mixture model for inferring the expressed state of genes in transcriptomes. PNAS 117(32): 19339-19346.

Trapnell C, Williams B, Pertea G, Mortazavi A, Kwan G, et al. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols 7: 562-578.

Unterberger C, Brunner L, Nabernegg S, Steininger KW, Steiner AK, et al. 2018. Spring frost risk for regional apple production under a warmer climate. PLoS One 13(7):e0200201.

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, et al. The genome of the domesticated apple (Malus × domestica Borkh.). Nat Genet. 2010;42(10): 833–839.

CHAPTER FOUR

Molecular identification and transcriptional patterning of *FLOWERING LOCUS T* genes in Apple (*Malus* × *domestica*)

ABSTRACT

The identification and characterization of *Flowering locus T* (*FT*) homologs in plants have played a critical role in understanding the evolution of unique flowering traits. Apple (*Malus* × *domestica* Borkh.) exhibits many flowering traits that are critical to production, including variation in bloom time and biennial bearing tendencies. Apple contains two presumed homologous copies of *FT, MdFT1* and *MdFT2,* of which *MdFT1* has been thoroughly identified and characterized. *MdFT2* has remained relatively uncharacterized due to an elusive identity. In this study, I identify a transcriptional product that exhibits conserved *FT* sequence and structure and a distinctive sequence from that of *MdFT1*. *MdFT1* expression was compared across a wide range of tissues and developmental stages using a gene expression atlas that resulted in the documentation that *MdFT1* is predominately expressed in apex and rachis tissues. *MdFT2* quantification was unsuccessful due to its absence from the gene expression atlas. Using paralog-specific probes, I conducted expression profiling of *MdFT1* and the *MdFT2* transcripts to characterize the daily and seasonal expression patterns. Both *MdFTs* were identified to be expressed in bourse shoot leaves but did not respond to floral inductive conditions. *MdFT1* daily expression was diurnal, a conserved pattern with Arabidopsis, and was cultivar dependent. Seasonal expression patterns and response to crop load conditions were also found to be cultivar dependent. However, expression patterns of both *MdFTs* were not differentially expressed in response to crop load conditions that are floral promotive. These results indicate that *MdFT* could have limited function in the bourse shoot leaves.

INTRODUCTION

The mechanism of floral initiation remained elusive until the identification and functional characterization of *FLOWERING LOCUS T* (*FT*) (Andrés and Coupland 2012). In Arabidopsis, *FT* is expressed in companion cells of the phloem within leaves and its translated product is loaded into the phloem stream for translocation to the apex (Corbesier et al. 2007, Andrés and Coupland 2012). In the apex, FT forms a transcription regulator complex with FLOWERING LOCUS D, which promotes the expression of meristem identity genes such as *APETALA 1* (Abe et al. 2005; Wigge et al. 2005). Thus, FT functions as the primary integrator of the flowering signal from the leaves to the apex. To ensure flowering occurs during optimal conditions, a diverse network of environmental and genetic mechanisms regulate the expression of *FT*. The primary mechanisms for Arabidopsis regulation of *FT* is by photoperiod and temperature. Under long-day conditions, *CONSTANS* (*CO*) and its translated peptide gradually accumulate throughout the day, reaching a peak 10-14 hours after daybreak (Sawa et al. 2007; Song et al. 2012; Golembeski et al. 2014). The accumulation of CO during these long-days leads to the promotion of *FT*, which triggers the transition to flowering (Song et al. 2012; Andrés and Coupland 2012; Golembeski et al. 2014). This CO-driven mechanism results in *FT* exhibiting a diurnal expression pattern when conditions are suitable for flowering. Under unfavorable conditions, such as short-days, CO is unable to accumulate due to its ubiquitination and resulting degradation, and transcription repression of the *CO* locus (Andrés and Coupland 2012; Golembeski et al. 2014). Temperature also regulates *FT* by the vernalization response that is conveyed by the *FLOWERING LOCUS C* (*FLC*) gene (Searle et al. 2006). Adequate cool temperatures will down regulate the expression of *FLC*, relieving its transcriptional repression of

FT. However, *FLC* mediated response is associated with a requirement for vernalization and many plants can initiation flowers without this requirement including apple.

The documentation of the conservation of the *FT* genes and their functions is critical to our understanding the diverse flowering process that are exhibited by other annual and perennial plants. In apple (*Malus* × *domestica* Borkh.), two *FT* genes have been identified *MdFT1* and *MdFT2* (Kotoda et al. 2010). *MdFT1* was found to be positioned on Chromosome (Chr.) 12 in both of the primary reference genomes (Velasco et al. 2010; Foster et al. 2014; Daccord et al. 2017). The second copy of *FT*, *MdFT2,* was mapped onto Chr. 4 (Kotoda et al. 2010). However, the two primary reference genomes have incompletely described the genomic coordinates and sequence of *MdFT2*. This difficulty to resolve the genomic position of *MdFT2* could be attributed to the striking similarity of the paralogs. Previous studies have reported that the *FT* paralogs have a >94% conserved coding sequencing identity (Kotoda et al. 2010). As a result of this high similarity, the misidentification of alleles from one *MdFT* gene as separate loci have occurred (Tränkner et al. 2011). This complexity to properly identify *MdFT2* has also resulted in *MdFT1* getting more attention related to the characterization of its function.

MdFT1 is predominantly expressed in the apical meristems of vegetative and reproductive shoots (Kotoda et al. 2010). The peak expression periods of *MdFT1* in the apex is during the anticipated periods of floral induction and initiation (Hättasch et al. 2008; Kotoda et al. 2010; Mimida et al 2011; Habermann et al. 2018; Guitton et al. 2016; Gottschalk et al. Chapter 2). The presence of fruit on the shoot, a floral repressive mechanism in some apple cultivars, has been reported to not influence *MdFT1* expression in the corresponding shoot apex (Guitton et al. 2016, Habermann et al. 2018, Gottschalk et al. Chapter 2). In contrast, the presumed *MdFT2* was found to be expressed in the apex, but its predominant expression is in

tissues of the floral organs (Kotoda et al. 2010). Seasonal expression patterns of *MdFT2* are closely associated with the period of floral development late in the growing seasons and following dormancy release prior to spring bloom (Kotoda et al. 2010).

Little is known if *MdFTs* exhibit the conserved daily expression pattern during floral induction that are characteristic of other long-day plants such as Arabidopsis. Additionally, many of these previous reports have relied on relative expression measures under narrow sets of conditions and tissues. In this study, I investigated the identity of the *MdFTs* using transcriptional models constructed in prior chapters of this dissertation. Using a gene expression atlas, I quantify the expression of *MdFT1* across a diverse set of tissues and developmental stages. Lastly, I investigated the daily and season expression patterns of *MdFT*, while addressing cultivar-specific expression in the bourse shoot leaves.

METHODS

Plant material and tissue sampling

Tissue collection was performed on trees maintained at the Michigan State University Clarksville Research Center (CRC) (Field: 42°52'24"N, 85°16'05"W – Station: 42°52'24"N, 85°15'30"W) in Clarksville, Michigan. All trees were 10 years old or older and grown in a vertical axis high-density system using standard commercial practices. The specific strains of the cultivars and rootstocks used are as follows: Gala “Brookfield” on Budagovsky 9 (Bud9), Honeycrisp on Bud9, and Red Delicious “Schlet Spur” on M26. Ten trees with similar, high bloom density were selected, and all blossoms were removed by hand at the full bloom stage (early May 2016). For 'Honeycrisp' and 'Gala', three additional trees with high bloom density were selected as non-thinned controls. Individual bourse shoots were tagged at full bloom, and were collected at dusk on the following time intervals: 2-3, 12-13, 26-27, 39-40, 55-56, 68-69,

and 83-84 days after full bloom (DAFB). For each biological replicate tree, ten tagged bourse shoots were collected per time point. The first mature leaf was excised from each collected bourse shoot and frozen in liquid N_2. To analyze daily gene expression, six 'Honeycrisp' and 'Red Delicious' trees with similar, high bloom density were selected at full bloom. Three of the six trees per cultivar were then thinned of blossoms by hand, while the other three trees remained nonthinned. Individual bourse shoots were collected at three-hour intervals over the course of one 24 hour period on June 13th, 2016 (33 DAFB). This date was selected as it represents a midpoint of the anticipated period of floral initiation. The first mature leaf was excised from each collected bourse shoot and frozen in liquid N_2.

MdFT sequence retrieval and analysis

All coding sequences were retrieved from the Genomic Database for Rosaceae (GDR) using the JBrowse (Buels et al. 2016) software. The *de novo* assembled *MdFT2*-like sequence was retrieved from previously published works (Gottschalk et al. Chapter 2). In short, the *de novo* transcripts were assembled from reads that failed to align to the GDDH13 v.1.1 reference genome sequence (Daccord et al. 2017). Alignments were performed using the HISAT2 aligner (v.2.1.0) invoking the --dta-cufflinks and --un-conc-gz options (Kim et al. 2015). Unaligned reads were then assembled into contiguous sequences using the Trinity *de novo* assembler using default settings (Haas et al. 2013). The *de novo* contigs were then constructed into gene models using the python program Trinity_gene_splice_modeler.py from the Trinity suite. Genomic sequences from Honeycrisp were obtained from an in-house dataset of Illumina sequence reads. Blastn was then used with an Expect (E)-value cutoff of $1e^{-12}$ to identify sequences that aligned to the *de novo* assembled *FT* transcript TRINITY_DN25413_c6_g1 (Camacho et al. 2008).

These sequences were then assembled into a *de novo* genomic sequence using the SPAdes assembler (v.3.11.1) with default assembly options (Bankevich et al. 2012).

Expression analysis

Total RNA was extracted from bourse shoot leaves using the Gasic et al. (2004) extraction protocol with the modification of using spermine instead of spermidine in the extraction buffer. Each extraction utilized 1 mm transverse section of five or more leaves per replicate extraction. For the seasonal expression experiment, two replicates per time point and treatment per cultivar were conducted. For the daily expression experiment, three replicates per time point per treatment per cultivar were used. Extracted RNA was then purified using a commercial clean-up kit (RNeasy Mini; QIAGEN, Germantown, MD). RNA quality was assessed using a 1.2% formaldehyde-agarose gels and quantified using a Nanodrop 2000c (Thermo Fisher Scientific, Waltham, MA). Gene expression was measured using a two-step quantitative real-time PCR. Previously designed primers and probes were used (Zhang et al. 2019). Each primer set was designed to be specific to selected *MdFT* sequences and overlapped with an exon junction. Specifically, *MdFT2* primers were designed using the *de novo* assembled transcript TRINITY_DN25413_c6_g1. A homolog of *MdACTIN* served as an internal control when calculating relative expression. We used an Agilent Technologies Stratagene Mx3005P (Santa Clara, CA) qPCR machine with cDNA derived from the extracted RNA. Each reaction was conducted using a mix of TaqMan® (Thermo Fisher Scientific, Waltham, MA) Gene Expression Master Mix (10 µl), 5x diluted cDNA template (2 µl), forward (1 µl) and reverse primers (1 µl), and probe (1 µl) for *ACTIN*, the primer-probe assay for the gene of interest (1 µl), and ddH2O (4 µl).

Statistical analysis

Statistics were conducted using R (v.3.6.2) (R Core Team 2017) inside the R Studio suite (v.1.1.423) (Rstudio Team 2015). R data analysis packages used in this study were ggplot2, ggpubr, and tidyverse (Wickham 2016, Wickham et al. 2019).

RESULTS AND DISCUSSION

Identification of FT-like sequences

With the release of the first apple genome and a subsequent high-quality genome derived from a double haploid individual, identification of *FT*-like sequences through sequence homology identification became possible (Velasco et al. 2010, Daccord et al. 2017). The first genome release (GDR v3.1.1; Velasco et al. 2010), had two annotated *FT*-like genes, MDP0000128821/MDP0000139278 and MDP0000132050 (Velasco et al. 2010), positioned on Chr. 4 and 12, respectively. The locus on Chr. 12 was previously annotated as *MdFT1* and exhibited an expected gene structure with four exons (Kotoda et al. 2010). The locus on Chr. 4 was previously annotated as *MdFT2* and its structure is less clear. This locus contained five predicted peptides with *FT*-like sequence but was annotated as separate genes. A later revised reference transcriptome annotated *MdFT* as a single transcript (MDP0000139278) that exhibited the conserved four exon structure as *MdFT1* (Fig 4.1) (Kotoda et al. 2010; Jung et al. 2019). In the subsequent high-quality genome (GDDH13 v.1.1), only *MdFT1* (MD12G1262000) was present in the genome assembly and featured overlapping alignments with both previously assembled *FT* genes (Fig 4.1) (Daccord et al. 2017).

To resolve the missing *MdFT2*, sequence reads from a previous apple transcriptome study that failed to map to the GDDH13 reference genome were *de novo* assembled (Gottschalk et al. Chapter 2). Within those assembled contigs, one identified gene contained *FT*-like

nucleotide sequence and translated into an FT-like peptide (TRINITY_DN25413_c6_g1). However, the gene appeared chimeric as it contained a second transcript positioned in the 5' direction to the *FT*-like sequence. To gain further evidence that the chimeric assembly is actual and not an artifact of the assembly process, we performed an independent *de novo* assembly of DNA sequence reads. We utilized a library of high-quality DNA sequences generated from the cultivar Honeycrisp and filtered the read libraries for sequences that contained *FT*-like homology using Blastn. These filtered reads were then assembled into eight contigs and aligned with the *de novo* assembled *FT*-like transcript. The alignments indicated that one of the assembled contigs shared >90% sequence identity to the *FT*-like region of the *de novo* transcript previously assembled (Fig 4.2). These assembled sequences also exhibited 100% translated identity to an NCBI cataloged *MdFT2* protein in the -1 frame. These results suggest that the *de novo* assembled transcript has a corresponding genomic sequence. However, these results do not establish that *MdFT2* is a separate locus as opposed to a divergent allele of *MdFT1*. From this point forward, the *de novo FT*-like sequence will be referred to as *MdFT2.*

Tissue-specific expression of MdFT1

In Arabidopsis, *FT* is expressed under promotive photoperiodic conditions in the phloem companion cells within the leaf (Corbesier et al. 2007, Andrés and Coupland 2012). The resulting peptide, FT, is then translocated from the leaf to the apex via the phloem stream, upon reaching the apex it triggers floral initiation (Corbesier et al. 2007, Andrés and Coupland 2012). To document if *FT* expression patterns are conserved between Arabidopsis and apple, we evaluated tissue-specific expression patterns using a gene expression browser constructed from an apple relative, *Malus fusca* (Gottschalk et al. *in prep*). *MdFT1* was found to be expressed in rachis and the shoot apical meristem (Fig 4.3). *MdFT1* expression was also found in the

cotyledons and leaves of seedlings. Little or no expression of *MdFT1* was observed in the leaves. However, this dataset did not assess gene expression in leaves from specific positions on the shoot or spur which could explain the absent expression in leaves. Unfortunately, *MdFT2* was absent in that dataset resulting our in ability to evaluate for tissue-specific expression.

Diurnal expression of MdFTs in bourse shoot leaves

Bourse shoot leaves are critical to facilitate floral initiation in apple, suggesting that a possible floral signal originates in those tissues (Fulford 1960; Elsysy and Hirst 2017, 2019). Moreover, FT in Arabidopsis is produced and exported from leaf tissues. To examine if *FT* regulation is conserved in apple, I evaluated the expression of *MdFTs* in the bourse shoot leaves. To induce flowering, blossoms were thinned from representative trees of 'Honeycrisp', 'Red Delicious' and 'Gala' to promote flowering. The thinning of blossoms was effective ($p < 0.05$) at promoting a return bloom in all three cultivars (Fig 4.4). This result suggests that spurs thinned of flowers initiated blossoms for the subsequent season, whereas the nonthinned remained predominately vegetative in identity. RNA was extracted from the first mature bourse shoot leaf over 24 hours at three-hour intervals from thinned and nonthinned spurs of 'Honeycrisp' and 'Red Delicious'. Using *MdFT* paralog-specific probes, relative expression was quantified in these samples. For *MdFT1*, peak expression was observed during the night time hours in 'Honeycrisp' indicative of a diurnal pattern (Fig 4.5). In comparing the effect of crop load, *MdFT1* in 'Honeycrisp' exhibited a stronger amplitude during its oscillation of expression in the thinned compared to the nonthinned (Fig 4.5 - left panels). In contrast, expression of *MdFT1* in 'Red Delicious' peaked around 7:00 PM in the thinned spurs (Fig 4.5 - bottom right panel). A oscillating pattern of expression in thinned 'Red Delicious' is evident but the amplitude was lower than in 'Honeycrisp' and it does not exhibit a diurnal pattern. However, no significant

differences were found between the expression *MdFTs* in the thinned compared to nonthinned leaves, for either cultivar at any time point.

For *MdFT2*, 'Honeycrisp' exhibited two peaks of expression in the nonthinned spurs (Fig 4.6 - top left panel). The peaks occurred between 10:00 AM and 1:00 PM and at 10:00 PM. In the thinned spurs, *MdFT2* had a single peak at 4:00 PM (Fig 4.6 - bottom left panel). In 'Red Delicious', *MdFT2* had a single peak at 4:00 PM under both crop load conditions similar to what was observed in the thinned 'Honeycrisp' leaves (Fig 4.6 - right two panels). However, the amplitude of the peak in the thinned spurs was more pronounced than nonthinned. No significant differences were found between the expression values of the thinned compared to nonthinned, for either cultivar at any time point.

I also investigated whether the expression pattern over time differed between cultivars and treatment for each of the *MdFTs*. The expression profiles were clustered using a Euclidian distance calculation to find similarities between cultivars and treatment (Fig 4.7). Differences in *MdFT1* expression profiles appear to be primarily associated with the cultivars, as they clustered together. For *MdFT2*, the clustering found that nonthinned Honeycrisp formed a clearly defined clade separate from the other expression profiles. The clade that contained the three other profiles of *MdFT2* were separated by cultivar. This result suggests that *MdFT2* expression profiles are distinct in a crop load and cultivar-specific fashion.

Seasonal expression patterns of MdFTs

In other perennial tree fruit, such as Citrus, differential expression of *FT*-like genes in leaves was associated with biennial bearing (Munoz-Fambuena et al. 2011; Ziv et al. 2014). I investigated if this finding was conserved in apple by measuring expression of the two *MdFT* paralogs in the bourse shoot leaves of two apple cultivars of opposing biennial bearing habits

throughout the growing season (Fig 4.8). ‘Honeycrisp’ is a biennial bearer whereas ‘Gala’ is an annual bearer (Embree et al. 2007, Elsysy and Hirst 2017). *MdFT1* exhibited higher expression in the nonthinned trees early in the growing season for ‘Honeycrisp’, with the expression decreasing after a peak around 39/40 DAFB. Although the seasonal expression pattens appear to differ, the differences between the thinned and nonthinned were found to not be significant. This result could be attributed to the low relative expression values and limited replications. In ‘Gala’, *MdFT1* expression patterning between the thinned and nonthinned was strikingly similar with relatively little change of the season. *MdFT2* in ‘Honeycrisp’ exhibited a responsive ($p < 0.01$) to thinning, with lower expression of *MdFT2* in the thinned trees compared to the than nonthinned (Fig 4.9). However, this difference in expression was associated with the nonthinned treatment which repressed floral initiation. *MdFT2* in ‘Gala’ exhibited no difference in expression between crop load treatments. When clustering the expression profiles of *MdFT1*, the clades were separated by crop load treatments (Fig 4.10). This result is indicative that *MdFT1* is more closely associated with the treatment than cultivar. The inverse situation was true for *MdFT2* where cultivar was the separating factor.

CONCLUSIONS

This study provides a greater understanding of the expression patterning of apple’s *FT* homologs. Our identification of *FT*-like sequences in apple provides confirming results for *MdFT1* but *MdFT2* is still not confirmed due to cryptic genomic position. The *de novo* assembled transcript that is *FT*-like exhibits different sequence and expression patterning than *MdFT1* suggesting its either *MdFT2* or *MdFT1* has unique allele-specific expression. Tissue- and developmental-specific expression was documented for *MdFT1* and was found to be primarily localized to the apex and rachis. The localization of *MdFT1* expression to the apex, suggests that

the *FT* signaling pathway could also be divergent in apple. *MdFT1* expression in bourse leaves was found to exhibit a diurnal pattern only for ‘Honeycrisp’ regardless of crop load conditions. However, this diurnal expression was not consistent for ‘Red Delicious’ where its peak expression occurred during the late evening. ‘Red Delicious’ *MdFT1* also exhibited no significant differences in expression in response to crop load. *MdFT2*, on the other hand, exhibited no conserved diurnal pattern for either cultivar. Seasonal expression patterns of *MdFT1* did not exhibit a significant response to the floral inductive thinning. In contrast, *MdFT2* exhibited an apparent increase in expression in the nonthinned trees of ‘Honeycrisp’. Although, this increase in expression was associated with the nonthinned treatment which repressed flowering. Thus, crop load does not appear to affect the daily or seasonal expression of *MdFT* in the bourse shoot leaves. In conclusion, apple’s *FT* homologs exhibit distinct expression patterns that are associated with genotypes, but are not influenced by crop load. *MdFTs* most likely do not function in a similar manner as Arabidopsis *FT* within leaf structures. This conclusion is also supported by previous works that identified a crop load specific response of *MdFT1* in the apex (Gottschalk et al. Chapter 2).

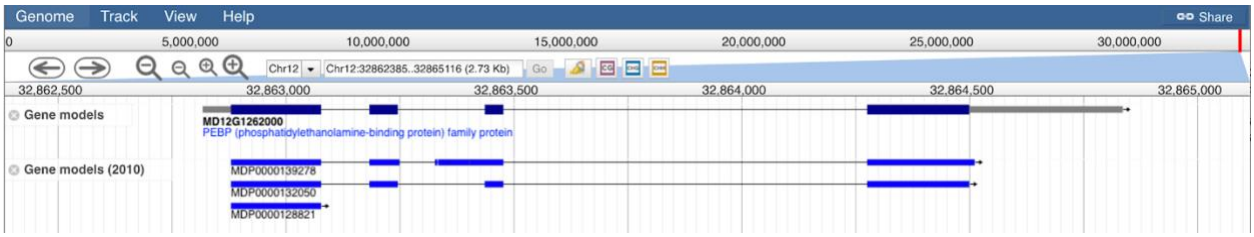


Figure 4.1. Genomic position of *MdFT1* in the GDDH13 genome and corresponding alignments with the two *FT* genes assembled in the GDR genome.

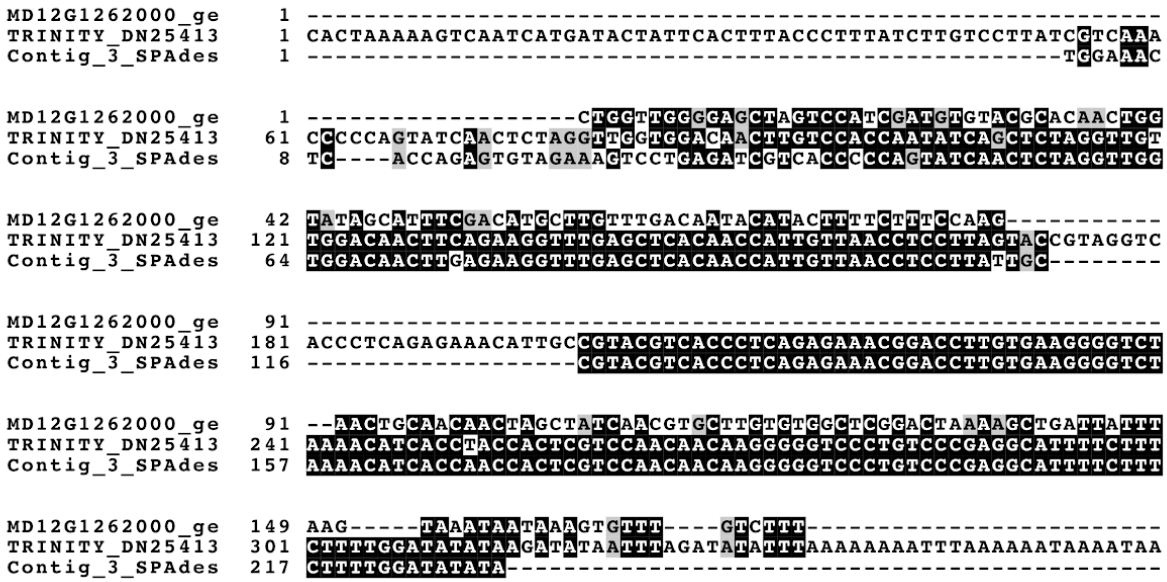
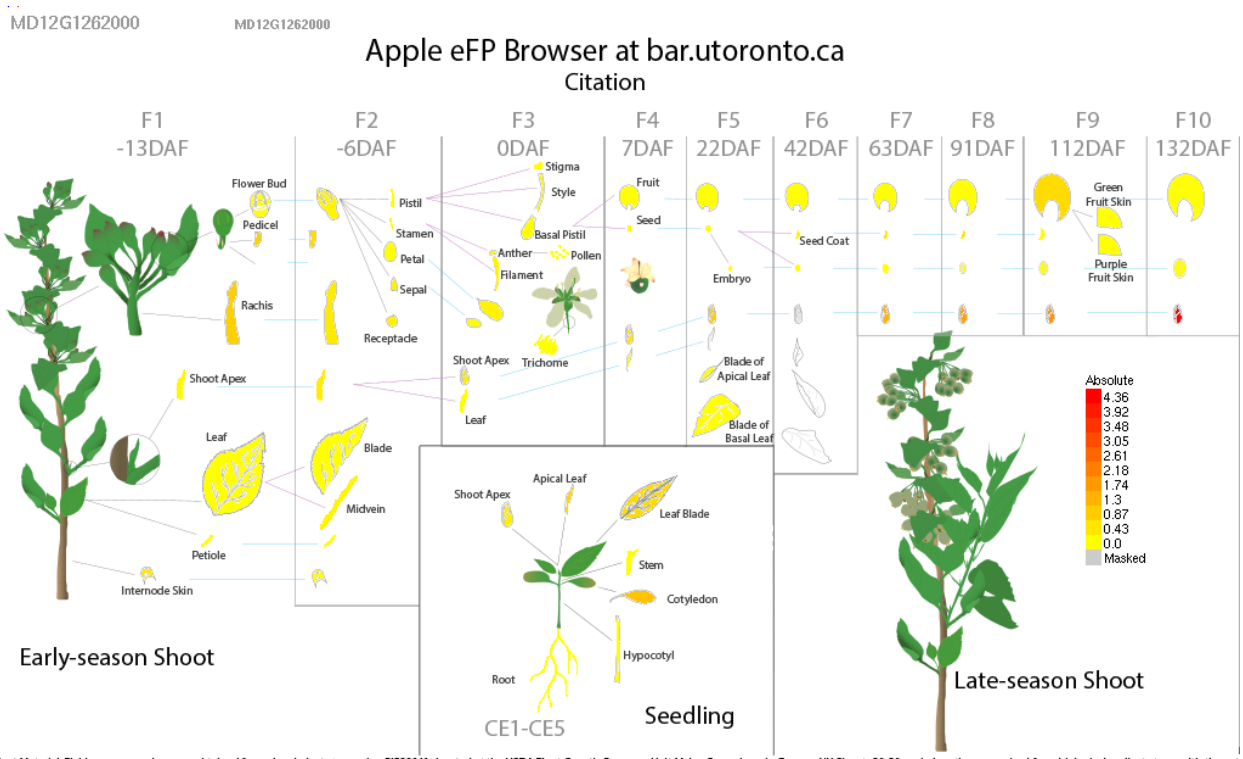


Figure 4.2. Alignment between the genomic sequence for *MdFT1* (MD12G1262000), segment of the de novo assembled *FT*-like transcript TRINITY_DN25413_c6.g1 (labelled as TRINITY_DN25413), and a de novo assembled *FT* DNA sequence contig.



Plant Material: Field-grown samples were obtained from clonal plants (accession PI589941) located at the USDA Plant Genetic Resource Unit Malus Germplasm in Geneva, NY. Shoots 20-30 cm in length were excised from biological replicate trees with the cut ends placed under water in a container. The ends were recut while submersed to reestablish vascular function. Shoots were then stored in the laboratory under ambient conditions for ~24 hours before tissue dissection was performed. For controlled-environment-grown plants, seeds were obtained from all biological replicate field-grown trees. Seeds were pooled and stratified for 12 weeks at 5 °C and then planted into a potting soil mix. Seedlings were grown at 24 °C with a 16 h photoperiod under fluorescent lighting. Independent groups of seedlings were then utilized as biological replicates for tissue dissection. Tissues sampled were cut into ~2 mm sections using razor blades and split into aliquots for RNA extractor or ChIP analysis.

For the RNA extraction, dissected tissues were cleaned and dried quickly prior to freezing using liquid nitrogen. RNA replicates 1 and 2 were from field-grown plants, and total RNA were extracted using a commercial kit (mirVana™ Ambion®) following the manufacturer's instructions. Replicate 3 was also from field-grown plants but total RNA was extracted using the Gasic et al. (2004) CTAB-based protocol with the spermine substituted for spermidine and cleaned using the RNeasy Mini Kit (QIAGEN®). Controlled-environment-grown plants utilized either of the previously described methods.

RNA-sequencing was performed on a Illumina TruSeq platform using a 101-b pair-end read format. Raw reads processed using the fastq-mcf software to remove adapters and primers, trim reads with Phred quality values <20, and remove reads shorter than 30 bases. Transcriptome assembly employed a genome-guided approach using the high-quality GDDH13 v1.1 reference apple genome (Daccord et al. 2017). Processed reads were mapped onto the reference genome using the HISAT2 aligner (Kim et al. 2015) under default parameters and were optimized for downstream processing using the --dta-cufflinks option. Mapped reads were then constructed into transcripts using StringTie (Pertea et al. 2015) while invoking the -G option to use reference annotation as a guide. Transcript and gene expression was then quantified and normalized using the Cufflinks software package (Trapnell et al. 2012).

Figure 4.3. Tissue and developmental stage specific expression of *MdFT1* in *Malus fusca*.

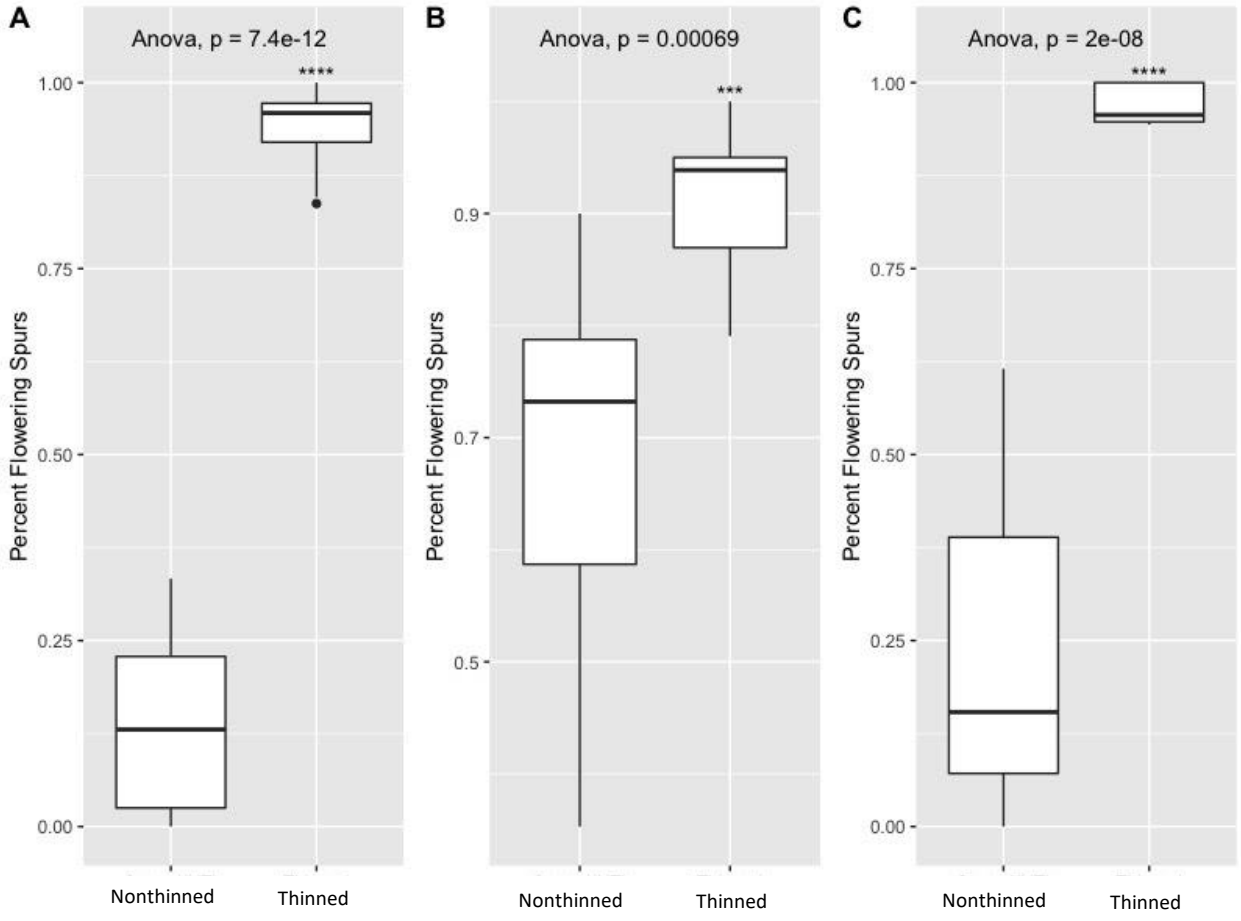


Figure 4.4. Percent return bloom in cultivars that received blossom thinning treatments. A) Honeycrisp, B) Red Delicious, C) Gala.

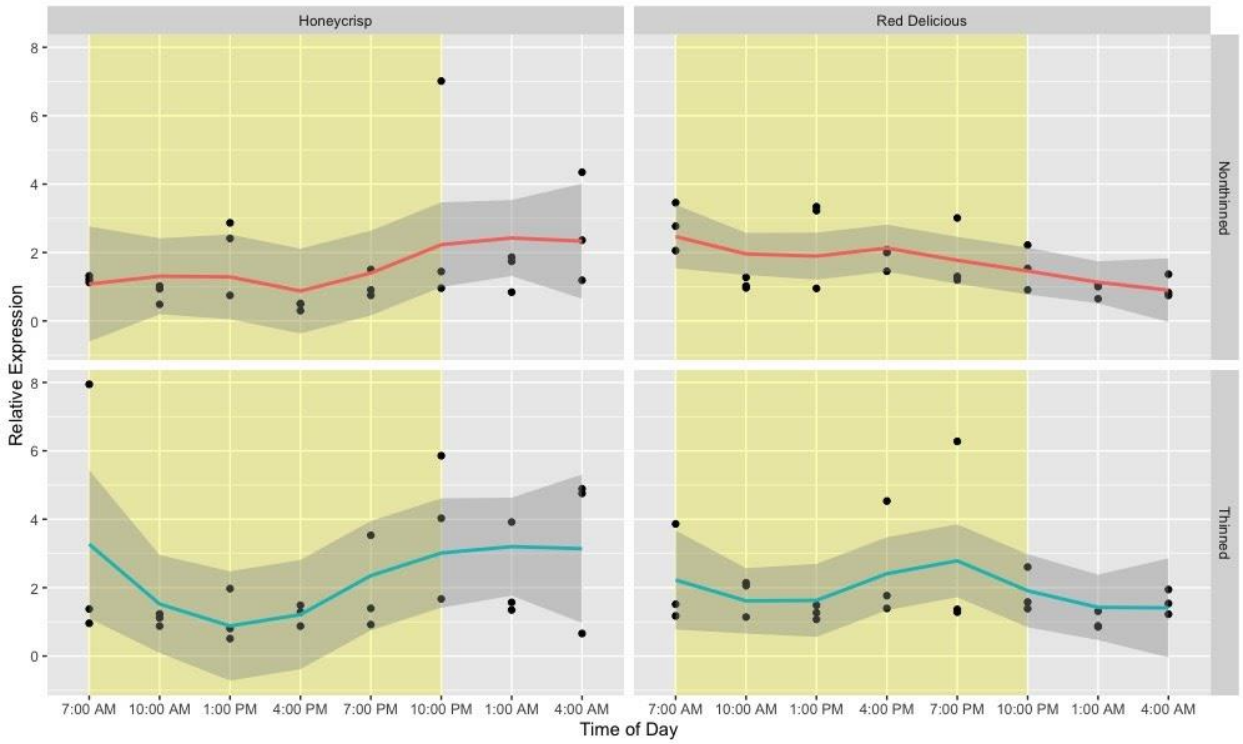


Figure 4.5. *MdFT1* relative expression in the bourse leaves of nonthinned and thinned shoots of two apple cultivars. Yellow shade is the estimated day light sampling hours. Generalized trend line and confidence interval are shown (grey shading).

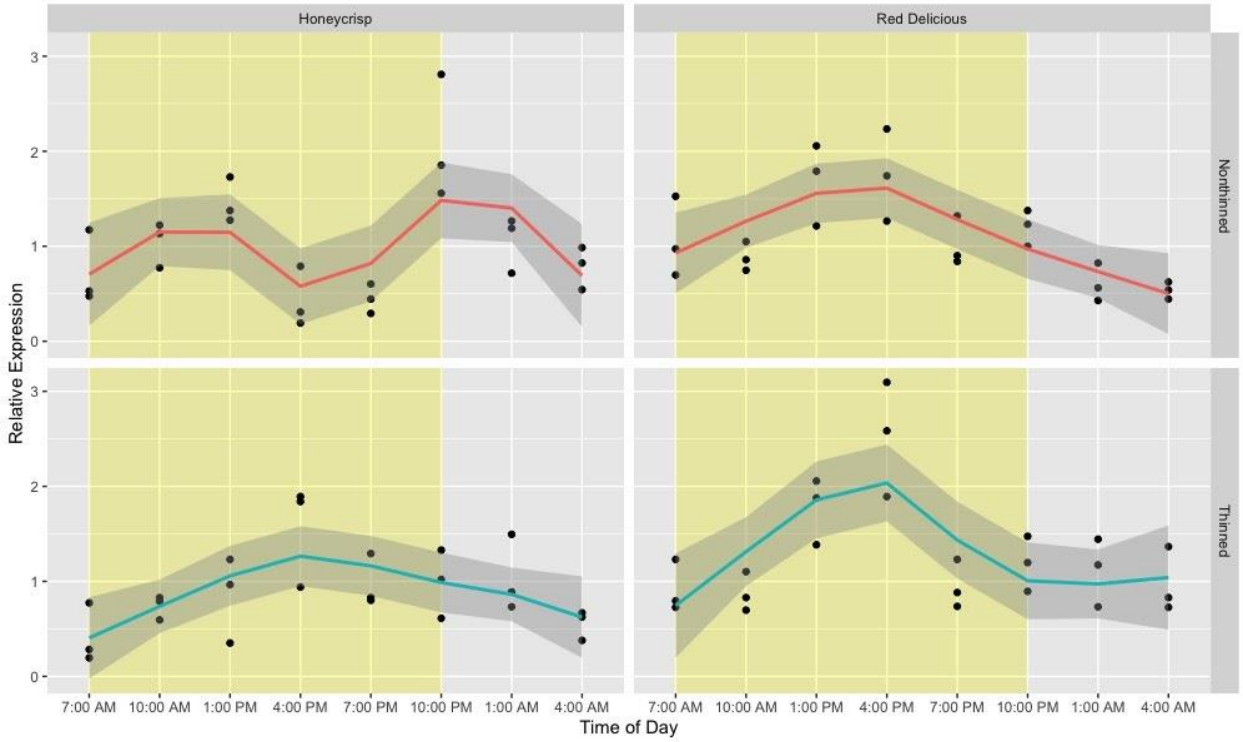


Figure 4.6. *MdFT2* relative expression in the bourse leaves of nonthinned and thinned shoots of two apple cultivars. Yellow shade is the estimated day light sampling hours. Generalized trend line and confidence interval are shown (grey shading).

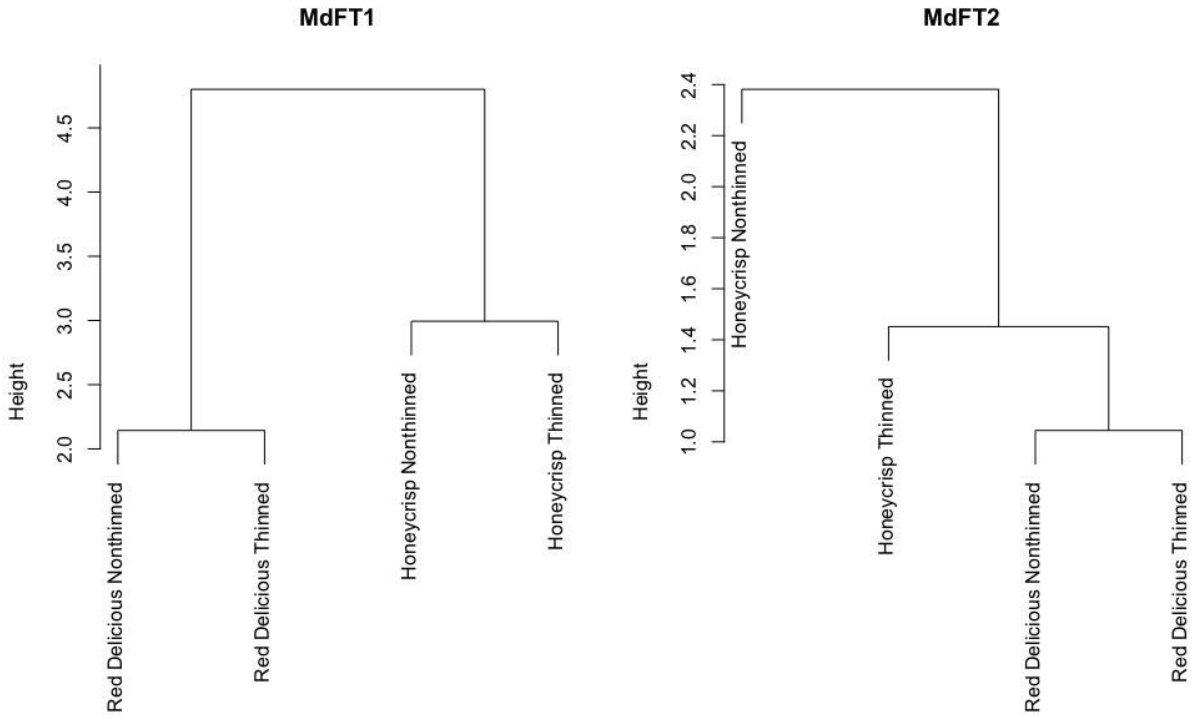


Figure 4.7. Clustering of *MdFT* homologs expression profiles over 24 hours from trees thinned or nonthinned of flowers.

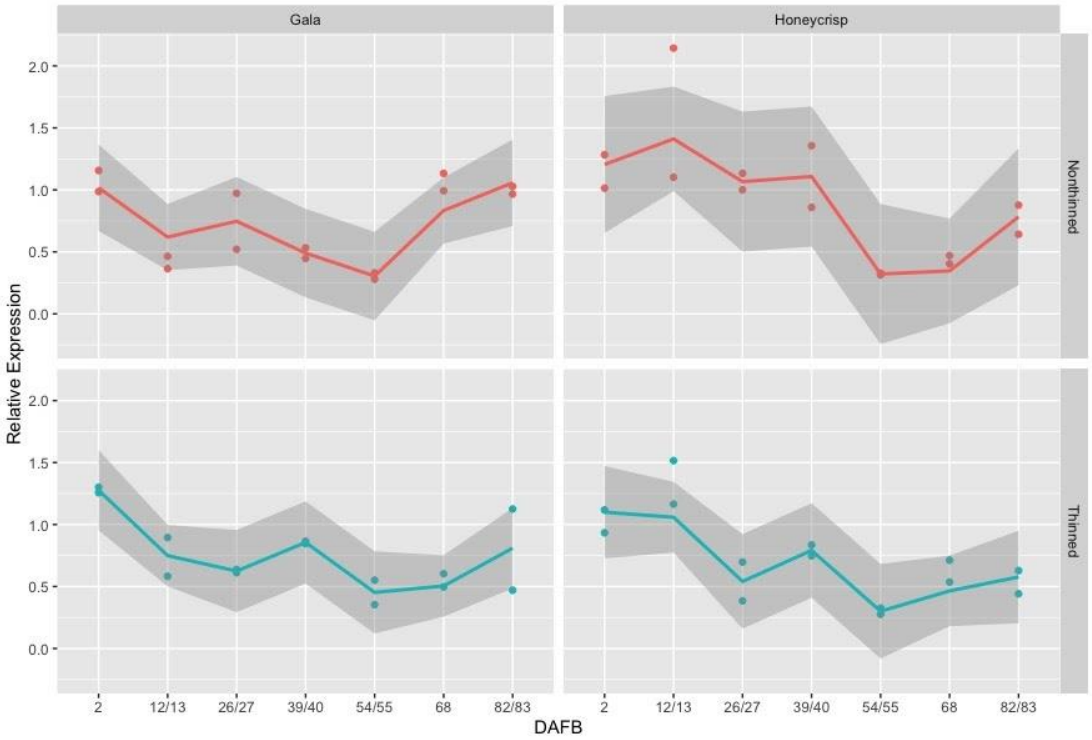


Figure 4.8. *MdFT1* relative seasonal expression in the bourse leaves of nonthinned and thinned shoots of two apple cultivars. Generalized trend line and confidence interval are shown (grey shading).

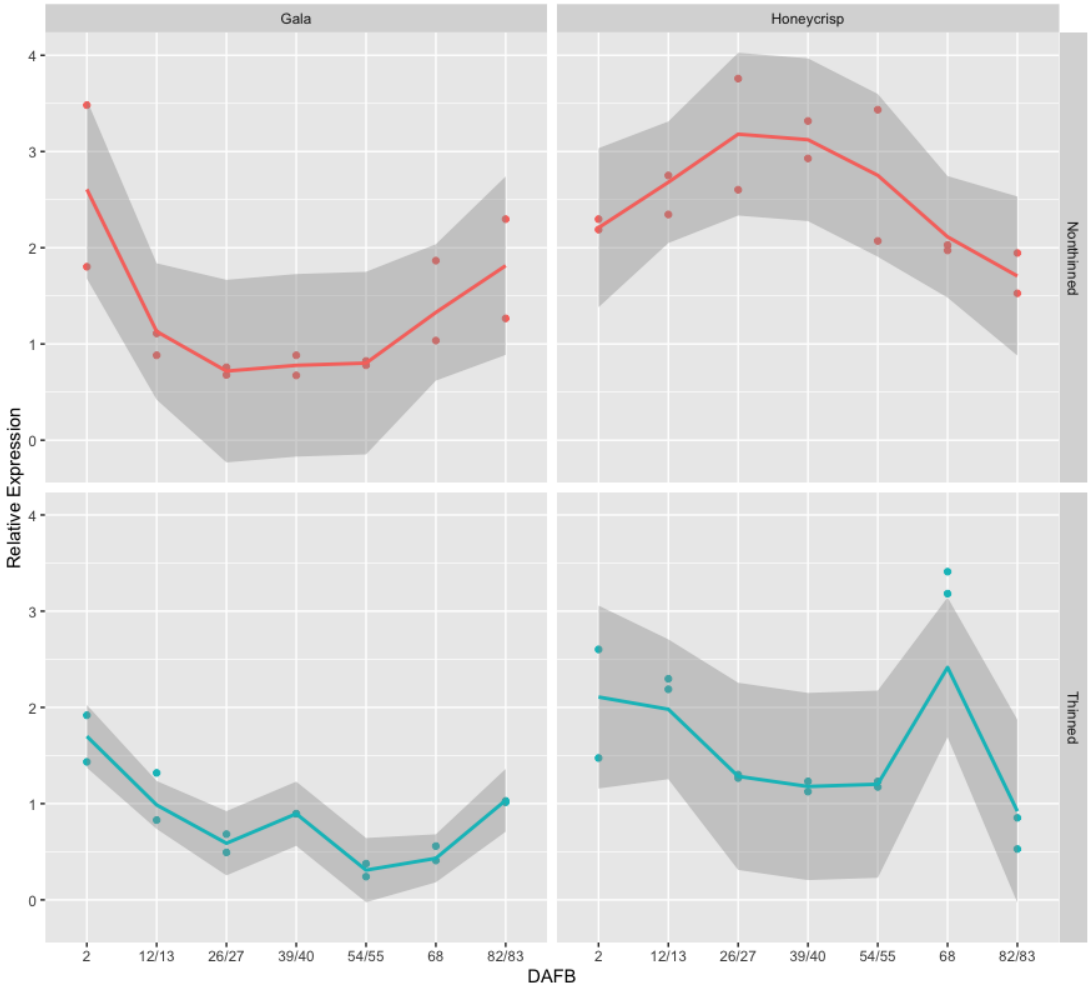


Figure 4.9. *MdFT2* relative seasonal expression in the bourse leaves of nonthinned and thinned shoots of two apple cultivars. Generalized trend line and confidence interval are shown (grey shading).

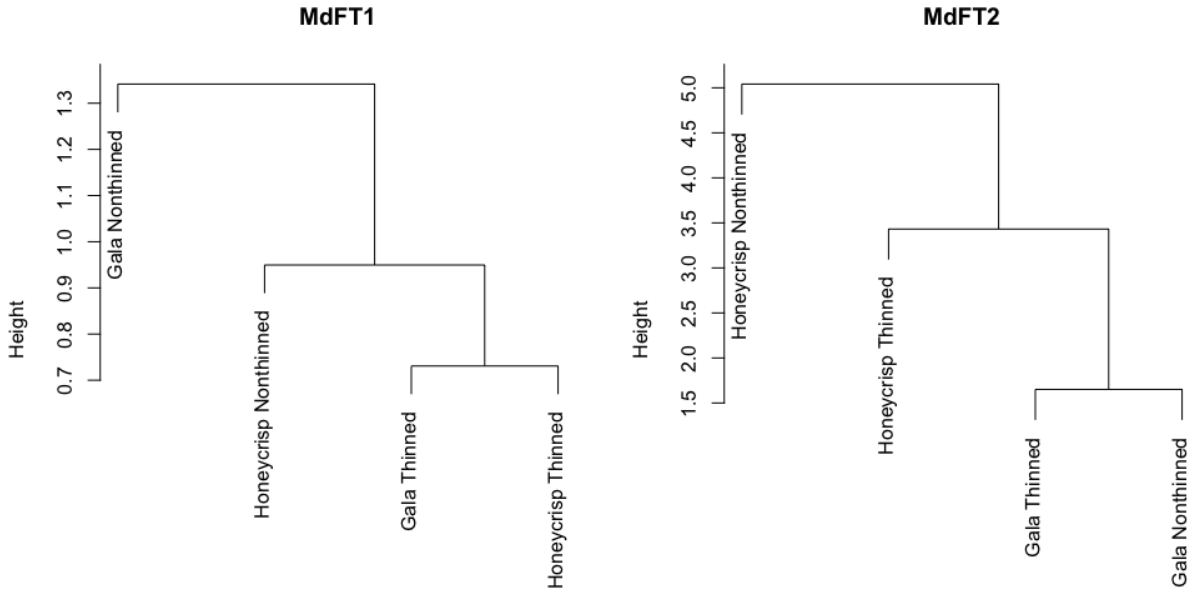


Figure 4.10. Clustering of *MdFT* homologs expression profiles over the season from trees thinned or nonthinned of flowers.

REFERENCES

REFERENCES

Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguichi A, et al. FD, a bZIP Protein Mediating Signals from the Floral Pathway Integrator FT at the Shoot Apex. Science 2005 309(5737): 1052-1056.

Andrés F and Coupland G. The genetic basis of flowering responses to seasonal cues. Nature Reviews Genetics 2012 13:627-639.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 2012 19(5): 455-477.

Buels R, Yao E, Diesh CM, Hayes RD, Munoz-Torres M, et al. JBrowse: a dynamic web platform for genome visualization and analysis. BMC Genome Biol 2016 16:66.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics 2008; 10:421.

Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, et al. FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of Arabidopsis. Science 2007 316(5827): 1030-1033.

Daccord N, Celton JM, Linsmith G, Becker C, Choisne N, et al. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. Nature Genetics 2017 49:1099-1106.

Elsysy M and Hirst P. Molecular Basis of Flower Formation in Apple Caused by Defoliation and Gibberellins. J Amer Soc Hort Sci 2019 144(6): 414-419.

Elsysy M and Hirst P. The role of spur leaves, bourse leaves, and fruit on local flower formation in apple: an approach to understanding biennial bearing. HortScience 2017 52(9): 1229-1232.

Embree CG, Myra MTD, Nichols DS, and Wright AH. Effect on blossom density and crop load on growth, fruit quality, and return bloom in ‘Honeycrisp’ apple. Hortscience 2007 42(7):1622-1625.

Fulford RM. The use of defoliating sprays for the control of biennial bearing in apples. HortScience 1960 35: 202-213.

Gasic K, Hernandez A, Korban SS. RNA extraction from different apple tissues rich in polyphenols and polysaccharides for cDNA library construction. Plant Mol Bio Rep. 2004 22: 437a-437g.

Golembeski GS, Kinmonth-Schultz HA, Song YH, and Imaizumi T. Photoperiodic flowering regulation in Arabidopsis thaliana. Adv Bot Res 2014 72: 1-28.

Gottschalk C, Xie Y, Flores M, Rogers S, Guan Q, and van Nocker S. Gene expression atlas of development in a wild apple species. *in prep*

Gottschalk C, Zhang S, Schwallier P, Rogers S, Bukovac MJ, and van Nocker S. Genetic mechanisms associated with the repressive effect of fruit on floral initiation in apple (Malus × domestica Borkh). *Submitted* to PLoS One

Guitton B, Kelner JJ, Celton JM, Sabau X, Renou JP, Chagne D and Costes E. Analysis of transcripts differentially expressed between fruited and deflowered 'Gala' adult trees: a contribution to biennial bearing understanding in apple. BMC Plant Bio 2016 16:55.

Haberman A, Ackerman M, Crane O, Kelner JJ, Costes E and Samach A. Different flowering response to various fruit loads in apple cultivars correlates with degree of transcript reaccumulation of a TFL1-encoding gene. The Plant J 2018 87(2): 161-173.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013 8(8): 1494-1512.

Hättasch C, Flachowsky H, Kapturska D, and Hanke MV. Isolation of flowering genes and seasonal changes in their transcript levels related to flower induction and initiation in apple (*Malus domestica*). Tree Physio 2008 28(10): 1459-1466.

Jung S, Lee T, Cheng CH, Buble K, Zheng P, et al. 15 years of GDR: new data and functionality in the Genome Database for Rosaceae. Nucleic Acids Res 2019 47:D1137–D1145.

Kim D, Langmead B, and Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015 12: 357-360.

Kotoda N, Hayashi H, Suzuki M, Igarashi M, Hatsuyama Y, et al. Molecular Characterization of FLOWERING LOCUS T-Like Genes of Apple (*Malus* × *domestica* Borkh.). Plant Cell Physiol 2010 51(4):561-575.

Mimida N, Ureshino A, Tanka N, Shigeta N, et al. Expression patterns of several floral gene during flower initiation in apical buds of apple (Malus x domestic Borkh.) revealed by in situ hybridization. Plant Cell Rep 2011 30:1485-1492.

Munoz-Fambuena N, Mesejo C, Gonzalez-Max MC, Primo-Millo E, Agusti M and Iglesias DJ. Fruit regulates seasonal expression of flowering genes in alternate-bearing 'Moncada' mandarin. Annals of Botany 2011 108:511-519.

R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www. R-project.org. 2013.

RStudio Team. 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Sawa M, Nusinow DA, Kay SA, and Imaizumi T. FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. Science 2007 318(5848): 261-265.

Searle I, He Y, Turck F, Vincent C, Fornara F, et al. The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. Gene and Development 2006 20: 898-912.

Song YH, Smith RW, To BJ, Millar AJ, and Imaizumi T. FKF1 conveys crucial timing information for constans stabilization in the photoperiodic flowering. Science 2012 336(6084: 1045-1049.

Tränkner C, Lehmann S, Hoenicka H, Hanke MV, Fladung M, et al. Note added in proof to: Over-expression of an FT-homologous gene of apple induces early flowering in annual and perennial plants. Planta 2011 233: 217-218.

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, et al. The genome of the domesticated apple (Malus × domestica Borkh.). Nature Genetics 2010 42:833–839.

Wickham H. ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York. 2016.

Wickham H, Averick M, Bryan J, Chang W, D'Agostino L et al. Welcome to the tidyverse. J Open Source Software 2019 4(43): 1686.

Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, et al. Integration of Spatial and Temporal Information During Floral Induction in Arabidopsis. Science 2005 309(5737): 1056-1059.

Zhang S, Zhang D, Fan S, Du L, Shen Y, Xing L, Li Y, Ma J, and Han M. Effect of exogenous hormones, and flowering-associated genes in 'Fuji' apple (Malus domestica Borkh.) Plant Physio Biochem 2016 107: 178-186.

Ziv D, Zviran T, Zezak O, Samach A and Irihimovitch V. Expression Profiling of FLOWERING LOCUS T-like Gene in Alternate Bearing 'Hass' Avocado Trees Suggests a Role of PaFT in Avocado Flower Induction. PloS One 2014 9(10).

CHAPTER FIVE

Efficacy of foliar-applied GA in controlling flowering

ABSTRACT

The precise control of flowering is a critical aspect of the production management for biennial bearing susceptible apple cultivars. Traditional methods to control flowering have relied on the use of flower and fruit thinning, which are imprecise and can result in undesirable off-target effects. An alternative approach is the use of plant growth regulators (PGRs) to influence the endogenous hormone pathways that control flowering. However, PGRs have failed to become a standard practice due to unpredictability and occasional ineffectiveness in eliciting a desirable response. In this study, I evaluated the use of exogenous foliar-applied gibberellic acid (GA) PGRs in controlling flowering in apple (*Malus* × *domestica* Borkh.) across three consecutive seasons. This study aimed to identify the most effective formulation of GA, timing of application, application rate, and cultivar-specificity. Two commercial available bioactive GAs, G_3 and GA_{4+7}, were evaluated for efficacy in five different apple cultivars. GA_3 exhibited the most promising results with applications in two of the three years resulting in the repression of flowering in the biennial cultivar Honeycrisp. In the one year that did not result in repression of flowering, a promotive effect was observed. The repressive responses to GA_3 were most effective when applied to low fruit load trees between 27 and 57 DAFB, which overlaps with the anticipated period of floral initiation in apple. The strongest repressive response was associated with the highest concentration of GA_3 evaluated here and yielded a reduction of 50% in the number of flowering spurs in the subsequent season. Additionally, the fruit harvest from the high concentration treated trees during the subsequent season exhibited earlier maturity and higher overall fruit quality without sacrificing the average yield per tree. GA_3 application during the summer in low-bearing years offers a potential PGR strategy to provide more consistent production in Honeycrisp with minimal loss in the quality of harvested fruit.

INTRODUCTION

Tree fruit growers have long sought an effective approach to manipulate flowering for the purpose of managing crop production and quality. The control of flowering is particularly important for many cultivars of apple (*Malus* × *domestica* Borkh.) that are prone to biennial (or alternate) bearing. In biennial cultivars, over- or under-cropping can drastically affect the amount of return bloom and the resulting crop load in the subsequent year (Jonkers 1979, Monselise and Goldschmidt 1982). Furthermore, many apple-growing regions are subject to blossom-killing spring frosts which can further amplify biennial bearing cycles.

Commercial apple producers typically manage biennial bearing through adjustments to the current year's crop load by manual or chemical thinning of flowers or fruit, prior to or during the early anticipated period of floral initiation during the current season (Dennis 2000). Although thinning increases return bloom and fruit size, when applied at appropriate times, those effects are not always associated with an increase in crop value (Davis et al. 2004). For example, chemical thinners can result in undesirable off-target effects such as foliar phytotoxicity and fruit russeting (Kon et al. 2018). The insecurities associated with chemical thinning practices have created a need to explore alternative management options. A promising alternative is the use of summer applications of plant growth regulator (PGR) to influence the endogenous mechanism(s) governing floral initiation.

The effects of exogenous application of PGRs on flowering in fruit trees has been the subject of numerous previous investigations (reviewed by Miller 1988 and Ionescu et al. 2017). Gibberellic acid (GA), in particular, has shown promise for inhibiting flowering when applied during the period of floral initiation in low bearing years. In apple, this anticipated period of floral initiation begins shortly after bloom, when vegetative meristems are competent to

transition to a reproductive meristem (Foster et al. 2003). The cessation of this competency period occurs roughly 100 days after full bloom (DAFB) (Foster et al. 2003). Guttridge (1962), was the first to report that the application of GA in low concentrations of 10 to 50 parts per million (ppm) could inhibit flowering in apple cultivars up to 40% compared to controls when applied during this period. However, the reported response was not consistent across all six cultivars evaluated (Guttridge 1962). Dennis and Edgerton (1966) evaluated the efficacy of GA_3 vs GA_{4+7} and found GA_{4+7} to be more effective at inhibiting flowering. This result further supported by work by Tromp (1982), who also reported that applications made at full bloom were effective at inhibiting flowering using GA_3, GA_7, or GA_{4+7}. Tromp (1982) also reported that GA_{4+7} was the most effective type of GA when applied at full bloom or later. Other reports have found GA_3 to be effective at inhibiting flowering when applied after full bloom, which contradicts some findings reported by Tromp (1982) (Luckwill and Silva 1979, Bertelsen and Tustin 2002). Many other previous reports have found GA to be effective at inhibiting flowering at various application times and concentrations (Bertelsen et al. 2002, McArtney 1994, McArtney and Li 1998, McArtney et al. 2007, Schmidt et al. 2009, Schmidt et al. 2010, Tromp 1987, Zhang et al. 2016). However, many of these studies on GA's efficacy have been undermined by inconsistent responses between cultivars, timings, and application rates.

Tromp (1987) evaluated four different bioactive GAs for effects on flowering. They reported that GA_3 and GA_4 did not influence flowering, a result that is contradictory to much of the previous reported effects. Moreover, Looney et al. (1985) demonstrated that GA_4 can increase flowering the subsequent year under particular circumstances. These inconsistent and contrasting responses to GA applications have limited its viability as a commercial approach to controlling flowering. In addition, many of the early works involving GA utilized applications at

full bloom, which can lead to the development of parthenocarpic fruit (Bukovac 1963, Galimba et al. 2019). Parthenocarpy can lead to changes in fruit size and quality (Galimba et al. 2019). Identifying a late-spring or summer application timing, in which GA is still an effective method to controlling flowering, would offset these concerns with bloom time applications.

In the following study, the efficacy of foliar-applied GAs were tested for controlling flowering in Honeycrisp and other high-value apple cultivars. This study aimed is to document the efficacy of early summer time applications, different GAs and application concentrations, inter-seasonal responses, and cultivar-specificity. In addition, the evaluation of fruit yield, maturity, and quality were assessed in trees that exhibited a significant response to GA.

METHODS

2016 PGR Experiments

All experiments were conducted at the Michigan State University Clarksville Research Center (CRC) (Field: 42°52'24"N, 85°16'05"W – Station: 42°52'24"N, 85°15'30"W) in Clarksville, Michigan. Trees were grown as a vertical axis with trellis support. All trees were greater than 10 years old and were randomly assigned to treatment dates or as controls. All PGR applications were made using a four-gallon pump-action backpack sprayer (GroundWork, Distributed by Tractor Supply Company, Brentwood, TN) with an application rate of 1 L of solution applied to all foliar surfaces on a per tree basis. After each application, spray tanks were triple rinsed before use in another experiment. Trees selected for the GA treatments and their controls underwent thinning of fruit 30 - 32 days after full bloom (DAFB), once fruit set was established.

For the evaluation of GA_{4+7}, we selected five apple cultivars that are commonly grown for commercial production in Michigan. The five cultivars selected were Red Delicious, Fuji,

Jonagold, Gala, and Honeycrisp. Twenty trees per cultivar, except Jonagold which had 16, were randomly assigned to one of four possible treatment application dates or to serve as controls. Using this design, three or four trees were assigned as replicates per treatment. Application were made at weekly intervals starting at 36 DAFB and continued through 57 DAFB (Table 5.1). For the evaluation of GA_{4+7}, commercially available Provide 10SG (Valent Biosciences, Libertyville, IL) at 188 ppm plus 0.1% Regulaid surfactant (Kalo, Overland Park, KS) was used. For evaluation of GA_3, 40 Honeycrisp trees were selected and randomly assigned to the same treatment dates as the GA_{4+7} experiment. Spray applications of GA_3 were made using commercially available ProGibb 4% (Valent Biosciences, Libertyville, IL) at 200 ppm plus 0.1% Regulaid. All control trees received water plus 0.1% surfactant application on 36 DAFB.

Assessment of the effect on flowering was conducted during bloom (tight cluster stage) of the following season. For each treatment date and PGR, a minimal of three representative trees were selected for flowering spur counting. Within each of the three selected trees, three random branches were selected and the total spurs and flowering spurs were counted on each branch.

2017 PGR Experiments

Field experiments were conducted at two different locations during the 2017 season. The first was at the CRC, in which the same orchards used in 2016 were used again in 2017. The second location was a commercial orchard located in Sparta, MI (Field: 43°06'13"N, 85°41'57"W). Applications performed at the CRC were conducted using the same equipment as described for the 2016 experiments. The commercial orchard applications were carried out using a commercial airblast sprayer (Blueline, Yakima, WA).

ProGibb was evaluated for efficacy at both field locations. At the CRC, four cultivars (Fuji, Gala, Honeycrisp, and Red Delicious) were selected for use, while two cultivars (Honeycrisp and Jonagold) were selected at the commercial orchard. For the CRC trials, 30 trees per cultivar were selected and hand-thinned of fruit prior to PGR application. Trees were grouped into blocks of five replicate trees and assigned to a specific treatment date or as a control. Applications of ProGibb at 400 ppm plus 0.1% Regulaid were made on a ~100 Growing Degree Days base 50 (GDD_{50}) intervals, starting at 400 and continuing until 800 GDD_{50} (Table 5.2). Control trees received an application of a mix of water and 0.1% Regulaid at 400 GDD_{50}. The commercial orchard had received a late spring frost during bloom resulting in an estimated 10 - 20% decrease in viable blossoms. Severely affected frost thinned rows were selected for use in new PGR trials. Two rows of each cultivar were assigned to one treatment date and a third row assigned as a control. ProGibb at 400 ppm plus 0.1% Regulaid applications were made at ~450 and ~550 GDD_{50}. Control trees received no applications.

Flowering data was collected again during bloom (tight cluster stage) in the following spring. At the CRC, every tree used in the experiment had three randomly selected branches used in counting total flowering and vegetative spurs. At the commercial orchard, ten randomly selected trees within the row were used in counting total flowering and vegetative spurs. In a similar fashion as the CRC experiment, three randomly selected branches from each of the ten trees were used.

2018 PGR Experiments

All field experiments were conducted at the CRC during the 2018 season and spray applications were performed using the same equipment as described for the 2016 experiments. Honeycrisp was selected as the focus cultivar for all trials during this season (Table 5.3). A

planting of third leaf trees was selected for evaluation of ProGibb at 400 ppm plus 0.1% Regulaid. Trees were planted in a high-density system consisting of 84 trees per row split into three blocks of 28 trees. All trees were thinned of flowers or fruit within two weeks after bloom. ProGibb was applied to three replicate blocks of five trees each per treatment date. PGR applications were applied at two-week intervals starting at 27 DAFB and continuing until 84 DAFB. Control trees received an application of water plus 0.1% Regulaid at 27 DAFB. In addition, ProGibb at 200 ppm and Provide at 100, 200, and 500 ppm plus 0.1% Regulaid were applied at 29 DAFB as a single timed treatment to three blocks of five trees each. One replication of the Provide at 200 ppm treatment was located in a separate row of thinned Honeycrisp trees used previously in 2016/17 experiments. For the Provide 500 ppm experiment, four floral/fruit thinned trees were assigned per treatment replicate. Application dates were the same as the previously discussed ProGibb 400 ppm experiment. The evaluation of flowering was carried out in the same fashion as described for the 2017 experiments. An additional assessment of visual bloom was determined by a trained profession using a rating scale of 1 (absent of flowers) to 10 (every spur is flowering).

Fruit Yield and Quality Experiments

For yield measurements, all fruit was harvested from each replicate tree from each of three treatment blocks during the commercial harvest date for Honeycrisp (September 15th, 2019). Harvested fruit was weighed (Model 4010 Scale Pelouze, Bridgeview, IL) and the yield recorded. Trunk diameter measurements were recorded for each tree in the experiment with a digital caliper (DKC-8050 GlowGeek, Guangdong, China) and recorded. For fruit quality, three representative fruit were collected from five replicate trees from each of the three treatment blocks. The fruit samples were then subjected to commercial quality assessment at Michigan

State University's Ridge Apple Quality Lab to be processed using a commercial apple sorting machine (Compac Spectrum System, Auckland, NZ). Firmness was measured using a penetrometer (QA Supplies, Norfolk, VA) mounted onto a moveable stand. Brix measurements were taken with a pocket brix-acidity meter (PAL-BX|ACID5, ATAGO, Tokyo, Japan).

Statistics

All statistics were performed using R (v.3.6.2) (R Core Team 2017) within the R Studio suite (v.1.1.423) (RStudio Team 2015). Plots were generated using the ggplots2 (Wickham 2016) and ggpubr packages.

RESULTS

Flowering response

In 2016, ProGibb (GA_3, 200 ppm) and Provide (GA_{4+7}, 188 ppm) were evaluated for their effectiveness at specific summer application timings that correspond to the anticipated period of floral initiation on fruit thinned trees. ProGibb was found to exhibit a repressive effect on return bloom in Honeycrisp, with effectiveness increasing at progressively later treatment dates (Figure 5.1A). The inhibition of flowering reached a statistically significant threshold when GA_3 was applied at 50 and 57 DAFB, with the later date having a stronger effect. However, the 57 DAFB GA_3 treatment exhibited greater variability in the flowering response. GA_{4+7} also exhibited a inhibitory effect on flowering in Honeycrisp but only when applied at 57 DAFB (Fig 5.1B). However, this inhibitory response was not statistically significant. For the four other high-value apple cultivars, the only significant effect on flowering in response to GA_{4+7} was observed in Red Delicious (Fig 5.1C-F). Here, GA_{4+7} promoted a return bloom when applied at 36 DAFB (Figure 5.1D).

Based on the significant effect observed with GA_3 applied in 2016, its evaluation was expanded during the 2017 season to include three additional cultivars. Additionally, a second location was used to evaluate an on-farm approach to the application of GA_3. GA_3 was found to have no inhibitory effect on flowering in any of the cultivars used at the CRC (Fig 5.2A-D). In contrast, Honeycrisp, Fuji, and Red Delicious exhibited an increase in flowering in response to the application of GA_3 (Fig 5.2A-C). Honeycrisp exhibited the strongest promotive response, with an increase in flowering of ~23% over the control at the final treatment date of 53 DAFB (Fig 5.2A). An increase in flowering of 15% and 12% were observed in 'Fuji' and 'Red Delicious', respectively (Fig 5.2B-C). 'Gala' exhibited no response to the application of GA_3 (Fig 5.2D). GA_3 was applied at the commercial orchard to full rows of frost-thinned 'Honeycrisp' and 'Jonagold' trees. A slight decrease in return bloom at the earlier application date was observed in Honeycrisp but the response did not to a significant threshold (Fig 5.2E). In contrast, 'Jonagold' exhibited a significant decrease in flowering at both treatment dates (Fig 5.2F).

In 2018, GA_3 and GA_{4+7} were reevaluated for effects on flowering in Honeycrisp due to the inconsistent results observed during the prior two seasons. For each GA, the number of application concentrations and timings were increased (Table 5.3). GA_3, when applied at 400 ppm, was found to be significantly effective at inhibiting flowering when applied at 27 DAFB, while also inhibiting flowering at 39 and 56 DAFB, but not significantly (Fig 5.3A-B). This result was consistent using both a quantitative count of flowering spurs per branch and a qualitative visual rating by a trained professional (Fig 5.3A-B**)**. The 27 DAFB treatment resulted in the inhibition of flowering at greater than 50% in the subsequent season compared to the control (Fig 5.3A). All of the GA_{4+7} treatments and the GA_3 200 and 300 ppm treatments were

found to be ineffective at inhibiting flowering in the spring of 2019 (Fig 5.3C-F). However, the GA_{4+7} 500 ppm treatment significantly increased flowering when applied at 39 and 56 DAFB (Fig 5.3C). Those promotive effects were less than a 20% increase over the controls.

Effects of GA on fruit maturity, size, and quality

In response to the flowering inhibitory effect of GA_3 in the 2018 trials, fruit maturity and quality was evaluated during the subsequent season's harvest. The GA_3 treatment that resulted in inhibited flowering were found to not have an effect on total yield per tree (Fig 5.4A). However, when the yield was normalized to trunk diameter (yield efficiency) a significant decrease was observed in response to the GA_3 application at 27 DAFB (Fig 5.4B). A decrease in the firmness was also observed with the treated trees (Fig 5.4C). Juice extracted from the fruit of the treated trees was found to have a higher sugar content (Fig 5.4D).

To evaluate if the GA_3 treatment resulted in larger, high-quality Honeycrisp, a commercial fruit sorting machine was used to quantify fruit-quality metrics. Firstly, the GA_3 treatment resulted in a significant increase in fruit size relative to the control (Fig 5.5A-C). The mass, length, and width were all significantly higher in the fruit of the treated trees regardless of application date. A significant increase in the percent red + pink peel coloration was also observed in the fruit of the treated trees (Fig 5.5D). Moreover, a higher percentage of fruit that met a "high-quality" and "fresh-market grade", as determined by a commercial sorting machine algorithm, were associated with the GA_3 treatments (Fig 5.5E-F).

Flowering in Honeycrisp two years after GA_3 treatment

In the spring of 2020, the trees used in the 2018 experiment were evaluated for variations in flowering following a year of no additional treatment or crop load management. 'Honeycrisp' is known to be an extremely biennial cultivar requiring extensive crop load management

(Embree et al. 2007). Thus, it was anticipated that the reduced flowering in 2019 of the treated trees would result in greater flowering in 2020. The trees treated in 2018 with GA_3 at 27 DAFB exhibited a return bloom in the spring of 2020 (Fig 5.6). The 27 DAFB treatment exhibited a statistically significant average >25% flowering spurs compared to the control, which was relatively absent of bloom (Fig 5.6A). The average flowering rating for the 27 DAFB was also statistically significant, with an average rating greater than six (Fig 5.6B).

DISCUSSION

In general, the results presented in this study reaffirmed many of the previously reported variability in the response to the application of GAs. Both promotive and inhibitory effects on flowering were found in response to foliar GA application, regardless of the GA type, timing of application, and cultivar. However, the most promising and consistent results was the use of GA_3 to inhibit flowering in the cultivar Honeycrisp. Here, an inhibition of flowering was observed in two out of the three years. Although a promising result, the most effective timing of application of GA_3 varied from year-to-year. This result underscores the need for more long-term studies on timing of application.

The results from the GA_3 application in 2018 was the most significant inhibitory effect observed. It was also found that the 27 DAFB GA_3 treatment did not significantly decrease total yield in the following season. However, there was a significant reduction in yield efficiency measurement, showing that there was less fruit per cm^2 of the trunk cross-sectional area of the treated trees. The treated trees were also found to produce larger fruit, which could explain the finding that total yield was similar, but the yield efficiency differed. In general, the treated fruit were found to be of higher quality compared to the controls. Taken together, these results

suggest GA_3 could offer a useful practice to controlling flowering in the biennial cultivar Honeycrisp when applied at 400 ppm at 27 DAFB.

To evaluate the potential of adaption of a GA_3 management strategy over chemical thinning the economics need to be compared. Applying GA_3 at a concentration of 400 ppm at a rate of .67L/tree will cost $420/acre for a moderately high-density orchard of 1,210 trees/acre (12’ x 3’ spacing). At this concentration, a grower can expect to obtain between 70 and 80% fresh-market quality graded fruit ($15 - 21/bushel). In comparison, using a blossom thinner will cost $40/acre per application with two to three applications expected, for a total cost of $80 - $120/acre for the same sized orchard. Using chemical thinners and manual fruit thinning, a grower can anticipate obtaining an average of 80% fresh-market quality graded fruit. The GA_3 approach is about 3.5 times more costly than traditional thinning. However, the additional cost of manual or chemical fruit thinning to establish appropriate crop loads could raise the costs of production closer to that of the GA_3 approach.

Table 5.1. Commercial GA products tested in 2016 for effectiveness at manipulating floral initiation.

Product Name (GA type)	GA Concentration	Application Timing (DAFB)	2016 Visual Bloom Density
ProGibb (GA_3)	200 ppm	36, 43, 50, 57	Low
Provide (GA_{4+7})	188 ppm	36, 43, 50, 57	Low

Table 5.2. Application dates during the 2017 season, the target GDD_{50} measurement, and the actual GDD_{50} measurement.

	Application Dates				
Application Date	5/27/17	6/06/17	6/12/17	6/16/17	6/20/17
GDD_{50} Target	400	500	600	700	800
GDD_{50} Measured	386.8	507.3	618.2	750.4	806.9
DAFB	29	39	45	49	53

Table 5.3. 2018 PGR application trial design.

PGR Product	GA Concentration	Application Interval	Application Timings	Number of Replicate Trees/Treatment Block	Replicate Blocks/Treatment Date	Total Trees
ProGibb	400 ppm	2 weeks	4 WAFB – 12 WAFB	5	3	75
	300 ppm	single application	4 WAFB	6	3	18
	200 ppm	single application	6 WAFB	5	3	15
Provide	500 ppm	2 weeks	4 WAFB – 12 WAFB	4	3	60
	200 ppm	single application	6 WAFB	5	3	15
	100 ppm	single application	6 WAFB	5	3	15
Control Groups	**Control Group and Associated Treatments**			**Number of Replicate Trees/Treatment Block**	**Replicate Blocks/Treatment Date**	**Total Trees**
Control 1	ProGibb 400			5	3	15
	ProGibb 200					
	Provide 200					
	Provide 100					
Control 2	ProGibb 300			6	3	18
Control 3	Provide 500			4	3	12

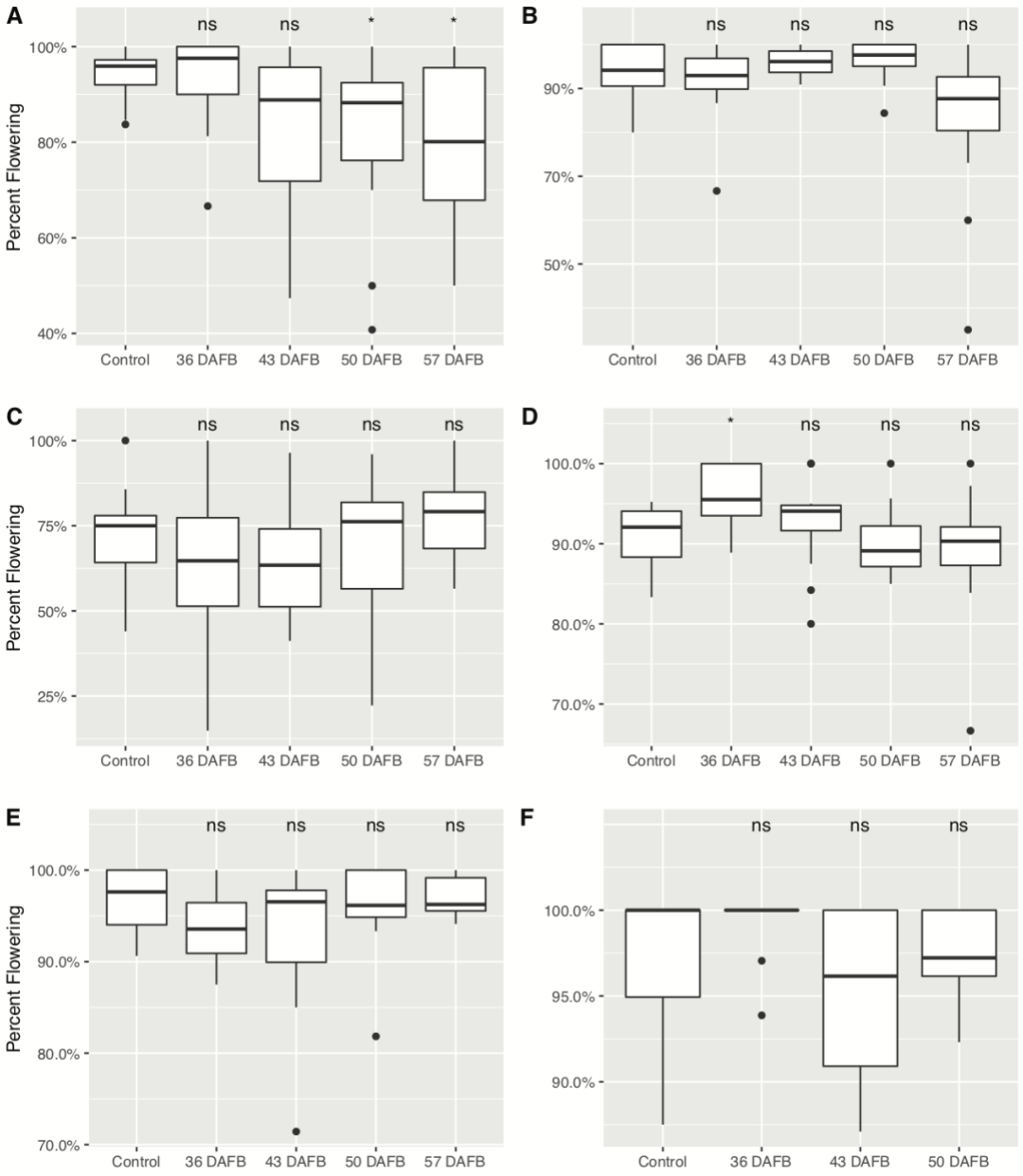


Figure 5.1. The percent flowering spurs in 2017 following application of different GAs to fruit thinned trees in 2016. A) GA_3 applied to Honeycrisp. B) GA_{4+7} applied to Honeycrisp. C) GA_{4+7} applied to Fuji. D) GA_{4+7} applied to Red Delicious. E) GA_{4+7} applied to Gala. F) GA_{4+7} applied to Jonagold. Black circles represent outliers in the dataset. Student's *t*-test , *P < 0.05, ns = not significant

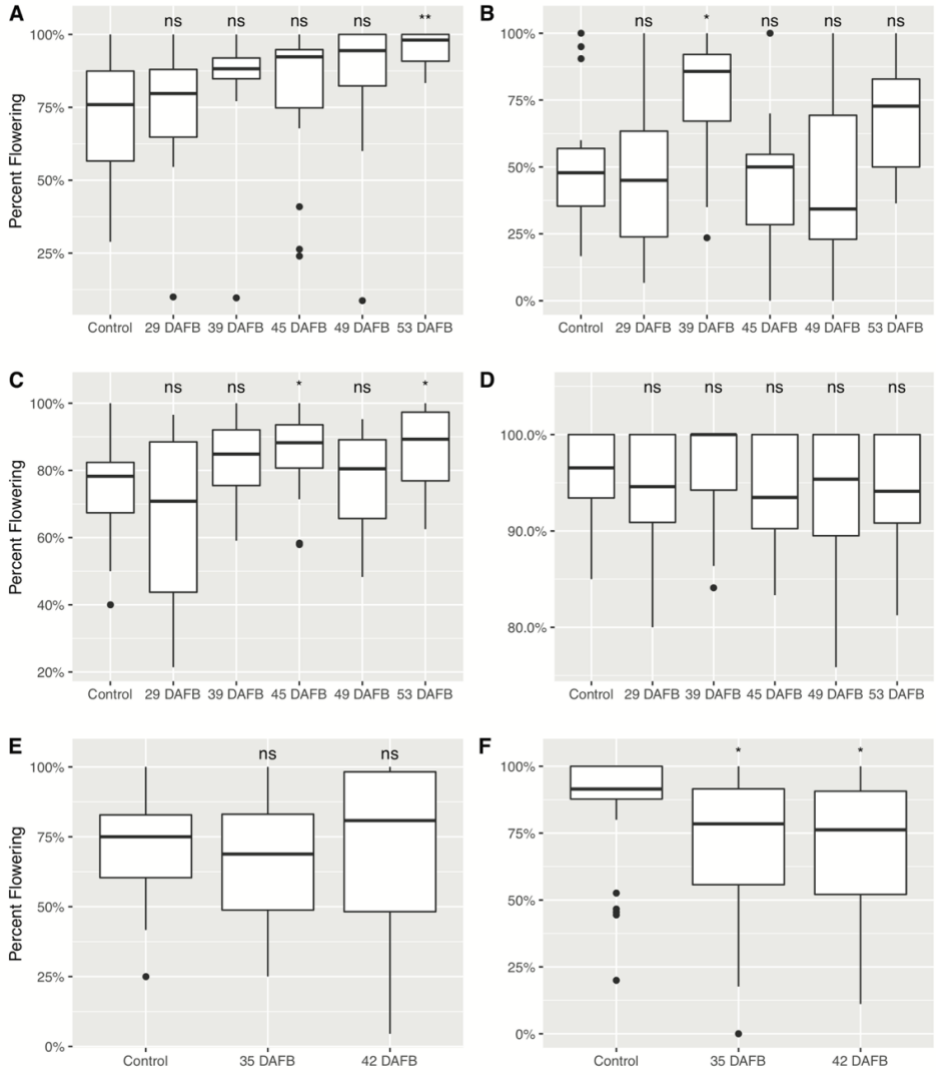


Figure 5.2. The percent flowering spurs in 2018 following application of GA_3 to thinned trees in 2017. A) GA_3 applied to fruit-thinned Honeycrisp. B) GA_3 applied to fruit-thinned Fuji. C) GA_3 applied to fruit-thinned Red Delicious. D) GA_3 applied to fruit-thinned Gala. E) GA_3 applied to frost-thinned Honeycrisp in a commercial orchard. F) GA_3 applied to frost-thinned Jonagold in a commercial orchard. Black circles represent outliers in the dataset. Student's *t*-test , **P < 0.01, *P < 0.05, ns = not significant

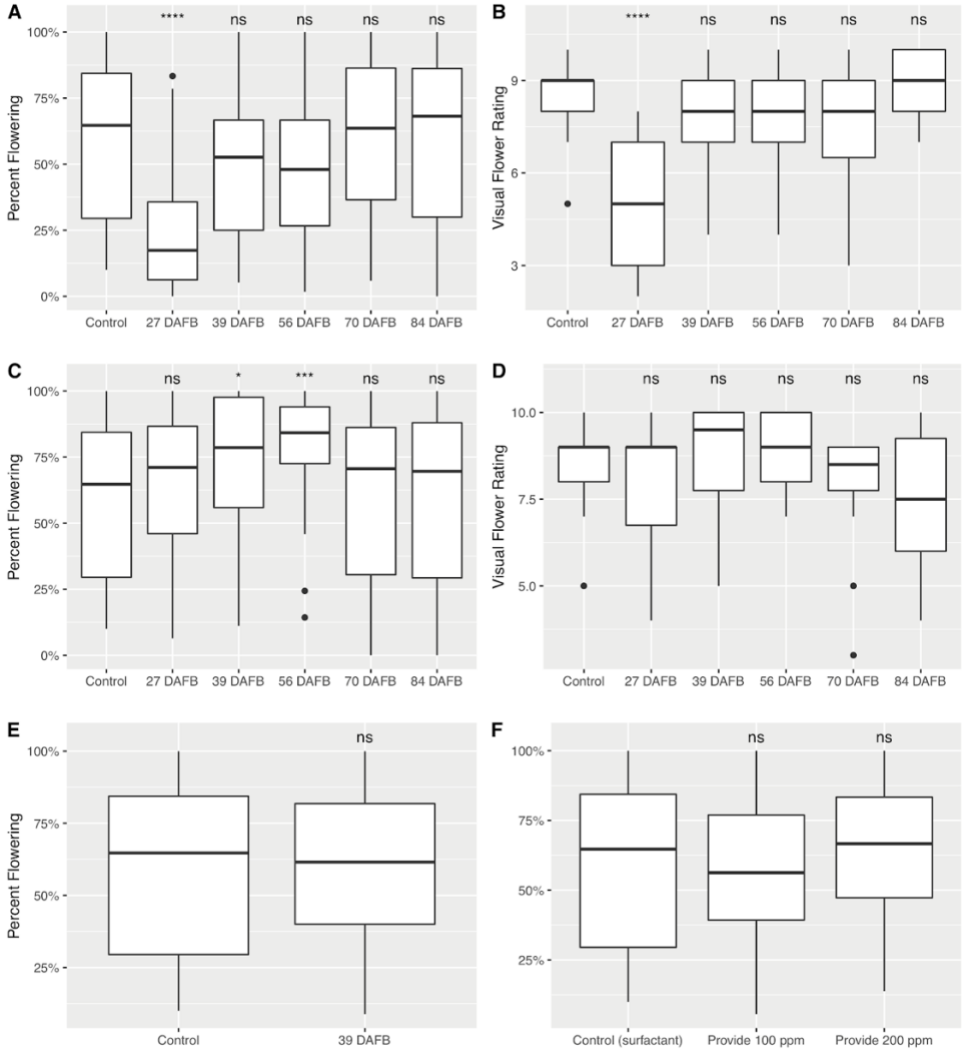


Figure 5.3. Flowering response in 2019 following application of different GAs to floral thinned Honeycrisp trees in 2018. A) Percent flowering spurs following application of GA_3 400 ppm. B) Visual flower rating following application of GA_3 400 ppm. C) Percent flowering spurs following application of GA_{4+7} 500 ppm. D) Visual flower rating following application of GA_{4+7} 500 ppm. E) Percent flowering spurs following application of GA_3 200 ppm F) Percent flowering spurs following application of different GA_{4+7} treatments 39 DAFB. Black circles represent outliers in the dataset. Student's *t*-test, **** P<= 0.0001, ***P < 0.001, *P < 0.05, ns = not significant

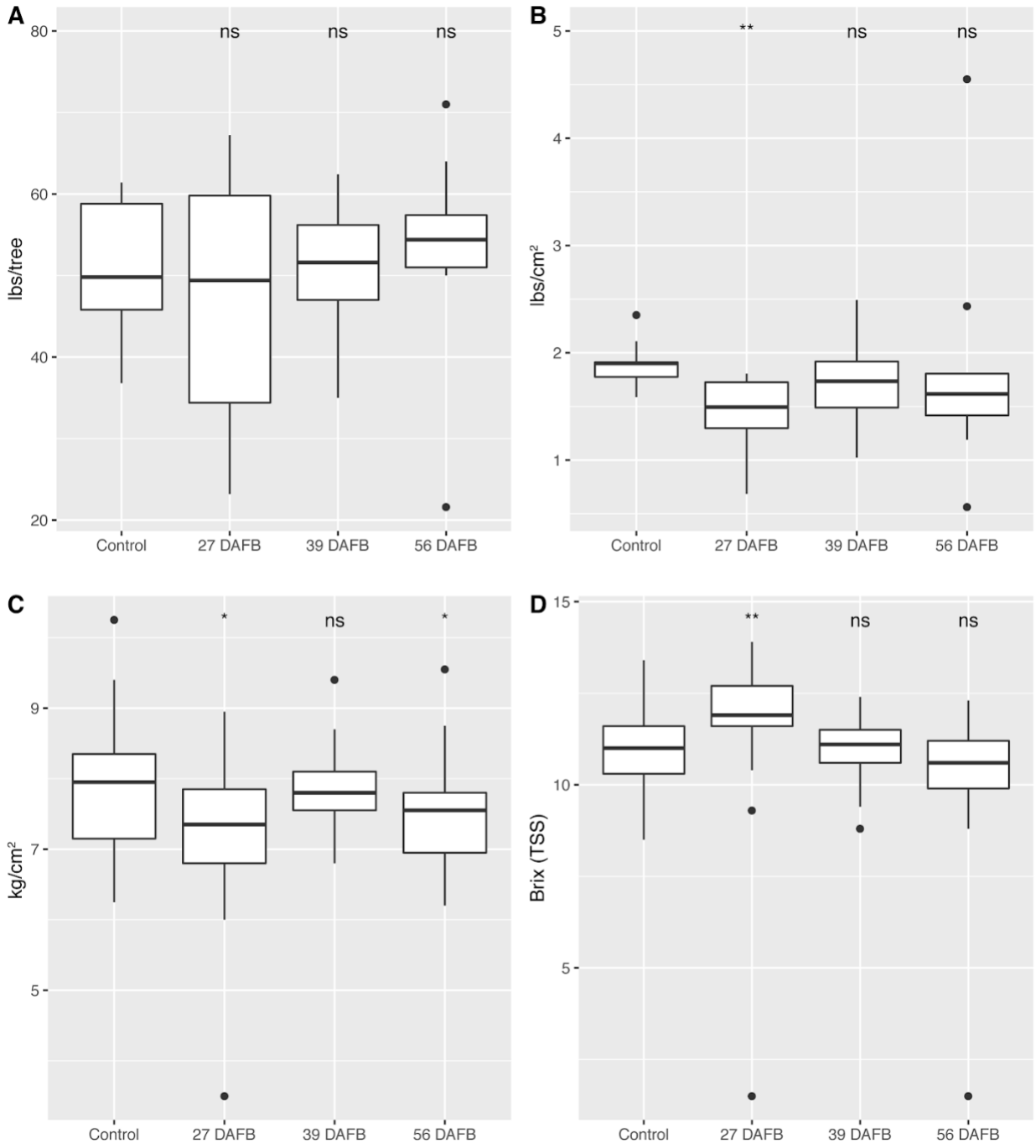


Figure 5.4. Yield and maturity metrics from the 2019 harvest following application of GA_3 to floral thinned Honeycrisp trees in 2018. A) Yield per tree. B) Yield per trunk cross sectional area. C) Fruit penetration force upon harvest. D) Sugar content. Black circles represent outliers in the dataset. Student's *t*-test, **P < 0.01, *P < 0.05, ns = not significant

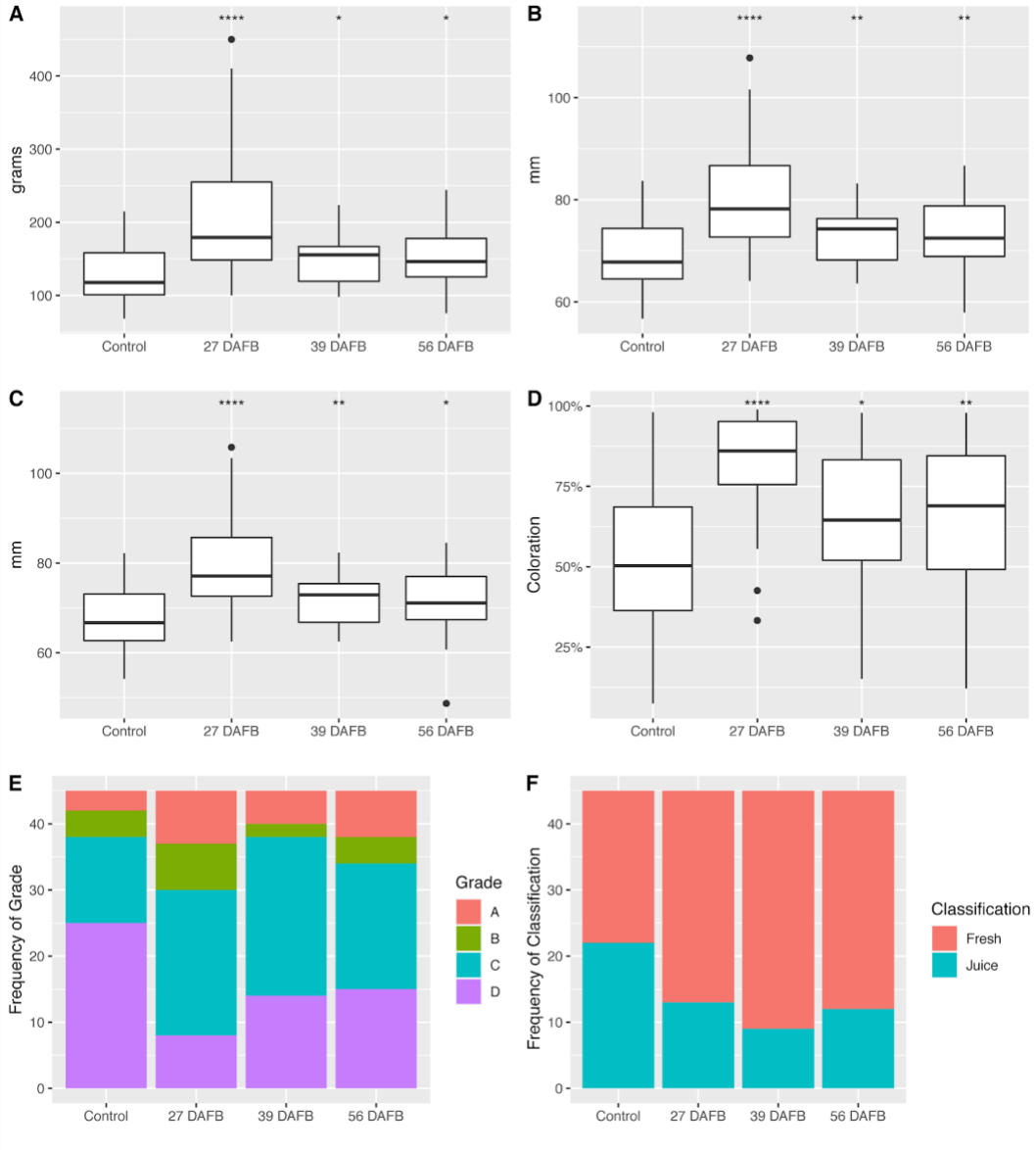


Figure 5.5. Fruit quality measurements from the 2019 harvest following application of GA_3 to floral thinned Honeycrisp trees in 2018. A) Mass. B) Major diameter. C) Minor diameter. D) Percent of skin coloration of Red + Pink. E) Frequency of fruit grades. F) Frequency of use classification. Black circles represent outliers in the dataset. Student's *t*-test, **** P<= 0.0001, ** P<= 0.01, * P< 0.05, ns = not significant

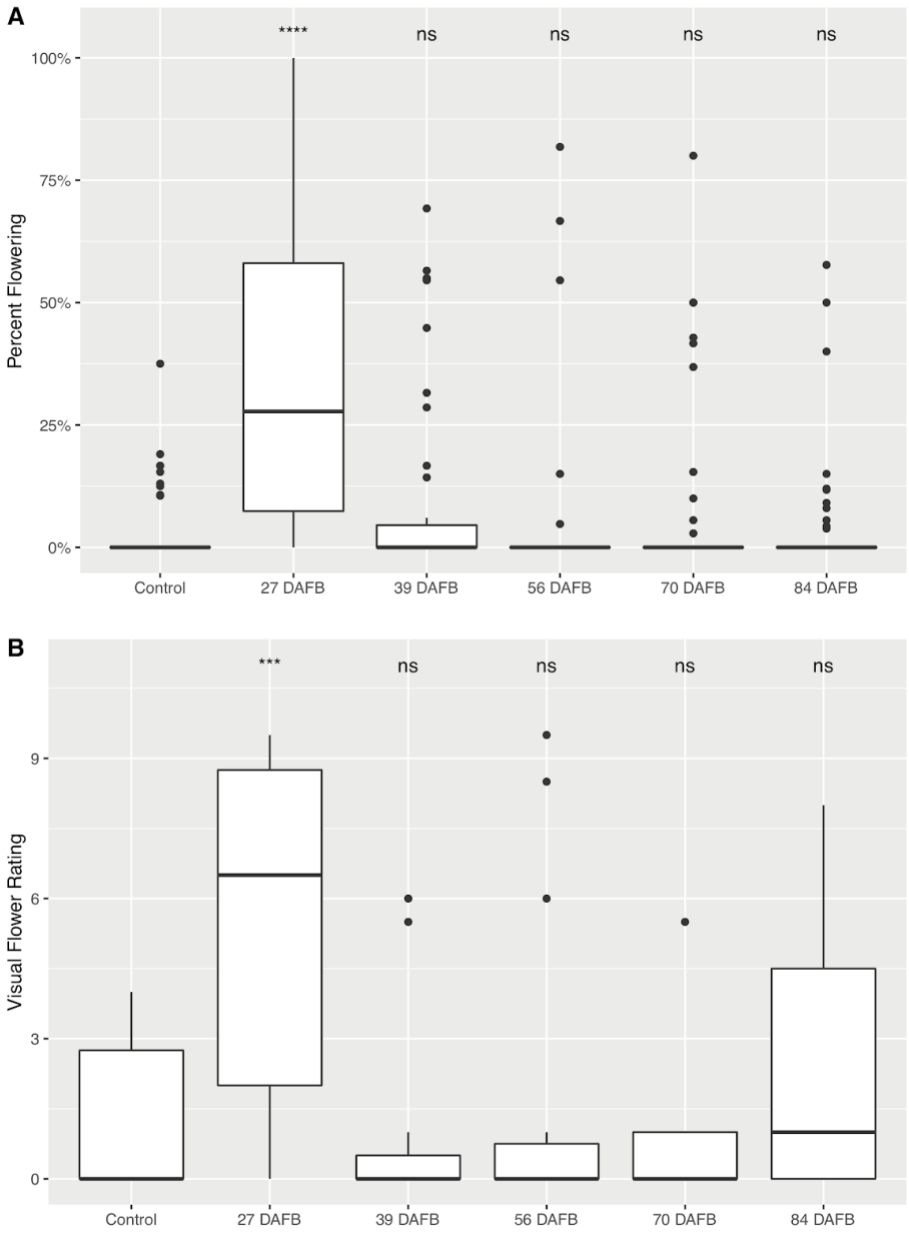


Figure 5.6. Flowering response in 2020 following application of GA_3 to floral thinned Honeycrisp trees in 2018. A) Percent flowering spurs following application of GA_3 400 ppm. B) Visual flower rating following application of GA_3 400 ppm. Black circles represent outliers in the dataset. Student's *t*-test, **** P<= 0.0001, *** P< 0.001, ns = not significant

REFERENCES

REFERENCES

Bertelsen MG and Tustin DS. 2002. Suppression of flower bud formation in light cropping trees of 'Pacific Rose' apple using gibberellin sprays. J. Hort. Sci. and Biotech. 77(6):753-757.

Bertelsen MG, Tustin DS, and Waagepetersen RP. 2002. Effects of GA_3 and GA_{4+7} on early bud development of apple. J. Hort. Sci. and Biotech. 77(1):83-90.

Bukovac, MJ. 1963. Induction of parthenocarpic growth of apple fruits with gibberellins A_3 and A_4. Botanical Gazzette 124(3): 191-195.

Bukovac MJ, Sabbatini P, and Schwallier PG. 2006. Modifying Alternate Bearing of Spur-Type 'Delicious' Apple with Ethephon. HortScience 41(7):1606-1611.

Davis K, Stover E, and Wirth F. 2004. Economics of Fruit Thinning: A Review Focusing on Apple and Citrus HortTechnology 14(2):282-289.

Dennis FG. 2000. The history of fruit thinning. Plant Growth Regulation 31:1-16.

Dennis FG, and LJ Edgerton. 1966. Effects of gibberellins and ringing upon apple fruit development and floral bud formation. Proc Amer Soci Hort Sci 88: 14-24.

Embree CG, Myra MD, Nichols DS, and Wright A. 2007. Effect of Blossom Density and Crop Load on Growth, Fruit Quality, and Return Bloom in 'Honeycrisp' Apple. HortScience 42(7): 1622-1625.

Foster T, Johnston R, and Seleznyova A. 2003. A morphological and quantitative characterization of early floral development in apple (*Malus* × *domestica* Borkh.). Annals of Botany 92(2): 199-206.

Galimba KD, Bullock DG, Dardick C, Liu Z, and Callahan AM. 2019. Gibberellic acid induced parthenocarpic 'Honeycrisp' apples (*Malus domestica*) exhibit reduced ovary width and lower acidity. Horticulture Research 6:41.

Guttridge CG. 1962. Inhibition of Fruit-Bud Formation in Apple with Gibberellic Acid. Nature 196: 1008.

Ionescu IA, Møller BL, and Sánchez-Pérez R. 2017. Chemical control of flowering time. J. Exp. Bot. 68(3):369-382.

Jonkers H. 1979. Biennial Bearing in Apple and Pear: A literature survey. Scientia Horticulturae 11:303-317.

Kon TM, Schupp JR, Yoder KS, Combs LD, and Schupp MA. 2018. Comparison of chemical blossom thinners using 'Golden Delicious' and 'Gala' pollen tube growth models as timing aids. HortScience 53(8): 1143-1151.

Luckwill LC and Silva JM. 1979. The effects of daminozide and gibberellic acid on flowering initiation, growth and fruiting of apple cv. Golden Delicious. J. Horticultural Sci. 54(3):217-223.

Looney NE, Pharis RP, and Noma M. 1985. Promotion of flowering in apple trees with gibberellin A_4 and C-3 epi-gibberellin A_4. Planta 165:292-294.

McArtney SJ. 1994. Exogenous gibberellin affects biennial bearing and the fruit shape of 'Braeburn' apple. New Zealand Journal of Crop and Horticulture Science 22:343-346.

McArtney SJ and Li S. 1998. Selective Inhibition of Flowering on 'Braeburn' Apple Trees with Gibberellins. HortScience 33(4):699-700.

McArtney S, Unrath D, Obermiller JD and Green A. 2007. Naphthaleneacetic Acid, Ethephon, and Gibberellin A4 + A7 Have Variable Effects on Flesh Firmness and Return Bloom of Apple. HortTechnology 17(1):32-38.

Miller SS. 1988. Plant bioregulators in apple and pear culture. Hort. Rev. 10:309-401.

Monselise S and Goldschmidt E. 1982. Alternate bearing in fruit trees. Horticultural Reviews 4:128-173.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

RStudio Team. 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Schmidt T, Elfving DC, Mcferson JR and Whiting MD. 2009. Crop Load Overwhelms Effects of Gibberellic Acid and Ethephon on Floral Initiation in Apple. HortScience 44(7):1900-1906.

Schmidt T, McFerson J, Elfving DC, and Whiting M. 2010. Practical Gibberellic Acid Programs for Mitigation of Biennial Bearing in Apple. Proc. XIth IS on Plant Bioregulators in Fruit Production (Ed. Costa G) Acta Hort. 884:663-670.

Tromp J. 1982. Flower-bud formation in apple as affected by various gibberellins. J. Horticultural Sci. 57(3):277-282.

Tromp J. 1987. Growth and flower-bud formation in apple as affected by paclobutrazol, daminozide, and tree orientation in combination with various gibberellins J. Horticultural Sci. 62(4):433-440.

Wickham H. 2016. ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York.

Zhang S, Zhang D, Fan S, Du L, Shen Y, Xing L, Li Y, Ma J, and Han M. 2016. Effects of exogenous GA_3 and its inhibitor paclobutrazol on floral formation, endogenous hormones, and flowering-associated genes in 'Fuji' apple (Malus domestica Borkh.). Plant Physiol Biochem. 107:178-186.

GENERAL CONCLUSIONS

This body of work provides a deeper understanding of the molecular mechanisms that underlie floral initiation in apple (*Malus* × *domestica* Borhk.) and how crop load interacts with that pathway to inhibit flowering. The second chapter begins with a census to identify homologs of flowering-related genes in apple. I used sequence homology, phylogenetic relationship, and syntenic organization to identify apple homologs of highly-studied flowering-related gene families. These gene families included *AP1, AGL24/SVP, FT/TFL1, FUL, LFY, SOC1,* and the *SPLs.* The identified apple homologs were then measured for transcriptional changes in the spur apex over the period of floral initiation in non-bearing, flowering-induced trees. Several of the homologs exhibited increases in expression over the sampling period, providing evidence of a presumed role in floral initiation. I also investigated how the presence of a crop load would impact the expression of these genes over the same period. A homolog of *TFL1, MdTFL1-2,* exhibited a strong and persistent upregulation in the apices of bearing trees compared to the non-bearing. Further investigations into genes that were co-expressed with *MdTFL1-2* identified potential upstream promoters or downstream targets. These included homologs of *FIE, APL, EFM, SPL3,* and *ATH1.* A final major finding from this chapter was that genes involved in gibberellic acid (GA) biosynthesis and deactivation were differentially expressed between the two crop load conditions. In particular, four *GA2ox* genes that function to deactivate bioactive GAs were upregulated by the presence of a crop load. In contrast, several of the *GA20ox* genes that are involved in GA biosynthesis were down-regulated by a crop load. The results presented in this chapter suggest that *MdTFL1-2* is associated with the repression of flowering in response to crop load. Additionally, the differential regulation of genes involved with the GA pathway could indicate a strong GA signal at the spur apex of bearing trees.

The third chapter of this dissertation explored the molecular diversity of floral initiation and development in apple species and cultivars. Previous research had identified a wide variation in spring bloom times of accessions maintained in the USDA *Malus* collection. Although the dormancy requirements of apple are known to play an important role in dictating spring bloom time, little attention has been given to documenting other factors that could contribute to the variations in spring bloom time. Six accessions from the USDA collection that represent extreme early- and late-blooming phenotypes were selected for transcriptional profiling at seasonal timings associated with floral initiation and development. In addition to the transcriptional profiling, a second census of flowering-related genes was conducted using the high-resolution transcriptomes of these *Malus* accessions. The floral initiation genes were found to be highly conserved across the genus, whereas the floral development genes were less conserved. The transcriptional profiling identified *MdTFL1* genes were differentially expressed among five of the six accessions during the presumed period of floral initiation. This differentially regulation of *MdTFL1* suggests a possible role in the onset of floral initiation in addition to its role in repressing flowering in response to a crop load. During the presumed period of floral development, the transcriptomes of the six accessions were relatively similar. Of note, homologs of *SEP1/2* and *SEP3* were differentially expressed in one accession that was found to be underdeveloped compared to three of the more developmentally advanced accessions. Although differences in expression were identified, these differences were not associated with the differences in spring bloom times. Based on these results, it can be concluded floral initiation and development do not contribute to the variation observed in spring bloom time.

The fourth chapter of this dissertation investigated the identity of *MdFTs* and characterized their expression. In Arabidopsis, *FT* is diurnally expressed in the leaf tissues and

its translated product is transported to the apex to trigger flowering. I investigated if Arabidopsis *FT* expression was conserved in apple leaves. In other tree fruit species, *FT* was identified to be differentially expressed in the leaves in response to crop load. This result suggests *FT* has a role in the mechanism of biennial bearing. Thus, I also investigated if crop load influences the expression patterning of *MdFT* in leaves. I confirmed that a *de novo* assembled *MdFT2* from the second chapter exists in apple as I was able to assemble corresponding DNA sequences that shared identity. However, I cannot conclusively determine that the *de novo* assembled *MdFT2* is a separate gene or just an allele of *MdFT1.* To identify tissue-specific expression of *MdFT,* I utilized a gene expression atlas constructed from 80+ different tissues and developmental stages of a wild apple species, *Malus fusca*. *MdFT1* was found to be expressed primarily in the apex and rachis tissues. I was unable to identify *MdFT2* tissue-specific expression, as the *de novo* assembled transcript is absent from the expression atlas. This tissue-specificity of expression differs from Arabidopsis as it has little to no expression in the leaves. However, the atlas fails to differentiate different leaf types that make up the shoot architecture of apple. Using probes specific to the *MdFT1* and the *de novo MdFT2,* I measured the expression of the two *MdFTs* in bourse shoot leaves. These specific leaves are necessary for the induction of flowering in apple, suggesting they have a potential role in the flowering pathway. First, I quantified expression over 24 hours to identify if a diurnal expression pattern was conserved under floral inductive conditions. *MdFT1* did exhibit a diurnal expression pattern but it was cultivar dependent. In addition, there was no significant difference in expression of either *MdFT* to floral inductive conditions (absence vs presence of a crop load). The finding that crop load did not affect the expression of either gene was also consistent when investigating for seasonal expression differences. The results from this chapter identified a possible *MdFT2* gene and demonstrated

that *MdFT* expression is not well-conserved between apple and Arabidopsis. Moreover, *MdFT* was not influenced by crop load in the bourse leaves suggesting that the biennial bearing mechanism that is hypothesized for other tree fruit species is not conserved in apple.

The fifth chapter investigated the efficacy of exogenously applied GAs in managing flowering. I demonstrated that Honeycrisp, a high-value and biennial bearing susceptible cultivar, was responsive to applications of GA_3. Two out of three years, GA was effective in repressing the bloom in Honeycrisp when a strong return bloom was expected. As a result of the repression of flowering, the treated trees produced larger, higher-quality, and earlier maturing fruit, without sacrificing total yield. These results suggest that GA offers a promising strategy to control flowering in Honeycrisp. However, the repression of flowering by GA was only observed in two out of the three years in this trial, which illustrates the unpredictability associated with foliar applications of plant growth regulators. When comparing the costs of a GA regiment vs traditional fruit thinning, the GA regiment is about 3.5 times more expensive, which is another contributing factor that limits on its adoption as a commercial practice.

The studies presented in this dissertation highlight the key role that *MdTFL1* has in inhibiting flowering in apple. In particular, the association between a crop load and the high expression of *MdTFL1-2* implicates it as the molecular mechanism causing biennial bearing in apple. In a previous study from our lab group, *MdTFL1-2* was identified as being upregulated in response to exogenous applications of GA. This previous finding and my finding that crop load influences the expression of *GA2ox* genes in the apex suggest that there is a strong GA signal in spur apex of high crop load trees. This strong GA signal could be the driving force behind the high expression of *MdTFL1-2* and the resulting repression of flowering in biennial cultivars. However, the mechanism by which fruit upregulate the possible GA content in the apex is

unknown. Additionally, the work presented in this dissertation does not address the potential competitive relationship between the developing fruit/seeds and the shoot apex for macro- or micronutrients, and/or sugars. Further experimentation is needed to explore how nutrient competition and biennial bearing are associated at the molecular level and if this competition is mutually exclusive from the GA theory. In conclusion, the information I present in this dissertation illuminates an association between *MdTFL1*, GA, and crop load as a molecular mechanism that drives biennial bearing tendency in apple.